

## A Hierarchical Bayes Error Correction Model to Explain Dynamic Effects of Price Changes

Dennis Fok, Richard Paap, Csilla Horváth and Philip Hans Franses

ERIM REPORT SERIES <i>RESEARCH IN MANAGEMENT</i>	
ERIM Report Series reference number	ERS-2005-047-MKT
Publication	September 2005
Number of pages	41
Persistent paper URL	
Email address corresponding author	dfok@few.eur.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus Universiteit Rotterdam P.O.Box 1738 3000 DR Rotterdam, The Netherlands Phone: + 31 10 408 1182 Fax: + 31 10 408 9640 Email: info@erim.eur.nl Internet: <a href="http://www.erim.eur.nl">www.erim.eur.nl</a>

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:  
[www.erim.eur.nl](http://www.erim.eur.nl)

REPORT SERIES  
*RESEARCH IN MANAGEMENT*

ABSTRACT AND KEYWORDS	
Abstract	<p>The authors put forward a sales response model to explain the differences in immediate and dynamic effects of promotional prices and regular prices on sales. The model consists of a vector autoregression rewritten in error-correction format which allows to disentangle the immediate effects from the dynamic effects. In a second level of the model, the immediate price elasticities, the cumulative promotional price elasticity and the long-run regular price elasticity are correlated with various brand-specific and category-specific characteristics. The model is applied to seven years of data on weekly sales of 100 different brands in 25 product categories. We find many significant moderating effects on the elasticity of price promotions. Brands in categories that are characterized by high price differentiation and that constitute a lower share of budget are less sensitive to price discounts. Deep price discounts turn out to increase the immediate price sensitivity of customers. We also find significant effects for the cumulative elasticity. The immediate effect of a regular price change is often close to zero. The long-run effect of such a decrease usually amounts to an increase in sales. This is especially true in categories characterized by a large price dispersion, frequent price promotions and hedonic, non-perishable products.</p>
Free Keywords	<p>Sales, Vector Autoregression, Marketing Mix, Promotional and Regular Price, Short and Long-term Effects, Hierarchical Bayes</p>
Availability	<p>The ERIM Report Series is distributed through the following platforms:</p> <p>Academic Repository at Erasmus University (DEAR), <a href="#">DEAR ERIM Series Portal</a></p> <p>Social Science Research Network (SSRN), <a href="#">SSRN ERIM Series Webpage</a></p> <p>Research Papers in Economics (REPEC), <a href="#">REPEC ERIM Series Webpage</a></p>
Classifications	<p>The electronic versions of the papers in the ERIM report Series contain bibliographic metadata by the following classification systems:</p> <p>Library of Congress Classification, (LCC) <a href="#">LCC Webpage</a></p> <p>Journal of Economic Literature, (JEL), <a href="#">JEL Webpage</a></p> <p>ACM Computing Classification System <a href="#">CCS Webpage</a></p> <p>Inspec Classification scheme (ICS), <a href="#">ICS Webpage</a></p>

# A Hierarchical Bayes Error Correction Model to Explain Dynamic Effects of Price Changes

Dennis Fok\*

*Econometric Institute  
Erasmus University Rotterdam*

Richard Paap

*Econometric Institute  
Erasmus University Rotterdam*

Csilla Horváth

*Econometric Institute  
Erasmus University Rotterdam*

Philip Hans Franses

*Econometric Institute and  
Department of Business Economics  
Erasmus University Rotterdam*

July 4, 2005

---

\*We thank Dick Wittink and three anonymous referees for their very helpful and extensive comments. We are indebted to participants of the Marketing Science conference 2004 in Rotterdam and to Marnik Dekimpe for their constructive comments on an earlier version of this paper. Finally, we thank Shuba Srinivasan for providing us with the data. Address for correspondence: Dennis Fok, Econometric Institute (H11-02), Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, phone: +31 10 4081315, fax: +31 10 4089162 email: [dfok@few.eur.nl](mailto:dfok@few.eur.nl). The model presented in this paper is implemented in Ox 3.4 (Doornik 1999).

# A Hierarchical Bayes Error Correction Model to Explain Dynamic Effects of Price Changes

## **Abstract**

The authors put forward a sales response model to explain the differences in immediate and dynamic effects of promotional prices and regular prices on sales. The model consists of a vector autoregression rewritten in error-correction format which allows to disentangle the immediate effects from the dynamic effects. In a second level of the model, the immediate price elasticities, the cumulative promotional price elasticity and the long-run regular price elasticity are correlated with various brand-specific and category-specific characteristics. The model is applied to seven years of data on weekly sales of 100 different brands in 25 product categories.

We find many significant moderating effects on the elasticity of price promotions. Brands in categories that are characterized by high price differentiation and that constitute a lower share of budget are less sensitive to price discounts. Deep price discounts turn out to increase the immediate price sensitivity of customers. We also find significant effects for the cumulative elasticity. The immediate effect of a regular price change is often close to zero. The long-run effect of such a decrease usually amounts to an increase in sales. This is especially true in categories characterized by a large price dispersion, frequent price promotions and hedonic, non-perishable products.

**key words:** sales; vector autoregression; marketing mix; promotional and regular price, short and long-term effects; Hierarchical Bayes

# 1 Introduction

There is substantial literature on dynamic price effects on sales, see for example, Kopalle et al. (1999), Paap & Franses (2000), Van Heerde et al. (2000) and Pauwels et al. (2002), among others. The term dynamic effect refers to the effect of a current change on future sales. In case of dynamic effects, the overall (net) effect of a price change obviously cannot completely be summarized by the immediate price elasticity. Ignoring the presence of dynamic effects will lead to erroneous conclusions. For example, Kopalle et al. (1999) find that promotions have positive contemporaneous effects on sales accompanied by negative future effects, and they rightfully emphasize that “models that do not consider dynamic promotional effects can mislead managers to overpromote”. Note that the net effect of promotions on sales can even be negative. For example, Jedidi et al. (1999) find that the long-term effects of promotions on sales are negative and, in an absolute sense, that they are about two-fifths of the magnitude of the positive short-run effect. Note that this estimate includes the (negative) effects of competitive reaction and of changes in consumer behavior.

In the literature there does not seem to be a consensus on the size of the dynamic effect relative to that of the immediate effect. Lack of consensus also appears in the literature on the so-called post-promotional dip. While several articles provide evidence of post-deal troughs, many others fail to find supporting empirical evidence (for example, Grover & Srinivasan 1992, Blattberg et al. 1995). This lack of consensus was first addressed by Neslin & Schneider Stone (1996). Later, Van Heerde et al. (2000) showed how one can estimate the pre- and postpromotion dip appropriately using time series models. The issue of the existence of postpromotion dips was then resolved by Macé & Neslin (2004) who consider the determinants of such dips.

An explanation for this initial disagreement in the marketing literature is that the effect varies across categories, stores, and brands (Macé & Neslin 2004). For example, people may not be inclined to stockpile brands with frequent discounts, but instead capitalize their storage capacity for other, less frequently promoted brands, as for the former category they may expect new discounts soon. Therefore, researchers may less likely find evidence of a post-promotion dip for brands with frequent price promotions. Also, there may be products that are more difficult to store as they are perishable or are large in size, and, there again, having evidence of a post-promotion dip can be rare.

Furthermore, it is sometimes not clear which marketing action is actually being consid-

ered. Most studies consider the effect of a change in actual price. This variable, however, captures two different pricing decisions, that is, price discounting and the regular price. Naturally, there will be a difference between the effects of a temporary price promotion (price discount) versus those of a permanent (regular) price change. All estimated effects will obviously be contingent upon the chosen definition. Bucklin & Gupta (1999) point out the need for separation of the two pricing decisions from the point of view of practitioners and researchers, and these authors classify the investigation and estimation of regular price elasticity as unresolved from both perspectives.

A third possible explanation concerns the definition of the dynamic effect. In the literature there seems to be some confusion about the definition (and the measurement) of concepts as “short-run”, “long-run”, “net”, “cumulative”, and “dynamic” effects. Some authors analyze, for example, the effect of frequent promotions on consumer brand perception and call this the long-term effect of promotion. In other articles, including ours, the focus is on the measurement of post-promotional dynamics in sales. In most mature product categories, sales are stationary. Temporary changes in price therefore cannot have a permanent effect on the future level of sales. But, permanent changes in regular price most likely do have a permanent effect. The immediate and dynamic effects of the two types of price changes are also expected to differ substantially (Bijmolt et al. 2005). Following this line of thought, it is important to distinguish between price promotions and changes in the regular price. Price promotions are by definition temporary, whereas changes in the regular price are permanent. Even though consumers only observe the actual price, they shall have no difficulty in identifying a price promotion from a regular price change. Regular price changes are usually much smaller in magnitude compared to price promotions. Furthermore, regular price changes will oftentimes be announced by the retailer.

To properly analyze the dynamic effects of promotions, we believe we have to distinguish between temporary price promotions and permanent regular price changes. For both price changes the immediate effect will be of interest. As discussed, temporary price promotions do not have a permanent effect. It therefore is not useful to consider the long-run effect of price promotions. The effect of a price promotion at time  $t$  on sales at time  $t + k$  for  $k \rightarrow \infty$  will in fact be zero for all (stationary) cases. For price promotions it is therefore more relevant to consider the total, or cumulative effect, on current and all future sales, that is, the sum of the effects over all future periods  $(t, t + 1, t + 2, \dots)$ . The sign of this cumulative effect will indicate whether the promotion may be beneficial

in the long run. Note that for permanent changes in the regular price it is likely that a permanent effect does exist. In this case, it is interesting to consider the long-run effect of such a change. The cumulative effect of a permanent change will not be informative.

In this paper we aim to identify the immediate and dynamic effects of temporary price promotions and permanent changes in the regular price. The term immediate effect is used for both price promotions and regular price changes. In sum, throughout this paper we will consider the *cumulative effect of a price promotion* and the *long-run effect of regular price change*. Finally, the term dynamic effect is used to capture all effects of current changes on future sales.

To investigate these dynamic effects we consider a new modeling methodology, which we apply to weekly data for a large amount of categories. We relate the effects of price changes of multiple brands in various product categories to observable product and category characteristics by using a Hierarchical Bayes (HB) - Error Correction Model (ECM). This model allows us to directly estimate the potentially differing immediate and dynamic effects of price changes on sales, where we relate these effects to characteristics of brands and categories. As mentioned, we explicitly distinguish between promotional price elasticities and regular price elasticities. Although many studies suggest that these elasticities differ, there is a gap in the marketing literature concerning the determinants of regular price elasticities. Our model provides important insights for brand managers and for retailers about these determinants, which can moderate the effect of price changes. Among other things, we analyze whether there are common determinants for the immediate promotional price elasticity and for the cumulative effects of price promotions.

It is important to emphasize that our analysis focuses on the cross-sectional, and not on the longitudinal determinants of price elasticities. That is, we analyze the differences in the immediate and cumulative effects of promotional and regular price changes across categories and brands, thereby assuming that the characteristics of the investigated markets do not change.

The remainder of this paper is organized as follows. In Section 2 we present a detailed overview of the literature. In this section we also discuss our hypotheses on the relationship between brand and category characteristics and the immediate and dynamic effects of promotional price and regular price. In Sections 3 and 4 we present our Hierarchical Bayes Error Correction Model in detail. Technical derivations of the estimation algorithm are relegated to Appendix A. The empirical results of this paper are presented in Section 5. We conclude in Section 6.

## 2 Literature

In the literature there are many articles that investigate the relationship between market characteristics and promotional price elasticities. The number of papers on regular price elasticities is much smaller. In Table 1 we compactly present a selection of studies on price elasticities. It is clear that most of these articles focus on relating *immediate promotional price elasticities* to brand, category, or consumer characteristics. Almost all studies use a *two-stage approach* for the empirical analysis and estimate *single-equation models*. Our approach, in contrast, uses a *HB-ECM* to relate the *immediate and cumulative effects* of *promotional price and regular price* to *brand and category* characteristics.

Among the listed in Table 1, we find only three articles that investigate the determinants of the immediate as well as of the dynamic effects of marketing actions. Foekens et al. (1999) use varying parameter models to investigate the effects of the properties of a particular discount (like the size of the discount and the time since the previous discount) on the intercept and price promotion parameters in a sales model. Nijs et al. (2001) consider the moderating effect of marketing intensity, competitive reactivity, and competitive structure on the category-demand effect of price promotions. Macé & Neslin (2004) investigate the determinants of pre- and postpromotion dips. As moderating factors they consider brand and category characteristics as well as characteristics concerning the store trading-area. Our approach is more in line with the latter two as we investigate heterogeneity in the immediate and dynamic effects of price across brands and categories, assuming constant parameters over time.

