OR Spectrum (2017) 39:403–445
DOI 10.1007/s00291-016-0459-6

CrossMark

REGULAR ARTICLE

A comparison of semiparametric and heterogeneous store sales models for optimal category pricing

Anett Weber[1] · Winfried J. Steiner[1] · Stefan Lang[2]

Received: 23 January 2015 / Accepted: 7 July 2016 / Published online: 25 August 2016

Abstract Category management requires sales response models helping to simultaneously estimate marketing mix effects for all brands of a product category. We, therefore, develop a general heterogeneity seemingly unrelated regression (SUR) model accommodating correlations between sales across brands. This model contains a latent class SUR model, the well-known hierarchical Bayesian SUR model and the homogeneous SUR model as special cases. We further propose a hierarchical Bayesian semiparametric SUR model based on Bayesian P-splines which comprises a homogeneous semiparametric SUR model as nested version. The results of an empirical application with store-level scanner data indicate that the flexible SUR approaches of modeling price response clearly outperform the various parametric (homogeneous and heterogeneous) SUR approaches with respect to not only predictive validity but also total expected category profits. In particular, functional flexibility turns out to be the primary driver for improving the predictive performance of a store sales model as heterogeneity pays off only once functional flexibility has been accounted for. Furthermore, since both flexible SUR models perform nearly equally well with respect to expected category profits, a uniform pricing strategy which is much less complex to implement than micromarketing can be recommended for our data.

✉ Anett Weber
anett.weber@tu-clausthal.de

Winfried J. Steiner
winfried.steiner@tu-clausthal.de

Stefan Lang
stefan.lang@uibk.ac.at

1 Department of Marketing, Institute of Management and Economics, Clausthal University of Technology, Julius-Albert-Str. 2, 38678 Clausthal-Zellerfeld, Germany

2 Department of Statistics, Faculty of Economics and Statistics, University of Innsbruck, Universitätsstrasse 15, 6020 Innsbruck, Austria

Keywords Seemingly unrelated regression models · Between-store heterogeneity · Functional flexibility · Profit maximization · Category pricing

1 Introduction

One of the most important challenges of a retailer in today's competitive environment is setting prices for the various brands in the product categories he offers in his retail chain (Nijs et al. 2007). Following recommendations in the marketing literature, e.g., by Basuroy et al. (2001), Chen et al. (1999), Chintagunta (2002), Hall et al. (2010), Levy et al. (2004) and Zenor (1994), retailers should adopt a category management approach to develop a profitable pricing strategy. This approach implies that demand, costs and prices of competing brands belonging to the same product category are accounted for when prices of the brand under consideration are determined (Zenor 1994). Empirical evidence that the profitability of a retailer can be increased considerably by moving on to profit maximization at the category-level instead of maximization at the brand-level is among others given by Basuroy et al. (2001), Hall et al. (2010) and Zenor (1994). The trend toward category management requires models helping to simultaneously estimate demand or sales of all brands belonging to the same product category.

Zellner (1962) introduced a seemingly unrelated regression (SUR) model that accounts for correlations between response variables. A nonparametric version was, e.g., suggested by Smith and Kohn (2000) who demonstrated in their simulation study that estimated nonlinear effects are biased and inefficient if separate univariate regressions instead of a multivariate regression system are applied (Lang et al. 2003). Lang et al. (2003) further have shown in a simulation study that a semiparametric SUR model is efficient even if the error terms of the individual regression equations are actually uncorrelated. However, if the error terms are correlated and design matrices are different across equations the benefits from adopting the semiparametric SUR approach are substantial compared to the use of single semiparametric regressions.

In the context of store sales modeling, there are only few studies applying SUR models that account for correlations between sales across brands. Along the lines of Hruschka (2006a) these studies can be characterized by two groups: while the first group accounts for heterogeneity across stores but models sales response parametrically (e.g., Hoch et al. 1995; Kamakura and Kang 2007; Montgomery 1997; Montgomery and Rossi 1999)[1] the second group allows for functional flexibility but assumes homogeneous marketing mix effects (e.g., Lang et al. 2003). The present study investigates advantages of both groups and consolidates the two streams of SUR store models. In particular, we introduce a hierarchical Bayesian semiparametric SUR model which simultaneously accounts for store heterogeneity, functional flexibility and correlations between sales across brands. Hence, this model estimates store-specific nonlinear price effects simultaneously for all brands of a product category and constitutes a generalization of the hierarchical Bayesian semiparametric model at

[1] Further studies either account for heterogeneity by using only store dummies (e.g., Reibstein and Gatignon 1984) or by estimating sales response for each store separately (e.g., Mulhern and Leone 1991) or do not account for heterogeneity (e.g., Hall et al. 2010).

the brand-level introduced by Lang et al. (2015). Special cases are the homogeneous semiparametric SUR model (e.g., Lang et al. 2003) representing the research stream which accommodates functional flexibility only, as well as the hierarchical Bayesian SUR model (e.g., Montgomery 1997) representing the research stream which accounts for store heterogeneity only, as well as the simple homogeneous SUR model (e.g., Greene 2008; Zellner 1962) which constrains marketing effects to be equal across stores. Concerning the latter stream, we additionally develop a general heterogeneity SUR model by extending the general heterogeneity model (e.g. Allenby et al. 1998; Frühwirth-Schnatter 2006; Lenk and DeSarbo 2000; Rossi et al. 2005; Verbeke and Lesaffre 1996). Accordingly, this model derives segments of stores and allows for heterogeneity within these segments by providing store-specific parameters. It further comprises the simple homogeneous SUR model , a hierarchical Bayesian SUR model, and a latent class SUR model which yields segment-specific parameters and does not allow for heterogeneity within segments. To the best of our knowledge, the hierarchical Bayesian semiparametric SUR model, the general heterogeneity SUR model as well as the latent class SUR model are new in the marketing literature.

Retailers typically use some form of price discrimination in order to increase their profitability (Khan and Jain 2005). The various models of our study enable different types of price discrimination depending on their type of representing heterogeneity. Hence, the homogeneous SUR models (parametric as well as semiparametric) lead to a uniform pricing policy setting the same prices in all stores of a retail chain. The hierarchical Bayesian SUR models (parametric as well as semiparametric) and the general heterogeneity SUR model entail a store-level pricing policy (referred to as micromarketing) where prices are allowed to vary across stores. That way, differences in price response across different store locations can be exploited (Hoch et al. 1995). In-between the extremes of uniform versus store-level pricing lies the segment-level approach which leads to the same prices within segments or groups of stores and can be accomplished by the latent class SUR model as well as by the general heterogeneity SUR model. A special case of the latter pricing policy is the so-called zone pricing policy which is widely used in the US grocery retailing industry (e.g., Chintagunta et al. 2003; Dobson and Waterson 2008; Hoch et al. 1995; Montgomery 1997). Here, stores are assigned to different price zones which are defined almost exclusively by the extent of local competition (Hoch et al. 1995). For example, from the study of Hoch et al. (1995), it becomes evident that no relation between the price zones of the large supermarket chain Dominick's Finer Foods (DFF) and price sensitivities of consumers exists. Thus, there should be huge potential for improving a retailer's zone pricing policy by determining zones or segments of stores based on differences in price response across stores.

