Chapter 4

Decomposing the Sales Effect of Promotions

4.1 Introduction

Given the important role of promotions in the marketing mix of European and American firms it is critical to fully understand the effects of promotions. It is known that promotions often result in large sales effects for the promoted brand.¹ For instance, a temporary price cut of 25 percent combined with a feature (mentioning the brand in a retailer’s newspaper advertisement) and a display may generate sales, in an average store, ten times as large as normal (Wittink et al. 1988). However, the key question is whether the sales increase is truly beneficial. A critical condition for our ability to determine the profitability of promotions is that a model of sales incorporates all relevant sources of sales increases. These sources differ in attractiveness to the manufacturer and retailer. For example, at the store level the sales increase for a promoted brand could come from other brands within the same store (brand switching), from the same brand in the same store in other time periods (stockpiling), from the same brand in other stores (store switching), etc. The retailer may not benefit much from brand switching within the store, and both the retailer and the manufacturer derive no benefit from sales borrowed from other time periods (unless, for example, higher inventories increase consumers’ consumption levels). Thus it is critical to decompose the unit sales effect of promotions.

To accomplish this we develop and estimate flexible econometric models based on store-level data. The models allow us to decompose sales increases into constituent sources including cross-brand effects, dynamic effects, and category expansion effects. We argue that models for the estimation of the

¹. Throughout this chapter we use the term “brand” instead of “item” for a product at the Stock Keeping Unit level. This terminology is consistent with extant decomposition research.
effectiveness of promotions should incorporate these effects because their inclusion is necessary for the determination of the net benefit of a unit sales effect. We use a new, flexible econometric model to accommodate the different sources of promotional effects on sales.

This chapter is organized as follows. In section 4.2 we provide a conceptual description of the different ways in which promotions can affect consumers. In section 4.3 we review the literature on the decomposition of unit sales effects at the individual- (household) level, and we argue that managers need to have access to decomposition models of aggregated data. In section 4.4 we identify the contributions of our approach, and in section 4.5 we show how we model the decomposition of the sales effect of promotions at the store level. We present empirical results for one product category in section 4.6, and provide our conclusions in section 4.7.

4.2 Conceptual decomposition


1. Brand switching;
2. Timing acceleration;
3. Quantity acceleration (sources 2 and 3 together are defined as purchase acceleration);
4. Anticipatory responses;
5. Store switching;
6. Deal-to-deal purchasing;
7. Increased consumption; and
8. Category switching.

Each source refers to one or more aspects that determine the difference between a consumer’s purchase behavior under promotional conditions and his or her purchase intention under non-promoted conditions. Brand switching means the consumer purchases a different brand than otherwise. Timing acceleration means purchasing sooner than otherwise, whereas quantity

---

2. Chintagunta (1993) provides an overview of selected empirical studies on three aspects of purchase behavior: category purchase, brand choice, and purchase quantity decisions.
acceleration involves purchasing more than otherwise. Timing- and quantity acceleration are defined together as purchase acceleration. Anticipatory responses are caused by learning effects. Households learn that promotions occur from time to time, and they sometimes defer their purchases until the week in which an item is promoted. Store switching means the purchase of the promoted brand in a store different from the regular outlet. Some households purchase a brand only on deal: deal-to-deal purchasing. In essence, deal-to-deal purchasing is an extreme form of a combined purchase acceleration- and anticipatory behavior, i.e., households restrict their purchases to promotional weeks. Put differently, purchase acceleration behavior and anticipatory responses are imperfect variants of deal-to-deal purchasing. To illustrate, Krishna (1994a) reports that one third of all American households only purchase coffee on promotion. Other households show increased consumption: they purchase accelerate due to promotions, have as a result more inventory at hand, but consume this inventory faster than the normal consumption rate, after which they resume their regular purchase behavior. Wansink (1996) conjectures that this effect shows up especially for promotions that encourage consumers to buy more: three-for-the-price-of-two and bonus packs. Category switching, finally, means that a household buys the brand on promotion, thereby foregoing a substitute product from another category.

We present the conceptual decomposition of promotional effects in the simplest possible manner. However, consumers may behave in a manner that reflects multiple sources. For example, consider a consumer who intends to buy the usual quantity of brand $B$, but in response to a promotion for brand $A$ purchases twice as much, half of which is stockpiled while the other half is consumed faster. In this case we observe a combination of brand switching, quantity acceleration and increased consumption. A second example is a consumer who purchases whichever brand is promoted, and does not purchase any brand off-deal. We then see a combined deal-to-deal- and brand switching purchase behavior. Nevertheless, for ease of understanding, we describe only the eight “main sources” of sales effects. These main sources are classified and described in Table 4.1 based on a hypothetical promotion for brand $A$ in week $t$ in store $i$. Each effect described above is defined as the difference between a household’s intention in the absence of the promotion and its behavior in the presence of the promotion.
Table 4.1: Short-term responses at the household level to a promotion for brand $A$ in store $i$ in week $t$, explained as the difference between intention and behavior

<table>
<thead>
<tr>
<th>Store choice</th>
<th>Category choice</th>
<th>Brand choice</th>
<th>Purchase quantity</th>
<th>Purchase timing</th>
<th>Speed of consumption</th>
<th>Name of effect</th>
<th>Net unit sales effect for manufacturer</th>
<th>Net unit sales effect in category for retailer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intention</td>
<td>$i$</td>
<td>yes</td>
<td>not $A$</td>
<td>normal</td>
<td>week $t$</td>
<td>normal</td>
<td>1. Brand switching</td>
<td>+</td>
</tr>
<tr>
<td>Behavior</td>
<td>$i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>week $t$</td>
<td>normal</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Intention</td>
<td>$i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>later</td>
<td>normal</td>
<td>2. Timing acceleration</td>
<td>0</td>
</tr>
<tr>
<td>Behavior</td>
<td>$i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>week $t$</td>
<td>normal</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Intention</td>
<td>$i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>earlier</td>
<td>normal</td>
<td>3. Quantity acceleration</td>
<td>0</td>
</tr>
<tr>
<td>Behavior</td>
<td>$i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>week $t$</td>
<td>normal</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Intention</td>
<td>not $i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>week $t$</td>
<td>normal</td>
<td>4. Anticipatory responses</td>
<td>0</td>
</tr>
<tr>
<td>Behavior</td>
<td>$i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>week $t$</td>
<td>normal</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Intention</td>
<td>$i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>promotion week</td>
<td>normal</td>
<td>5. Store switching</td>
<td>0</td>
</tr>
<tr>
<td>Behavior</td>
<td>$i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>week $t$</td>
<td>normal</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Intention</td>
<td>$i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>week $t$</td>
<td>normal</td>
<td>6. Deal-to-deal purchasing</td>
<td>+</td>
</tr>
<tr>
<td>Behavior</td>
<td>$i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>week $t$</td>
<td>normal</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Intention</td>
<td>$i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>week $t$</td>
<td>normal</td>
<td>7. Increased consumption</td>
<td>+</td>
</tr>
<tr>
<td>Behavior</td>
<td>$i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>week $t$</td>
<td>normal</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Intention</td>
<td>$i$</td>
<td>no</td>
<td>$A$</td>
<td>normal</td>
<td>-</td>
<td>normal</td>
<td>8. Category switching</td>
<td>+</td>
</tr>
<tr>
<td>Behavior</td>
<td>$i$</td>
<td>yes</td>
<td>$A$</td>
<td>normal</td>
<td>week $t$</td>
<td>normal</td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

$^a$ Intention is a household’s intended purchase behavior, in the absence of a promotion for brand $A$ in store $i$ in week $t$.