Table 1 also shows that most studies consider moderators of price elasticity. As discussed before, price promotions and changes in regular price are likely to have different effects on sales. In this study we explicitly distinguish between these two price effects.

### 2.1 Overview of findings

In this subsection we give a brief overview of the literature on the effects of price promotions and of regular price. First we consider the relative size of the different effects. Next, we consider possible determinants of the price elasticities in more detail.

Based on the extensive literature on the post-promotion dip we expect the cumulative effect of price promotions to be smaller than the immediate effect. The long-run elasticity of regular price can be larger (in absolute value) than the immediate effect. Due to its lasting feature, a regular price cut does not induce consumers to accelerate their purchase



and stockpile. So, a permanent price reduction has its effect over a longer time period. Consumers decide to buy the price-reduced brand at the regular shopping trip causing no post-promotion dip.

We expect short-run promotional elasticity to be stronger than the immediate elasticity of regular price (Bijmolt et al. 2005, Blattberg et al. 1995). The relationship between the cumulative promotional elasticity and the long-run elasticity of regular price is less clear. In fact, the cumulative effect of a price promotion measures something completely different from the long-run effect of a decrease in regular price.

Below we discuss the expected relationship between brand and category characteristics and the price elasticities. We make a selection of variables based on the properties of the available data and the existing literature that relates promotional elasticities to market, category, brand, and/or consumer characteristics (see, for example, Bell et al. 1999, Hoch et al. 1995, Narasimhan et al. 1996, Raju 1992). We extend the set of already studied explanatory characteristics with additional variables and provide a discussion of the determinants of dynamic effects. While the marketing literature on theorizing and analyzing the effects of price promotions is quite extensive, there are significantly fewer papers that deal with the determinants of the effects of regular price (see, for example, Blattberg & George 1991, Guadagni & Little 1983, Jedidi et al. 1999, Mitra & Lynch 1995, Shankar & Krishnamurthi 1996). We therefore mainly focus on price promotions. To develop our hypotheses we take as a starting point a utility-maximizing consumer who operates under a budget constraint (Bell et al. 1999, Varian 1992). We provide insights about the expected marginal effects of brand and category characteristics.

As we will see, in several cases we cannot formulate hypotheses about how brand- and category-specific characteristics influence cumulative effects. In these cases our research can be considered as a quest for empirical evidence on whether and how certain category and brand-specific characteristics can be related to the dynamic effects of price. We distinguish three groups of variables, that is, category-specific variables, brand-specific variables, and variables that can be defined at both levels. For example, promotion frequency can be defined at the category level. However, within a category the level of promotion may also be relevant. Note that this distinction is in contrast with most of the existing literature. For example, Macé & Neslin (2004) only consider the relative promotion frequency within a category, and they do not consider the general frequency for the category. For each characteristic, we summarize the literature and give, if possible, hypotheses for the sign of the immediate and cumulative effect of a price promotion.

## Category-specific characteristics

### Average budget share

The budget share of a category captures two distinct dimensions: (i) the general price level in the category, and (ii) the relative purchase frequency and purchase quantity. The first dimension leads to the conclusion that expensive product categories are likely to exhibit a smaller immediate effect and a lower and shorter post-promotion dip. The difference between the immediate effect and the cumulative effect will therefore be relatively small. A consumer, facing a budget constraint, will probably feel more inclined to buy extra quantity of a less expensive product as the additional purchase will lead to a smaller grocery bill (Raju 1992). In addition, price promotions may be less effective for expensive products because higher-income (and, hence, less price-sensitive) shoppers may constitute a greater ratio of the consumer population.

The second dimension results in a large promotional price effect and a large promotional dip. The latter is also found empirically by Macé & Neslin (2004). The argument here is that as the consumer will certainly be in need of the product in the near future, s/he will be inclined to purchase additional items of the high budget share category.

**Utilitarian** Bell et al. (1999) argue that relatively more necessity (i.e., nonimpulse) products are expected to exhibit lower primary demand effects and higher secondary effects. However, when facing a promotion for necessity products, consumers, knowing that they will surely be in the need of such products in the future, may be more inclined to pile up at home, even by postponing purchases of products in other categories. Moreover, shoppers with higher income (and, hence, who are less price sensitive) may constitute a smaller proportion of the consumer population of necessity goods. So, we conjecture that necessity products have higher immediate effects and also a possibly larger difference between the immediate and the cumulative effect. Wakefield & Imman (2003) make a distinction along a similar dimension and find consumers to be less price sensitive in categories that are perceived as primarily hedonic in nature.

A regular price increase in a utilitarian product category is likely to have a lower effect than a similar increase in a more hedonic category. In case of a price rise of a necessity product, households with a binding budget constraint may be inclined to reduce purchases of other, less essential products.

**Perishability** Consumers may favor price promotions in a category with easily storable goods, that is, a category where one can allow for purchases at irregular intervals as a response to deals (Raju 1992, Narasimhan et al. 1996). In case of easily storable goods,

consumers are more likely to substitute future purchases with the consumption of stored items and are also more likely to utilize the accumulated stock during a longer period after the promotion. This suggests that promotions in a category with less perishable products lead to higher sales during the offer but also to a larger and longer dip after the promotion than in a more perishable one. Therefore, here we expect the cumulative effect to be smaller than the immediate effect.

**Competitive intensity** Nijs et al. (2001) argue that in less competitive environments, characterized by a smaller number of brands, price-promotion effectiveness is expected to be higher. Moreover, more competition implies that brands are less segmented from each other. Therefore, there is a higher probability that a promoted brand will have a close substitute from which consumers will easily switch to the promoted brand. This results in a higher immediate effect and possibly a higher cumulative effect in a more competitive category.

Narasimhan et al. (1996) argue that brand proliferation may signal the existence of many market segments. Such a differentiation protects the brands from competitors' actions. At the same time, brand proliferation has also been identified as a potential cause for weaker brand loyalty (Narasimhan et al. 1996).

In this study, we use two measures of competitive intensity, that is, the market concentration index and price dispersion in a category. The first is a sophisticated measure of the market concentration. Price dispersion captures aspects of brand assortment.

## **Category- and brand-specific characteristics**

We measure the intensity of the use of marketing variables at the category level as well as at the brand level. Some categories may be promoted more often than others, but also within a category there may be differences in promotion frequency. To separate these two effects, we measure the characteristics at the category level and use a relative measure to distinguish brands within a category.

**Frequency of price promotional activity** Theory suggests mixed effects of the frequency of promotions. The theory on price consciousness (Kopalle et al. 1999, Mela et al. 1997, 1998), leads to a positive effect of the price-promotion frequency on the immediate effect and on the size of the post-promotion dip. In categories where price promotional activity is high, consumers may become more price conscious. As a consequence, they tend to purchase the products on deals and they may develop a habit of stockpiling.

On the other hand, the use of discounts may reduce consumers' reference prices (Kalaynaram & Winer 1995) resulting in a *lower* level of effectiveness of discounts. Furthermore, if a category is promoted infrequently, consumers are more likely to use these opportunities to stock-up for future consumption (Raju 1992).

Within a category other processes may also play a role. Brands with relatively frequent price promotions are often considered to be of lower quality than similar, rarely promoted brands. Intense promotional activity may also influence the mix of consumers for a brand. More specifically, frequently promoted brands may draw a larger proportion of the price-sensitive consumer base (Zenor et al. 1998).

Empirical findings about this relation seem to be ambiguous. Blattberg et al. (1995) emphasize the last two points and state that "the greater the frequency of deals, the lower the height of the deal spike". Bolton (1989) finds no significant relationship between category price activity and price elasticity of brands in the category. Raju (1992) concludes that sales in a product category in which the brands are promoted relatively often exhibit lower variability. Zenor et al. (1998) find that brands with higher levels of promotional activity are associated with more elastic demand than those that promote less. Nijs et al. (2001) discover that the frequency of price-promotional activity in the category is positively related to the short-run effect of price promotions. However, in the long run this effect appears to disappear. Finally, Macé & Neslin (2004) report a positive correlation between the relative price-promotion frequency and the size of the postpromotion dip.

These ambiguous findings may partly be due to the fact that some researchers tried to figure out the consequences for price elasticities of more frequent promotional usage in a *category*, while others focused on the result of frequent promotional usage for *brands*. In this study we use the price promotion frequency at the category level as well as at the brand level.

**Average depth of price promotion** In categories where price reductions are relatively large, consumers, who expect to get a high reward, are probably inclined to accelerate their purchase. As such, they draw sales from the weeks following the promotion, unless consumption increases correspondingly (Foekens et al. 1999). This would result in a high immediate increase in sales and large difference between the immediate effect and the cumulative effect of the promotion.

In categories where consumers are used to deep price cuts, a small change in regular price may not trigger a reaction by consumers. So, we expect a lower effect of regular price changes in categories with relatively large price discounts.

**Frequency of display/feature activity** Bolton (1989) argues that display activity may be systematically related to own price elasticities. The relative frequency of displays may influence consumers' belief about the popularity and quality of market offerings. This effect is more pronounced within a category than across categories. On the one hand, display activity may encourage customers to apply choice roles that rely less on search for price information, arriving at less price-elastic sales. On the other hand, it may lead customers to compare prices, which would result in higher price-elastic sales. Bolton (1989) finds support for the first case, that sales are more inelastic with respect to their prices for categories and brands that are frequently displayed.

Feature activities are often used to provide information about the prices and about price promotion in a retail outlet. So, frequent feature activity in a category is likely to make current consumers more aware of the prices and the occurrence of promotional activities in the category (Bolton 1989, Moriarty 1985). This suggest that brands in categories with frequent retailer advertising activity should have high immediate price elasticity and also a large drop in sales after the price promotion. A similar distinction can be made for more often promoted *brands* within a category.

Another theory is based on the fact that feature and display activity may increase brand salience as well as price salience (Shankar & Krishnamurthi 1996). The former may induce consumers to differentiate brands more and thereby increasing relative preferences and reducing the consideration set. This would lead to lower regular price elasticity (Mitra & Lynch 1995). On the other hand, increased price salience may lead to more price comparison within the category.

## **Brand-specific characteristics**

**Brand size** Bolton (1989) documents that brands with a relatively high market share tend to be operating on the flat proportion of their sales response functions. Hence, larger brands tend to be less own price elastic. Blattberg et al. (1995) mention this relationship as the second amongst the empirical generalizations for promotions.

**Price segment of a brand** A discount may attract several types of consumers: (i) consumers who usually purchase a competing brand; (ii) consumers who would otherwise find the brand too expensive; and (iii) already loyal consumers (Raju 1992). The promotion of a brand in an expensive price category may induce all three types of consumers to buy the promoted product. However, the promotion of a lower priced product is unlikely to attract consumers from the second category, suggesting the immediate effect to be lower

(given that the regular consumer base is equal across the different price segments), and the post-promotion dip to be larger. In addition, lower-income (and hence, more price-sensitive) shoppers may constitute a larger fraction of the consumer population of less expensive brands (Raju 1992). These people are more inclined to buy the brand on discount and save money by stockpiling. This indicates opposite immediate and cumulative effects.

Finally, brands in a lower price segment probably attract more consumers with tight budget constraints who therefore are more focused on prices and also are more price sensitive. So, we expect the immediate and the long-run effects of regular price to be higher for lower priced brands within a category.

### 3 Analyzing immediate and dynamic effects

In this section we present a modeling framework for estimating (the determinants of) the dynamic effect of price promotions on log sales, when the logarithm of sales is unit-root stationary. We first consider a model for one product category. We focus the discussion on price effects, but the methodology holds for any set of marketing instruments. In Section 4 the model will be extended to capture multiple categories.

In recent literature on market structures it has been shown that marketing efforts, such as temporary price promotions, do not have permanent effects on sales. A prerequisite for permanent effects of temporary promotions is the non-stationarity of sales. Srinivasan et al. (2000), Nijs et al. (2001), and Pauwels et al. (2002), among others, have shown that, in the categories considered, almost all log sales series for fast moving consumer goods are stationary. This result is not surprising as a unit root in log sales implies that all frequent temporary price promotions will lead to permanent increases in sales, which seems an unrealistic assumption. Hence, to study dynamic effects of temporary price promotions, it is more interesting to examine the *cumulative effect* of a temporary price promotion on current and future log sales than to study the permanent effect.