Montgomery (1997) estimates a hierarchical Bayesian SUR model which relates possible store-specific differences in marketing mix effects with characteristics concerning demographics and local competition of the stores. His study demonstrates the superiority of micromarketing with respect to expected profits compared to a uniform pricing policy. He further shows that the current zone pricing policy of DFF is indeed better than a policy without price discrimination but could be improved by modified zones. Chintagunta et al. (2003) report similar results in the context of a mixed logit demand model with random coefficients. Thus, profits can be increased considerably if the retail chain switches to a store-level pricing policy. However, this could lead to significant losses in customer welfare. Based on the store-level results of their micromarketing approach, they divided the stores into five zones which lead to higher expected profits compared to the existing zones of DFF emphasizing the potential for improving these zones, as well. Khan and Jain (2005) use an aggregate mixed logit approach and provide empirical evidence of significant profit increases when prices are allowed to vary across stores. Finally, Hruschka (2007) specifies his sales response model as a multilayer perceptron (neural network) with store-specific coefficients. The optimization procedure used simultaneously determines optimal prices and the number of clusters (segments of stores). That way, Hruschka shows that expected profits increase with the number of clusters and are highest for a store-level pricing policy.

The central concern of the present study is the comparison of sales response models that imply different pricing policies at different aggregation levels (chain-level vs. segment-level vs. store-level). The main difference compared to previous studies is, however, that models are already estimated at that level at which pricing decisions should be made. For example, we estimate a latent class SUR model yielding segment-specific price effects in order to determine segment-specific optimal prices, while a hierarchical Bayesian SUR model yielding store-specific price effects is estimated in order to determine store-specific optimal prices. So far, with the exception of Montgomery (1997) there is no study which simultaneously estimates store sales models for all brands of a product category *and* simultaneously determines optimal prices for those brands. However, the optimization results of Montgomery (1997) are based on only one model which provides coefficients at the store-level. In addition, to the best of our knowledge determining optimal prices based on a sales response model that allows for both functional flexibility in terms of a heterogeneous semiparametric specification of price effects *and* correlations between sales across brands has not yet been proposed in the relevant literature.[2]

Accordingly, the main objectives of our study can be summarized by the following questions:

1. Does a category-level sales response model that allows for correlations between sales across brands benefit from accommodating either store heterogeneity, or functional flexibility, or both features in terms of fit and predictive validity? And if, which of those features pays off more?
2. Do category-level store sales models with different representations of heterogeneity and/or flexibly modeled price effects provide different implications with respect to expected profits and optimal prices?
3. Should retailers adopt a store-level, segment-level or chain-level pricing policy?

To answer these questions, we follow Lang et al. (2015) and explore whether a store sales model benefits from accommodating either heterogeneity, functional flexibility, or both features by comparing parametric sales response models accounting for heterogeneity only, a homogeneous semiparametric sales response model that allows for functional flexibility only, and a heterogeneous semiparametric model accommodating

[2] Optimal prices based on a homogeneous semiparametric SUR model were already determined in Weber (2015).

both features to a simple model capturing none of these features. Our study, however, differs from that of Lang et al. (2015) in the following aspects: first, while Lang et al. (2015) estimate store sales models at the brand-level (i.e., separately for each brand), estimation of our models is conducted at the category-level (i.e., simultaneously for all brands) accounting for correlations between sales across brands. Second, in our parametric models store heterogeneity is accounted for by different representations of heterogeneity yielding store-specific effects (hierarchical Bayesian SUR model), segment-specific effects (latent class SUR model) as well as store-specific marketing effects within segments (general heterogeneity SUR model), whereas Lang et al. (2015) only address store-specific differences in sales response using the hierarchical Bayes model. Third, profit and pricing implications are obtained in Lang et al. (2015) through maximization of expected profits at the brand-level. In contrast, we maximize the category profit and determine optimal prices simultaneously for all brands of the product category.

Our results will indicate that the flexible SUR approaches of modeling price response clearly outperform the various parametric (homogeneous and heterogeneous) SUR approaches not only with respect to predictive validity but also with respect to total expected category profits. In particular, functional flexibility turns out to be the primary driver for improving the predictive performance of a store sales model as heterogeneity pays off only once functional flexibility has been accounted for. Furthermore, since the flexible SUR models perform nearly equally well with regard to expected profits a uniform pricing strategy which is much less complex to implement than micromarketing can be recommended for the retailer considered in our empirical application. Of course, profit and pricing implications may change when extending our model framework from the current category management point of view to a cross-category management approach, for example. We address this latter point in the summary and discussion section of the paper.

The rest of the paper is organized as follows: the next section introduces the general heterogeneity SUR model as well as the hierarchical Bayesian semiparametric SUR model and all nested versions of these models. Moreover, the procedure to determine optimal prices and expected profits for the whole product category is described. Subsequently, the various model versions are compared in an empirical application using store-level scanner data from both a statistical and a managerial point of view. A final discussion of results and an outlook on further research perspectives complete the paper.

2 Methodology

2.1 Modeling sales response

The general heterogeneity SUR model Our proposed model is a combination of the general heterogeneity model introduced by Verbeke and Lesaffre (1996) and the well-known seemingly unrelated regression model introduced by Zellner (1962). That way, we are able to model store-specific sales response within segments of stores and simultaneously allow for correlations between brands. To be more precise, for each brand the log unit sales are modeled as a sum of store-specific own- and cross-item price and own-item display effects as well as 9- and 99-ending price effects on the one hand and fixed holiday effects on the other hand:

$$y_{mit} = \ln Q_{mit} = \gamma_{mi0} + \sum_{j=1}^{M} \beta_{mij} \ln(P_{jit}) + \sum_{l=1}^{L} \gamma_{mil} D^{l}_{mit} + \delta_m E_t + \epsilon_{mit}, \quad \epsilon_{mit} \sim N(0, \sigma^2), \tag{1}$$

with

- Q_{mit} unit sales of brand m ($m = 1, \ldots, M$) in store i ($i = 1, \ldots, N$) and week t ($t = 1, \ldots, T_i$);
- P_{jit} observed unit price for brand j in store i and week t;
- D^{l}_{mit} further observed covariates ($l = 1, \ldots, L$) including own display and price ending variables;
- E_t dummy variable indicating if there is a holiday in week t (=1) or not (=0);
- γ_{mi0} store effect for brand m accounting for differences in baseline sales across stores (e.g., due to their spatial location);
- β_{mij} own-item price effect ($j = m$) and cross-item price effects ($j \neq m$) for brand m in store i;
- γ_{mil} effects of further covariates ($l = 1, \ldots, L$) for brand m in store i;
- δ_m holiday effect for brand m and
- ϵ_{mit} disturbance term.

Using matrix notation model (1) can be rewritten as

$$y_{mi} = X_{mi}\alpha_m + W_{mi}\beta_{mi} + \epsilon_{mi}, \qquad \epsilon_{mi} \sim N(0, \sigma^2_m I_{T_i}), \tag{2}$$

where y_{mi} is the vector of T_i observed log unit sales of brand m in store i, X_{mi} is a $T_i \times d_m$ design matrix for the fixed effects of brand m which are equal across stores, and W_{mi} is a $T_i \times r_m$ design matrix for the store-specific effects of brand m. The corresponding effects are contained in the $d_m \times 1$ vector α_m and in the $r_m \times 1$ vector β_{mi}, respectively. ϵ_{mi} is a $T_i \times 1$ vector of error terms following a multivariate normal distribution with mean 0 and covariance matrix $\sigma^2_m I_{T_i}$. The $M \times 1$ vector of error terms $(\epsilon_{1it}, \ldots, \epsilon_{Mit})'$ across brands follows a multivariate normal distribution $N(0, \Sigma)$, too.[3] Note that σ^2_m corresponds to the mth diagonal element of Σ.

