A Hierarchical Bayes Error Correction Model to Explain Dynamic Effects of Promotions on Sales

Dennis Fok[*], Csilla Horváth, Richard Paap, and Philip Hans Franses

Econometric Institute, Erasmus University Rotterdam

Econometric Institute Report EI-2004-27

June 29, 2004

Abstract

For promotional planning and market segmentation it is important to understand the short-run and long-run effects of the marketing mix on category and brand sales. In this paper we put forward a sales response model to explain the differences in short-run and long-run effects of promotions on sales. The model consists of a vector autoregression rewritten in error-correction format which allows us to disentangle the long-run effects from the short-run effects. In a second level of the model, we correlate the short-run and long-run elasticities with various brand-specific and category-specific characteristics. The model is applied to weekly sales of 100 different brands in 25 product categories.

Our empirical results allow us to make generalizing statements on the dynamic effects of promotions in a statistically coherent way.

key words: sales; vector autoregression; marketing mix; short and long-term effects; Hierarchical Bayes

[*] We thank participants of the Marketing Science conference 2004 in Rotterdam for their comments. We are indebted to Marnik Dekimpe for his 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`. The model presented in this paper is implemented in Ox 3.3 (Doornik 1999).

1 Introduction

A well-known finding in sales promotion research is that there is a trough after the deal (Blattberg et al. 1995, p. G127). The reason for this so-called post-promotion dip is that consumers purchase a larger quantity in response to promotions. They consume a proportion of these goods in the period of purchase and stockpile the rest (Mela et al. 1998). As a consequence, purchases on future trips (when the products are offered at the regular price again) are substituted by utilizing the stored products.

The literature presents the existence of dynamic effects of promotions (Kopalle et al. 1999, Paap & Franses 2000, Van Heerde et al. 2000, Pauwels et al. 2002), and hence the overall (net) effect of promotions cannot only be described by the immediate promotional elasticity. Kopalle et al. (1999) find that promotions have positive contemporaneous effects on sales accompanied by negative future effects. They also emphasize that "models that do not consider dynamic promotional effects can mislead managers to overpromote". In line with this, Jedidi et al. (1999) find that the long-term effects of promotions on sales are negative and, in an absolute sense, they are about two-fifths of the magnitude of the positive short-run effect.

While several articles found evidence of post-deal troughs, many others fail to find empirical evidence of a post-promotion dip (for example, Grover & Srinivasan 1992). Hence, findings are mixed (Blattberg et al. 1995). An explanation for this disagreement in the marketing literature may be that the dynamic effect varies across categories, stores, and brands. 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 expect new discounts soon. Therefore, researchers may be less likely to 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 large in size, and there again finding evidence of a post-promotion dip might be rare.

In the present paper we use a Hierarchical Bayes (HB) - Error Correction Model (ECM), which allows us to directly estimate the potentially differing short-run and long-run marketing mix effects on sales, where we relate these effects with characteristics of brands and categories. The proposed model provides an opportunity to find explanations for the above mentioned mixed evidence concerning the dynamic effect of promotions. In addition, our model provides important insights for brand managers and for retailers about the determinants that affect the length and the magnitude of the dynamic effects and the net effects of promotions. Among other things, we analyze whether there are common determinants for the immediate price elasticity and for the dynamic effects of price promotions.

In the literature there are a number of articles that investigate the relationship between market characteristics and promotional price elasticities. In Table 1 we compactly present a selection of such studies. It is clear that most of these articles focus on relating *immediate price elasticities* to brand, and/or category, and/or consumer specific 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 dynamic effects* of *price and other promotional activities* to *brand and category* characteristics.

Among the listed articles we find only two that investigate the determinants of short-term as well as long-term 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. Our approach is more in line of Nijs et al. (2001) as we investigate heterogeneity in the immediate and dynamic effects of promotions across brands and categories, assuming constant parameters over time.

The differences and advantages of our approach to that of Nijs et al. (2001) can be summarized as follows. First of all, it is important to emphasize the differences in the measures of cumulative effects between the two approaches. Nijs et al. (2001) obtain the cumulative effects from Impulse Response Functions (IRFs) of the estimated Vector Autoregression (VAR). This incorporates possible competitive and feedback effects of competitors. In contrast, our measure of long-run effect in the ECM excludes these effects. This enables us to have a better focus on the determinants of *dynamic demand reactions*.

Second, the HB - ECM allows us to disentangle the short-run and long-run effects into separate parameters. This enables us to relate the short-run and long-run effects directly to store and brand characteristics. Hence, we do not have to rely on derivative measures like the IRF.

Third, Nijs et al. (2001) use a two-step approach to relate short-run and long-run effects to moderating variables. This approach does not appropriately account for the uncertainty in the first-level parameter estimates when obtaining the parameter estimates and standard errors in the second stage. In finite samples, this leads to underestimation of the standard errors of the parameters in the second-stage regression.

Fourth, Nijs et al. (2001) compute the long-run (or, net) effects of promotions based on accumulated impulse response functions. As these accumulated impulse response functions are non-linear functions of the model parameters, the uncertainty in the estimated net effect is usually quite large when compared with the uncertainty in the model parameter estimates. From an efficiency point of view, it is therefore more reliable to estimate a parameter representing the long-run effects than the impulse responses.

Fifth, we not only consider category-specific variables but also brand-specific variables.

Sixth, besides focusing on the determinants of price promotion elasticities, we also measure determinants of the effectiveness of other promotional variables (such as display and feature). Most of marketing literature so far mainly focuses on the determinants of price or price promotion elasticities.

Finally, in a multiplicative model (that is the most frequently used model in marketing and the one applied by Nijs et al. 2001) the sum of the impulse responses over the dust-settling period has no straightforward interpretation, see Wieringa & Horváth (2004). They cannot be interpreted as long-run elasticities as Nijs et al. (2001) claim. This is simply due to the fact that the sum of logs is not equal to the log of sums.

The remainder of this paper is organized as follows. In Section 2 we present a more detailed overview of the literature. In this section we discuss hypotheses on the relationship between brand and category characteristics and the long-run and short-run effects of marketing instruments. 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 Hypotheses

In this section we identify category-specific and brand-specific characteristics that may be related to the short- and the long-term effects of price promotions. 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 variables with additional variables and provide a discussion on the determinants of dynamic effects. In the discussion we rely on the notion of a utility-maximizing consumer who operates under a budget constraint (Bell et al. 1999, Varian 1992) and 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 long-run (or net) 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 dynamic effects of marketing actions.