For the regular price, changes are likely to be permanent. Even if sales are stationary, permanent changes in price may lead to permanent changes in the sales in the long run. Hence, for such variables it is most relevant to consider the effect of a permanent change on sales in the *long run*.

To describe the dynamic pattern in sales of brands in a product category we start with a vector autoregression with explanatory variables (VARX). Denote the sales of brand  $i$

at time  $t$  by  $S_{it}$ , for  $i = 1, \dots, I$  and  $t = 1, \dots, T$ , where  $I$  is the number of brands in the market. To model the vector of sales  $S_t = (S_{1t}, \dots, S_{It})'$ , we consider a VARX(1) model

$$\log S_t = \mu + \Gamma \log S_{t-1} + \sum_{k=1}^K (A_k \log X_{kt} + C_k \log X_{k,t-1}) + \varepsilon_t, \quad (1)$$

where  $\varepsilon_t \sim N(0, \Sigma)$  and  $\mu$  denotes a vector of intercept parameters. The vector  $X_{kt} = (X_{k1t}, \dots, X_{kIt})'$ ,  $k = 1, \dots, K$ , denotes an  $I$ -dimensional vector of a potential explanatory variable. For example,  $X_{kit}$  denotes the  $k$ th marketing-mix variable of brand  $i$  at time  $t$  (e.g. promotional price or regular price).  $A_k$  and  $C_k$  are  $I \times I$  parameter matrices. The diagonal elements of these matrices describe the own effect of the marketing variables, while the off-diagonal elements represent the cross effects.

It follows from (1) that the immediate effect of a change in  $X_k$  on the log sales is given by the elasticity

$$\frac{\partial S_t}{\partial X_{kt}} \frac{X_{kt}}{S_t} = \frac{\partial \log S_t}{\partial \log X_{kt}} = A_k. \quad (2)$$

Hence, the immediate effect is equal to  $A_k$  and does not depend on whether the change is permanent or temporary.

### Dynamic effect of permanent changes

To determine the dynamic effects of a marketing instrument ( $X_{kt}$ ) on sales, we solve (1) for  $\log S_t$  by repeated substitution. This results in

$$\log S_t = \Gamma^\tau \log S_{t-\tau} + \sum_{j=0}^{\tau-1} \Gamma^j (\mu + \sum_{k=1}^K (A_k \log X_{k,t-j} + C_k \log X_{k,t-j-1}) + \varepsilon_{t-j}). \quad (3)$$

Under the stationarity condition (that is, the eigenvalues of  $\Gamma$  are within the unit circle), the influence of log sales at time  $t - \tau$  on current log sales disappears for large  $\tau$  as  $\lim_{\tau \rightarrow \infty} \Gamma^\tau = 0$ . Next, if we set the explanatory variables at fixed values, that is,  $X_{kt} = X_k$  for all  $t$  and  $k = 1, \dots, K$ , it holds for  $\tau \rightarrow \infty$  that

$$\log S_t = (\mathbf{I} - \Gamma)^{-1} \mu + \sum_{k=1}^K (\mathbf{I} - \Gamma)^{-1} (A_k + C_k) \log X_k + \sum_{j=0}^{\infty} \Gamma^j \varepsilon_{t-j}, \quad (4)$$

where  $\mathbf{I}$  denotes the identity matrix. As  $E[\varepsilon_{t-j}] = 0$  for all  $j$ , the long-run expectation, assuming a constant marketing mix, of the vector of log sales given  $X_1, \dots, X_K$  equals

$$E[\log S | X_1, \dots, X_K] = (\mathbf{I} - \Gamma)^{-1} \mu + \sum_{k=1}^K (\mathbf{I} - \Gamma)^{-1} (A_k + C_k) \log X_k. \quad (5)$$

This expectation denotes the long-run relation between log sales and the explanatory variables. This relation is especially useful to determine the long-run effect of a permanent change in one of the explanatory variables. The size of the absolute values of the eigenvalues of  $\Gamma$  translates into the speed of convergence to the long-run equilibrium. The long-run elasticity of  $X_k$  on  $S$  is given by

$$\frac{\partial S}{\partial X_k} \frac{X_k}{S} = \frac{\partial \log S}{\partial \log X_k} = (\mathbf{I} - \Gamma)^{-1}(A_k + C_k) \equiv B_k. \quad (6)$$

The diagonal elements of  $B_k$  represent the elasticity of marketing-mix variable  $k$  of brand  $i$  on brand  $i$ , while the off-diagonal elements represent the cross elasticities. Using the terminology introduced before, if  $X_k$  gives the regular price then  $B_k$  gives the effect of a permanent price change on sales in the long run.

### Dynamic effect of temporary changes

It follows immediately from (3) that, under stationarity, a temporary change in one of the  $X_{kt}$  variables at time  $t$  has no impact on the sales at time  $t + j$  in the long run. This is due to the fact that the term  $\Gamma^j$  will be zero for large  $j$ . Only a permanent change in the value of a marketing instrument can have a permanent long-run effect on the sales, which then of course depends on the values of the relevant parameters. To summarize the dynamics for a temporary promotion, it is more interesting to measure the cumulative effect on future sales.

Under stationarity, it is straightforward to show that the cumulative effect of a temporary change in  $\log X_{kt}$  on current and future log sales is given by

$$\sum_{j=0}^{\infty} \frac{\partial \log S_{t+j}}{\partial \log X_{kt}} = \sum_{j=0}^{\infty} \Gamma^j (A_k + C_k) = (\mathbf{I} - \Gamma)^{-1} (A_k + C_k) = B_k. \quad (7)$$

To sum up, for temporary (price) promotions we interpret  $B_k$  as the cumulative effect, while for permanent actions, like permanent changes in the regular price, it measures the long-run effect. In the application below, we will see that in general the immediate effects of price promotions are larger in size than the cumulative effects. In other words, some of the immediate increases in sales due to a promotion might be compensated by lower sales in future periods. For the regular price it turns out that the immediate effect is smaller than the long-run effect.



## Error-correction specification

To analyze the dynamic effects of (permanent) changes in  $X_k$  the VARX representation in (1) is not directly suitable. Although we assume stationarity (eigenvalues of  $\Gamma$  within the unit circle), it is still difficult to directly interpret the parameters in a vector autoregressive model with current and lagged exogenous variables. Indeed, the parameters combine the immediate and dynamic effects of the explanatory variables on the dependent variables. Therefore, it is more useful to write the VARX model in an error-correction format, as this provides a direct link between the various effects of a marketing instrument on log sales and the relevant model parameters.

In the marketing literature the error-correction model has been used by, for example, Franses (1994) and Paap & Franses (2000) to distinguish the immediate from the dynamic effects. This approach is in contrast with studies by, for example, Mela et al. (1998) and Jedidi et al. (1999), where the dynamics enter through the model parameters. In these studies the preferences and marketing sensitivity of households may change as a consequence of (intensified) promotional activities. In this case the dynamic effect is defined as the impact of a promotion on the future, while accounting for the changes in individual behavior. In this paper we take a different approach and consider (aggregate) household behavior to be constant. The dynamics in sales are directly caused by feedback loops in household behavior.

To disentangle the immediate effects of  $X_{kt}$  on the log sales from the dynamic effects, that is, to allow for direct estimation of these effects, it is convenient to rewrite (1) in the format of an error-correction model (ECM), see Hendry et al. (1984), that is,

$$\Delta \log S_t = \mu + \sum_{k=1}^K A_k \Delta \log X_{kt} + \Pi \left[ \log S_{t-1} - \sum_{k=1}^K B_k \log X_{k,t-1} \right] + \varepsilon_t, \quad (8)$$

where  $\Pi = (\Gamma - \mathbf{I})$ ,  $B_k = (\mathbf{I} - \Gamma)^{-1}(A_k + C_k)$ , and  $\Delta$  where denotes the first-differencing operator, that is,  $\Delta y_t = y_t - y_{t-1}$ . Note that this transformation only involves a rearrangement of terms, that is, (8) is exactly equivalent to (1). The advantage of the ECM representation is that we can directly link explanatory variables to the immediate effects  $A_k$  and the dynamic effects  $B_k$ , as will be discussed in the next section. Under stationarity, parameter estimation in the ECM representation is hardly more difficult than in the VARX specification (1) as it just involves a parameter transformation.

Although the ECM in (8) only models the relation between two consecutive time periods, error-correction models are very well suited to analyze the long-run, see Granger

(1993) for a discussion. The long-run relation between log sales and the log  $X_{kt}$  variables is put in the error correction term  $(\log S_{t-1} - \sum_{k=1}^K B_k \log X_{k,t-1})$  and hence the long-run effects are given by  $B_k$ . That is, this parameter gives the marginal effect of a permanent change of log  $X_{kt}$  on the log sales in the long run. The parameter matrix  $\Pi$  contains the adjustment parameters and determines the speed of convergence to the long-run relation.

Finally, the autoregressive structure of our model does not automatically imply that changes in  $X_k$  have a dynamic effect on sales. A special case of the model is where  $A_k = B_k$  for all  $k$ . The ECM (8) then simplifies to a common factor model (Hendry et al. 1984), that is, (8) reduces to

$$(\log S_t - \sum_{k=1}^K A_k \log X_{kt}) = \mu + \Gamma(\log S_{t-1} - \sum_{k=1}^K A_k \log X_{k,t-1}) + \varepsilon_t. \quad (9)$$

A temporary change in  $X_{kt}$  now only has an effect on current log sales and not on future sales. Hence, the immediate effects are equal to the cumulative effects.

So far, we only considered an error-correction model for sales in a single product category. In the next section, we discuss the analysis for more than one category.

## 4 HB-Error Correction Model

In this section we discuss our model for the case of a large number of categories. This model also allows us to estimate the moderating effects of brand and category characteristics on price elasticities. In a separate subsection, we discuss the differences between our approach and the existing literature, where we take Nijs et al. (2001) as a representative paper.

### 4.1 Hierarchical Bayes Analysis

Let  $S_{ct}$  denote the  $I_c$ -dimensional vector of sales for category  $c$  in week  $t$ . Note that categories are allowed to have different numbers of brands  $I_c$ . The  $I_c$ -dimensional vectors  $X_{ckt}$  contain the  $k$ -th marketing mix variables for the brands in category  $c$  in week  $t$ . The error-correction model (8) for category  $c$  is given by

$$\Delta \log S_{ct} = \mu_c + \sum_{k=1}^K A_{ck} \Delta \log X_{ckt} + \Pi_c (\log S_{c,t-1} - \sum_{k=1}^K B_{ck} \log X_{ck,t-1}) + \varepsilon_{ct}, \quad (10)$$

with  $\varepsilon_{ct} \sim N(0, \Sigma_c)$  for  $c = 1, \dots, C$  and  $t = 1, \dots, T_c$ . Note that we allow for different intercepts  $\mu_c$ , immediate  $A_c$  and dynamic  $B_c$  effects and variance matrices for the error

terms  $\Sigma_c$ , across categories. The adjustment parameters  $\Pi_c$  are also allowed to be different across categories. The categories may even have a different number of brands and observations.

To relate the immediate and dynamic elasticity parameters to explanatory variables we collect the parameters describing the effects of marketing-mix variables of brand  $i$  on the sales of brand  $i$  (as we focus on the own effects) in the  $I_c$ -dimensional vectors  $\alpha_{ck} = \text{diag}(A_{ck}) = (\alpha_{1ck}, \dots, \alpha_{I_c ck})'$  and  $\beta_{ck} = \text{diag}(B_{ck}) = (\beta_{1ck}, \dots, \beta_{I_c ck})'$  for  $k = 1, \dots, K$ . The immediate and dynamic elasticities will obviously differ across brands and across categories. Some of these differences can be attributed to observable characteristics of the brand and/or category, such as depth and frequency of promotion or perishability of the product, as we have discussed in Section 2.1. Another part of the differences across elasticities cannot be explained. In sum, we propose to describe the immediate and dynamic elasticity parameters by

$$\alpha_{ick} = \lambda'_{1k} z_{ic} + \eta_{ick} \quad (11)$$

$$\beta_{ick} = \lambda'_{2k} z_{ic} + \nu_{ick}, \quad (12)$$

where  $z_{ic}$  is an  $L$ -dimensional vector containing an intercept and  $L - 1$  explanatory variables for brand  $i$  in category  $c$ , like frequency and depth of promotion and category competitiveness. The  $L$ -dimensional vectors  $\lambda_{1k}$  and  $\lambda_{2k}$  describe the effects of the brand characteristics on the immediate and the dynamic elasticities, respectively. The error terms  $\eta_{ick}$  and  $\nu_{ick}$  have zero mean and are assumed to be uncorrelated across brands and categories. We do however allow for correlation in the error terms across the  $k$  marketing-mix variables, that is, we assume that  $\eta_{ic} = (\eta_{ic1}, \dots, \eta_{icK})' \sim N(0, \Sigma_\eta)$  and  $\nu_{ic} = (\nu_{ic1}, \dots, \nu_{icK})' \sim N(0, \Sigma_\nu)$ .