We obtain the general heterogeneity SUR model by combining the sales response models of all brands into one system of equations:

$$\begin{bmatrix} y_{1i} \\ \vdots \\ y_{Mi} \end{bmatrix} = \begin{bmatrix} X_{1i} & & \\ & \ddots & \\ & & X_{Mi} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_M \end{bmatrix} + \begin{bmatrix} W_{1i} & & \\ & \ddots & \\ & & W_{Mi} \end{bmatrix} \begin{bmatrix} \beta_{1i} \\ \vdots \\ \beta_{Mi} \end{bmatrix} + \begin{bmatrix} \epsilon_{1i} \\ \vdots \\ \epsilon_{Mi} \end{bmatrix}, \tag{3}$$

[3] For simplicity, we assume that correlations between brands are the same for all stores, i.e. $\Sigma_i = \Sigma \; \forall i$.

or in a more compact form specified as

$$y_i = X_i\alpha + W_i\beta_i + \epsilon_i, \quad \epsilon_i \sim N\left(0, \Sigma \otimes I_{T_i}\right). \tag{4}$$

The general heterogeneity (SUR) model assumes that each store-specific coefficient β_i can be assigned to one of K segments and follows a multivariate normal distribution within the segment (Lenk and DeSarbo 2000). Hence, the distribution of heterogeneity is modeled as a mixture of multivariate normal distributions

$$\beta_i \sim \sum_{k=1}^{K} \eta_k N\left(\beta_k^G, Q_k^G\right) \tag{5}$$

with unknown segment means $\beta_1^G, \ldots, \beta_K^G$, unknown covariance matrices $Q_1^G, \ldots, Q_K^G$ and unknown segment probabilities $\eta = (\eta_1, \ldots, \eta_K)'$ (Frühwirth-Schnatter et al. 2005). If class membership is known β_i is multivariate normally distributed with mean β_k^G and covariance matrix Q_k^G if store i belongs to segment k. In this case, η_k is the proportion of all stores belonging to segment k (Lenk and DeSarbo 2000). By introducing a latent segment indicator S_i, $i = 1, \ldots, N$, taking values in $\{1, \ldots, K\}$ with probabilities $P(S_i = k) = \eta_k$ model (5) can be written as (Frühwirth-Schnatter et al. 2004):

$$\beta_i \sim \begin{cases} N\left(\beta_1^G, Q_1^G\right), & \text{if } S_i = 1, \\ \quad\vdots \\ N\left(\beta_K^G, Q_K^G\right), & \text{if } S_i = K. \end{cases} \tag{6}$$

Depending on the number of segments K and the covariance matrices Q_k^G $(k = 1, \ldots, K)$ there are some special cases of the general heterogeneity SUR model:

1. For $K = 1$ and $Q_k^G = 0$, we obtain the conventional homogeneous SUR model as introduced by Zellner (1962). Consequently, all store-specific coefficients β_i correspond to the single mean β_1^G and are equal across stores.
2. For $K = 1$ and $Q_k^G \neq 0$, a hierarchical Bayesian SUR model results, e.g., as it was applied by Montgomery (1997). Here, all store-specific coefficients β_i follow the same multivariate normal distribution.
3. For $K \neq 1$ and $Q_k^G = 0$ $(k = 1, \ldots, K)$, we obtain a latent class SUR model. Accordingly, all store-specific coefficients are equal to the corresponding segment mean but vary across segments.

Within a fully Bayesian framework the unknown parameters $\phi = (\alpha, \beta_1^G, \ldots, \beta_K^G, \eta, Q_1^G, \ldots, Q_K^G, \Sigma)$ have to be supplemented with prior distributions. According to common practice, we use conjugate priors (following e.g. Allenby et al. 1998; Frühwirth-Schnatter et al. 2005; Koop 2003; Lenk and DeSarbo 2000). Hence, we use a Dirichlet prior $D(e_{01}, \ldots, e_{0K})$ for the segment probabilities η and place multivariate normal priors $N(\mu_{0\alpha}, \Sigma_{0\alpha})$ and $N(\mu_{0\beta_k^G}, \Sigma_{0\beta_k^G})$ on the fixed effects α and the segment-specific effects β_k^G $(k = 1, \ldots, K)$, respectively. Finally, the covariance matrices Q_k^G $(k = 1, \ldots, K)$ and Σ are supplemented with inverse Wishart priors $IW(a_{0Q_k^G}, B_{0Q_k^G})$ and $IW(a_{0\Sigma}, B_{0\Sigma})$, respectively.

Estimation of the general heterogeneity SUR model is based on the partly marginalized random permutation Gibbs sampling algorithm which was suggested by Frühwirth-Schnatter et al. (2004) in the context of single regression models. We adapted this algorithm to our more general SUR approach.[4] Details on further prior specifications and the sampling scheme are provided in Appendix 1.

The hierarchical Bayesian semiparametric SUR model We consider a hierarchical Bayesian semiparametric SUR approach by combining the hierarchical Bayes SUR model nested in the general heterogeneity SUR model with a semiparametric SUR model similar to that of Lang et al. (2003).[5] Contrary to the parametric approach, own- and cross-item price effects are now modeled nonparametrically leading to the following heterogeneous semiparametric sales response model for one brand m:

$$y_{mit} = \ln Q_{mit} = \gamma_{mi0} + \sum_{j=1}^{M}(1 + \alpha_{mij}) f_{mj}(P_{jit}) + \sum_{l=1}^{L} \gamma_{mil} D_{mit}^{l} + \delta_m E_t + \epsilon_{mit}, \quad \epsilon_{mit} \sim N(0, \sigma^2), \tag{7}$$

with unknown smooth functions $f_{mj}(P_{jit})$, $m = 1, \ldots, M$ and $j = 1, \ldots, M$, multiplied by store-specific random effects. Following Steiner et al. (2007), we apply a Bayesian P(enalized)-spline approach. Accordingly, the unknown smooth functions f_{mj} can be written as a linear combination of B-spline basis functions (e.g. Brezger and Steiner 2008) which is given by

$$f_{mj}(x) = \sum_{o=1}^{O_{mj}} \beta_{mjo} B_{mjo}(x) \tag{8}$$

for the jth price effect of brand m. Penalization of too large deviations between adjacent regression coefficients is accomplished by second order random walks (e.g., Lang and Brezger 2004):

$$\beta_{mjo} = 2\beta_{mj,o-1} - \beta_{mj,o-2} + u_{mjo}, \quad u_{mjo} \sim N(0, \tau_{mj}^2). \tag{9}$$

Furthermore, following Brezger and Steiner (2008), monotonicity constraints are imposed on own- and cross-item price effects.

We obtain a hierarchical Bayesian semiparametric SUR model by combining the sales response models of all brands into one system of equations:

4 Identification problems due to label switching (Celeux et al. 2000) are considered by determining identifiability constraints as described in Frühwirth-Schnatter et al. (2004). Subsequently, we use a constrained permutation sampler as suggested in Frühwirth-Schnatter (2001).