In our discussion below we focus on the determinants of price promotion elasticities. We use price indices (ratio of actual to regular prices) to capture only the promotional price effects and regular prices to capture changes in price levels. This also allows for a comparison across categories. In the discussion we distinguish three groups of variables, that is, category-specific variables, brand-specific variables, and variables that can be defined at both levels. For each, we summarize the literature and give, if possible, hypotheses for the sign of the short-run and long-run effect of a price promotion. We will denote the difference between the long-run and short-run effect as the dynamic effect. In this section we do not consider the determinants of the effectiveness of other types of promotions as they have received very little attention in the literature.

Category-specific variables

Storability Consumers may favor price promotions in a category with easily storable goods, that is, a category where one can more easily manage inventory to allow for purchases at irregular intervals in response to deals (Raju 1992, Narasimhan et al. 1996). This means that consumers can buy more in the period of the promotion and are not compelled to buy on a particular trip. In general, people purchase more items of storable products on deal because they are able to store them in the next period and consume them when the next purchase occasions would otherwise occur. On the one hand, this arises because consumers are more likely to substitute future purchases with the consumption of stored items in case of easily storable goods and are also more likely to utilize the accumulated stock during a longer period after the promotion. This suggests that promotions in a more storable category lead to higher sales during the offer but a larger and longer dip after the promotion than in a less storable one. Therefore, immediate and dynamic effects have opposite signs. We do not have strong expectations about whether brands in more storable categories should have higher or lower overall effect than in less storable ones.

Average share of budget of a category It is important to stress that the share of budget of a category can capture two distinct effects. The first effect is that more expensive categories may represent a higher share of the budget. Raju (1992) found that the (relative) expensiveness of the product category may affect variability in category sales for the following reasons. A consumer, facing a budget constraint, will probably feel more inclined to buy extra quantity of the less expensive product as the additional purchase will lead to a smaller grocery bill. In addition, price promotions may be less effective for more expensive products because higher income (and, hence, less price sensitive) shoppers may constitute a greater ratio of the consumer population of expensive products. So, both theories arrive at the conclusion that expensive product categories are likely to exhibit smaller variability in sales, inducing a smaller immediate effect and a lower and shorter post-promotion dip. At the same time, a 1 % discount of a more expensive product category means a higher absolute discount, and hence, more money to be saved and higher price promotional elasticity in a more expensive product category.

The second effect is that categories in which consumers purchase relatively more may also constitute a higher share of budget of the customers. In this case the same price discount would encourage consumers to buy more items of such product category, as in the future the consumer will certainly be in need of this product. In a relatively short time period money can be saved by consuming the stored products.

Necessity Bell et al. (1999) argue that relatively more necessity (i.e., nonimpulse) products are expected to exhibit lower primary demand effect and higher secondary effect. 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. At the same time, shoppers with higher income (and, hence, less price sensitive consumers) may constitute a greater proportion of the consumer population of non-necessity goods. So, we conjecture that necessity products have higher immediate effects and also a larger and possibly longer dip. Wakefield & Imman (2003) made a distinction between functional and hedonic products (a categorization along a similar dimension) and found consumers to be less price sensitive in categories that are perceived as primarily hedonic in nature. In the empirical analysis we capture the necessity of a product category along two dimensions;

the hedonic and utilitarian nature of the category.

Competitive intensity Most studies use the number of brands in a category as a measure for competitive intensity and provide different interpretations to this measure. Some studies use it to reflect brand assortment. Narasimhan et al. (1996), for example, argue that brand proliferation within a category may signal the existence of many market segments in the category and hence, room for product differentiation. This 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). Nijs et al. (2001) argue that in less competitive environments, that are characterized by a smaller number of brands, price-promotion effectiveness is expected to be higher. This is due to a better opportunity for cooperative activities among the brands to restrict output and raise prices in order to reach a more elastic area of the demand curve and due to lower searching costs.

We use two other measures of competitive intensity, that is, the market concentration index and price dispersion in a category. The concentration measure is a more sophisticated measure of the market concentration than the number of brands as it also takes into account the relative differences in size among brands in a category. Price dispersion captures aspects of brand assortment, with a higher price dispersion meaning more brand assortment in the category.

Category- and brand-specific variables

Frequency of price promotional activity Theories suggest mixed effects of frequent promotions. First, in categories where price promotional activity is frequent, consumers may become more price-conscious (Kopalle et al. 1999, Mela et al. 1997, 1998). Consumers who expect future promotions, tend to purchase the products on deals. They may be more inclined to postpone their purchase and to get the habit of stockpiling. This suggests that in categories that are characterized by frequent price promotions, price discounts have higher immediate increases in own sales and larger and longer post-promotion dips to occur. Second, increased use of discounts may reduce consumers' reference prices (Kalaynaram & Winer 1995) resulting in a lower level of utility in a category for a discount of the same size. Third, if a category is promoted infrequently, consumers are more likely to use these opportunities to stock-up for future consumption than in categories with regular discounts, where consumers are not likely to benefit much from stocking up on any promotional action (Raju 1992). These theories lead to contrasting links between the frequency of price promotions in a category and the immediate and dynamic effects of price elasticities.

The same arguments hold at the brand level. In addition, brands with frequently lowered prices are often considered to be of lower quality than similar, rarely promoted brands. Increased promotional activity may also affect 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. On the one hand, 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" as one point (point 4) of the empirical generalizations on the effects of promotions. 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. Nijs et al. (2001) discover that a key determinant of the short-run effect of price promotion is the frequency of price promotional activity in the category. A higher price-promotional frequency is found to result in a larger increase in sales in the short-run. However, in the long-run this effect appears to have vanished. Zenor et al. (1998) find that brands with higher levels of promotional activity are associated with more elastic demand than those that promote less. These ambiguous findings may 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 on the brand level.

Average depth (or magnitude) of promotion In categories where price reductions are on average larger than in similar categories, consumers, who expect to get a higher reward, are probably more 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 higher immediate increase in sales and a larger and longer lasting post-promotion dip. These arguments also hold at the brand level, that is for *brands* with deeper discounts.

Frequency of display activity Bolton (1989) argues that display activity may be systematically related to own price elasticities and the relationship may be two-way. The relative frequency of displays may influence consumers' belief about the popularity and quality of market offerings. And, this effect is more pronounced across brands 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 in the store.

Frequency of feature activity 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 and to attract new price sensitive shoppers to purchase the product (Bolton 1989, Moriarty 1985). These theories suggest that brands in categories with more frequent retailer advertising activity should have higher immediate price elasticity and also a larger drop in sales after the price promotion than in less featured ones. A similar distinction can be made between more often promoted *brands*.

Brand-specific variables

Brand size Bolton (1989) argues (supported by empirical evidence) that brands with higher brand shares tend to be operating on the flat proportion of their sales response functions, where "flatness" reflects consumer preferences. 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 marketing.