$^b$ Behavior is a household’s actual purchase behavior, given the presence of a promotion for brand $A$ in store $i$ in week $t$.

$^c$ From a category manager’s perspective, category switching is beneficial due to an increase in the net number of units sold in the category. For the retailer, this increase comes at the cost of reduced sales in other categories. For the manufacturer this effect is likely to be beneficial as long as the category in which sales decrease is not represented in the manufacturer’s product portfolio.
4.2. Conceptual decomposition

We show in Table 4.1 that these effect types, measured in units sold, differ in attractiveness to the manufacturer and retailer. Brand switching is beneficial for the manufacturer, at least in the absence of competitive reactions from other manufacturers, but not for the retailer. Timing acceleration and quantity acceleration are unattractive for both manufacturers and retailers. Both effects imply a shift of future unit sales of brand A in store i to the sales spike for brand A in week t in store i. Of course, by filling the consumer’s pantry, manufacturers can prevent consumers from switching to other brands in subsequent weeks. Similarly, anticipatory responses imply a shift of past unit sales for a given brand into the current sales spike, and do not generate net unit sales effects. Store switching is beneficial for the retailer, in the absence of competitive reactions by other retailers, but not for the manufacturer.

The proper classification of deal-to-deal purchase behavior is subject to disagreement. It depends on whether households would have bought brand A at all, if it were never promoted. Our conjecture is that in a market with close substitutes and frequent promotions, the “deal-to-deal households” would buy other brands only in the absence of promotions for brand A. Hence, we classify the deal-to-deal effect as beneficial for both manufacturers and retailers. Increased consumption is beneficial for both parties too. And finally, within the category considered, category switching may suit both manufacturers and retailers. Nevertheless, the retailer loses sales of substitute products and so might the manufacturer (see footnote c in Table 4.1). To summarize, the decomposition of the promotion effect on unit sales is critical for managers to determine whether a (gross) sales effect is beneficial. Based on the magnitudes of the decomposition sources we can compute the net unit sales effect for both manufacturers and retailers. We return to this issue in section 4.6.

---

3. We do not provide profit implications, since we lack data on margins and costs.
4. We define this effect, however, as a “multiple source” since it involves the substitution of a future cross-brand purchase with a current own-brand purchase. We do not find evidence for this effect in our data; see also footnote 10.
5. We note, however, that in the absence of promotions for all brands, these households would probably purchase brand A. In that case, the deal-to-deal effect would imply zero additional sales. Note however, that in Table 4.1 we show the net effects from a promotion for brand A ceteris paribus, which means that other brands continue to offer promotions.
4.3 Past research on decomposition

There are five major studies of the decomposition issue based on household panel data: Gupta (1988), Chiang (1991), Chintagunta (1993), Bucklin, Gupta, and Siddarth (1998), and Bell, Chiang, Padmanabhan (1999). Gupta (1988) provided the first decomposition of the effect of promotions (with data from the coffee category). He studied the effects of promotion on when, what, and how much households buy. He used different models for each of the three components: an Erlang-2 interpurchase time model (when), a multinomial logit model for brand choice (what), and a cumulative logit model for purchase quantity (how much). The brand choice decision accounts for 84 percent of the promotion effect, shorter interpurchase time represents about 14 percent of the sales increase, whereas stockpiling is a negligible phenomenon accounting for about 2 percent of the sales increase (see also Table 4.2).

Chiang (1991) posed a slightly different question: what are the effects of promotion on whether (i.e. purchase incidence), what and how much to buy. He used a simultaneous model for the three components, and provided a decomposition into these three sources. Chiang’s decomposition percentages (also for coffee) are very similar to the ones found by Gupta (1988). For feature his results are (display results in parentheses): 81 (85) percent brand switching, 13 (10) percent increases in purchase incidence (i.e., purchase timing), and the remaining 6 (5) percent increases in purchase quantity (see also Table 4.2).

Chintagunta (1993) extends the approach of Gupta (1988) and Chiang (1991) by including unobserved household heterogeneity in the models. Gupta (1988) and Chiang (1991) however, do include observed heterogeneity. For example, Gupta (1988) uses a household’s purchase history as a means for accounting for heterogeneity. Chiang (1991) also uses income, to capture variations in purchase behavior across households. Nevertheless, Chintagunta (1993) argues that it is important to include unobserved heterogeneity as well, such as a household’s intrinsic preferences for the different brands in a product category. He demonstrates that not accounting for the effects of unobserved heterogeneity leads to biased estimates of the effects of the promotion variables. We display his decomposition results in Table 4.2.

---

6. Studies conducted in the 1990’s appear to prefer to model the “whether” over the “when” question. This change may have been inspired by Wheat and Morrison (1990) who show that the modeling of purchase incidence is almost always preferable to the modeling of interpurchase time.
Bucklin, Gupta, and Siddarth (1998) extend the analyses by Gupta (1988), Chiang (1991), and Chintagunta (1993) by not only accounting for household heterogeneity but also deriving latent consumer segments. They develop a joint approach to segment households on the basis of their responses to price and promotion in brand choice, purchase incidence, and purchase quantity. The authors model brand choice (what to buy) by multinomial logit, incidence (whether to buy) by nested logit, and quantity (how much to buy) by poisson regression. Response segments are determined probabilistically using a mixture model. The authors apply the approach to household panel data for the yogurt category and find substantial differences across segments in the relative impact of choice, incidence, and quantity decisions on overall sales response to price. Their results are contained in Table 4.2.

Bell, Chiang, and Padmanabhan (1999) provide an empirical generalization of the decomposition of promotional response. The authors use a similar model as Chiang (1991) (which does not account for unobserved heterogeneity) and obtain results for 173 brands across 13 different product categories. They regress brand-specific decomposition elasticities on: (1) category-specific factors, (2) brand-specific factors, and (3) consumer characteristics. They conclude that these factors explain a significant amount of the variance in the three components of promotional response for a brand. For example, storability of a product category causes the purchase acceleration effects (incidence and quantity) to increase. The authors also find instances where some predictor variables do not affect total elasticities but do affect individual components of the total elasticity.

We summarize the decomposition results from the five studies in Table 4.2. In Table 4.2 we see that the average brand switching component is by far the largest (74 percent), followed by purchase quantity (15 percent), and purchase timing (11 percent). However, the percent decompositions differ substantially across categories. Categories for which inventories tend to be restricted, such as margarine and ice cream, show relatively small purchase quantity percentages. For more detail on reasons for differences in decomposition sources across categories and brands, see Bell, Chiang, and Padmanabhan (1999).