5 Thus, our proposed model constitutes an extension of the hierarchical Bayesian semiparametric approach introduced by Lang et al. (2015).

$$\begin{aligned}
\begin{bmatrix} y_1 \\ \vdots \\ y_M \end{bmatrix} &= \begin{bmatrix} A_{11}X_{11} & & \\ & \ddots & \\ & & A_{M1}X_{M1} \end{bmatrix} \begin{bmatrix} \beta_{11} \\ \vdots \\ \beta_{M1} \end{bmatrix} + \cdots \\
&+ \begin{bmatrix} A_{1M}X_{1M} & & \\ & \ddots & \\ & & A_{MM}X_{MM} \end{bmatrix} \begin{bmatrix} \beta_{1M} \\ \vdots \\ \beta_{MM} \end{bmatrix} \\
&+ \begin{bmatrix} W_{10}Z & & \\ & \ddots & \\ & & W_{M0}Z \end{bmatrix} \begin{bmatrix} \gamma_{10} \\ \vdots \\ \gamma_{M0} \end{bmatrix} + \cdots + \begin{bmatrix} W_{1L}Z & & \\ & \ddots & \\ & & W_{ML}Z \end{bmatrix} \begin{bmatrix} \gamma_{1L} \\ \vdots \\ \gamma_{ML} \end{bmatrix} \\
&+ \begin{bmatrix} V_1 & & \\ & \ddots & \\ & & V_M \end{bmatrix} \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_M \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_M \end{bmatrix}, \quad \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_M \end{bmatrix} \sim N\left(0, \Sigma \otimes I_n\right). \qquad (10)
\end{aligned}$$

Here, $n = \sum_{i=1}^{N} T_i$ is the number of all observations with respect to one brand m (observations, i.e., weeks and stores, are the same across brands). X_{mj} $(m, j = 1, \ldots, M)$ corresponds to the design matrix for the jth price effect of brand m with elements $X_{mj}(g, o) = B_{mjo}(x_{jg})$ where B_{mjo} is the oth B-spline basis function for the jth nonparametric effect of brand m and x_{jg} denotes the gth observation of the jth price variable. $A_{mj} = \text{diag}(1 + \alpha_{mi_1 j}, \ldots, 1 + \alpha_{mi_n j})$ $(m, j = 1, \ldots, M)$ constitutes an $n \times n$ diagonal matrix of random scaling factors where $i_g \in \{1, \ldots, N\}$ denotes the store to which the g-th observation belongs. W_{ml} $(m = 1, \ldots, M$ and $l = 0, \ldots, L)$ is an $n \times n$ diagonal matrix of the observations which belong to the lth parametrically and store-specifically modeled variable (with W_{m0} being an $n \times n$ identity matrix), and Z is a matrix indicating if observation g belongs to store i (in this case $Z(g, i)$ equals 1, otherwise it equals 0). Finally, V_m $(m = 1, \ldots, M)$ is the usual design matrix for the homogeneous parametric effects of brand m.

To accomplish the Bayesian specification, the random scaling factors contained in $\alpha_{mj} = (\alpha_{m1j}, \ldots, \alpha_{mNj})'$ as well as the store-specific parametric effects $\gamma_{ml} = (\gamma_{m1l}, \ldots, \gamma_{mNl})'$ are a priori assumed to be multivariate normal distributed with $N(0, \phi_{mj}^2 I_n)$ and $N(0, \psi_{ml}^2 I_n)$, respectively. The variance parameters ϕ_{mj}^2 $(m, j = 1, \ldots, M)$ of the multiplicative random effects and ψ_{ml}^2 $(m = 1, \ldots, M$ and $l = 1, \ldots, L)$ of further additive random effects are supplemented with inverse gamma priors $IG(0.001, 0.001)$. Inverse gamma priors $IG(0.001, 0.001)$ are also assigned to all variance parameters τ_{mj}^2 $(m, j = 1, \ldots, M)$ which control the amount of smoothness of the P-splines (Brezger and Steiner 2008). Choosing the value 0.001 for both hyperparameters of the inverse gamma distribution leads to noninformative priors for all variance parameters. A diffuse prior $\delta \propto$ const is used for the parameter vector $\delta = (\delta_1, \ldots, \delta_M)'$ and an inverse Wishart distribution $IW(1, 0.005 I_M)$ is placed on the covariance matrix Σ.

Note, that the proposed model reduces to a homogeneous semiparametric SUR model similar to that of Lang et al. (2003) if $\alpha_{mij} = 0$ for $m, j = 1, \ldots, M$ and $i = 1, \ldots, N$, and $\gamma_{mil} = \overline{\gamma}_{ml}$ for $m = 1, \ldots, M, i = 1, \ldots, N$ and $l = 1, \ldots, L$, to a hierarchical Bayesian parametric SUR model if all nonlinear functions $f_{mj}, m, j = 1, \ldots, M$, are replaced by linear effects, and to a simple homogeneous parametric SUR model if *both* $\alpha_{mij} = 0$ for $m, j = 1, \ldots, M$ and $i = 1, \ldots, N$, $\gamma_{mil} = \overline{\gamma}_{ml}$ for $m = 1, \ldots, M, i = 1, \ldots, N$ and $l = 1, \ldots, L$, *and* all nonlinear functions f_{mj}, $m, j = 1, \ldots, M$, are replaced by linear effects (compare Eq. (7)).

Estimation is fully Bayesian. Details on Bayesian inference are provided in Appendix 2.

2.2 Optimal category pricing

The main objective of a retailer is to maximize the chain-level profit Π_t of a product category in each week[6] t (e.g., Chintagunta et al. 2003; Hruschka 2007; Khan and Jain 2005; Kim et al. 1995; Montgomery and Bradlow 1999; Vilcassim and Chintagunta 1995) which is given by

$$\Pi_t = \sum_{m=1}^{M} \sum_{i=1}^{N} (price_{mit}^{\text{opt}} - c_{mt}) \hat{Q}_{mit}(\mathbf{p}_{it}^{\text{opt}}). \tag{11}$$

Here, c_{mt} denotes the wholesale cost of brand m in week t and $\hat{Q}_{mit}$ are the predicted unit sales of brand m in store i and week t which depend on the (optimal) prices of all brands. Thus, optimal prices $\mathbf{p}_{it}^{\text{opt}} = (price_{1it}^{\text{opt}}, \ldots, price_{Mit}^{\text{opt}})$ are determined for each store i, $i = 1, \ldots, N$, in a given week t. To compare pricing scenarios at different levels of aggregation (chain-level vs. segment-level vs. store-level) and following Hruschka (2007), we define

$$price_{mit}^{\text{opt}} = \sum_{k=1}^{K} \mathbb{1}(S_i = k) price_{mkt}^{\text{opt}}, \quad m = 1, \ldots, M, \ i = 1, \ldots, N. \tag{12}$$

Accordingly, the optimal price $price_{mit}^{\text{opt}}$ of store i is equal to the optimal price $price_{mkt}^{\text{opt}}$ of segment k if store i belongs to segment k indicated by $\mathbb{1}(S_i = k)$ (segment-level pricing). If there is only one segment, i.e., $K = 1$, optimal prices are the same for all stores of the retail chain (chain-level or uniform pricing). If the number of segments is equal to the number of stores, i.e., $K = N$, optimal prices vary across stores[7] (store-level pricing or micromarketing).