Price segment of a brand A discount may attract several types of consumers. It may induce consumers who usually purchase a competing brand, to switch to the promoted brand. It may also attract consumers who would otherwise find the brand too expensive. Finally, among the loyal consumers of the promoted brand, a discount may induce stockpiling or an increase in consumption (Raju 1992). The promotion of a brand in an expensive price category may induce all the 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 greater fraction of the consumer population of less expensive brands (Raju 1992). These people are more inclined to buy the brand on discount and save additional money by stockpiling. This indicates opposite short-run and long-run effects.

We summarize in Table 2 our hypotheses based on marketing theory about the relationship of dynamic effects with the above mentioned brand- and category-specific features. We have put various question marks in the table for cases where the literature does not yield a prediction for the sign of the effect.

3 Analyzing short and long-run effects

In this section we present a modeling framework for estimating (the determinants of) the dynamic effect of marketing instruments on log sales, when the logarithm of sales is unit-root stationary. We first consider a model for one product category. 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 frequent temporary price promotions will lead to permanent increases in sales, which is an unrealistic assumption in the long run. Hence, to study dynamic effects of the marketing mix, it is more interesting to examine the cumulative effect of a temporary price promotion on current and future log sales instead of the permanent effect. We will denote this cumulative effect as the long-run effect. Below we show that in a vector autoregression this cumulative effect is equal to the effect of a permanent price change on log sales in the long run.

To describe the sales of brands in a product category we start with a vector autoregression with explanatory variables (VARX). Later on, we will rewrite this model in an error-correction format. Denote the sales of brand i at time t by S_{it}, for $i = 1, \ldots, I$ and $t = 1, \ldots, T$, where I stands for the number of brands in the market. To model the vector of sales $S_t = (S_{1t}, \ldots, S_{It})'$, we consider a VARX(1) model

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

where $\varepsilon_t \sim N(0, \Sigma)$ and μ denotes a vector of intercept parameters. The vectors $X_{kt} = (X_{k1t}, \ldots, X_{kIt})'$, $k = 1, \ldots, K$ denote I-dimensional vectors of possible explanatory variables, for example, the kth marketing-mix variable of brand i at time t. 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.

The VARX model is stationary if the eigenvalues of Γ are within the unit circle. Although we assume stationarity, it is still difficult to interpret the parameters in a vector autoregressive model with current and lagged exogenous variables. The parameters combine the short-run and long-run effects of the explanatory variables on the dependent variables. We can rewrite the VARX model in the error-correction format and the resulting parametrization leads to a direct interpretation of the long-run and the short-run effect of a marketing instrument on log sales.

In the marketing literature the error-correction model has been used by, for example, Franses (1994) and Paap & Franses (2000) to distinguish the short-run from the long-run 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 long-run effect is defined as the impact of a promotion on the future 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 determine the pattern of the dynamic effects of a marketing instrument (X_{kt}) on sales we solve (1) backwards for $\log S_t$ by repeated substitution. This results in

$$\log S_t = \Gamma^\tau \log S_{t-\tau} + \sum_{j=0}^{\tau-1} \Gamma^j (\mu + \sum_{k=1}^{K} (A_k \log X_{k,t-j} + C_k \log X_{k,t-j-1}) + \varepsilon_{t-j}). \tag{2}$$

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

$$\log S_t = (\mathbf{I} - \Gamma)^{-1}\mu + \sum_{k=1}^{K} (\mathbf{I} - \Gamma)^{-1}(A_k + C_k) \log X_k + \sum_{j=0}^{\infty} \Gamma^j \varepsilon_{t-j}, \tag{3}$$

where $\mathbf{I}$ denotes the identity matrix. As $\mathrm{E}[\varepsilon_{t-j}] = 0$ for all j, the long-run expectation of the log sales given $X_1, \ldots, X_K$ equals

$$\mathrm{E}[\log S | X_1, \ldots, X_K] = (\mathbf{I} - \Gamma)^{-1}\mu + \sum_{k=1}^{K} (\mathbf{I} - \Gamma)^{-1}(A_k + C_k) \log X_k. \tag{4}$$

This expectation denotes the long-run relation between log sales and the explanatory variables. The size of the absolute values of the eigenvalues of Γ translates to the speed of convergence to the long-run equilibrium. The long-run elasticity of X_k on S is given by

$$\frac{\partial S}{\partial X_k} \frac{X_k}{S} = \frac{\partial \log S}{\partial \log X_k} = (\mathbf{I} - \Gamma)^{-1}(A_k + C_k) \equiv B_k, \tag{5}$$

where 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. The long-run variance is given by

$$\mathrm{V}[\log S | X_1, \ldots, X_K] = \sum_{j=0}^{\infty} \Gamma^j \Sigma (\Gamma')^j \equiv V, \tag{6}$$

which is finite if the eigenvalues of Γ are within the unit circle, that is, in case of stationarity.

It follows immediately from (2) that under stationarity a temporary change in one of the X_k variables at time t has no impact on the sales at time $t + j$ in the long run, as the term Γ^j will be zero for large j. Only a permanent change in the value of a marketing instrument may have a permanent long-run effect on the sales. The long-run effect on the log sales is measured by the parameters in the matrix $B_k = (\mathbf{I} - \Gamma)^{-1}(A_k + C_k)$. A temporary change of X_k does however have a short-run effect on sales. The direct short-run effect of X_{kt} on the log sales is measured by A_k. To disentangle the long-run effects from the short-run effects of X_k on the log sales, 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,

$$\Delta \log S_t = \mu + \sum_{k=1}^{K} A_k \Delta \log X_{kt} + \Pi(\log S_{t-1} - \sum_{k=1}^{K} B_k \log X_{k,t-1}) + \varepsilon_t, \tag{7}$$

where $\Pi = (\Gamma - \mathbf{I})$, $B_k = (\mathbf{I} - \Gamma)^{-1}(A_k + C_k)$, and Δ denotes the first-differencing operator, that is, $\Delta y_t = y_t - y_{t-1}$.

The short-run, or instantaneous, effects are given by

$$\frac{\partial S_t}{\partial X_{kt}} \frac{X_{kt}}{S_t} = \frac{\partial \log S_t}{\partial \log X_{kt}} = A_k. \tag{8}$$

Although the ECM in (7) 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 so-called error correction term 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 Π contains the adjustment parameter and determines the speed of convergence to the long-run equilibrium.