We will abbreviate the model above as HB-ECM. Akçura et al. (2004) and Montgomery et al. (2004) use vector autoregressive structure in a Bayesian setting for latent utility variables, but as far as we know we are the first to use a vector autoregressive model in error correction format for sales in a Hierarchical Bayes setting. To estimate the parameters in the model (10) with (11)–(12), we use a Bayesian approach. Bayesian estimation provides exact inference in finite samples. To obtain posterior results we use the Markov Chain Monte Carlo (MCMC) simulation method. In Appendix A we derive the likelihood function of the model together with the full conditional posterior distributions which are necessary in the Gibbs sampler.

Another estimation strategy which is often applied in practice, is a two-step proce-

ture in which, first, individual market-level models are estimated and, in a second stage regression, the parameters from the market-level models are related to brand and market characteristics, see, for example, Nijs et al. (2001). This method is however theoretically less elegant as the uncertainty in the first-level parameter estimates is not correctly accounted for in the second stage, and vice versa. In finite samples, this leads to underestimation of the uncertainty in the parameter estimates in the second stage.

## 4.2 Comparison with existing literature

Before we continue with a discussion of our empirical results, we list the main differences between our approach and the current standard approach in the literature, where we take Nijs et al. (2001) as a representative study.

First of all, Nijs et al. (2001) obtain the dynamic effects from Impulse Response Functions [IRFs] based on estimated Vector Autoregression [VAR] models. This incorporates possible competitive and feedback effects of competitors. In contrast, our measure of the dynamic effect in the ECM excludes these effects. This enables us to have a better focus on the determinants of *dynamic demand reactions*. In our view, competitive reactions and demand (consumer) reactions to promotions are two topics that need to be addressed separately. A proper understanding of consumer response to changes in marketing variables should precede the analysis of competitive reactions. In the end the consumer will see a competitive reaction just as another promotion. Based on the outcomes of our model and given a likely competitive reaction, one can easily judge the net effect of the promotion.

Second, the HB-ECM allows us to disentangle the immediate and dynamic effects into separate estimable parameters. This enables us to relate these effects directly to store and brand characteristics. Hence, we do not have to rely on derivative measures like the IRF, or on a two-step approach to relate immediate and dynamic effects to moderating variables. Through this error-correction specification, we also avoid one of the disadvantages of the two-step procedure. This two-step procedure does not appropriately account for the uncertainty in the first-level parameter estimates when estimating the second stage model. In finite samples, this leads to underestimation of the standard errors of the parameters in the second-stage regression. Furthermore, the ECM model allows us to judge the accuracy of our dynamic effects estimates, as standard errors for these estimates are easily obtained. Nijs et al. (2001), for example, compute the cumulative effects of promotions using accumulated impulse response functions. As these accumulated impulse response functions are non-linear functions of the model parameters,

the uncertainty in the estimated effect is usually difficult to compute. More importantly, the uncertainty is often quite large when compared to the uncertainty in model parameter estimates. From an efficiency point of view, it is therefore more reliable to directly estimate a parameter representing the long-run effects than to rely on the impulse responses.<sup>1</sup>

Another difference is that we investigate category *and* brand characteristics that may affect price elasticities. Finally, besides focusing on the determinants of promotional price elasticities, we also consider the effects of changes in the regular price. Most of the marketing literature so far mainly focuses on the determinants price promotion elasticities.

## 5 Empirical results

In this section, we use our HB-ECM to explain differences in immediate and dynamic effects of promotional price and regular price on sales across brands and product categories. In Section 5.1 we discuss the product categories we consider in our analysis and the available data. Section 5.2 contains the estimation results.

### 5.1 Data and Variables

The data we consider are weekly sales volumes of fast moving consumer goods in 25 product categories. The data are obtained from the database of a large supermarket chain, Dominick's Finer Foods, which are collected in the Chicago area in the period September 1989 to May 1997. Sales are aggregated from SKU to brand level using static weights as described in Srinivasan et al. (2004), who use the same data set. In a number of cases aggregation leads to biased results. For example, Christen et al. (1997) show that one cannot straightforwardly obtain unbiased estimates of promotion elasticities from data that is aggregated across stores. In our case we use a similar aggregation, that is, from SKUs to brands. However, as we are explicitly interested in brand-level elasticities the aggregation does not pose problems here. Although using brand-level data one does not retrieve the average promotion effectiveness across all SKUs, the brand-level estimate still provides useful information. This estimate can be interpreted as a promotion elasticity,

---

<sup>1</sup>Note that in a multiplicative model, which is the most frequently used model in marketing and the one also applied by Nijs et al. (2001), the sum of the impulse responses to changes in the logs of variables over the dust-settling period has no straightforward interpretation, see Wieringa & Horváth (2005) for a recent account.

it gives the percentage change in brand sales due to a percentage price change. One can of course argue about the specification of the brand-level model, as it does not match the aggregation of separate log-linear SKU-level models.

Our data concern the product categories: bottled juice, cereals, cheese, cookies, crackers, canned soup, dish detergent, front-end candies, frozen dinners, frozen juice, fabric softener, laundry detergents, oatmeal, paper towels, refrigerated juice, soft drinks, shampoos, snack crackers, toothbrushes, canned tuna, toothpaste and bathroom tissue.

In each product category we take only the top four brands, and hence we have  $4 \times 25 = 100$  different brands. We specify 25 error-correction models as in (10). The dependent variable  $S_t$  consists of the total weekly sales of the brands in the separate product categories. As explanatory variables we consider the marketing-mix variables, display, feature, regular price and promotional price indexes.

The original database only contains the actual price. To decompose the actual price series into regular and promoted price we smooth the actual price series using cubic splines with asymmetric weights. In the smoothing algorithm positive errors are weighted ten times stronger than negative errors. In this way we construct a series that follows the actual price in case of no promotion and it does not follow temporary drops in price. Sustained drops however are reflected in the regular price. For some categories the actual price shows seasonal variation, and we then include seasonal dummies in the smoothing algorithm for these categories. To measure price discounts we use a price index, that is, the actual price divided by the regular price. This price index is a natural measure for the size of a promotion and it also allows for a comparison across categories.

The immediate and dynamic own effects of the price index and the regular price, denoted by  $\alpha_{ick}$  and  $\beta_{ick}$ , are explained by characteristics of the brand and product category in the second stage of the model as explained in (11) and (12).

In principle, the display and feature variables reflect the percentage of stock-keeping units of the brand that are promoted in a given week. However, these variables turn out not to be consistently monitored in this database. In some cases a display or feature may not have been recorded. The variables are therefore an imperfect measure for the actual promotional actions. To account for the effects of display and feature as much as possible, we do however include these variables in the first layer of our model but the effects are not included as dependent variables in the second layer. We will not discuss their estimates as their interpretation can be unclear. Concerning cross effects of marketing instruments, we only allow for cross promotional price effects.

Finally, we control for seasonal variation in the sales series. To account for possible seasonality in the sales series, we include 13 seasonal dummies in the model, which each cover 4 consecutive weeks. The starting point of the period of 4 consecutive weeks is chosen in such a way that it produces the best fit. If necessary, we also include dummies to capture special holidays (Easter, Memorial Day, Christmas) and lagged values of these dummies to deal with the dynamic effects of these events. This pre-analysis is done per category. Unit root analysis shows that all sales series are (trend) stationary after correcting for possible seasonality and possible breaks in the regular price series.

To summarize, we only include the  $\alpha_{ick}$  and  $\beta_{ick}$  of the promotional price index and regular price in the second layer of the HB-ECM. As explanatory variables of  $\alpha_{ick}$  and  $\beta_{ick}$ , we use brand-level characteristics and category-level characteristics. Note that to construct these variables we need to make use of the sales data. Strictly speaking this may induce an endogeneity problem. However this will not be a problem as only summary statistics of sales are used and we also only use the model for descriptive purposes.<sup>2</sup> Furthermore, although the recording of display and feature in the database is not perfect, we do use these variables to obtain a measure of the relative frequency of these promotions. The underlying assumption here is that the reporting process is the same across brands and categories. A summary and the formal definition of these variables can be found in Appendix B.

## 5.2 Estimation results

The HB-ECM is analyzed using Gibbs sampling as presented in Appendix A. Posterior results are based on 200,000 draws of which the first 100,000 are used as burn in. To remove correlation in the chain we only consider every 10th draw to compute posterior results. The unreported plots of the draws of the model parameters of the second layer (11) and (12) show that the Markov Chain has converged.

First, we summarize the posterior means of the effects of the (log) price index and the log regular price in graphs. Figure 1 presents the distribution of the posterior means of the immediate effect of price promotions and of regular price changes, the cumulative effect of price promotions, and the long-run effect of a regular price change. These histograms show

---

<sup>2</sup>Our model is a descriptive model. Following Franses (2005) we apply diagnostics on residual autocorrelation for the ECM models, prior to estimating the HB two-level specification. There are no strong indications that we need to modify the dynamic structure of the models. For the sake of coherence and interpretation, we specify first-order dynamics for all equations.

the posterior means across all brands and all product categories. Overall the dispersion in the dynamic effects tends to be smaller than that in the immediate effects. For price promotions the sign of all posterior means is according to our expectations. Both the immediate and the cumulative effects are negative for all brands in all product categories. Overall, the cumulative effect tends to be smaller in absolute value than the immediate effect. In general, some of the positive effects of a price promotion are compensated in the periods following the promotion by, for example, the effects of stockpiling. This finding can be quantified by considering the posterior probability that the magnitude of the immediate effect is larger than the cumulative effect. This posterior probability, over all brands and categories, is 0.75.

For the regular price the graphs show much more dispersion in the posterior means. For the immediate and the long-run effect there are brands with a positive regular price effect as well as brands with a negative effect. However, the mean immediate effect and the mean cumulative effect over all brands are negative. The graphs also seem to indicate that one cannot a priori indicate if the immediate effect of regular price is larger or smaller than the long-run effect. The posterior probability that the immediate effect is larger in magnitude than the long-run effect is 0.57, which confirms our initial idea.

When we compare the promotional price elasticity with the regular price elasticity we find that overall the regular price elasticity tends to be closer to 0. For the immediate effect, the posterior probability that the promotional price elasticity has a larger magnitude compared to the regular price elasticity is 0.72. The corresponding probability for the dynamic effect is 0.78.

### **Moderating factors of price promotion elasticities**

Now we turn to the second layer of our HB-ECM, where we explain differences in the effects of price promotions and regular price on sales. Table 2 presents the posterior means and posterior standard deviations of the parameters in the second level of our model, that is, (11) and (12). This table gives the determinants of the immediate and cumulative effects of price promotions and the immediate and long-run effects of regular price changes.

First of all, we focus on the determinants of the promotional price effects, as in the literature this has received almost exclusive attention. There are quite some characteristics that significantly influence the effectiveness of price promotions. The price segment, for example, has a negative influence on the immediate promotion price elasticity. This means



that brands in a higher price segment will have stronger promotional price effects. The effect of brand size corresponds with our conjecture in Section 2.1. Larger brands tend to have smaller immediate and cumulative price effectiveness.

An interesting finding is that while price promotion frequency for a category does not seem to influence the immediate or the cumulative effectiveness of a price promotion, the relative frequency of price promotions of a brand within a category does influence the cumulative promotional price elasticity. It seems that the promotion frequency is most important within a category. Note that Nijs et al. (2001) do find a significant effect across categories, while Bolton (1989) does not. In general, brands with a high frequency tend to have a small cumulative elasticity. The average depth of the price promotions in a category however does have a strong influence on the price promotion elasticity. Deeper price promotions correspond to stronger price effects. Even after controlling for differences in the depth of promotions across categories, the relative depth of promotion for a particular brand enforces the elasticity. This again holds for the immediate as well as for the cumulative effect. These findings correspond to our hypothesis and to the findings of Raju (1992) and of Foekens et al. (1999) for their brand B.