We note that all studies of household data restrict the attention to three of the eight decomposition effects. Since the three effects reported in Table 4.2 add up to 100 percent, it is appropriate to call this a restricted decomposition. It would be useful to quantify the other five effects as well and obtain an unrestricted decomposition from models of household data. However,
Table 4.2: Results from household-level decomposition research

<table>
<thead>
<tr>
<th>Study</th>
<th>Category</th>
<th>Brand</th>
<th>Timing</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>switching</td>
<td>acceleration</td>
<td>acceleration</td>
</tr>
<tr>
<td>Gupta (1988)</td>
<td>Coffee</td>
<td>84</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>Chiang (1991)</td>
<td>Coffee (feature)</td>
<td>81</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Coffee (display)</td>
<td>85</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Chintagunta (1993)</td>
<td>Yogurt</td>
<td>40</td>
<td>15</td>
<td>45</td>
</tr>
<tr>
<td>Bucklin, Gupta, Siddarth (1998)</td>
<td>Yogurt</td>
<td>58</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Margarine</td>
<td>94</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Softdrinks</td>
<td>86</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Sugar</td>
<td>84</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Paper towels</td>
<td>83</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Bathroom tissue</td>
<td>81</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Dryer Softeners</td>
<td>79</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Yogurt</td>
<td>78</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Ice Cream</td>
<td>77</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Potato Chips</td>
<td>72</td>
<td>5</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Bacon</td>
<td>72</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Liquid Detergents</td>
<td>70</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Coffee</td>
<td>53</td>
<td>3</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>Butter</td>
<td>49</td>
<td>42</td>
<td>9</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>74</td>
<td>11</td>
<td>15</td>
</tr>
</tbody>
</table>

Different studies may find different decomposition percentages for the same category due to model differences, data differences, etcetera.

since store-level data are the typical source for management decisions about promotions, we pursue the topic of decomposition on those data.

### 4.4 Contributions of this research

We extend the decomposition research as follows:
1. We obtain an *unrestricted* decomposition;
2. We use *store-level* scanner data;
3. We let the decomposition depend on the *size of the price discount*;
4. We let the decomposition depend on the *type of support* offered for the price discount;
5. We provide the first application of *structured semiparametric regression* methods in marketing research.
We elaborate on each of these points below.

**Unrestricted decomposition**

The household-level decomposition research has focused on three of the eight decomposition sources: brand switching, timing acceleration, and quantity acceleration. The implicit assumption in this research is that the total own-brand sales effect can be accounted for by just these three factors. Instead, we estimate the total own-brand sales effect and show that it is generally larger than the sum of the three decomposition sources. However, we cannot provide unique attribution to the remaining sources. Therefore, our approach measures the total size of the remaining sources together, and we define it as category expansion (see below). 7

**Use of store-level data for decomposition**

We use store-level data rather than household-level data. In theory, household data provide the best opportunities for decomposition. However, for business decisions about promotional expenditures, household data are used infrequently. Store-level data are more likely to be used by managers (Bucklin and Gupta 2000) and they are more appropriate for low-incidence product categories. 8 Therefore, it is important to have methods suitable for the decomposition of sales effects based on store-level data. However, the primary drawback of store-level data is that we cannot assign purchases to individual households, which makes it impossible to measure household-level responses. For that reason, Bucklin, Gupta, and Siddharth (1998) say that decomposition analyses are impossible to conduct with store-level data. However, we present a store-level approach that can produce managerially useful decomposition effects. In Table 4.3 we show the decomposition effects that can be measured with each of the two data types. Although all effects are in theory measurable with household-level data (first column), only the first three effects have actually been considered in published decomposition studies so far. Note that we use a different terminology for the decomposition sources at the store level.

We now consider the measurement of decomposition effects at the store level. Brand switching effects at the household level are equivalent to negative cross-brand effects at the store level. Timing acceleration, quantity

---

7. Given the type of data used, we cannot disentangle the remaining sources.
8. For other arguments in favor of store-level data, see section 1.6.
acceleration, and anticipatory responses lead to negative dynamic effects as we show in Chapter 3. Store switching at the household level leads to negative cross-store effects at the store level (or: interstore sales displacement, using the terminology of Walters 1991). It is possible to measure cross-store effects with store data if the stores in the data set are direct competitors, as Kumar and Leone (1988) and Walters (1991) have shown. However, since our data represent a random sample from a specific region and we lack geographical information on the stores, we cannot infer cross-store effects. Deal-to-deal purchasing is measurable with household data but not with store data (see also Chapter 3). Increased consumption is measurable indirectly with household data as Ailawadi and Neslin (1998) have shown, but not with store data. Category switching is measurable with both household- and store data, although it is a nontrivial task.

There are three differences between our store-level approach for decomposition and the household-level approach. The first difference is due

---

9. In the case of a display extension effect, the total dynamic effect may be positive, as we also note in Table 3.6.c

10. We have included dynamic cross-brand effects in a previous study (Van Heerde, Leeﬂang, Wittink 1997). The results indicated that these effects are negligible in the product categories studied.

11. In a previous study (Van Heerde, Leeﬂang, Wittink 1996) we have accommodated cross-store effects at the chain level, but were unable to ﬁnd signiﬁcant effects in one product category studied.
4.4. Contributions of this research

Some household-level studies explicitly model the dependency between the three components of consumer purchase behavior (what, whether/when, and how much to buy) (Chiang 1991, Chintagunta 1993), whereas other studies assume independency (Gupta 1988, Bucklin, Gupta, and Siddarth 1998, Bell, Chiang, and Padmanabhan 1999). Our store-level model accommodates automatically any dependency between decisions about brand, timing, and quantity since our criterion variable is the sales of a specific brand at a specific point in time, which combines the three aspects. Hence our model is consistent with the household-level models that assume dependency, but not with the other models, and this may cause differences in our decomposition results, relative to some of the results reported in Table 4.2.

The second difference is due to the data. Whereas we have data on all purchases within a category in a specific store and in a specific week, household data sets only contain a subset of these purchases since only a subset of all store customers are included in household panels. If these customers are a probability sample of all customers, there is no problem. However, the way the data are selected in household studies may cause differences with the store-level approach. Some studies use “household selection” and others use “purchase selection”. In “household selection”, researchers first select those households from the database who restrict their purchases to the specific set of brands to be modeled. Or, only the “heavy users” are selected, omitting infrequent buyers. In “purchase selection”, researchers first select perhaps the same set of brands, and next select all purchases of these brands from the database. All purchases by panel members of the brands selected for inclusion are used in the analysis, independent of whether panel members also purchased other brands in the category. The problem with models based on data obtained by household selection is that the own-brand effect of price promotions may be severely biased downward. Gupta et al. (1996) report a 42 percent larger own-brand price elasticity under purchase selection than under household selection. The reason for this bias is that data obtained by household selection exclude purchases of (potentially price-promoted) brands by households excluded from the selected sample. Most clearly, the excluded households tend to have a larger consideration set of brands. On the other hand, purchase selection leads to estimates of own-brand price effects that are only slightly different from the ones obtained from store-level models, at least in one particular setting (Gupta et al. 1996). Of the five household decomposition studies reviewed in this
chapter, Chintagunta (1993) uses “household selection”, whereas the others use “purchase selection”. Interestingly, his study has the smallest (40 percent) brand switching component.