As already discussed before, it can be shown using (2) that B_k in the error correction model (7) is also equal to the cumulative effect of a temporary change in $\log X_{kt}$ on current and future log sales, that is, under stationarity the following property holds

$$\sum_{j=0}^{\infty} \frac{\partial \log S_{t+j}}{\partial \log X_{kt}} = \sum_{j=0}^{\infty} \Gamma^j (A_k + C_k) = B_k. \tag{9}$$

Finally, a special case of the model is where $A_k = B_k$ for all k. In this case the short-run effects are equal to the long-run effects. The ECM (7) then simplifies to a common factor model, see Hendry et al. (1984). A temporary change in $\log X_{kt}$ now has only an effect on current log sales and not on future sales, which, from a marketing perspective, could be implausible. In the application below, we will see that in general the short-run effects are larger in size than the long-run 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 a large number of categories.

4 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 (7) for category c is given by

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

with $\varepsilon_{ct} \sim N(0, \Sigma_c)$ for $c = 1, \ldots, C$ and $t = 1, \ldots, T_c$. Note that we allow for different intercepts μ_c, short-run A_c and long-run B_c effects and variance of the error term Σ_c in each category. The adjustment parameters Π_c are also allowed to be different across categories. The categories may even have a different number of brands and observations.

To relate the short-run and long-run 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 $\alpha_{ck} = \text{diag}(A_{ck}) = (\alpha_{1ck}, \ldots, \alpha_{I_c ck})'$ and $\beta_{ck} = \text{diag}(B_{ck}) = (\beta_{1ck}, \ldots, \beta_{I_c ck})'$ for $k = 1, \ldots, K$. The long-run and short-run 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 discussed in Section 2. Another part of the elasticity cannot be explained, either by the fact that it is specific to the brand or the category. In sum, we propose to describe the short-run and long-run 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 λ_{1k} and λ_{2k} describe the effects of the brand characteristics on the short-run and the long-run elasticities, respectively. The error terms η_{ick} and ν_{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 $\eta_{ic} = (\eta_{ic1}, \ldots, \eta_{icK})' \sim \text{N}(0, \Sigma_\eta)$ and $\nu_{ic} = (\nu_{ic1}, \ldots, \nu_{icK})' \sim \text{N}(0, \Sigma_\nu)$.

We will abbreviate the model above as HB-ECM. As far as we know we are the first to use this model in marketing. 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 procedure 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.

5 Empirical results

In this section, we use our HB-ECM to explain differences in short-run and long-run effects of marketing-mix variables on sales across brands and product categories. In Section 5.1 we discuss the product categories and marketing-mix variables we consider in our analysis. Section 5.2 discusses the estimation results.

5.1 Data and Variables

The data we consider are weekly sales 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 as described in Srinivasan et al. (2004), who use the same data set. It concerns the product categories: bottled juice, cereals, cheese, cookies, crackers, canned soup, dish detergent, frond-end candies, frozen diners, 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. Hence, we end up with $4 \times 25 = 100$ 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 price indexes. The display and feature variables reflect the percentage of stock-keeping units of the brand that are promoted in a given week. 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 does not follow temporary drops in price. Sustained drops 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 short-run and long-run own-effects of the marketing-mix variables, denoted by

α_{ick} and β_{ick}, are explained by characteristics of the brand and product category in the second stage of the model as explained in (11) and (12). The variation in regular price across brands and product categories is relatively small. Therefore we choose not to include the own-effect of regular price in the second stage of our model. Hence, for this variable we will not consider the determinants of the dynamic effects. Instead, we include the variable only to control for possible changes in the regular price. We do however allow this variable to have another long-run effect than a short-run effect.

Concerning cross effects of marketing instruments, we only allow for cross promotional price effects. For the regular price we do not include cross effects for the above-mentioned reasons. For feature and display the correlation between different brands is too high to yield appropriate estimates of cross 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 cover 4 consecutive weeks. The starting point of the period of 4 consecutive weeks is chosen in a way that it produces the best fit. 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.

To summarize, we only include the α_{ick} and β_{ick} of display, feature and actual price in the second layer of the HB-ECM. As explanatory variables of α_{ick} and β_{ick}, we use brand-level characteristics and category-level characteristics. A summary and the formal definition of these variables can be found in Appendix B. Further details on this classification can be obtained from the authors.

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 50,000 are used as burn in. Furthermore, to remove correlation in the chain we only consider every 15th draw for our results.

First, we summarize the posterior means of the effects of the marketing mix variables (these are log price index, feature and display) in several graphs. Figure 1 presents the distribution of the posterior means of the long-run, short-run and dynamic effect of the three marketing instruments across the brands and product categories. The signs of the short-run and long-run effects are all according to expectations. The dispersion in the long-run effects tends to be smaller than that in the short-run effects. The graphs on the bottom row in Figure 1 give the distribution of the dynamic effect, that is, the difference between the long-run and the short-run effect. For feature and display the dynamic effect is mostly negative, for the price index this effect is mostly positive. In all three cases this indicates that there are negative dynamic effects of promotions. Overall, the magnitude of the long-run effect of price is smaller than the short-run effect. In general, some positive effects of a price cut are compensated in the periods following a promotion by, for example, the effects of stockpiling.

In Figure 2 we give a scatter plot of the posterior means of the long-run effect versus the posterior means of the short-run effect. For all three marketing instruments we find a positive correlation. This implies that, in general, brands that have a large short-run effect of a particular marketing instrument also have a large long-run effect of that instrument.

In Figure 3 we present the same type of scatter plots, only now we compare the effectiveness of different marketing instruments. Overall we do not find strong correlations between the effectiveness of different marketing instruments. We only find a strong positive relationship between the short-run effects of feature and display.

Now, we turn to the second layer of our HB-ECM, where we explain differences in dynamic effects of the marketing-mix variables on sales. Table 3 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 long-run) effects of promotions.

Moderating factors of price promotion elasticities

First of all, we focus on the determinants of the price effects, as in the literature this has received almost exclusive attention. The results for the price index should be compared to the hypotheses summarized in Table 2.

There are quite some characteristics that significantly influence the effectiveness of price promotions. The price segment for example has a negative influence on the short-run price parameter. This means that brands in a higher price segment will have stronger price effects. The effect of brand size corresponds with our conjecture in Table 2. Larger brands tend to have smaller price effectiveness, on the short-run as well as in the long-run.

An interesting finding is that while price promotion frequency of a category does not seem to influence the immediate or the long-run effectiveness of a price promotion, which is in accordance with the findings of Bolton (1989) but not with those of Nijs et al. (2001), the frequencies of feature and display usage in a category do have significant influence. On the brand level, we find that more frequent price promotions correlate with less strong (short- and long-run) price effects. This probably results from the fact that in the analyzed markets frequent price promotion of a brand strongly damages the brand image and results in lower reference prices for the brands. At the same time, it is more difficult to harm the image of an entire category, and also, if all brands of a category are more often on promotion, consumers expect their brand to be on promotion soon and a price discount is less likely to induce brand switching.