In Section 2.1 we hypothesized that brand sales are more elastic for categories and brands that are frequently featured in flyers and newspapers. This turns out to be true only for the category feature frequency and the immediate promotional price elasticity. Again, the results on the immediate effects coincide with the findings of Bolton (1989). Consumers are more aware of the prices in a category with frequent price-oriented advertising. Such a category attracts price-sensitive consumers, and households will be more inclined to stockpile (as the higher immediate effect is partly offset in the long-run). Note that within a category there is no effect of the price promotion frequency.

Interestingly, only the relative display activity within a category is relevant. Brand sales are more elastic for brands with a relatively high display frequency (opposite to the findings of Bolton 1989). This holds for the immediate as well as the cumulative effect.

We also find that more perishable product categories or categories with a low budget share tend to have a smaller immediate promotional price elasticity. Categories that are characterized by a large dispersion in prices in general have smaller immediate and cumulative promotional price effects. Finally, categories with a high market concentration tend to have stronger cumulative price elasticities.

## Moderating factors of regular price elasticities

For the relation between brand/category characteristics and the regular price elasticity we cannot rely on previous literature. We therefore will only list some interesting findings.

The first interesting finding, but probably not unexpected, is that across all brands, the mean immediate effect of regular price is close to zero. In fact, we hardly find any characteristics that correlate with the immediate regular price effect. In practice, the regular price will only change gradually. Such changes will not attract new customers immediately. Only in the long run customers will get used to the new price and the effect of regular price changes are therefore expected to be relevant only in the long run.

For utilitarian categories we find a strong (negative) regular price elasticity. For this type of product, decreases in regular price lead to immediate increases in sales. Two other category characteristics are marginally important. The price dispersion in a category is positively correlated with the immediate regular price effect. Categories with a high price dispersion may even have a positive regular price elasticity. In this case differences in regular price may reflect differences in product quality. An increase in regular price may be a signal for a higher quality brand, as such a regular price increase may lead to higher sales. Finally, more categories in which the product are highly perishable tend to have stronger immediate regular price elasticities.

For the long-run effect we find a substantial mean effect, that is, the value of zero is not contained in the 99% highest posterior density region of the intercept. We also find more relevant characteristics. Surprisingly, again none of the brand level characteristics turn out to be relevant. Four of the category characteristics are relevant for the long-run regular price elasticity.

Utilitarian products tend to have smaller long-run regular price elasticities. Earlier, we found a large immediate effect for these products. Although the immediate effect may be quite large, the long-run effect of decreases in regular price is much smaller. In categories that have a high price dispersion, or have relatively frequent price promotions, the long-run effects of regular price tend to be stronger. Also for the long-run regular price effect, we find perishability to be important. Contrary to the immediate effect, we find that, for the long run, perishable products have smaller regular price effects.

## 6 Conclusions

In this paper we proposed a Hierarchical Bayes - Error Correction Model to explain the differences in immediate and dynamic effects of price on sales. We explicitly distinguished between the effects of *temporary* price promotions and *permanent* changes in regular price. The model is applied to weekly sales for 100 different brands in 25 product categories. In the second layer of the model the immediate and dynamic effects of price promotions and changes in regular price are related to brand-level and category-level characteristics. Parameter estimates were obtained using MCMC.

The HB approach allows us to analyze the dynamic effects of price in a statistically coherent way. Our results show that price elasticities can be explained by several brand-specific and category-specific factors. We find that many of the results for the immediate effect of price promotions are in line with previous literature and that, although in most cases the influence of these factors on cumulative effects is somewhat lower, it is statistically significant and has the same sign as the effect on immediate elasticities. The dynamic effects are smaller but these do not cancel out from the relationship between price promotion elasticity and category and brand characteristics.

We find mostly significant moderating effects on the elasticity of price promotions. We also find significant effects for the cumulative elasticity. Brands in categories that are characterized by high price differentiation and that constitute a lower share of budget are less sensitive to price discounts. Deep price discounts in a category or for a brand turn out to increase the immediate price sensitivity of customers. Another interesting finding is that while the relative display frequency within a category influences the promotional price elasticity, the feature frequency only matters across categories. A possible explanation is that while the use of feature mainly generates a general need for the product category, display attracts the attention of the buyer when making the actual brand choice.

For changes in regular price we hardly find any relevant explanatory variables. As perhaps expected, the immediate effect of a regular price change is often close to zero. The long-run effect of such a change usually leads to an increase in sales. This is especially true for categories characterized by a large price dispersion, frequent price promotions and hedonic, non-perishable categories.

Our study can be extended in several ways. One possibility is to consider moderating factors of cross-brand elasticities. This would be especially interesting with data on brands of the same manufacturer or data at the SKU level. Second, as pointed out earlier, we do

not consider competitive reactions or feedback effects in our ECM model, which allows us to focus purely on the determinants of dynamic demand reactions. It might however be interesting to see how the results would change if we were to build a model in which these relationships were considered.

## A Bayes estimation

To analyze the HB-ECM, we consider the exact likelihood function. We put the first observations in each category equal to the long-run equilibrium, that is,

$$\log S_{c1} = -\Pi_c^{-1}\mu_c + \sum_{k=1}^K B_{ck} \log X_{ck1} + \varepsilon_{c1} \quad (13)$$

with  $\varepsilon_{1c} \sim N(0, V_c)$ , where  $V_c$  is the long-run variance<sup>3</sup>

To derive the likelihood function, we summarize the elements of  $A_k$  and  $B_k$  which we relate to explanatory variables, in the  $K$ -dimensional row vectors  $\alpha_{ic} = [\alpha_{ick}]_{k=1}^K$  and  $\beta_{ic} = [\beta_{ick}]_{k=1}^K$ . The equations (11) and (12) can be written in matrix notation

$$\alpha_{ic} = \Lambda'_1 z_{ic} + \eta_{ic} \quad (14)$$

$$\beta_{ic} = \Lambda'_2 z_{ic} + \nu_{ic} \quad (15)$$

for  $i = 1, \dots, I_c$ , where the  $L \times K$  matrices  $\Lambda_1$  and  $\Lambda_2$  contain the vectors  $\lambda_{1k}$  and  $\lambda_{2k}$ , respectively. The likelihood function of the model is given by

$$\prod_{c=1}^C \int_{\alpha_c, \beta_c} \phi(\varepsilon_{c1}; 0, V_c) \prod_{t=2}^{T_c} \phi(\varepsilon_{ct}; 0, \Sigma_c) \prod_{i=1}^{I_c} \phi(\alpha_{ic}; \Lambda'_1 z_{ic}, \Sigma_\eta) \phi(\beta_{ic}; \Lambda'_2 z_{ic}, \Sigma_\nu) d\alpha_c d\beta_c, \quad (16)$$

where  $\phi(x; \mu, \Sigma)$  is the density function of the multivariate normal distribution with mean  $\mu$  and variance  $\Sigma$  evaluated at  $x$ , and where  $\alpha_c = (\alpha'_{1c}, \dots, \alpha'_{I_c c})'$  and  $\beta_c = (\beta'_{1c}, \dots, \beta'_{I_c c})'$ .

To obtain posterior results, we use the Gibbs sampling technique of Geman & Geman (1984) with data augmentation, see Tanner & Wong (1987). An introduction into the Gibbs sampler can be found in Casella & George (1992), see also Smith & Roberts (1993) and Tierney (1994). Hence, the latent variables  $\alpha_c$  and  $\beta_c$  are sampled alongside the model parameters  $\{\{A_{ck}\}_{k=1}^K, \{B_{ck}\}_{k=1}^K, \mu_c, \Pi_c, \Sigma_c\}_{c=1}^C, \Lambda_1, \Lambda_2, \Sigma_\eta$  and  $\Sigma_\nu$ . The Bayesian analysis is based on uninformative priors for the model parameters. To improve convergence of the MCMC sampler we impose inverted Wishart priors on the  $\Sigma_\eta$  and  $\Sigma_\nu$  parameter with scale parameter  $\kappa_1 \mathbf{I}_K$  and degrees of freedom  $\kappa_2$ . We set the value of  $\kappa_1$  to  $\frac{1}{1000}$  and  $\kappa_2$  equal to 1 such that the influence of the prior on the posterior distribution is marginal, see Hobert & Casella (1996) for a discussion.

In the remainder of this appendix we derive the full conditional posterior distributions of the model parameters and the latent variables  $\alpha_c$  and  $\beta_c$ . In deriving the sampling distributions we build on the results in Zellner (1971, Chapter VIII).

---

<sup>3</sup>The long-run variance follows from (3) and is given by  $V_c = \sum_{j=0}^{\infty} \Gamma_c^j \Sigma_c (\Gamma_c')^j$ , where  $\Gamma_c = \mathbf{I} + \Pi_c$ . The variance is finite if the eigenvalues of  $\Gamma_c$  are within the unit circle, that is, in case of stationarity.

## Sampling of $\Pi_c$

The full conditional posterior distribution of  $\Pi_c$  is not of a known family. To sample  $\Pi_c$  we use a Metropolis-Hastings sampler of Metropolis et al. (1953) and Hastings (1970). To obtain a candidate sampling distribution we use that (10) is just a multivariate regression model. We rewrite (10) as

$$\Delta \log S_{ct} - \mu_c - \sum_{k=1}^K A_{ck} \Delta \log X_{ckt} = \Pi_c (\log S_{c,t-1} - \sum_{k=1}^K B_{ck} \Delta \log X_{ck,t-1}) + \varepsilon_{ct}. \quad (17)$$

This equation is a multivariate regression model with a normal distributed error term and regression parameter matrix  $\Pi_c$ . Hence, if we neglect the model for the initial observation (13) the full conditional posterior distribution of  $\Pi_c'$  will be matrix normal with mean

$$\hat{\Pi}'_c = \left( \sum_{t=2}^{T_c} W_{ct} W_{ct}' \right)^{-1} \left( \sum_{t=2}^{T_c} W_{ct} Y_{ct}' \right) \quad (18)$$

and variance

$$\hat{\Sigma}_{\Pi'_c} = \left( \Sigma_c \otimes \left( \sum_{t=2}^{T_c} W_{ct} W_{ct}' \right)^{-1} \right), \quad (19)$$

with  $Y_{ct} = \Delta \log S_{ct} - \mu_c - \sum_{k=1}^K A_{ck} \Delta \log X_{ckt}$  and  $W_{ct} = \log S_{c,t-1} - \sum_{k=1}^K B_{ck} \Delta \log X_{ck,t-1}$ . We will use this distribution as the candidate for the M-H sampler. We denote the sampled candidate by  $\Pi_c^{\text{cand}}$ .

As we cannot neglect the model for the first observation, the true full conditional posterior density of  $\Pi_c$  is proportional to the matrix normal candidate density provided above and the density of the first observation (13). This allows us to construct a particular form of the Metropolis-Hastings sampler which is known as the Independent sampler. As the candidate density is part of the target density (full conditional posterior density), the acceptance-rejection probability simplifies to

$$\frac{\phi(\varepsilon_{1c}; 0, V_c)|_{\Pi_c = \Pi_c^{\text{cand}}} \phi(\Pi_c^{\text{cand}'}; \hat{\Pi}'_c, \hat{\Sigma}_{\Pi'_c}) \phi(\Pi_c^{\text{old}'}; \hat{\Pi}'_c, \hat{\Sigma}_{\Pi'_c})}{\phi(\varepsilon_{1c}; 0, V_c)|_{\Pi_c = \Pi_c^{\text{old}}} \phi(\Pi_c^{\text{old}'}; \hat{\Pi}'_c, \hat{\Sigma}_{\Pi'_c}) \phi(\Pi_c^{\text{cand}'}; \hat{\Pi}'_c, \hat{\Sigma}_{\Pi'_c})} = \frac{\phi(\varepsilon_{1c}; 0, V_c)|_{\Pi_c = \Pi_c^{\text{cand}}}}{\phi(\varepsilon_{1c}; 0, V_c)|_{\Pi_c = \Pi_c^{\text{old}}}}, \quad (20)$$

where  $\Pi_c^{\text{old}}$  denotes the previous draw and  $\varepsilon_{c1} = \log S_{c1} + \Pi_c^{-1} \mu_c - \sum_{k=1}^K B_{ck} \log X_{ck1}$ , see Chib & Greenberg (1995) for a similar approach in an exact likelihood analysis of an autoregressive model.