We further account for model uncertainty when determining optimal prices. In particular, expected profits in week t are considered as the arithmetic mean across U Gibbs samples of coefficients[8] (e.g., Hruschka 2007; Montgomery 1997). Accordingly, the expected profit of the whole product category in week t can be written as

6 According to, e.g., Chintagunta et al. (2003), Kadiyali et al. (2000) and Sudhir (2001), retailers (as well as manufacturers) make their pricing decisions every week.

7 Note that store-specific or micromarketing pricing strategies can be derived not only from the hierarchical Bayesian SUR model but also from the general heterogeneity SUR model since price coefficients are store-specific in the latter model as well.

8 We set $U = 200$.

$$E[\Pi_t] = \frac{1}{U}\sum_{u=1}^{U}\sum_{m=1}^{M}\sum_{i=1}^{N}(price_{mit}^{\text{opt}} - c_{mt})\hat{Q}_{mit}^{u}(\mathbf{p}_{it}^{\text{opt}}) \tag{13}$$

where $\hat{Q}_{mit}^{u}(\mathbf{p}_{it}^{\text{opt}})$ are the predicted sales of brand m in store i and week t computed by using the coefficient sample drawn in iteration u of the Gibbs sampling algorithm (after the burn-in period) instead of estimates of conditional means. Since there are no interactions between segments or stores (Montgomery 1997), optimization can be done separately for each segment or store, respectively. Therefore, we rewrite Eq. (13) as

$$E[\Pi_t] = \sum_{k=1}^{K} E[\Pi_{kt}], \tag{14}$$

where

$$E[\Pi_{kt}] = \frac{1}{U}\sum_{u=1}^{U}\sum_{m=1}^{M}\sum_{i=1}^{N}\mathbb{1}(S_i = k)(price_{mkt}^{\text{opt}} - c_{mt})\hat{Q}_{mit}^{u}(\mathbf{p}_{kt}^{\text{opt}}). \tag{15}$$

Finally, total expected profits result by summing over weeks $t = 1, \ldots, 89$:

$$E\left[\Pi\right] = \sum_{t=1}^{89} E\left[\Pi_t\right]. \tag{16}$$

According to Hruschka (2007), Khan and Jain (2005) and Montgomery (1997), the average price level of the product category (after brand prices were optimized) should not deviate too much from its currently observed level in order to leave the retailer's image unchanged. In all three studies, price levels are defined as the market share weighted average price across all brands in the product category. Contrary to them, we consider the retailer's current price image as the simple (unweighted) average price across all brands in the product category which reflects the price structure as perceived by consumers. While Hruschka (2007) accounts for that price level restriction at the store-level, Khan and Jain (2005) and Montgomery (1997) impose this constraint at that level at which pricing decisions are made. Here, we require that optimized price levels (chain-, segment- and store-level) must not deviate too much from the average observed price level across all stores of the chain in a given week t which can be written as

$$pl_t^{\text{obs}} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{M}\sum_{m=1}^{M} price_{mit}, \tag{17}$$

where $price_{mit}$ is the observed price of brand m in store i and week t.

The new price level of segment k in a given week t based on optimal prices can be defined as

$$pl_{kt}^{\text{new}}(\mathbf{p}_{kt}^{\text{opt}}) = \frac{1}{M}\sum_{m=1}^{M} price_{mkt}^{\text{opt}}. \tag{18}$$

Note that the price level is the same for all stores of segment k since optimal prices are the same, as well. If there is only one segment we receive the same price level for all stores of the chain and if there are as much segments as stores store-specific new price levels result.

Montgomery (1997) claims that observed and new price levels have to be equal which is a really severe restriction. Khan and Jain (2005) constrain the new price levels to be smaller than or at least equal to the observed ones. Similar to Hruschka (2007), we allow for small deviations between new and observed price levels. In particular, we constrain new price levels to deviate not more than 1 % from observed price levels in a considered week:

$$a_{kt} := \max\left(1 - \frac{pl_{kt}^{\text{new}}(\mathbf{p}_{kt}^{\text{opt}})}{pl_t^{\text{obs}}}, 1 - \frac{pl_t^{\text{obs}}}{pl_{kt}^{\text{new}}(\mathbf{p}_{kt}^{\text{opt}})}\right) \leq 0.01 \tag{19}$$

Finally, optimal prices for a certain brand are assumed to lie within the price range as observed for this brand.

The inequality constraints (19) are incorporated into the optimization procedure by using an objective function including a penalty function (e.g., Weicker 2007). Hence, we obtain the following maximization problem for segment k ($k = 1, \ldots, K$) in week t ($t = 1, \ldots, 89$):

$$\max_{\mathbf{p}_{kt}^{\text{opt}}} \; Z = \begin{cases} E\left[\Pi_{kt}\right], & \text{if } a_{kt} \leq 0.01 \\ -10000, & \text{otherwise} \end{cases} \tag{20}$$

$$\text{subject to} \quad \min_{i,t}\{price_{mit}\} \leq price_{mkt}^{\text{opt}} \leq \max_{i,t}\{price_{mit}\}, \quad m = 1, \ldots, M.$$

Since it is common practice that retailers base their pricing decisions on a sales promotion calendar (e.g., Silva-Risso et al. 1999), optimization is carried out conditionally on the given promotional strategy (e.g., Chintagunta et al. 2003; Kim et al. 1995), relating to the use of displays in our data. Of course, it would also be possible to determine optimal prices and expected profits for different scenarios of display usage. Thus, the explanatory power of our study is not limited from a managerial point of view (Vilcassim and Chintagunta 1995).

3 Empirical application

3.1 Data and variable specification

The data set used for estimation comes from a major supermarket chain named Dominick's Finer Foods (DFF) in the Chicago metropolitan area and was provided by the James M. Kilts Center, University of Chicago. It refers to eight brands of the refrigerated orange juice category and contains weekly collected observations from 81 stores over a time span ranging from 67 to 85 weeks including unit sales, retail prices and display activities for the various brands[9] (Brezger and Steiner 2008; Steiner et al. 2007). Descriptive statistics are summarized in Table 1.

The mean prices indicate the differences in quality across the three quality tiers of premium brands (Florida Natural and Tropicana Pure), national brands (Citrus Hill, Florida Gold, Minute Maid, Tree Fresh and Tropicana) and the supermarket's private label brand (Dominick's). Thus, the premium brands are the most expensive brands and the private label brand Dominick's is the cheapest brand. The same applies to the mean costs (wholesale costs) which significantly differ across the three quality tiers, as well. Price and cost variation of a brand in one store are measured by the coefficient of variation which corresponds to the ratio of the standard deviation of actual prices or costs and the mean actual price or cost, respectively (Bolton and Shankar 2003; Shankar and Krishnamurthi 1996). Averages of the coefficients of variation across stores are provided in Table 1, too. From the third and fifth column, it becomes evident that prices as well as costs are substantially different over time. Moreover, DFF varies prices of a brand across stores in most weeks setting up to 27 different price levels across the 81 stores. On average five different price levels of a brand are found in one week while weekly wholesale costs are nearly constant across stores. Finally, we investigated the pricing behavior of DFF. In Table 1, the correlations between prices and costs are displayed for the various brands. Correlations are rather small for the premium brand Florida Natural as well as the store brand Dominick's and moderate for the remaining brands. This finding indicates that DFF does not use a simple cost plus approach for setting prices.