In Section 2, we hypothesized that brand sales are more elastic for categories and brands that are frequently featured in flyers and newspapers. This notion is supported by our empirical findings. There are short- and long-run effects of category level feature activity, while the effect of brand level feature frequency seems to have no influence on sales in the short or in the long run. Again, the results on the immediate effects coincide with the findings of Bolton (1989). So, frequent price-oriented advertising in a category appears to make consumers more aware of the prices, attracts price sensitive consumers, and also inclines people to stockpile more (as the higher immediate effect is partly offset in the long-run) but this does not hold on the brand-level.

Interestingly, the use of display activity in a category and for a brand appears to influence the effects of price promotions in opposite directions. Brand sales are more inelastic in categories that are frequently displayed in a store (which is in line with Bolton 1989), but more elastic for brands with frequent display (opposite to the findings of Bolton 1989). So, whether intensified display activity of a brand results in higher or lower price-promotional elasticity depends on the relative importance of these two factors. If the brand-specific and the category-specific effects are equally important (when for example there is only one brand in the category or when only one brand is on display), the effects might cancel out, while if there are several brands in a category that are promoted frequently, the intensified display usage of a brand will result in higher price elasticity (the brand effects dominate).

The depth of price promotions also correlates with price promotion effectiveness. Brands and categories characterized by deep promotions tend to have larger short-run price effects, corresponding to our hypothesis and to the findings of Raju (1992) and of Foekens et al. (1999) for brand B. This also holds for the long-run effects, although the long-run effect becomes smaller than the short-run effects, possibly due to larger and/or longer post-promotion dip, especially for the category-specific variable.

We find that price discounts in categories, that constitute a high average share of budget, have high elasticities. Among the two dimensions (utilitarian and hedonic) that we choose to describe the necessity of a product category, the hedonic aspect turns out to have a significant effect on the short-run price promotion elasticity. Price promotions in hedonic categories appear to have lower sales effects.

Finally, larger price dispersion (and hence, a more segmented category) goes together with less strong (short-term and long-term) price promotion effects.

Moderating factors of non-price promotion (feature and display) elasticities

For the relation between brand/category characteristics and the effectiveness of feature and display we cannot rely on previous literature. We therefore will only list some interesting findings. For feature promotions we find an effect of the price promotion frequency on the brand level. More frequent price promotions by a brand are correlated with larger long-run feature effects. Larger brands tend to have larger short-run feature effects. For display promotions we find more significant relationships. Again, brands with frequent price promotions have larger display effectiveness (on the short term as well as the long term).

Frequent use of price promotion by a brand is positively correlated with the effectiveness of display. This relationship is the opposite for the frequent use of display activity of a brand. This result is similar to the findings for the price elasticity, that more frequent use of a promotional tool results in a lower average effect of that instrument. Consumers get used to having the impulses from the instrument and need stronger or different types of stimuli to notice and buy the brand.

At the category level, however, the frequency of displays has the opposite effect. Display is relatively effective in categories in which displays are used relatively frequently. Note that we do not give statements about causal relationships. It could also be that displays are used often in categories in which this instrument is presumed to be effective. Finally, more utilitarian categories have a larger effect size of displays in the short run.

6 Conclusions

In this paper we propose a Hierarchical Bayes - Error Correction model to explain the differences in short-run and long-run effects of promotions on sales. The model is applied to weekly sales of 100 different brands in 25 product categories. In the second layer of the model the short-run and long-run parameters are related brand-level and category-level characteristics. Parameters estimates are obtained using MCMC.

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

We find mostly significant effects on the elasticity of price discounts. Brands in categories that are characterized by high price differentiation and that constitute a lower share of budget are less sensitive to price discounts, in the short- and in the long-run. Deep price discounts in a category or of a brand turn out to increase the immediate price sensitivity of customers. We find significant and somewhat lower effects in the long-run for the brand-specific variable, however, the positive short-run effect it is dissipated for the category-specific variable in the long-run. Frequent use of price promotion of a brand decreases the effectiveness of discounts (in the short-run and in the long-run), while the frequent use of discounts in a category appear not to have significant effects on the price promotion elasticity of a brand. The effectiveness of price promotion is also influenced by the frequent use of non-price promotions. Frequent use of feature and rare display activity in a category increases the elasticity of price promotions, however, this result is only significant in the short-run for display. Frequent display activity of a brand increases the effectiveness of its price discounts.

We also find a few significant relationships between the elasticities of non-price promotional activities (feature and display) and some category- and brand-specific variables, especially concerning the effectiveness of display. Frequent price discount and infrequent display activity of a brand increases the effectiveness of display. This result is similar to the findings for price elasticity, that more frequent use of a promotional tool results in a lower average effect of that instrument. This may be due to that consumers get used to having the impulses from the instrument around and need stronger or different types of stimuli to make them notice and buy the given brand. Frequent use of display and shallow price promotions in a category increase display effectiveness.

Our study can be extended in several ways. First, if would be interesting to measure the immediate and long-run effects of changes in regular price and to correlate these to brand- and category specific components. This is not possible for our database due to the low variation in regular price across categories and brands. Second, despite allowing for cross-promotional price effects we focus on measuring moderating factors on own-brand elasticities. Capturing characteristics that influence cross-brand elasticities would provide support for the development of strategic competitive reactions of brand managers. This would be especially interesting with data available about brands of the same manufacturer or data on SKU-level because that would make it possible to measure whether the factors have different cross effects between brands of the same and of competing manufacturers and to capture characteristics that decrease the cannibalizing effect. Third, as pointed out earlier, we do not consider competitive reactions or feedback effects in our VEC 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 built a model in which these relationships were considered, for example, to examine whether the results would become more in line with those of Nijs et al. (2001). Fourth, we assumed the elasticities to be constant over time. It is, however, possible that, for example, a price promotion that soon follows another discount may have lower immediate effect, due to the products still kept in stock. So, modeling this would provide ideas for managers about how frequently they should plan the promotions. A relating issue would be to capture whether some brand-and category-specific characteristics have any influence of the speed of adjustment of the sales to a discount.

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,

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

with $\varepsilon_{1c} \sim N(0, V_c)$, where V_c follows from (6) with $\Gamma = \Pi_c + \mathbf{I}$ and $\Sigma = \Sigma_c$.