## Sampling of $\Sigma_c$

The full conditional posterior distribution of  $\Sigma_c$  is not of a known family. To sample  $\Sigma_c$  we use again a Metropolis-Hastings sampler. To obtain a candidate sampling distribution

we use that (10) is just a multivariate regression model. Hence, if we neglect the model for the first observation (13), the full conditional posterior distribution of  $\Sigma_c$  is an inverted Wishart distribution with scale parameter  $\sum_{t=2}^{T_c} \varepsilon_{ct} \varepsilon'_{ct}$  and  $T_c - 1$  degrees of freedom, where  $\varepsilon_{ct} = \Delta \log S_{ct} - \mu_c - \sum_{k=1}^K A_k \Delta \log X_{ckt} - \Pi_c(\log S_{c,t-1} - \sum_{k=1}^K B_{ck} \Delta \log X_{ck,t-1})$ . This will be the candidate distribution of the M-H step, which provides us with  $\Sigma_c^{\text{cand}}$ .

As we cannot neglect the first observation, the true full conditional posterior density of  $\Sigma_c$  is proportional to the Inverted Wishart candidate density and the density of the first observation (13). This leads again to the Independent sampler variant of the Metropolis-Hastings sampler. As the candidate density is part of the target density (full conditional posterior density), the acceptance-rejection probability simplifies to

$$\frac{\phi(\varepsilon_{1c}; 0, V_c)|_{\Sigma_c = \Sigma_c^{\text{cand}}}}{\phi(\varepsilon_{1c}; 0, V_c)|_{\Sigma_c = \Sigma_c^{\text{old}}}}, \quad (21)$$

where  $\Sigma_c^{\text{old}}$  denotes the previous draw of  $\Sigma_c$ .

## Sampling of $\Lambda_1$ and $\Lambda_2$

To sample  $\Lambda_1$ , we note that we can write (14)

$$\alpha'_{ic} = z'_{ic} \Lambda_1 + \eta'_{ic}, \quad (22)$$

and hence it is a multivariate regression model with regression matrix  $\Lambda_1$ . Hence, the full conditional posterior distribution of  $\Lambda_1$  is a matrix normal distribution with mean

$$\left( \sum_{c=1}^C \sum_{i=1}^{I_c} z_{ic} z'_{ic} \right)^{-1} \left( \sum_{c=1}^C \sum_{i=1}^{I_c} z_{ic} \alpha_{ic} \right), \quad (23)$$

and covariance matrix

$$\left( \Sigma_\eta \otimes \left( \sum_{c=1}^C \sum_{i=1}^{I_c} z_{ic} z'_{ic} \right)^{-1} \right). \quad (24)$$

The derivation of the sampling distribution of  $\Lambda_2$  proceeds in the same manner. The full conditional posterior distribution of  $\Lambda_2$  is a matrix normal distribution with mean

$$\left( \sum_{c=1}^C \sum_{i=1}^{I_c} z_{ic} z'_{ic} \right)^{-1} \left( \sum_{c=1}^C \sum_{i=1}^{I_c} z_{ic} \beta_{ic} \right), \quad (25)$$

and covariance matrix

$$\left( \Sigma_\nu \otimes \left( \sum_{c=1}^C \sum_{i=1}^{I_c} z_{ic} z'_{ic} \right)^{-1} \right). \quad (26)$$

## Sampling of $\Sigma_\eta$ and $\Sigma_\nu$

To sample  $\Sigma_\eta$  we note that (14) is a multivariate regression model. Hence the full conditional posterior distribution of  $\Sigma_\eta$  is an inverted Wishart distribution with scale parameter  $\kappa_1 \mathbf{I}_K + \sum_{c=1}^C \sum_{i=1}^{I_c} (\alpha_{ic} - \Lambda'_1 z_{ic})(\alpha_{ic} - \Lambda'_1 z_{ic})'$  and degrees of freedom  $\kappa_2 + \sum_{c=1}^C I_c$ . The  $\kappa$  terms results from the inverted Wishart prior on  $\Sigma_\eta$  which is used to improve convergence of our Gibbs sampler, see Hobert & Casella (1996) for a discussion.

The sampling of  $\Sigma_\nu$  can be done in exactly the same manner. The parameter  $\Sigma_\nu$  is sampled from an inverted Wishart distribution with scale parameter  $\kappa_1 \mathbf{I}_K + \sum_{c=1}^C \sum_{i=1}^{I_c} (\beta_{ic} - \Lambda'_2 z_{ic})(\beta_{ic} - \Lambda'_2 z_{ic})'$  and degrees of freedom  $\kappa_2 + \sum_{c=1}^C I_c$ .

## Sampling of $\mu_c$ and cross effects in $A_{ck}$ and $B_{ck}$

To sample  $\mu_c$ , and the parameters measuring the cross effect in  $A_{ck}$  and  $B_{ck}$  we first split up  $X_{ckt} = (X_{ck1t}, \dots, X_{ckI_{ct}})'$  for  $k = 1, \dots, K$  into two parts  $X_{ct}^{\text{own}} = [X_{ckit}]_{k=1}^K$  and  $X_{ct}^{\text{cross}} = [[X_{ckjt}]_{j=1 \neq i}^{I_c}]_{k=1}^K$  to disentangle the own effects from the cross effects.

Define  $X_{ct}^{\text{own}} = \text{diag}(X_{c1t}^{\text{own}}, \dots, X_{cI_{ct}}^{\text{own}})'$  and  $X_{ct}^{\text{cross}} = \text{diag}(X_{c1t}^{\text{cross}}, \dots, X_{cI_{ct}}^{\text{cross}})'$ . Equation (13) and (10) can now be written as

$$\begin{aligned} \log S_{c1} - \log X_{c1}^{\text{own}} \beta_c &= -\Pi_c^{-1} \mu_c + \log X_{c1}^{\text{cross}} b_c + \varepsilon_{c1} \\ \Delta \log S_{ct} - \Delta \log X_{ct}^{\text{own}} \alpha_c & \\ -\Pi_c (\log S_{c,t-1} - \log X_{c,t-1}^{\text{own}} \beta_c) &= \mu_c + \Delta \log X_{ct}^{\text{cross}} a_c - \Pi_c \log X_{c,t-1}^{\text{cross}} b_c + \varepsilon_{ct}, \end{aligned} \quad (27)$$

where  $a_c$  and  $b_c$  capture the cross-effects in the matrices  $A_{ck}$  and  $B_{ck}$  for  $k = 1, \dots, K$ . This system can be written in a multivariate regression model

$$Y_{ct} = W_{ct} \gamma + \varepsilon_{ct}, \quad (28)$$

where  $Y_{ct}$  contains the left-hand side of (27),  $W_{ct}$  contains  $(-\Pi_c^{-1}; 0; \log X_{c1}^{\text{cross}})$  for the first observation and  $(\mathbf{I}_c; \Delta \log X_{ct}^{\text{cross}}; -\Pi_c \log X_{c,t-1}^{\text{cross}})$  for the remaining observations, and where  $\gamma = (\mu'_c, a'_c, b'_c)'$ . The error term is normal distributed with mean 0 and covariance matrix  $\Sigma_c$  ( $V_c$  for the first observation). Hence, the full conditional distribution of  $\gamma$  is normal with mean

$$\left( W'_{c1} V_c^{-1} W_{c1} + \sum_{t=2}^{T_c} W'_{ct} \Sigma_c^{-1} W_{ct} \right)^{-1} \left( W'_{c1} V_c^{-1} Y_{c1} + \sum_{t=2}^{T_c} W'_{ct} \Sigma_c^{-1} Y_{ct} \right) \quad (29)$$

and covariance matrix

$$\left( W'_{c1} V_c^{-1} W_{c1} + \sum_{t=2}^{T_c} W'_{ct} \Sigma_c^{-1} W_{ct} \right)^{-1}. \quad (30)$$



## Sampling of $\alpha_c$

To sample  $\alpha_c$  we rewrite the second equation of (27) as

$$\Delta \log S_{ct} - \mu_c - \Delta \log X_{ct}^{\text{cross}} a_c - \Pi_c (\log S_{c,t-1} - \sum_{k=1}^K B_{ck} \log X_{c,t-1}) = \Delta \log X_{ct}^{\text{own}} \alpha_c + \varepsilon_{ct}, \quad (31)$$

which can be written in matrix notation

$$Y_{ct} = W_{ct} \alpha_c + \varepsilon_{ct}, \quad (32)$$

where  $Y_{ct} = \Delta \log S_{ct} - \mu_c - \Delta \log X_{ct}^{\text{cross}} a_c - \Pi_c (\log S_{c,t-1} - \sum_{k=1}^K B_{ck} \log X_{c,t-1})$  and  $W_{ct} = \Delta \log X_{ct}^{\text{own}}$ . Furthermore, we write the  $I_c$  equations of (14) as

$$-U_c = -\mathbf{I}_{KI_c} \alpha_c + \eta_c, \quad (33)$$

where  $U_c$  is a  $(KI_c)$ -dimensional vector containing the terms  $\Lambda_1' z_{ic}$ ,  $i = 1, \dots, I_c$ , and where  $\mathbf{I}_{KI_c}$  is a  $(KI_c)$ -dimensional identity matrix. The error term  $\eta_c$  is normal distributed with mean 0 and covariance matrix  $(I_c \otimes \Sigma_\eta)$ . To sample  $\alpha_c$ , we combine (32) and (33)

$$\begin{aligned} \Sigma_c^{-1/2} Y_{ct} &= \Sigma_c^{-1/2} W_{ct} \alpha_c + \Sigma_c^{-1/2} \varepsilon_{ct} \\ -(\mathbf{I}_c \otimes \Sigma_\eta^{-1/2}) U_c &= -(\mathbf{I}_c \otimes \Sigma_\eta^{-1/2}) \alpha_c + (\mathbf{I}_c \otimes \Sigma_\eta^{-1/2}) \eta_c. \end{aligned} \quad (34)$$

Hence, the full conditional posterior distribution of  $\alpha_c$  is normal with mean

$$\left( (\mathbf{I}_c \otimes \Sigma_\eta^{-1}) + \sum_{t=2}^{T_c} (W_{ct}' \Sigma_c^{-1} W_{ct}) \right)^{-1} \left( (\mathbf{I}_c \otimes \Sigma_\eta^{-1}) U_c + \sum_{t=2}^{T_c} (W_{ct}' \Sigma_c^{-1} W_{ct}) \right), \quad (35)$$

and covariance matrix

$$\left( (\mathbf{I}_c \otimes \Sigma_\eta^{-1}) + \sum_{t=2}^{T_c} (W_{ct}' \Sigma_c^{-1} W_{ct}) \right)^{-1}. \quad (36)$$

## Sampling of $\beta_c$

To sample  $\beta_c$ , we rewrite (27) as

$$\begin{aligned} \log S_{c1} - \log X_{c1}^{\text{cross}} b_c - \Pi_c^{-1} \mu_c &= \log X_{c1}^{\text{own}} \beta_c + \varepsilon_{c1} \\ \Delta \log S_{ct} - \mu_c - \sum_{k=1}^K A_{ck} \Delta \log X_{ct} - \Pi_c (\log S_{c,t-1} - \log X_{c,t-1}^{\text{cross}} b_c) &= \log X_{c,t-1}^{\text{own}} \beta_c + \varepsilon_{ct}, \end{aligned} \quad (37)$$

which can be written in matrix notation

$$\begin{aligned} V_c^{-1/2} Y_{c1} &= V_c^{-1/2} W_{c1} \beta_c + V_c^{-1/2} \varepsilon_{c1} \\ \Sigma_c^{-1/2} Y_{ct} &= \Sigma_c^{-1/2} W_{ct} \beta_c + \Sigma_c^{-1/2} \varepsilon_{ct}, \end{aligned} \quad (38)$$

for  $t = 1, \dots, T_c$ , where  $Y_{ct}$  denotes the left-hand side of (37) and  $W_{ct}$  the right-hand side. Again, we write the  $I_c$  equations of (15) as

$$-(\mathbf{I}_c \otimes \Sigma_\nu^{-1/2})U_c = -(\mathbf{I}_c \otimes \Sigma_\nu^{-1/2})\beta_c + (\mathbf{I}_c \otimes \Sigma_\nu^{-1/2})\nu_c, \quad (39)$$

where  $U_c$  is a  $(KI_c)$ -dimensional vector containing the terms  $\Lambda'_2 z_{ic}$ ,  $i = 1, \dots, I_c$ . The distribution of the error term  $\nu_c$  is normal with mean 0 and covariance matrix  $(I_c \otimes \Sigma_\nu)$ . If we combine (38) with (39) it is easy to see that the full conditional posterior distribution of  $\beta_c$  is normal with mean

$$\left( (\mathbf{I}_c \otimes \Sigma_\nu^{-1}) + W'_{c1} V_c^{-1} W_{c1} + \sum_{t=2}^{T_c} (W'_{ct} \Sigma_c^{-1} W_{ct}) \right)^{-1} \left( (\mathbf{I}_c \otimes \Sigma_\nu^{-1})U_c + W'_{c1} V_c^{-1} Y_{c1} + \sum_{t=2}^{T_c} (W'_{ct} \Sigma_c^{-1} Y_{ct}) \right), \quad (40)$$

and covariance matrix

$$\left( (\mathbf{I}_c \otimes \Sigma_\nu^{-1}) + W'_{c1} V_c^{-1} W_{c1} + \sum_{t=2}^{T_c} (W'_{ct} \Sigma_c^{-1} W_{ct}) \right)^{-1}. \quad (41)$$