To avoid the problem of multicollinearity, cross-item price effects are specified at the tier level. Precisely, we define new independent variables $price\_national_{it}$ and $price\_premium_{it}$ capturing the lowest price of any competing national or premium brand in store i and week t, respectively, while $price\_private_{it}$ corresponds to the price of the only private label brand Dominick's.[10] Furthermore, 9- ($end9$) and 99-ending ($end99$) own-item price effects are taken into account. Hence, within our SUR framework and exemplarily for the general heterogeneity SUR model, the sales response model of one brand in store i and week t is given by[11]:

$$\begin{aligned}\ln Q_{it} = {} & \gamma_{i0} + \beta_{i1} \cdot \ln(price_{it}) + \beta_{i2} \cdot \ln(price\_premium_{it}) \\ & + \beta_{i3} \cdot \ln(price\_national_{it}) + \beta_{i4} \cdot \ln(price\_private_{it}) \\ & + \gamma_{i1} \cdot display_{it} + \gamma_{i2} \cdot end9_{it} + \gamma_{i3} \cdot end99_{it} + \delta \cdot E_t + \epsilon_{it}. \quad (21)\end{aligned}$$

Note that the term $\beta_{i4} \cdot \ln(price\_private_{it})$ has to be omitted if the sales equation is specified for the private label brand Dominick's.

9 Originally, the data is provided at the UPC-level and consists of 18 UPCs. In order to handle this number of products we aggregated highly correlated UPCs to 8 brands accounting for a market share of about 96 % in the refrigerated orange juice category (64 oz) during the considered time span.

10 For details compare Steiner et al. (2007, p. 387).

11 The notation which is not explained here is adopted from model (1).

Table 1 Descriptive statistics for prices, costs, market shares and display activity of brands in the refrigerated orange juice category

Brand	Prices ($)		Costs ($)		Correlation between prices and costs	Average weekly market shares (%)	Average display activity (%)
	Mean [min; max]	Coeff. of variation	Mean [min; max]	Coeff. of variation			
Florida Natural (FN)	2.86 [1.57; 3.33]	11.39	1.91 [0.24; 2.33]	13.33	0.30	4	34
Tropicana Pure (TP)	2.95 [1.49; 3.87]	18.16	2.08 [0.83; 2.71]	15.89	0.57	12	49
Citrus Hill (CH)	2.31 [1.13; 3.07]	14.09	1.73 [0.37; 2.01]	11.39	0.68	8	34
Florida Gold (FG)	2.17 [0.99; 3.08]	18.59	1.46 [0.54; 1.95]	16.42	0.63	6	38
Minute Maid (MM)	2.22 [1.29; 3.17]	18.64	1.64 [0.27; 2.07]	11.83	0.64	10	47
Tree Fresh (TF)	2.16 [0.99; 2.69]	12.94	1.53 [0.95; 1.92]	17.10	0.53	8	33
Tropicana (T)	2.19 [1.49; 2.99]	16.73	1.53 [0.49; 1.91]	10.83	0.58	18	43
Dominick's (D)	1.76 [0.99; 2.69]	22.48	1.19 [0.48; 1.57]	16.32	0.41	35	28

3.2 Benchmark models and model performance

We estimate all versions of the general heterogeneity SUR model as well as the hierarchical Bayesian semiparametric SUR model and its nested homogeneous version, i.e., we estimate

1. a parametric homogeneous[12] SUR model (PHomSM) that constrains marketing effects to be equal across stores,
2. a parametric latent class SUR model (PLCSM) with different numbers of segments (two to six) yielding segment-specific marketing effects,
3. a parametric hierarchical Bayesian SUR model (PHBSM) that reveals store-specific marketing effects,
4. a parametric general heterogeneity SUR model (PHetSM) with two segments[13] allowing for heterogeneous store-specific marketing effects within segments,
5. a flexible (semiparametric) homogeneous SUR model (FHomSM) that allows for flexible nonlinear price effects which are equal across stores, and
6. a flexible hierarchical Bayesian semiparametric SUR model (FHBSM) providing store-specific nonlinear price effects.

Comparison of estimation results is based on model fit in terms of the deviance information criterion (Fahrmeir et al. 2007) and the log model likelihood (Frühwirth-Schnatter 2006; Rossi et al. 2005), predictive performance measured by the root mean squared sales prediction error (RMSE), estimated price elasticities as well as estimated price effects. To assess the predictive validity, we randomly split the data (of each store) into an estimation sample including about 75 % of the observations and a validation sample for the remaining 25 % of observations, and subsequently computed the RMSE values in both the estimation (in-sample RMSE) and the validation sample (out-of-sample RMSE). Finally, expected profits and optimal prices are obtained by determining the solution of the maximization problem (20). This provides uniform prices equal across stores for the homogeneous SUR models (PHomSM, FHomSM), segment-specific prices equal within segments for the latent class SUR model (PLCSM), and store-specific prices for the heterogeneous SUR models (PHBSM, PHetSM, FHBSM). The optimization procedure is based on an evolutionary algorithm with a derivative-based (quasi-Newton) method (called GENOUD = GENetic Optimization Using Derivatives) which is implemented in the statistical software R[14] (see Sekhon and Mebane 1998 for the description of the algorithm and Mebane and Sekhon 2011 for details about the implementation in R using the function *genoud()* of the R package *rgenoud*). More details about the evolutionary algorithm, the specification of the arguments of the function *genoud()* that control its performance, and the size of our optimization problems are given in Appendix 4.

[12] Homogeneity refers to the effects of all independent variables. The model, however, contains store-specific intercepts in order to account for differences in baseline sales across stores (like all other models do). A table summarizing all model specifications is given in Appendix 3.

[13] We will explain below why we abstain from estimating the PHetSM for a higher number of segments.

[14] The software R is free and available at http://cran.r-project.org/.

Table 2 Fit and predictive validity for all SUR model versions (number of segments in parentheses)

Model	DIC	logML	Out-of-sample RMSE	Out-of-sample improvement (%)	In-sample improvement (%)
PHomSM	99446	−52474	146	–	–
PLCSM (2)	98988	−52516	147	0.36	−0.50
PLCSM (3)	99282	−52863	147	0.49	−0.97
PLCSM (4)	99230	−53059	147	0.69	−0.98
PLCSM (5)	99236	−53208	147	0.64	−1.04
PLCSM (6)	99316	−53431	149	1.49	−1.32
PHBSM	99184	**−52464**	147	0.39	−3.58
PHetSM (2)	100973	−55329	161	9.74	−6.03
FHomSM	93499	–	138	−5.73	−9.26
FHBSM	**91730**	–	**131**	−10.48	−16.01

3.3 Estimation results

Fit and predictive validity The estimation results of our models are provided in Table 2 (with best models indicated in **bold**). With respect to model fit measured in terms of the deviance information criterion (DIC) and the log model likelihood (logML), the less complex parametric SUR models clearly outperform PHetSM, with PLCSM with two segments and PHBSM being the best models, respectively. This finding can be explained as follows: First, although PHetSM is the most flexible parametric model the DIC as well as the model likelihood penalize for model complexity (Frühwirth-Schnatter 2006; Spiegelhalter et al. 2002) leading to a larger (worse) DIC value and a smaller (worse) logML-value, respectively. Second, the PHetSM two-segment solution yields one empty segment with no stores assigned to it and a corresponding estimated segment size of almost zero. For that reason, we did not increase the number of segments for that model. Hence, according to this most complex parametric SUR model the data do not support two (or more) segments with additional inner-segment variation. We therefore exclude this model in the subsequent analyses. The differences in DIC values as well as logML-values between the less complex parametric model versions are, moreover, rather moderate. The flexible SUR models, however, lead to a considerably better model fit compared to the parametric models providing much smaller DIC values.[15]

Concerning predictive validity, PHomSM turned out to be the best parametric model. Again, differences between the parametric model versions are rather small. In contrast, predictive validity could be improved considerably by FHomSM and FHBSM with impressing relative improvements of about 6 and 10 % compared to PHomSM, respectively.