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

$$\alpha_{ic} = \Lambda_1' z_{ic} + \eta_{ic} \tag{14}$$

$$\beta_{ic} = \Lambda_2' z_{ic} + \nu_{ic} \tag{15}$$

for $i = 1, \ldots, I_c$, where the $L \times K$ matrices Λ_1 and Λ_2 contain the vectors λ_{1k} and λ_{2k}, respectively. The likelihood function of the model is given by

$$\prod_{c=1}^{C} \int_{\alpha_c,\beta_c} \phi(\varepsilon_{c1}; 0, V_c) \prod_{t=2}^{T_c} \phi(\varepsilon_{ct}; 0, \Sigma_c) \prod_{i=1}^{I_c} \phi(\alpha_{ic}; \Lambda_1' z_{ic}, \Sigma_\eta)\phi(\beta_{ic}; \Lambda_2' z_{ic}, \Sigma_\nu) d\alpha_c d\beta_c, \tag{16}$$

where $\phi(x; \mu, \Sigma)$ is the density function of the multivariate normal distribution with mean μ and variance Σ evaluated at x, and where $\alpha_c = (\alpha_{1c}', \ldots, \alpha_{I_c c}')'$ and $\beta_c = (\beta_{1c}', \ldots, \beta_{I_c c}')'$.

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 α_c and β_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 Σ_ν. 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 Σ_η and Σ_ν parameter with scale parameter $\kappa_1 \mathbf{I}_K$ and degrees of freedom κ_2. We set the value of κ_1 to $\frac{1}{1000}$ and κ_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 α_c and β_c. In deriving the sampling distributions we build on the results in Zellner (1971, Chapter VIII).

Sampling of Π_c

To sample Π_c we rewrite (10) as

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

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

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

and variance

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

with $Y_{ct} = \Delta \log S_{ct} - \mu_c - \sum_{k=1}^{K} A_{ck} \Delta \log X_{ckt}$ and $W_{ct} = \log S_{c,t-1} - \sum_{k=1}^{K} B_{ck} \Delta \log X_{ck,t-1}$.

The model for the initial observation involves Π_c in the mean and variance, and hence the full conditional posterior distribution is not normal. To sample Π_c we use the Metropolis-Hastings algorithm of Metropolis et al. (1953). We use the matrix normal distribution discussed above to sample the candidate Π_c^{cand}. As the candidate density is part of the target density, the acceptance-rejection probability simplifies to

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

where Π_c^{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 Σ_c

To sample Σ_c we notice that (10) is just a multivariate regression model. If we neglect the model for the first observation (13), the full conditional posterior distribution of Σ_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 $\varepsilon_{ct} = \Delta \log S_{ct} - \mu_c - \sum_{k=1}^{K} A_k \Delta \log X_{ckt} - \Pi_c(\log S_{c,t-1} - \sum_{k=1}^{K} B_{ck} \Delta \log X_{ck,t-1})$. Again, we cannot neglect the model for the initial observation (13) as Σ_c ends up in the variance of the error term. We use the Metropolis-Hastings sampler to simulate Σ_c. As candidate density we take the inverted Wishart distribution discussed above, which provides us Σ_c^{cand}. The candidate density is again part of the target density and hence the acceptance-rejection probability is

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

where Σ_c^{old} denotes the previous draw of Σ_c.

Sampling of Λ_1 and Λ_2

To sample Λ_1, we note that we can write (14)

$$\alpha'_{ic} = z'_{ic}\Lambda_1 + \eta'_{ic}, \tag{22}$$

and hence it is a multivariate regression model with regression matrix Λ_1. Hence, the full conditional posterior distribution of Λ_1 is a matrix normal distribution with mean

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

and covariance matrix

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

The derivation of the sampling distribution of Λ_2 proceeds in the same manner. The full conditional posterior distribution of Λ_2 is a matrix normal distribution with mean

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

and covariance matrix

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

Sampling of Σ_η and Σ_ν

To sample Σ_η we note that (14) is a multivariate regression model. Hence the full conditional posterior distribution of Σ_η 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 κ

terms results from the inverted Wishart prior on Σ_η which is used to improve convergence of our Gibbs sampler, see Hobert & Casella (1996) for a discussion.

The sampling of Σ_ν can be done in exactly the same manner. The parameter Σ_ν 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 μ_c and cross effects in A_{ck} and B_{ck}

To sample μ_c, and the parameters measuring the cross effect in A_{ck} and B_{ck} we first split up $X_{ckt} = (X_{ck1t}, \ldots, X_{ckI_ct})'$ for $k = 1, \ldots, K$ into two parts $X_{cit}^{\text{own}} = [X_{ckit}]_{k=1}^{K}$ and $X_{cit}^{\text{cross}} = [[X_{ckjt}]_{j=1 \neq i}^{I_c}]_{k=1}^{K}$ to disentangle the own effects from the cross effects. Define $X_{ct}^{\text{own}} = \text{diag}(X_{c1t}^{\text{own}}, \ldots, X_{cI_ct}^{\text{own}})'$ and $X_{ct}^{\text{cross}} = \text{diag}(X_{c1t}^{\text{cross}}, \ldots, X_{cI_ct}^{\text{cross}})'$. Equation (13) and (10) can now be written as

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

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

$$Y_{ct} = W_{ct} \gamma + \varepsilon_{ct}, \tag{28}$$

where Y_{ct} contains the left-hand side of (27), W_{ct} contains $(-\Pi_c^{-1} \vdots 0 \vdots \log X_{c1}^{\text{own}})$ for the first observation and $(\mathbf{I}_c \vdots \Delta \log X_{ct}^{\text{cross}} \vdots -\Pi_c \log X_{c,t-1}^{\text{own}})$ 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 Σ_c (V_c for the first observation). Hence, the full conditional distribution of γ is normal with mean

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

and covariance matrix

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

Sampling of α_c

To sample α_c we rewrite the second equation of (27) as

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

which can be written in matrix notation

$$Y_{ct} = W_{ct}\alpha_c + \varepsilon_{ct}, \tag{32}$$

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

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

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

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

Hence, the full conditional posterior distribution of α_c is normal with mean

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

and covariance matrix

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

Sampling of β_c

To sample β_c, we rewrite (27) as

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

which can be written in matrix notation

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

for $t = 1, \ldots, T_c$, 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

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

where U_c is a (KI_c)-dimensional vector containing the terms $\Lambda_2' z_{ic}$, $i = 1, \ldots, I_c$. The distribution of the error term ν_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 β_c is normal with mean

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

and covariance matrix

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

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 on 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 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

B.1 Category-specific variables

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

Average budget share of the category To measure the budget share of a category we use the average total expenditures in the category over time, that is, $\frac{1}{T_c} \sum_{t=1}^{T_c} \sum_{i=1}^{I_c} S_{ict} P_{ict}$.