The third difference between our approach and the household-level approach for decomposition is due to the process of aggregation. We measure the sales responses aggregated across heterogeneous households. We do not account for household heterogeneity, whereas household-level studies typically do in a variety of ways. Importantly, the key question our approach answers is: “what are the sizes of the decomposition effects?”, rather than: “why are the sizes of the decomposition effects as large as they are?”. The latter issue can be studied in future research.

Decomposition depends on the price discount magnitude

The relative sizes of the decomposition sources should depend on the magnitude of the price discount offered. As we have seen in Chapter 2, the sizes of the cross-brand effects of a promotion are influenced by the amount of the discount (see Figures 2.4-2.6). Analogously, we propose that the size of each decomposition source depends on the monetary incentive. For example, households may only be willing to accelerate the quantity purchased if the price reduction exceeds their transportation- and holding costs of the increased inventory. Hence we may only observe (negative) dynamic effects for price discounts that are sufficiently high. Finally, category expansion effects may also occur only for large price discount levels. Thus, while previous research assumes the same decomposition percentages for different price discount levels, we relax this assumption.

Decomposition depends on the type of support

We also propose that the sizes of the decomposition sources depend on the type of support for the price discount. We distinguish four types: feature-only, display-only, feature and display together, and neither feature nor display. For example, store switching may be induced especially by feature promotions, because features are the device by which stores reach potential store switchers (Narasimhan, Neslin, and Sen 1996). Since store switching effects are (implicitly) captured in the category expansion effect, we expect the category expansion effect to be larger for featured price cuts than for nonfeatured price cuts. A second example is from Chapter 3, in which we see that the dynamic effects of price discounts depend strongly on the support type (see Tables 3.5.c
4.5. Modeling strategy

and 3.6.c). Although Gupta (1988) and Chiang (1991) provide decomposition results for feature separately from display, we believe it is appropriate to think of promotions as having a price reduction as its core, and a communication device which is feature-only, display-only, feature and display, or neither.

**Application of structured semiparametric regression**

Finally, we present the first application of a “structured semiparametric regression model” in marketing. This model includes the partial effect of one predictor nonparametrically, the partial- and interaction effects of other predictors nonparametrically, and the effects of yet other predictors parametrically. The rationale for this approach is outlined in section 4.5.3.

**Summary of model requirements**

Summarizing, we need a store-level model to isolate the following three sources of the unit sales effect of promotions:12

- cross-brand effects;
- dynamic effects;
- category expansion effects,13

and we want to use a flexible model so that the sizes of these three effects depend on:

- the size of the price discount offered;
- the type of support offered.

Based on Table 4.1, we argue that for the manufacturer, the cross-brand effects and category expansion effects, measured in unit sales, are beneficial. For the retailer, only category expansion effects are beneficial. We measure net promotion effects for both parties in an application presented in section 4.6.

4.5 Modeling strategy

In this section we first present a visual representation of our modeling strategy for decomposition. We next provide the model specification, and outline the

12. Table 4.3 shows how these store-level decomposition sources relate to household-level decomposition sources.

13. Note that the category expansion effect also captures cross-brand effects for brands that belong to the product category but for which we have no data.
model estimation. We finally explain how we apply the model to obtain the decomposition of the sales effect of a promotion.

### 4.5.1 Visual model representation

We provide an illustration of the basic idea behind our model in Figure 4.1. The upper part of this figure shows the impact of a promotion for brand A in week \( t \) on its own-brand sales: a 300 unit increase (300 = 400–100). In this hypothetical example, we see that the brand looses 30 units in the prepromotion weeks (30 = 100–70), specifically in week \( t - 1 \). In addition, it looses 100 units in the postpromotion weeks (60 (= 100–40) in week \( t + 1 \) plus 40 (= 100–60) in week \( t + 2 \)). The total dynamic sales effect summed across pre- and postpromotion weeks equals –130 units. In the lower part of Figure 4.1 we observe that the cross-brand effect is equal to –150 units, i.e., the other brands loose 150 units in week \( t \). We can then compute the category expansion effect as the unit sales increase plus dynamic effects and cross-brand effects, i.e., 300–130–150 = 20. Hence in this example the decomposition is: cross-brand effects 50 percent (150/300); dynamic effects 43 percent (130/300) and category expansion effects 7 percent (20/300).14 To summarize, the model described below is designed to allow us to attribute part of the promotion effect on sales to cross-brand effects and to dynamic effects. The remainder is called category expansion.

### 4.5.2 Model specification

We need models that capture own-brand effects, cross-brand effects, and dynamic effects. We measure own-brand effects by including own-brand price instruments as predictors in the model for own-brand sales. This model also contains prices for other brands but these do not convey the cross-brand effects we need. Rather we need the effect of own-brand price on the sales of each of the other brands. These effects are obtained by including the own-brand price instrument as a predictor in the models for sales of each of the other brands. Finally, dynamic effects are inferred by including lead- and lagged own-brand price instruments as predictors in the model for own-brand sales. We include the latter predictors in an unrestricted form rather than in an Almon polynomial or Exponential decay model, since (1) this is the easiest way to

---

14. If we neglect the category expansion effect as household-level studies do, we would have a decomposition into \((50/93)\times100\) percent = 54 percent cross-brand effects and \((43/93)\times100\) percent = 46 percent dynamic effects.
implement dynamic effects, and (2) the model results from the unrestricted specification are almost indistinguishable from the results for the optimal dynamic specification in Chapter 3 (see Tables 3.5.c and 3.6.c). Given that this procedure is used for each brand, the model for own-brand sales includes as predictors own-brand price instruments, cross-brand price instruments, and own-brand lead- and lagged price instruments (where the cross-brand effects are used as mentioned above). Thus, in a category with \( J \) brands, we need \( J \) models.

We need a model that allows the magnitude of each decomposition effect to depend on the level of price discount. Since we do not know a priori the exact nature of this dependency, we use nonparametric regression methods. In addition, we need a model that also allows the magnitude of
each decomposition effect to depend on the type of support offered. We use the four different price instruments (price indices) that were introduced in Chapter 3. Hence we have, for each brand, the following variables: (1) price discounts without support, (2) price discounts with feature-only support, (3) price discounts with display-only support, and (4) price discounts with feature-and-display support.

Our model is a combination of the nonparametric elements from the model of Chapter 2 (equation (2.1)), and the current- and dynamic effects for each promotion condition from the model of Chapter 3 (equation (3.4)). For brand $k, k = 1, \ldots, J$:

$$
\ln S_{ik,t} = m \left( \left\{ \ln PI_{ijl,t} \right\}_{j=1,l=1}^{j=J,l=4}, \{ \ln PI_{ikl,t-\delta} \}_{u=1,l=1}^{u=sk,l=4}, \{ \ln PI_{ikl,t+\gamma} \}_{v=1,l=1}^{v=s'k,l=4} \right) + \alpha_F F_{ik,t} + \alpha_D D_{ik,t} + \alpha_{FD} FD_{ik,t} + \psi_{ik} R_t + \xi_{ik} W_t + u_{ik,t}
$$

$i = 1, \ldots, N; \quad k = 1, \ldots, J; \quad t = 1, \ldots, T.$

(4.1)

In equation (4.1), $m(.)$ is a nonparametric function. The definitions of the variables are the same as those for model (3.4) in Chapter 3. We emphasize that $F$, $D$, and $FD$ are defined as dummies for feature-only, display-only, and feature and display without accompanying price cuts. We do not expect cross-brand effects nor dynamic sales effects for these non-price promotions.