## B The definition of explanatory variables ( $z_{ic}$ )

In this appendix we list the category and brand characteristics that are used in our empirical section to explain the dynamic effects of promotions. We give a description of each variable and, if necessary, a formal, mathematical definition. The characteristics are organized based on the level at which they are defined (category level, brand level or both) and the concept they measure (eg. competitive intensity). In this appendix we use the following notation:

$S_{ict}$	Sales volume of brand $i$ in category $c$ at time $t$
$M_{ict} = S_{ict} / \sum_{i=1}^{I_c} S_{ict}$	Market share of brand $i$ at time $t$
$\overline{M}_{ic} = \frac{1}{T_c} \sum_{t=1}^{T_c} M_{ict}$	(Time) average market share
$P_{ict}$	(Actual) price of brand $i$ in category $c$ at time $t$
$RP_{ict}$	Regular price of brand $i$ in category $c$ at time $t$
$\overline{RP}_c = \frac{1}{I_c T_c} \sum_{t=1}^{T_c} \sum_{i=1}^{I_c} RP_{ict}$	Average regular price in category $c$
$PI_{ict} = P_{it} / RP_{it}$	(Promotional) Price index

**Average budget share** To measure the budget share of a category we use the average total expenditures in the category over time, that is,  $\frac{1}{T_c} \sum_{t=1}^{T_c} \sum_{i=1}^{I_c} S_{ict} P_{ict}$ . Note that we do not need a denominator (total budget) as we will standardize all characteristics in our empirical analysis.

**Utilitarian** To characterize the type of a product category we measure the utilitarian nature of the category. This characteristic is operationalized by three levels (low, middle, high) and the measures are obtained using a number of experts. The three levels are coded by the numbers 0, 0.5, and 1. The team of experts consisted of a food marketing expert and two marketing research experts. These experts had no relation to this research project. The product categories involved in this study were easy to classify, there was hardly any disagreement among the experts. When there was a disagreement the authors voted on the final classification. Further details on the classification can be obtained from the authors.

**Perishability** An important component of the storability of a category is the *perishability*. This characteristic is again defined using the opinions of a number of experts. This variable has three levels (low, middle, high).

**Market concentration** The market concentration in category  $c$  is measured using:

$$\sum_{i=1}^{I_c} \overline{M}_{ic} \log \overline{M}_{ic} \text{ (Raju 1992).}$$

**Price dispersion** The price dispersion in a category can be obtained by comparing the highest regular price in a category with the lowest regular price, that is, for category  $c$ :

$$\sum_{t=1}^{T_c} \left( \max_i (RP_{ict}) - \min_i (RP_{ict}) \right) / (T_c \cdot \overline{RP}_c).$$

**Price promotion frequency** We measure the price promotion frequency of a brand by counting the number of times that the price index within the category is below 0.95 (i.e. when there is at least 5% discount). See, for example, Mulhern et al. (1999), where a similar measure is used. The price promotion frequency of the category is obtained by counting the number of times that at least one brand's price index is below 0.95 (see Raju 1992). As explanatory variables we use the frequency of the category and the relative brand-level frequency, that is, brand frequency / category frequency.

**Depth of price promotions** For a brand we define the average depth of a promotion as  $\frac{\sum_{i=1}^{T_c} \log(PI_{it})}{FREQ_{ic}}$ , where  $FREQ_{ic}$  denotes the price promotion frequency of brand  $i$  in category  $c$ . On the category level we use the mean of the average brand-level depth of promotion. Raju (1992) uses very similar measures, however he uses the difference between the regular and actual price instead of a price index.

Again we include the category-specific measure as well as a relative measure of depth of price promotion as explanatory variables.

**Feature/Display frequency** The frequency of display or feature of brand  $i$  can simply be obtained by taking the average of the percentage of SKUs promoted by the brand over time.

Denoting the percentage of SKUs promoted by brand  $i$  in category  $c$  at time  $t$  by  $x_{ict}$ , we define the category level frequency of promotion in category  $c$  by

$$\frac{\sum_{t=1}^{T_c} 1 - \prod_{i=1}^{I_c} (1 - x_{it})}{T_c}.$$

The term  $(1 - x_{ict})$  can be interpreted as the probability that a SKU of brand  $i$  is *not* on promotion in week  $t$ . Following this reasoning (and assuming independence in the timing of promotions across brands) the probability that no SKU is on promotion is  $\prod_{i=1}^{I_c} (1 - x_{it})$ . Our measure for the frequency of promotion can therefore be seen as the average probability that at least one SKU is on promotion.

**Brand size** To measure the relative size of a brand in the category we use the average market share, where market shares are defined using the units sold.

**Price segment** As an indication of the price segment to which a brand belongs in a certain category we construct an index based on the regular price, that is,

$$\frac{1}{T_c} \sum_{t=1}^{T_c} \frac{RP_{ict}}{\sum_{i=1}^{I_c} RP_{ict}/I_c}.$$

A high value of the index corresponds with a relatively expensive brand within the category.

Table 1: Overview of the literature on the determinant of price promotion effectiveness

Study	Explanatory variables	Immediate and/or dynamic effects	Dependent variable	Approach
Bolton (1989)	Category+brand char.	Immediate	Price elasticity	Single equation (2-step)
Fader & Lodish (1990)	Category characteristics	Immediate	% of volume purchased on -price cut-feature activity -display activity -store coupon offer	Factor and cluster analysis
Raju (1992)	Category characteristics	Immediate	Variability in category sales	Single equation (1-step <sup>1</sup> )
Hoch et al. (1995)	Consumer+competitor char.	Immediate	Price elasticity	Single equation (2-step)
Shankar & Krishnamurthi (1996)	Retailer pricing policy and promotional variables	Immediate	Regular price elasticity	Single equation (3-step)
Narasimhan et al. (1996)	Category characteristics.	Immediate	Non-supported price elasticities Supported price elasticities with: feature & display activity	Single equation (1-step <sup>2</sup> )
Montgomery (1997)	Demographic+competitive characteristics of stores	Immediate	Price sensitivity	System (HB model)
Mulhern et al. (1999)	Brand and consumer char.	Immediate	Price elasticity	Single equation (2-step)
Bell et al. (1999)	Category, brand, and consumer char.	Immediate (primary vs. secondary demand effects)	Price elasticity	Single equation (2-step <sup>3</sup> )
Foekens et al. (1999)	Char. of price discount	Immediate+dynamic	Price, display/feature sensitivity	Single equation Varying parameters
Nijs et al. (2001)	Category characteristics	Immediate+dynamic	Price elasticity	VAR (2-step)
Wakefield & Imman (2003)	Product char. and social context of purchase	Immediate	Price elasticity	Single equation (1-step <sup>2</sup> )
Macé & Neslin (2004)	Product+category+store trading area char.	Immediate+ Pre- and postpromotion dips	Price elasticity	Single equation (2-step <sup>3</sup> )
<i>This study</i>	Category+brand char.	Immediate+dynamic	Promotional and regular price elasticity	VAR (HB model)

<sup>1</sup>They do not need a two-step approach because they do not need to build a statistical model to capture variability in category sales.

<sup>2</sup>This is actually a two-stage approach because they use price elasticities obtained from another study.

<sup>3</sup>They explicitly account for uncertainty in the first-level estimates in the second stage.

Table 2: Posterior means of the effects of covariates on immediate and dynamic effects of price promotions and regular price changes ( $\lambda_1$  and  $\lambda_2$  in 11 and 12), posterior standard deviations in parentheses.

Characteristic <sup>†</sup>	Price index				Regular price			
	Immediate effect		Cumulative effect		Immediate effect		Long-run effect	
Intercept	-2.291	(0.067) ***	-1.907	(0.078) ***	-0.542	(0.369)	-0.826	(0.128) ***
<i>Category level characteristics</i>								
Avg. budget share	-0.232	(0.088) ***	-0.191	(0.106) *	0.291	(0.669)	-0.259	(0.183)
Utilitarian	-0.090	(0.085)	0.131	(0.111)	-1.017	(0.488) **	0.503	(0.159) ***
Perishability	0.145	(0.085) *	-0.006	(0.097)	-0.723	(0.448) *	0.273	(0.157) *
Market concentration	-0.068	(0.075)	-0.193	(0.099) *	0.163	(0.364)	-0.173	(0.143)
Price dispersion	0.271	(0.077) ***	0.224	(0.093) **	0.760	(0.426) *	-0.308	(0.139) **
Price promotion frequency	-0.051	(0.094)	-0.118	(0.124)	0.569	(0.442)	-0.374	(0.172) **
Depth of price promotions	-0.234	(0.076) ***	-0.186	(0.074) **	-0.198	(0.361)	-0.167	(0.143)
Feature frequency	-0.317	(0.100) ***	-0.116	(0.104)	-0.007	(0.439)	0.047	(0.177)
Display frequency	0.094	(0.085)	-0.033	(0.093)	0.697	(0.560)	0.120	(0.187)
<i>Brand level characteristics</i>								
Rel. price promotion frequency	0.152	(0.089)	0.221	(0.107) **	-0.481	(0.345)	0.247	(0.163)
Rel. depth of price promotions	-0.111	(0.067) *	-0.126	(0.054) **	0.125	(0.230)	0.117	(0.121)
Rel. feature frequency	0.030	(0.080)	0.045	(0.098)	0.035	(0.334)	-0.058	(0.160)
Rel. display frequency	-0.169	(0.071) **	-0.181	(0.082) **	-0.266	(0.263)	-0.105	(0.130)
Brand size	0.162	(0.076) **	0.178	(0.091) **	-0.319	(0.245)	0.030	(0.143)
Price segment	-0.183	(0.070) ***	-0.082	(0.084)	0.121	(0.298)	0.029	(0.132)

\*\*\*, \*\*\*, zero not contained in 90%, 95% and 99% highest posterior density region, respectively.

<sup>†</sup> Characteristics are standardized to have mean 0 and variance 1.