[15] It is not possible to compute the log model likelihood for the flexible models since this measure of fit is not applicable in models with improper priors as are used for the nonlinear functions of FHomSM and FHBSM.

Analysis of in-sample improvements of RMSE values over PHomSM reveals that the more complex the model the better is the in-sample fit. This finding is reasonable since this measure of fit ignores model complexity. Combining the results of in- and out-of-sample performance in terms of RMSE values we can conclude that the more complex parametric model versions, especially the PHetSM two-segment model, lead to overfitting.

Taking a closer look at the predictive performance of the various models at the brand-level which is summarized in Table 3 we find that the incorporation of heterogeneity into a parametric model can lead to small improvements in RMSE values for some brands. In particular, PHBSM reveals a slightly better predictive validity than PHomSM for 6 out of 8 brands. However, the predictive performance of PHBSM is considerably worse for the national brands Citrus Hill and Florida Gold. PHetSM fails to provide better results than PHomSM for all brands supporting the overfitting argument at the brand-level, too. In contrast, allowing for functional flexibility yields substantial improvements for nearly all brands (except Citrus Hill[16]). Importantly, combining heterogeneity and functional flexibility enables a still much better predictive accuracy for all brands except for Tropicana. Lang et al. (2015) have recently illustrated why addressing heterogeneity alone in a store sales model may not be sufficient to improve the predictive performance substantially (if at all), and that one may need nonlinearity (functional flexibility) in the model before one can find evidence for heterogeneity. The reason for this phenomenon is that parametric models may be simply not able to capture the strong nonlinearity in price response contained in sales data. As a consequence, this reason is also responsible for the observation that the more complex heterogeneity specifications considered here result in overfitting, as far as parametric models are concerned (due to the lack of functional flexibility, the models do not get better by accommodating more heterogeneity or different representations of heterogeneity). In contrast, once nonlinearity has been considered (FHomSM), accounting for heterogeneity can provide an additional improvement in predictive accuracy (FHBSM).

Altogether, the findings demonstrate that the consideration of heterogeneity in a parametric SUR model leads to only small improvements in fit and predictive validity whereas the accommodation of functional flexibility solely as well as accounting for both features reveals a superior model fit and forecasting accuracy. In the subsequent analysis, we will show that the superiority of the flexible models can be explained by the fact that complex nonlinearities in price response exist which cannot be captured adequately by the parametric models.

Price elasticities Table 4 shows posterior estimates of segment-specific price coefficients of the parametric SUR models contained in β_k^G (compare equation (5)) which correspond to mean price elasticities since the underlying model (1) is a multiplicative model. Mean price elasticities resulting from PHomSM and PHBSM do not differ

[16] For Citrus Hill, PHomSM reveals a sum of absolute errors of 46912, FHomSM a much smaller one of 44230. The considerably worse RMSE value of FHomSM can be traced back to one single week in the validation sample that obtains a much higher weight in case of squared errors as compared to absolute errors. In order to stay conservative with this flexible model, we omit this week later in our comparison of expected profits across models.

Table 3 Brand-level results of predictive validity for all SUR model versions

Model	RMSE values (and relative improvements over PHomSM) for brand															
	FN		TP		CH		FG		MM		TF		T		D	
PHomSM	25.97		57.69		92.49		157.88		55.40		71.34		117.12		334.87	
PLCSM (2)	25.98	0	58.55	1	94.80	2	156.91	−1	55.15	0	71.28	0	117.04	0	336.45	0
PLCSM (3)	25.57	−2	56.45	−2	90.83	−2	160.80	2	55.24	0	70.72	−1	118.04	1	336.53	0
PLCSM (4)	26.44	2	57.34	−1	91.48	−1	156.32	−1	54.75	−1	71.42	0	117.47	0	339.37	1
PLCSM (5)	26.80	3	57.05	−1	93.49	1	155.89	−1	54.91	−1	71.73	1	115.31	−2	339.45	1
PLCSM (6)	25.52	−2	56.60	−2	94.79	2	157.42	0	54.56	−2	70.98	−1	117.41	0	342.40	2
PHBSM	25.40	−2	56.10	−3	99.29	7	170.77	8	54.51	−2	69.75	−2	116.60	0	329.53	−2
PHetSM (2)	29.36	13	58.77	2	100.38	9	168.36	7	56.52	2	72.24	1	158.06	35	361.11	8
FHomSM	21.87	−16	54.20	−6	102.84	11	108.18	−31	52.97	−4	61.31	−14	111.83	−5	328.06	−2
FHBSM	18.65	−28	47.51	−18	79.36	−14	85.26	−46	50.32	−9	45.91	−36	112.59	−4	322.39	−4

Table 4 Mean price elasticities of the SUR model versions

Brand	PHomSM	PHBSM	PLCSM 2-segment	
			Segment 1	Segment 2
FN	−3.52	−3.52	−3.81	−3.28
TP	−2.85	−2.87	−3.17	−2.61
CH	−3.73	−3.76	−4.21	−3.36
FG	−3.61	−3.58	−3.90	−3.40
MM	−3.29	−3.28	−3.41	−3.23
TF	−2.44	−2.52	−2.72	−2.23
T	−3.50	−3.42	−3.71	−3.34
D	−3.69	−3.70	−4.02	−3.45

Brand	FHomSM				FHBSM			
	Mean	I1	I2	I3	Mean	I1	I2	I3
FN	−3.88	−0.59	−4.63	−3.53	−3.86	−0.95	−4.71	−3.39
TP	−2.74	−2.71	−3.40	−2.38	−2.76	−3.20	−3.38	−2.40
CH	−3.70	−6.73	−3.61	−1.92	−3.66	−7.31	−3.57	−1.43
FG	−3.79	−5.30	−3.30	−2.38	−3.93	−5.45	−3.48	−2.07
MM	−3.25	−2.38	−3.57	−2.70	−3.18	−2.28	−3.53	−2.44
TF	−1.61	−3.62	−1.49	−1.37	−1.40	−2.95	−1.32	−1.12
T	−3.44	−3.70	−3.29	−5.91	−3.31	−3.43	−3.22	−5.11
D	−3.32	−2.68	−3.83	−11.32	−3.32	−2.69	−3.81	−11.69

I1: ≤1.50$; I2: (1.50$; 2.50$]; I3: >2.50$

significantly which is supported by their overlapping 95 % credible intervals (not displayed). Furthermore, these elasticities lie in-between the mean segment-specific elasticities obtained from the two-segment PLCSM. Indicated by 95 % credible intervals, they significantly differ from segment-specific elasticities in case of Tropicana Pure, Citrus Hill, Tree Fresh (this applies only for the elasticities resulting from PHomSM) and Dominick's. For seven out of eight brands, credible intervals of the two segment-specific coefficients suggested by PLCSM do not overlap indicating that differences in price response between the two segments really exist (the only exception is the national brand Minute Maid). In contrast, we do not find significant differences in store-level price elasticities resulting from PHBSM since credible intervals pairwise overlap across all stores. Moreover, store-level coefficients do not significantly differ from the mean price effect across stores since credible intervals overlap, too.