4.5.3 Model estimation

The difficulty with model (4.1) is that it suffers from the curse of dimensionality. There may be many predictors in the nonparametric function $m(.)$ in (4.1), even with just five brands, since each can have up to four different price promotion variables. In addition, model (4.1) also includes potentially many lead- and lagged effects of the four own-brand price promotion variables. In a product category with five brands, four lead periods and five lag periods, the number of predictors with nonparametric effects is $(5 \times 4) + (9 \times 4) = 56$. As a result, nonparametric estimation would be highly unreliable or impossible.

The key challenge is to estimate this model but control the amount of unreliability. We do this by using a “structured semiparametric regression model”. It is a variant of the structured nonparametric regression model (Linton and Nielsen 1995). This model offers an approach (as is available through semiparametric models – see Chapter 2) to reduce the curse of
dimensionality, specifically by forcing many interaction effects to be zero. Hence it is a multivariate regression model with a nonparametric partial effect for one predictor (scalar $z_1$), and nonparametric partial- and interaction effects for other predictors (vector $z_2$): $y = m_1(z_1) + m_2(z_2) + u$, where $y$ is the criterion variable, $m_1(.)$ a univariate nonparametric function of scalar predictor $z_1$, $m_2(.)$ a multivariate nonparametric function of the vector of predictors $z_2$, and $u$ the disturbance term. By letting each predictor in turn be $z_1$, we can obtain all nonparametric partial effects sequentially. Note that we only set the interactions between the focal variable and the other variables equal to zero. However, the partial nonparametrically estimated effect of $z_1$ on $y$ holds all effects of the variables in $z_2$ constant both in a flexible nonlinear and interactive way (within the set $z_2$).

Since our model includes yet other predictors parametrically (vector $x$), we have a “structured semiparametric regression model”. A simple version of this model is:

$$y = m_1(z_1) + m_2(z_2) + x' \beta + u,$$

(4.2)

with $\beta$ a vector of regression parameters. By interchanging the contents of $z_1$ and $z_2$ we can estimate all partial nonparametric effects for all own-brand-, cross-brand-, lead-, and lagged price indices included in the $m(.)$ function in (4.1). We outline the estimation procedure for this model in Appendix H. 16

4.5.4 Model application for decomposition

We use model (4.1) to estimate the decomposition effects (cross-brand effects, dynamic effects, and category expansion effects) as outlined below. Let $\Delta$ be the difference operator. We then define $\Delta S_{ij,u}[\{\Delta PI_{iklt} = d-1\}]$ as the difference in brand $j$’s unit sales in store $i$ in week $u$ from its base level given a decrease in the price index with support $l$ for brand $k$ in week $t$ in store $i$ from its regular level (1) to level $d$. Hence

$$\Delta S_{ij,u}[\{\Delta PI_{iklt} = d-1\}] \equiv S_{ij,u}[\{PI_{iklt} = d, \text{other price indices} = 1\}] - S_{ij,u}[\{PI_{iklt} = 1, \text{other price indices} = 1\}]$$

In this definition, $u$ may run from $t - s'_k$ (the first prepromotion effect week) to $t + s_k$ (the last postpromotion effect week). Of course, it also includes $u = t$.

16. In sum, we distinguish in this thesis the following regression models, in increasing order of flexibility: the parametric model ($y = x' \beta + u$); the structured semiparametric model ($y = m_1(z_1) + m_2(z_2) + x' \beta$); the semiparametric model ($y = m(z) + x' \beta + u$); the structured nonparametric model ($y = m_1(x_1) + m_2(x_2) + u$); and the nonparametric model ($y = m(z) + u$).
Suppose there are \( J \) brands. We can write for brand \( k \):

\[
\Delta \sum_{u=-s_k'}^{s_k} \sum_{j=1}^{J} S_{ij,t+u}[\Delta PI_{ikl,t} = d - 1] = \Delta S_{i,k,t}[\Delta PI_{ikl,t} = d - 1] \\
+ \Delta \sum_{j=1 \atop j \neq k}^{J} S_{ij,t}[\Delta PI_{ikl,t} = d - 1] \\
+ \Delta \sum_{u=-s_k'}^{s_k} S_{ik,t+u}[\Delta PI_{ikl,t} = d - 1] \\
+ \Delta \sum_{u=-s_k'}^{s_k} \sum_{j=1 \atop u \neq 0}^{J} S_{ij,t}[\Delta PI_{ikl,t} = d - 1]. \quad (4.3)
\]

The term on the left-hand side of (4.3) represents the increase in category sales in store \( i \) across relevant time periods due to the use of \( PI_{ikl,t} \). We do not try to model this quantity directly. Instead, we derive it from its constituent parts on the right-hand side. The first term on the right-hand side in (4.3) is the own-brand effect (expected to be positive); the second term is the sum of the cross-brand effects (expected to be negative); the third term is the within-brand total dynamic sales effect (expected to be negative in the case of purchase acceleration and positive in the case of a display extension effect, see Chapter 3); the fourth effect is the cross-brand total dynamic effect (expected to be close to zero, see footnote 10). By setting this fourth effect equal to zero and interchanging sums and delta’s we obtain:

\[
\sum_{u=-s_k'}^{s_k} \sum_{j=1}^{J} \Delta S_{ij,t+u}[\Delta PI_{ikl,t} = d - 1] = \Delta S_{i,k,t}[\Delta PI_{ikl,t} = d - 1] \\
+ \sum_{j=1 \atop j \neq k}^{J} \Delta S_{ij,t}[\Delta PI_{ikl,t} = d - 1] \\
+ \sum_{u=-s_k'}^{s_k} \Delta S_{ik,t+u}[\Delta PI_{ikl,t} = d - 1]. \quad (4.4)
\]
The left-hand side of equation (4.4) is equal to zero only if the current own-brand sales effect plus the cross-brand effect and the total dynamic effect equals zero. In other words, if the own-brand effect can be fully explained by cross-brand effects and dynamic effects. This is the assumption in all household-level studies. If the left-hand side of equation (4.4) is larger than zero, then there is a part of the own-brand effect that cannot be attributed to current reductions in sales of other brands or to losses in own-brand pre- and postpromotion sales. We define this part as category expansion. Hence category expansion equals the left-hand side of (4.4), and may include other sources, such as store switching, increased consumption, deal-to-deal purchasing, and category switching.

We use the notation \( C E_{kl} \) for the category expansion effect for instrument \( l \) used by brand \( k \); \( O B_{kl} \) for the own-brand effect; \( C B_{kl} \) for cross-brand effects; and \( T D_{kl} \) for total dynamic effects. Equation (4.4) then becomes:

\[
CE_{kl} = OB_{kl} + CB_{kl} + TD_{kl}.
\]  

(4.5)

Based on model (4.1), estimated with the strategy of (4.2), and based on the logic of (4.4), we estimate the sizes of the decomposition sources for a specific instrument \( PI_{kl} \) in an average store and in an average week as follows\(^18\):

- Own effect: take \( \ln S_{ik,t} \) as the criterion variable in model (4.1). Next compute for an average store and an average week, the log sales estimate for all price indices at the regular levels (=1), and also the log sales estimate for lower values of \( PI_{kl} \) while the other price indices stay constant. Specifically, include \( \ln PI_{kl} \) as \( z_1 \) and the other predictors as \( z_2 \) (see equation (4.2)). Next transform log sales into sales.\(^19\) Finally, compute the own-brand effect estimate as the difference between the

\(^{17}\) We now omit the indices \( i \) for store and \( t \) for week since we estimate this quantity using data pooled across stores and weeks. Hence we are interested in (4.4) for an average store and an average week. Since (4.4) is a linear equation, taking averages does not affect its validity.