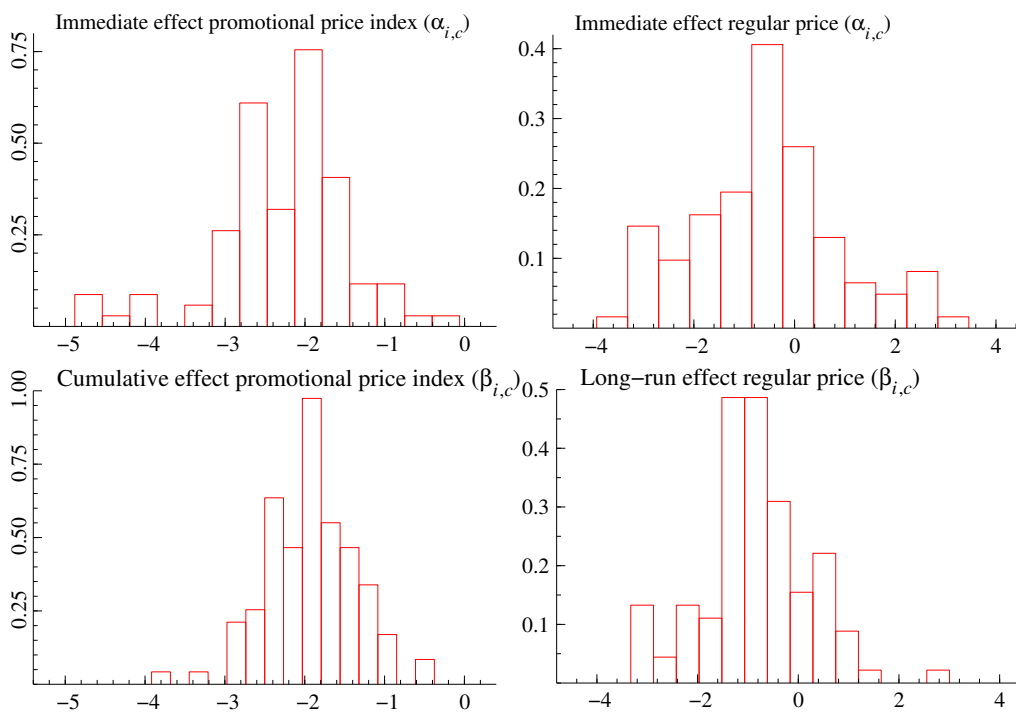


Figure 1: Histogram of posterior means of price effects for all 100 brands

## References

- Akçura, M. T., Gönül, F. F. & Petrova, E. (2004), ‘Consumer learning and brand valuations: An application on over-the-counter drugs’, *Marketing Science* **23**(1), 156–169.
- Bell, D. R., Chiang, J. & Padmanabhan, V. (1999), ‘The decomposition of promotional response: an empirical generalization’, *Marketing Science* **18**, 504–526.
- Bijmolt, T. H. A., van Heerde, H. J. & Pieters, R. G. M. (2005), ‘New empirical generalizations on the determinants of price elasticity’, *Journal of Marketing Research* **42**(2), 141–156.
- Blattberg, R. C., Briesch, R. & Fox, E. J. (1995), ‘How promotions work’, *Marketing Science* **14**, G122–G132.
- Blattberg, R. C. & George, E. I. (1991), ‘Shrinkage estimation of price and promotional elasticities - seemingly unrelated equations’, *Journal of the American Statistical Association* **86**, 304–315.
- Bolton, R. N. (1989), ‘The relationship between market characteristics and promotional price elasticities’, *Marketing Science* **8**, 153–169.
- Bucklin, R. E. & Gupta, S. (1999), ‘Commercial use of upc scanner data: Industry and academic perspectives’, *Marketing Science* **18**(3), 247–273.
- Casella, G. & George, E. I. (1992), ‘Explaining the gibbs sampler’, *American Statistician* **46**, 167–174.
- Chib, S. & Greenberg, E. (1995), ‘Understanding the metropolis-hastings algorithm’, *American Statistician* **49**(4), 327–335.
- Christen, M., Gupta, S., Porter, J. C., Staelin, R. & Wittink, D. R. (1997), ‘Using market-level data to understand promotion effects in a nonlinear model’, *Journal of Marketing Research* **34**(3), 322–334.
- Doornik, J. A. (1999), *Object-Oriented Matrix Programming Using Ox*, 3rd edn, London: Timberlake Consultants Press and Oxford: [www.nuff.ox.ac.uk/Users/Doornik](http://www.nuff.ox.ac.uk/Users/Doornik).
- Fader, P. S. & Lodish, L. M. (1990), ‘A cross-category analysis of category structure and promotional activity for grocery products’, *Journal of Marketing* **54**, 52–65.
- Foekens, E. W., Leeflang, P. S. H. & Wittink, D. R. (1999), ‘Varying parameter models to accommodate dynamic promotion effects’, *Journal of Econometrics* **89**, 249–268.



- Franses, P. H. (1994), 'Modeling new product sales: An application of cointegration analysis', *International Journal of Research in Marketing* **11**, 491–502.
- Franses, P. H. (2005), 'On the use of econometric models for policy simulation in marketing', *Journal of Marketing Research* **42**, 4–14.
- Geman, S. & Geman, D. (1984), 'Stochastic relaxations, gibbs distributions, and the bayesian restoration of images', *IEEE Transaction on Pattern Analysis and Machine Intelligence* **6**, 721–741.
- Granger, C. W. J. (1993), 'What are we learning about the long-run?', *The Economic Journal* **103**(417), 307–317.
- Grover, R. & Srinivasan, V. (1992), 'Evaluating the multiple effects of retail promotions on brand loyal and brand switching segments', *Journal of Marketing Research* **39**, 76–89.
- Guadagni, P. M. & Little, J. D. C. (1983), 'A logit model of brand choice calibrated on scanner data', *Marketing Science* **2**(3), 203–238.
- Hastings, W. K. (1970), 'Monte Carlo sampling methods using Markov chains and their applications', *Biometrika* **57**(1), 97–109.
- Hendry, D. F., Pagan, A. R. & Sargan, J. D. (1984), Dynamic specification, in Z. Griliches & M. Intriligator, eds, 'Handbook of Econometrics', Vol. 2, North-Holland, Amsterdam, chapter 18, pp. 1023–1100.
- Hobert, J. P. & Casella, G. (1996), 'The effect of improper priors on gibbs sampling in hierarchical linear mixed models', *Journal of the American Statistical Association* **91**(436), 1461–1473.
- Hoch, S. J., Kim, B., Montgomery, A. L. & Rossi, P. E. (1995), 'Determinants of store-level price elasticity', *Journal of Marketing Research* **32**, 17–29.
- Jedidi, K., Mela, C. F. & Gupta, S. (1999), 'Managing advertising and promotion for long-run profitability', *Marketing Science* **18**, 1–22.
- Kalaynaram, G. & Winer, R. S. (1995), 'Empirical generalizations from reference price research', *Marketing Science* **14**, G161–G169.
- Kopalle, P. K., Mela, C. F. & Marsh, L. (1999), 'Dynamic effect of discounting on sales: Empirical analysis and normative pricing implications', *Marketing Science* **18**, 317–332.
- Macé, S. & Neslin, S. A. (2004), 'The determinants of pre- and postpromotion dips in sales of frequently purchased goods', *Journal of Marketing Research* **41**, 339–350.

- Mela, C. F., Gupta, S. & Jedidi, K. (1998), 'Assessing long-term promotional influences on market structure', *International Journal of Research in Marketing* **15**, 89–107.
- Mela, C. F., Gupta, S. & Lehmann, D. R. (1997), 'The long-term impact of promotion and advertising on consumer brand choice', *Journal of Marketing Research* **34**, 248–261.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, N. H. & Teller, E. (1953), 'Equation of state calculations by fast computing machines', *Journal of Chemical Physics* **21**(6), 1087–1092.
- Mitra, A. & Lynch, J. G. (1995), 'Toward a reconciliation of market power and information theories of advertising effects on price elasticity', *Journal of Consumer Research* **21**(4), 644–659.
- Montgomery, A. L. (1997), 'Creating micro-marketing pricing strategies using supermarket scanner data', *Marketing Science* **16**, 315–337.
- Montgomery, A. L., Li, S., Srinivasan, K. & Liechty, J. C. (2004), 'Modeling online browsing and path analysis using clickstream data', *Marketing Science* **23**(4), 579–595.
- Moriarty, M. M. (1985), 'Transfer function analysis of the relationship between advertising and sales: A synthesis of prior research', *Journal of Business Research* **13**, 247–257.
- Mulhern, F. J., Williams, J. D. & Leone, R. P. (1999), 'Variability of brand price elasticities across retail stores: Ethnic, income, and brand determinants', *Journal of Retailing* **74**, 427–466.
- Narasimhan, C., Neslin, S. A. & Sen, S. K. (1996), 'Promotional elasticities and category characteristics', *Journal of Marketing* **60**, 17–30.
- Neslin, S. A. & Schneider Stone, L. (1996), 'Consumer inventory sensitivity and the postpromotion dip', *Marketing Letters* **7**(1), 77–94.
- Nijs, V. R., Dekimpe, M. G., Steenkamp, J.-B. E. M. & Hanssens, D. M. (2001), 'The category-demand effects of price promotions', *Marketing Science* **20**, 1–22.
- Paap, R. & Franses, P. H. (2000), 'A dynamic multinomial probit model for brand choice with different long-run and short-run effects of marketing-mix variables', *Journal of Applied Econometrics* **15**, 717–744.
- Pauwels, K., Hanssens, D. M. & Siddarth, S. (2002), 'The long-term effects of price promotions on category incidence, brand choice, and purchase quantity', *Journal of Marketing Research* **39**, 421–439.

- Raju, J. S. (1992), ‘The effect of price promotions on variability in product category sales’, *Marketing Science* **11**, 207–220.
- Shankar, V. & Krishnamurthi, L. (1996), ‘Relating price sensitivity to retailer promotional variables and pricing policy: An empirical analysis’, *Journal of Retailing* **72**, 249–272.
- Smith, A. F. M. & Roberts, G. O. (1993), ‘Bayesian computing via the Gibbs sampler and related Markov Chain Monte Carlo methods’, *Journal of the Royal Statistical Society B* **55**, 3–23.
- Srinivasan, S., Popkowski Leszczyc, P. T. L. & Bass, F. M. (2000), ‘Market share response and competitive interaction: The impact of temporary, evolving and structural changes in prices’, *International Journal of Research in Marketing* **17**(4), 281–305.
- Srinivasan, S., Pauwels, K., Hanssens, D. M. & Dekimpe, M. G. (2004), ‘Do promotions benefit manufacturers, retailers, or both?’, *Management Science* **50**(5), 617–629.
- Tanner, M. & Wong, W. (1987), ‘The calculation of posterior distributions by data augmentation’, *Journal of the American Statistical Association* **82**, 528–550.
- Tierney, L. (1994), ‘Markov chains for exploring posterior distributions’, *Annals of Statistics* **22**, 1701–1762.
- Van Heerde, H. J., Leeflang, P. S. H. & Wittink, D. R. (2000), ‘The estimation of pre- and postpromotion dips with store- level scanner data’, *Journal of Marketing Research* **37**, 383–395.
- Varian, H. (1992), *Microeconomic Analysis*, 3rd Edition, Norton, New York.
- Wakefield, K. L. & Imman, J. J. (2003), ‘Situational price sensitivity: the role of consumption occasion, social context and income’, *Journal of Retailing* **79**, 199–212.
- Wieringa, J. E. & Horváth, C. (2005), ‘Computing level-impulse responses of log-specified VAR systems’, *International Journal of Forecasting* **21**(2), 279–289.
- Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York.
- Zenor, M. J., Bronnenberg, B. J. & McAlister, L. (1998), ‘The impact of marketing policy on promotional price elasticities and baseline sales’, *Journal of Retailing and Consumer Services* **5**, 25–32.

## Publications in the Report Series Research\* in Management

### ERIM Research Program: "Marketing"

2005

*Assessing Customer Evaluation and Revenue Consequences of Component Sharing Across Brands in the Vertical Product Line*

Peter C. Verhoef and Koen H. Pauwels

ERS-2005-007-MKT

<http://hdl.handle.net/1765/1936>

*Intra- and Inter-Channel Competition in Local-Service Sectors*

Kathleen Cleeren, Marnik G. Dekimpe And Frank Verboven

ERS-2005-018-MKT

<http://hdl.handle.net/1765/1925>

*The European Consumer: United In Diversity?*

Aur lie Lemmens, Christophe Croux and Marnik G. Dekimpe

ERS-2005-022-MKT

<http://hdl.handle.net/1765/1942>

*Managing Product-Harm Crises*

Harald J. van Heerde, Kristiaan Helsen and Marnik G. Dekimpe

ERS-2005-044-MKT

<http://hdl.handle.net/1765/6887>

*Why Consumers Buy Lottery Tickets When the Sun Goes Down on Them. The Depleting Nature of Weather-Induced Bad Moods.*

Sabrina Bruyneel, Siegfried Dewitte, Philip Hans Franses and Marnik G. Dekimpe

ERS-2005-045-MKT

<http://hdl.handle.net/1765/6890>

*A Hierarchical Bayes Error Correction Model to Explain Dynamic Effects of Price Changes*

Dennis Fok, Richard Paap, Csilla Horv th and Philip Hans Franses

ERS-2005-047-MKT

*When are CRM Systems Successful? The Perspective of the User and of the Organization*

Gerrit H. Van Bruggen and Berend Wierenga

ERS-2005-048-MKT

---

\* A complete overview of the ERIM Report Series Research in Management:

<https://ep.eur.nl/handle/1765/1>

ERIM Research Programs:

LIS Business Processes, Logistics and Information Systems

ORG Organizing for Performance

MKT Marketing

F&A Finance and Accounting

STR Strategy and Entrepreneurship