Necessity The necessity of a product category is measured using two dimensions, that is, the *Utilitarian* and the *Hedonic* nature of the category. These characteristics are again operationalized with three levels (low, middle, high) and they are obtained using experts.

Competitive intensity We specify the competitive intensity using a market concentration index and the price dispersion.

Market concentration index for category c: $\sum_{i=1}^{I_c} \overline{M}_{ic} \log \overline{M}_{ic}$ (Raju 1992).

(Average) price dispersion in category c: $\sum_{t=1}^{T_c} \left(\max_i \left(RP_{ict}\right) - \min_i \left(RP_{ict}\right)\right) / \left(T_c \cdot \overline{RP}_c\right)$.

B.2 Category- and brand-specific variables

Frequency of price promotions

Brand-specific: Percentage of weeks where the price index 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.
Category-specific: The percentage of weeks in which at least one brand is on (at least 5%) promotion, that is, the percentage of weeks for which for at least for one i it holds that $PI_{ict} < 0.95$ (see Raju 1992).

Average depth of promotions

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

Frequency of display and feature promotions

Brand-specific: 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.
Category-specific: 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 depth of promotion can therefore be seen as the average probability that at least one SKU is on promotion.

B.3 Brand-specific variables

Brand size To measure the relative size of a brand in the category we use the market share.

Price segment As an indication of the price segment to which a brand belong in a certain category we construct a index based on the regular price, that is, $\frac{1}{T_c}\sum_{t=1}^{T_c}\frac{RP_{ict}}{\sum_{i=1}^{I_c} RP_{ict}/I_c}$.

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[1])
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[2])
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[3])
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[2])
This study	Category+brand char.	Immediate+dynamic	price elasticity, display/feature sensitivity	VAR (HB model)

[1]They do not need a two-step approach because they do not need to build a statistical model to capture variability in category sales.
[2]This is actually a two-stage approach because they use price elasticities obtained from another study.
[3]They explicitly account for uncertainty in the first-level estimates in the second stage.

Table 2: Overview of hypotheses on dynamic effects

	Short-run effect	Dynamic effect	Long-run effect
Storability	+	-	?
Average budget share of a category	?	?	?
Necessity	+	-	?
Number of brands	-	?	-/?
Frequency of price promotion	?	?	?
Average depth of promotion	+	-	?
Frequency of display activity	?	?	?
Frequency of feature activity	+	-	?
Frequency of coupon activity	?	?	?
Brand size	-	?	-/?
Price segment of a brand	?	?	?

Table 3: Posterior means of the effects of covariates on short-run and long-run effects of the marketing mix (λ_1 and λ_2 in 11 and 12), posterior standard deviations in parentheses.

Characteristic[†]	log Price index				Feature				Display			
	Short run		Long run		Short run		Long run		Short run		Long run	
Intercept	-2.303	(0.066)***	-1.952	(0.070)***	0.454	(0.033)***	0.258	(0.035)***	0.982	(0.053)***	0.810	(0.063)***
Brand level characteristics												
Price segment	-0.169	(0.068)**	-0.076	(0.080)	-0.040	(0.032)	-0.013	(0.035)	-0.005	(0.050)	-0.034	(0.056)
Brand size	0.139	(0.072)**	0.201	(0.077)**	0.099	(0.038)**	0.037	(0.040)	0.029	(0.058)	-0.168	(0.070)**
Depth price prom	-0.152	(0.074)*	-0.150	(0.055)***	-0.007	(0.036)	-0.038	(0.031)	0.035	(0.054)	-0.006	(0.041)
Price prom freq.	0.256	(0.118)**	0.222	(0.125)*	0.076	(0.061)	0.147	(0.065)**	0.287	(0.093)***	0.396	(0.104)***
Feature freq.	-0.046	(0.111)	0.099	(0.112)	-0.033	(0.057)	-0.043	(0.058)	-0.060	(0.089)	-0.117	(0.093)
Display freq.	-0.333	(0.145)**	-0.287	(0.141)**	-0.104	(0.072)	-0.032	(0.071)	-0.353	(0.108)***	-0.392	(0.115)***
Category level characteristics												
Avg. budget share	-0.301	(0.092)***	-0.177	(0.087)**	-0.021	(0.049)	-0.030	(0.051)	-0.130	(0.075)*	-0.356	(0.082)***
Price dispersion	0.257	(0.075)***	0.235	(0.081)***	-0.053	(0.037)	-0.014	(0.039)	-0.088	(0.059)	-0.098	(0.069)
Competitive index	-0.089	(0.072)	-0.129	(0.078)	0.031	(0.037)	-0.019	(0.039)	0.032	(0.057)	0.003	(0.060)
Depth price prom	-0.237	(0.091)**	-0.112	(0.087)	-0.006	(0.045)	0.025	(0.042)	-0.117	(0.070)*	-0.157	(0.069)**
Price prom freq	-0.093	(0.118)	-0.128	(0.134)	0.056	(0.060)	-0.035	(0.067)	-0.030	(0.090)	-0.174	(0.106)*
Feature freq.	-0.326	(0.130)***	-0.237	(0.128)*	-0.009	(0.066)	0.052	(0.069)	0.051	(0.104)	0.017	(0.112)
Display freq.	0.323	(0.148)**	0.216	(0.145)	0.098	(0.075)	0.031	(0.073)	0.243	(0.113)**	0.388	(0.117)***
Utilitarian	0.125	(0.122)	0.136	(0.128)	0.063	(0.061)	-0.029	(0.060)	0.212	(0.096)**	0.146	(0.111)
Hedonic	0.325	(0.134)**	0.085	(0.147)	-0.028	(0.066)	-0.093	(0.068)	-0.031	(0.102)	0.010	(0.119)
Perishability	0.061	(0.089)	-0.028	(0.103)	-0.070	(0.046)	-0.004	(0.047)	-0.104	(0.067)	0.003	(0.071)

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

[†] Characteristics are standardized to have mean 0 and variance 1.