\(^{18}\) Since we would like to obtain an estimate for an average store and week, we replace the indices \( i \) for store and \( t \) for week by dots.

\(^{19}\) To do this, we follow a similar approach as in (2.4). Nevertheless, estimation of the nonparametrically estimated correction factor \( \exp(\frac{1}{2}\hat{\sigma}^2_k) \) would take too much time on a Pentium II 266 Mhz. personal computer (approximately 12 days for each equation). Therefore, we use a surrogate correction factor, obtained from the parametric version of (4.1): \( \exp(\frac{1}{2}\hat{\sigma}^2_k) \).
sales estimate at each price index level $d$ and the sales estimate when all prices are at their regular levels:

$$\hat{OB}_{kl}|d = \Delta \hat{S}_{k..} | \{\Delta P_{kl..} = d - 1\}$$

$$= \hat{S}_{k..} | \{P_{kl..} = d, \text{ other price indices} = 1\}$$

$$- \hat{S}_{k..} | \{P_{kl..} = 1, \text{ other price indices} = 1\}. \quad (4.6)$$

We note that we implicitly impute expected values of zero for the error terms in (4.6).

- **Cross-brand effect:** take $\ln S_{j..}, j \neq k$, each in turn, as the criterion variable in (4.1). Compute the log sales estimates for all price indices at regular levels ($=1$), and also the estimates for varying values of $P_{kl..}$, again keeping other predictors at regular levels. Now $\ln P_{kl..}$ takes the role of $z_1$ in these other brand log sales equations. Transform log sales into sales (see also footnote 19) and add the resulting values for $j \neq k$, for each price index level $d$ to obtain the total cross-brand sales estimate for each own-brand price index level. Finally, estimate the cross-brand effect as the total cross-brand sales estimate minus the cross-brand sales for all prices at regular levels:

$$\hat{CB}_{kl}|d = \sum_{j=1}^{J} \Delta \hat{S}_{j..} | \{\Delta P_{kl..} = d - 1\}$$

$$= \sum_{j=1}^{J} \{\hat{S}_{j..} | \{P_{kl..} = d, \text{ other price indices} = 1\}$$

$$- \hat{S}_{j..} | \{P_{kl..} = 1, \text{ other price indices} = 1\}\}. \quad (4.7)$$

- **Total dynamic effect:** we compute the total dynamic effect as the difference between the sum over $u$ of own-brand sales at each price index level $d$ and the regular sales level:

$$\hat{TD}_{kl}|d = \sum_{u=-s'_{k}}^{s_{k}} \Delta \hat{S}_{k..+u} | \{\Delta P_{kl..} = d - 1\}. \quad (4.8)$$

In equation (4.1) the dynamic effect is captured by leading and lagging own-brand price indices. For consistency with equation (4.1), we rewrite
(4.8) in the same form used to actually compute the total dynamic effect. To accomplish this, we sum the current sales effects across relevant lead- and lagged price index variables and obtain:

\[ \hat{T}\Delta|d = \sum_{u=-s_k}^{s_k} u \Delta\hat{\Delta} \]

\[ = \sum_{u=-s_k}^{s_k} (\hat{\Delta} \Delta + \Delta \hat{\Delta}) \]

\[ = \sum_{u=-s_k}^{s_k} (\hat{\Delta} \Delta + \Delta \hat{\Delta}) \]

\[ = \sum_{u=-s_k}^{s_k} (\hat{\Delta} \Delta + \Delta \hat{\Delta}) \]

\[ = \sum_{u=-s_k}^{s_k} (\hat{\Delta} \Delta + \Delta \hat{\Delta}) \]

(4.9) Hence we compute the log sales estimate for all price indices at regular level (=1), and the log sales estimate for lower values \(d\) of each specific lead or lagged variable \(\Delta\), \(u = -1, \ldots, s_k\) (lagged effects) and \(u = 1, \ldots, s_k\) (lead effects), while the other price indices stay at one. Hence the variables \(\Delta\) for varying \(u\) each take in turn the role of \(z_1\) in the own-brand log sales equation. Next transform log sales to sales (see footnote 19), and take the sum as in (4.9).

- We obtain the category expansion effect at price index level \(d\) as the own-brand effect plus the sum of the total cross-brand effect and the total dynamic sales effect:

\[ CE_k|d = \hat{\Delta}B_k|d + CE_{kl}|d + CE_{kl}|d \]

(4.10)

### 4.6 Results

To estimate the model we employ two years of weekly scanner data for five brands in the 6.5 oz. canned tuna fish product category \((J = 5)\). We use three of the brands for which we have results in Chapters 2 and 3 (brands 1-3). We include the other two brands (brands 4 and 5) to account for cross-brand effects as completely as possible.\(^{20}\) The data cover one supermarket chain in one

---

\(^{20}\) We note, however, that our definition of the category may be too narrow, since there may be other tuna varieties households consider to be substitutable for the five brands we use. The problem of category definition is one all researchers face.
metropolitan area with 28 stores, and we use all 104 weeks for estimation.21 We present descriptive statistics for the brands in Table 4.4. Brands 1-3 are premium-priced brands with similar regular prices, whereas brands 4 and 5 are in a lower price tier.22 We can see in Table 4.4 that all brands except brand 5 use price promotions frequently.

Before the estimation of (4.1) is possible, we have to determine the optimal lead and lag lengths \( s' \) and \( s_k \) for each of the four brands that engage in price promotions (brands 1-4). We do this by using a grid search over \( s' \) and \( s_k \) for the parametric version of (4.1). This grid search is infeasible due to an excessive amount of computer time required for (4.1) itself. Analogous to the grid search in Chapter 3, we vary \( s' \) and \( s_k \) independently from zero to six weeks, and select the model specification that minimizes AIC. The results are in Table 4.5.23

21. We do not split the data into estimation- and validation samples as in Chapter 2, because we want to use all observations to measure multiple nonparametric effects as efficiently as possible.

22. Since brands 4 and 5 have lower regular prices, we do not expect that households who normally purchase any of the other three brands will switch to brands 4 or 5 if either is price promoted. The reverse, however, is more probable. Such asymmetry in switching between price tiers has been documented by Blattberg and Wisniewski (1989). Therefore, we expect the cross-brand price effects for brands 4 and 5 to be small in the sales equations for brands 1-3. For this reason, we also think the results in Chapters 2 and 3 should not depend on the inclusion c.q. exclusion of brands 4 and 5.

23. We see that the optimal lead and lag weeks are different from the ones obtained in Chapter 3. This difference may be attributable to the use of a second year of weekly data, which may differ in the promotional timing for the brands.
4.6. Results

Table 4.5: Optimal lead and lag weeks for parametric version of (4.1)

<table>
<thead>
<tr>
<th>Brand</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of lead weeks ($s'_k$)</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Number of lag weeks ($s_k$)</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>AIC</td>
<td>$-1.830$</td>
<td>$-2.286$</td>
<td>$-1.698$</td>
<td>$-2.381$</td>
</tr>
</tbody>
</table>

$^a$Since brand 5 does not offer price promotions at all, we cannot model lead or lagged effects for this brand.

Table 4.6: Number of predictors and observations for model (4.1)

<table>
<thead>
<tr>
<th>Model for brand</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nonparametric current price effects</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Number of nonparametric lead price effects</td>
<td>12</td>
<td>16</td>
<td>20</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>Number of nonparametric lagged price effects</td>
<td>24</td>
<td>20</td>
<td>4</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Number of parametric predictors</td>
<td>125</td>
<td>125</td>
<td>128</td>
<td>126</td>
<td>131</td>
</tr>
<tr>
<td>Total number of predictors</td>
<td>177</td>
<td>177</td>
<td>168</td>
<td>174</td>
<td>147</td>
</tr>
<tr>
<td>Number of observations</td>
<td>2660</td>
<td>2660</td>
<td>2744</td>
<td>2688</td>
<td>2912</td>
</tr>
</tbody>
</table>

Based on the information in Tables 4.4 and 4.5, we summarize the number of nonparametric price effects to be estimated in Table 4.6. We estimate model (4.4) for each price variable separately, to obtain a nonparametric estimate of its partial effect. We have sixteen current effects: four brands with four price instruments each. The number of lead (lag) effects equals the number of lead (lag) weeks times four instruments. For completeness, we also display the number of predictors included parametrically and the number of observations available for each model.

The next step is the estimation of (4.1) with the approach briefly outlined in (4.2) and explained in further detail in Appendix H. $^{24}$ We show the decomposition results in Figures 4.2 and 4.3. Figure 4.2 contains the absolute sizes of the three decomposition effects, for varying levels of price discounts, and for varying levels of support, averaged across the brands. For each support type, we take the range of price discounts in which all brands have price promotions. Hence we see differences in the x-axes: the minimum price index is 0.70 for both unsupported price discounts (Figure 4.2.a) and price discounts supported with display-only (Figure 4.2.c), whereas the minimum is 0.80 for

$^{24}$ Our estimation approach does not account for possible autocorrelation, since there are no methods available to accomplish this for the type of nonparametric models we use. Note, however, that the parametric model results in Chapter 3 changed very little when accounting for autocorrelation relative to the OLS results.
feature-only price discounts (Figure 4.2.b) and 0.65 for price discounts with feature and display support (Figure 4.2.d).

Figure 4.2: Absolute decomposition effects

(a) No support  (b) Feature-only support

(c) Display-only support  (d) Feature-and-display support

The four graphs in Figure 4.2 show what appear to be inconsistencies. For a given price discount level, we expect the largest unit sales increase if it is supported by feature and display, and the lowest effect for price discounts without support. The effects for price discounts with feature-only and display-only should be somewhere in between, and we do not know a priori which of these two will have the largest effects. However, the empirical results are not consistent with these expectations. For example, for a 20 percent price discount (i.e., a price index of 0.80), we observe the largest sales effect for display-only support, next largest for feature and display support, followed by no support, and finally feature-only support.
There are several possible explanations for these findings. First, the nonparametric estimation method might somehow produce results that are entirely different from extant methods, which would cast doubt on the usefulness of the former method. To consider this possibility, we also computed the average parametric price elasticities for each support type. These elasticities produce the same pattern: the most negative price elasticity is for display-only supported price discounts (−3.61), followed by feature-and-display supported discounts (−2.95), the unsupported price elasticity (−2.26), and finally the elasticity for feature-only supported price discounts (−1.64). Importantly, the (parametric) standard errors are too low for these differences to be due to chance. Hence the unexpected pattern in promotion effects is not unique to our new estimation method.

A second possible explanation for the inconsistencies is that they are due to averaging across brands. For example, brands differ in the relative frequency with which the different supports are used, and if the brands are heterogeneous in parameters, the process of averaging can produce seemingly inconsistent results. However, all brands, except for brand 2, show at least one inconsistency in the order of the effects.

A third possibility is the timing of promotions with each support type. For instance, if feature-only price discounts would be offered primarily in week(s) after price discounts with other support types, they could produce effects that are biased towards zero if consumers have stocked up. However, since our model accounts for the reduced effectiveness of promotions due to these phenomena, this explanation should not apply. Nevertheless, if promotions with feature always occur after multiple weeks with other support types, the feature-based promotion is at a disadvantage. The sample data, however, do not support this explanation.

A fourth possibility is the joint occurrence of own-brand- and cross-brand promotions, causing the own-brand effect to be small (especially for feature-only discounts). Our model does account for cross-brand effects, but not for interaction effects between discounts for different brands. However, the interaction effects would have to be very large to produce the observed differences in the graphs in Figure 4.2. A fifth possibility is that out-of-stock

25. Unfortunately, Linton and Nielsen 1995 do not provide confidence bounds for higher-dimensional cases of their structured nonparametric model.
26. The semiparametric model (2.1) does account for these interactions, see also Figures 2.7 and 2.8.
conditions occur more frequently in the presence of features. This possibility cannot be explored, since we do not have data on the availability of the brands.

Notwithstanding the inconsistencies, the graphs in Figure 4.2 show that the absolute sizes of the three decomposition effects depend heavily on both the level of the price discount and the support type. For instance, an unsupported price discount on average generates category expansion effects only for price indices below 0.85 (Figure 4.2.a). However, price discounts with feature and display generate a large absolute category expansion effect for all price discount levels (Figure 4.2.d). An attractive aspect of all graphs in Figure 4.2 is that the absolute sizes of all decomposition effects increase monotonically for all support types, with the exception of the category expansion effect for featured-only price discounts. However, the nonmonotonicity for the latter case is very modest and is easily attributable to unreliability.

We show in Figure 4.3 the relative sizes of the decomposition effects. We have expressed each effect as a percent of the own-brand effect so that their sum equals 100 percent.

The main conclusions we derive from Figures 4.2 and 4.3 are the following:

- The absolute and relative sizes of the decomposition effects (cross-brand effects, dynamic effects, and category expansion effects) depend on the type of support offered for the price discount;
- These absolute and relative sizes also depend on the level of price discount, given the type of support offered;
- For unsupported price discounts, the contribution from dynamic effects is the largest (45-60 percent) for all price discounts, followed by cross-brand effects (from about 20 percent to about 40 percent). Category expansion effects only occur at the largest price discount levels and then contribute 35 percent;
- For price discounts with feature-only support, the relative contribution from cross-brand effects is the largest for low price discount levels (more than 60 percent); at intermediate price discount levels cross-brand effects and category expansion effects both contribute 50 percent; at the highest price discounts dynamic effects become especially strong. For example, when the price index is 0.80, dynamic effects account for 55 percent, cross-brand effects for 30 percent, and category expansion effects for 15 percent;
4.6. Results

Figure 4.3: Relative decomposition effects

- For price discounts with display-only support, the dynamic effects equal 70 percent and cross-brand effects 30 percent for very low price discount levels. With increasing price discount levels category expansion grows, and the three effects are almost equally important for price indices less than 0.80;
- For price discounts with feature and display support, the category expansion effect contributes approximately 75 percent and cross-brand effects 25 percent at most discount levels, although there are also small dynamic effects for larger discount levels (10-15 percent). One possible explanation for the relatively small dynamic effects is the cancellation of pre- and post promotion dips by the display-extension effect (see also Chapter 3): a display for a promotional item may continue in the store in
the week(s) after the price promotion supported by feature and display, causing positive postpromotion effects;

- The dynamic effects are relatively much smaller for discounts with feature support (Figures 4.3.b and 4.3.d) than for discounts without feature support (Figures 4.3.a and 4.3.c). In addition, there are threshold levels for the dynamic effects for this type of discounts;

- The category expansion effect is clearly larger than zero in many cases. Therefore it is inappropriate to assume it away, as household-level decomposition studies implicitly do.