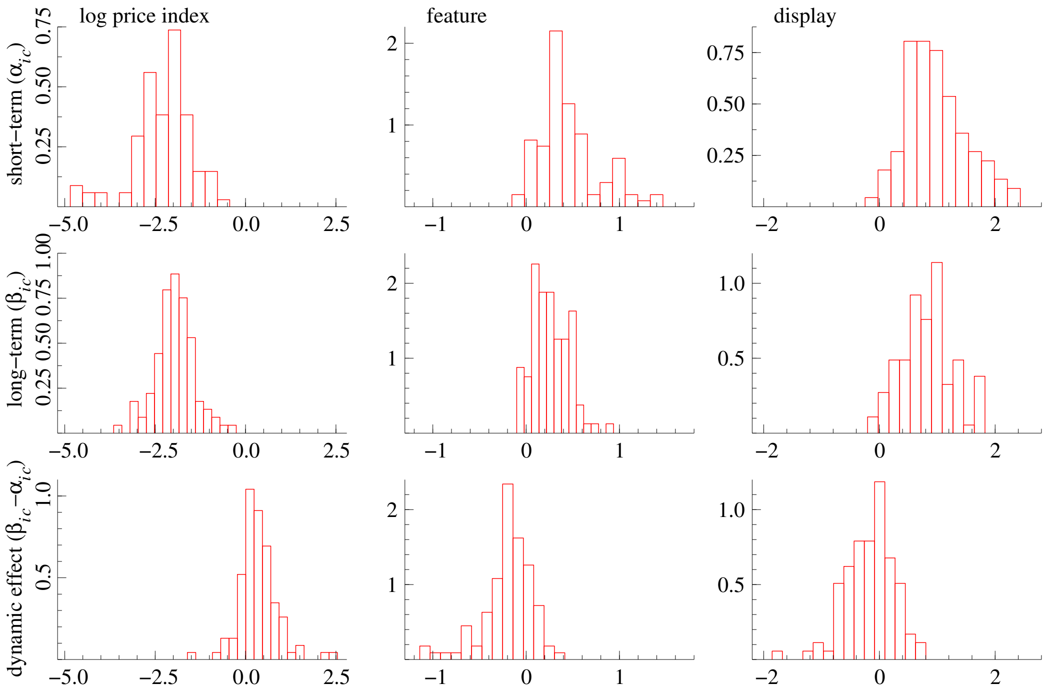

Figure 1: Histogram of posterior means of marketing-mix effectiveness for all 100 brands

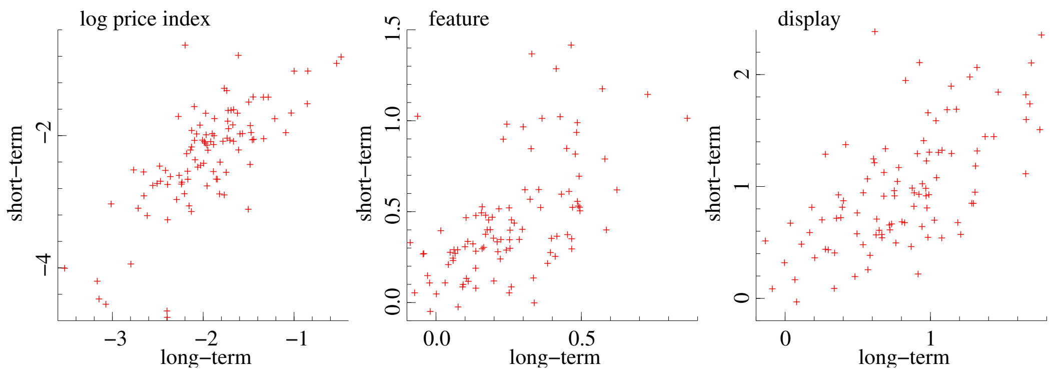

Figure 2: Scatter plots of long-run effects versus short-run effects (posterior mean per brand) for all 100 brands

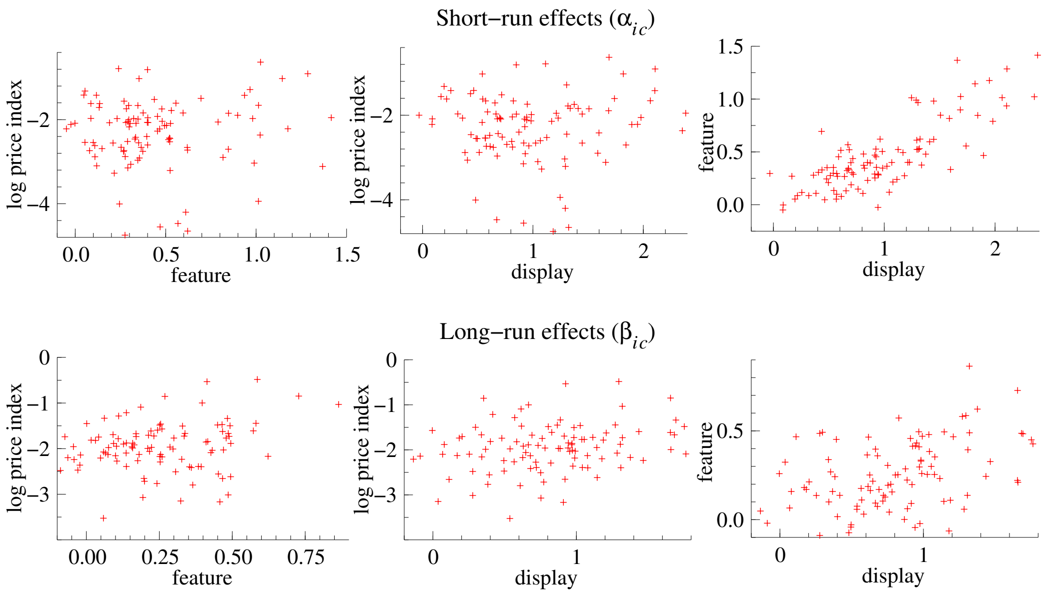

Figure 3: Scatter plots of posterior means or marketing-mix effectiveness

References

Bell, D. R., Chiang, J. & Padmanabhan, V. (1999), 'The decomposition of promotional response: an empirical generalization', *Marketing Science* **18**, 504–526.

Blattberg, R. C., Briesch, R. & Fox, E. J. (1995), 'How promotions work', *Marketing Science* **14**, G122–G132.

Bolton, R. N. (1989), 'The relayionship between market characteristics and promotional price elasticities', *Marketing Science* **8**, 153–169.

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**, 327–335.

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.

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 accomodate 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.

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. (1993), 'What do we learn about the long-run?', *Economic Journal* **103**, 307–317.

Grover, R. & Srinivasan, V. (1992), 'Evaluating the multiple effects of retail promotions on brand lowal and brand switching segments', *Journal of Marketing Research* **39**, 76–89.

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.

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', *Jurnal of Marketing Research* **34**, 248–261.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953), 'Equations of state calculations by fast computing machines', *Journal of Chemical Physics* **21**, 1087–1092.

Montgomery, A. L. (1997), 'Creating micro-marketing pricing strategies uning supermarket scanner data', *Marketing Science* **16**, 315–337.

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.

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* **forthcoming**.

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), *Macroeconomic 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. (2004), 'Impulse response analysis of structural vector autoregressive log-transformed time series', *International Journal of Forecasting* **forthcoming**.

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.