So far, we have shown average effects across the brands. We could also provide a decomposition of the own-brand effect for each brand separately. To save space, we present instead one case that illustrates the managerial use of our decomposition. Suppose that the manager of brand 2 and the manager of the retail chain believe it is instructive to evaluate the gross- and net unit sales effects from promotions for brand 2. The gross unit sales effect equals the own-brand effect in the promoted week. We have seen that the net unit sales effect for the manufacturer equals the category expansion effect plus the absolute cross-brand effect, whereas for the retailer, only the category expansion effect is beneficial. We show in Figure 4.4 the gross unit sales effect as well as the net unit sales effects separately for manufacturer and retailer, from promotions for brand 2 with various types of support and varying discount levels.

The gross unit sales effect suggests that for a (say) 25 percent price discount, the order of attractiveness is: (1) feature and display support, (2) display-only support, (3) feature-only support, and (4) no support. However, based on the net unit sales effect for the retailer, we see that a 25 percent price discount with display-only is the least attractive option. Almost all of the gross effect is lost due to cross-brand effects and dynamic effects. Instead, the retailer has a strong incentive to use feature and display support. This combination provides also the highest net sales gain for the manufacturer. Importantly, these insights could not have been obtained from models without decomposition effects.

27. We take brand 2 on purpose, since this is the only brand for which the order of estimated effects is consistent with our expectations, i.e., the largest sales effects are due to price discounts with feature and display support, and the lowest due to unsupported price discounts.
4.7 Conclusions

In this chapter we provide both a conceptual and an empirical decomposition of the unit sales effect due to promotions. The conceptual decomposition indicates that the sources of incremental sales due to a promotion differ in attractiveness to the manufacturer and retailer. Our model is based on store-level data, and it allows us to decompose the sales effect of price promotions into cross-brand effects, dynamic effects, and category expansion effects.

Our decomposition approach encompasses five contributions relative to previous decomposition approaches (Gupta 1988, Chiang 1991, Chintagunta 1998, Bucklin, Gupta, and Siddarth 1998, and Bell, Chiang, and Padmanabhan 1999). In the first place, we calibrate our model with store-level sales data, whereas previous studies used household-level purchase data. We argue
that it is important to develop decomposition models based on store-level data, because these data are used more frequently in practice. Previous household-level research has restricted the decomposition to three effects: brand switching, purchase quantity, and purchase incidence (timing). This relates to the second contribution: we provide an unrestricted decomposition, i.e., we quantify “category expansion” effects, which include sales increases due to store switching, deal-to-deal purchasing, increased consumption, and/or category switching. We find several instances in which category expansion is nonzero. The third and fourth contributions are that we allow the size of each of the three decomposition effects to depend on the level of price discount and on the type of support offered. Our results indicate that there are strong dependencies. The final contribution is that we use a flexible “structured semiparametric model”, whereas previous decomposition research used parametric models. One limitation of our results is an inconsistency in the order of the effect sizes of price discounts with different support types.

Our results indicate that the magnitudes of the net sales effects due to a promotion can be very different for manufacturers and retailers. Especially in an age of increased cooperation between manufacturer and retailer (see also section 1.3.5), it is important that the same model indicates how specific activities affect each of these parties. Our model can be used as a powerful tool for implementing promotions that yield high net unit sales effects for manufacturer and retailer simultaneously. This is exactly the purpose of tailor-made promotions. Of course, to do this one needs to go beyond the net unit sales measure. For example, to determine the profitability of a promotion, various costs have to be subtracted from the difference in revenues between non-promoted and promoted conditions. See section 5.3 for a discussion of profitability calculations.

We note that our net effects assume that competing manufacturers do not react. In fact, the benefit of brand switching is ephemeral if competitive reactions are intense. The sparse empirical evidence suggests that managers react to each other’s actions even if the cross-brand effects are not significant (e.g., Leeftang and Wittink 1996).

An important future research issue is the development of household-level models that provide a decomposition into all relevant sources. The fact that the category expansion effect is larger than zero in many cases implies that it does not suffice to only answer the “what”, “whether”, and “how much to purchase” questions (brand choice, purchase timing, and purchase quantity). A complete decomposition of the unit sales effect of a promotion can be obtained
by modeling a consumer’s decision concerning what, when, how much, and where to buy (store switching), and subsequently, at what speed to consume the quantity purchased. Although it is in principle possible to measure cross-store effects of promotions using store-level data, we think it requires household data to measure increased consumption effects of promotions.

A related open question in decomposition research is the actual size of the deal-to-deal purchasing segment. Krishna (1994a) states that this segment consists of one third of all households, at least for coffee. Mela, Jedidi, and Bowman (1998) study the long-term effects of promotions on consumer stockpiling. They hypothesize that households develop price expectations on the basis of their prior exposure to promotions over a long period of time, such as months or years. They find that increasing expectations of future promotions lead to: (1) a reduced likelihood of purchase incidence on a given shopping trip, but (2) an increase in the quantity bought when households do decide to buy, mostly using a promotional offer. This suggests that with increases in promotional activity, a larger number of households show the deal-to-deal purchase behavior displayed in Figure 3.2. Thus, it is useful to include the purchases by the deal-to-deal segment explicitly as a new decomposition source for promotion effects in future (household-level) research. In current household-level research, these purchases are hidden in brand switching-, purchase incidence-, or purchase quantity effects. Explicit measurement of the deal-to-deal segment is important because it makes a large difference whether a price promotion attracts consumers who are normally loyal to other brands, or consumers who are loyal to deals, including those of other brands.

Another important future research issue in the area of decomposition is a comparison of household-level- and store-level decomposition results. We could generate a household-level data set and estimate the percentages attributable to brand choice, purchase incidence, and purchase quantity based on, for example, the model by Bell, Chiang, and Padmanabhan (1999). Next we can aggregate the purchases of the households to the store-week level and estimate a model similar to the one proposed in this chapter. To improve comparability, we would maximize the similarities in the models. We can then compare the decomposition results, such as the percent due to brand choice based on the household model with the percent attributed to cross-brand effects based on the store model. In addition, we can compare the percent attributed to purchase quantity and -timing at the household level with the total dynamic effect measured at the store level. Any differences can be explored
by considering, for example, differences in: (1) models, (2) data, and (3) aggregation.