While the price elasticities of the parametric SUR models are constant over the whole price range, FHomSM and FHBSM reveal different price elasticities for different price levels. Specifically, we divided the observed price range into a low (I1), a medium (I2) and a high price interval (I3) and computed mean price elasticities for each interval as well as for the entire price range. The results are displayed in Table 4, as well. Interestingly, the overall mean price elasticities suggested by FHomSM and

FHBSM are quite similar to those suggested by PHomSM. Only the price elasticities of Tree Fresh resulting from the flexible models turn out to be clearly lower (in absolute terms) than that resulting from PHomSM. In contrast, if we take a look at the mean price elasticities within the three price intervals we find considerable differences compared to the overall mean especially in the low and high price intervals. For example, the premium brand Florida Natural even reveals an inelastic demand in the low price interval while demand is elastic in the medium and high price intervals. Price elasticities are further particularly high (in absolute terms larger than 5) in the low price interval for Citrus Hill and Florida Gold and in the high price interval for Tropicana and Dominick's.[17] While mean price elasticities resulting from the flexible models are quite similar in the medium price interval we find noticable differences in the low and high price intervals. Furthermore, price elasticities obtained from FHBSM significantly vary across stores indicated by the number of coefficients that significantly differ from mean price elasticities (at most 40 %). Altogether, we find that differences in mean price elasticities turned out much larger across different price ranges (FHomSM, FHBSM) than across stores (PHBSM, PLCSM, FHBSM). Hence, this kind of variation of price elasticities seems to play a more important role when modeling price response than differences in price elasticities across stores or groups of stores. However, variation of price elasticities across stores becomes more relevant if functional flexibility is already accounted for. These findings are further in line with the results on predictive performance which could be improved considerably by accommodating functional flexibility only as well as by accommodating functional flexibility and heterogeneity jointly while only slightly or not at all by accounting for heterogeneity only (compare Tables 2 and 3).

With respect to the remaining model parameters, estimated holiday effects turned out to be negative for the national brands Minute Maid and Tropicana as well as for the store brand. Apparently, these brands are less preferred in weeks with a holiday. This effect is further not significant at 5 % for the premium brand Florida Natural (FHomSM and FHBSM additionally reveal a nonsignificant holiday effect for Florida Gold). For the remaining brands, the model versions reveal significant positive holiday effects. Estimated display effects are positive as expected for all national brands, not significant for the premium brand Florida Natural as well as for the store brand Dominick's, and slightly negative for the premium brand Tropicana Pure. 9- and 99-ending price effects turned out to be positive and (mostly) significant for all premium and national brands. FHomSM even suggests positive and significant price ending effects for all brands while the parametric model versions yield a negative 99-ending effect for Dominick's and a nonsignificant 99-ending effect for Tree Fresh. FHBSM reveals nonsignificant 99-ending effects for Florida Gold and Dominick's. These findings indicate the importance of involving price ending effects in price response models and in particular for related price setting decisions.

Estimated price effects Figure 1 illustrates estimated own-item price effects for the brands Florida Gold, Tree Fresh and Tropicana Pure representing 3 of the 4 brands

[17] Please note that the high price interval consists of only one observed price in case of the store brand Dominick's. Hence, the respective price elasticity (in absolute terms) is probably overestimated for high prices of Dominick's.

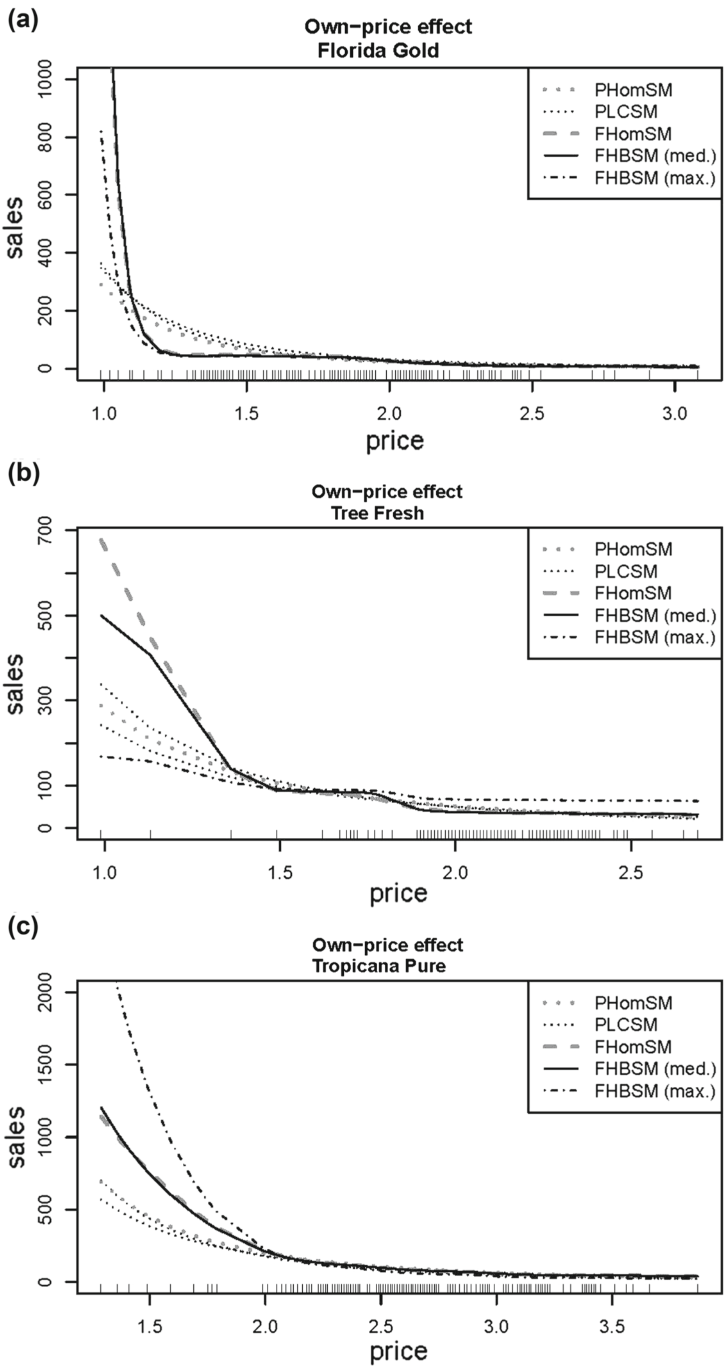

Fig. 1 Estimated own-item price effects resulting from PHomSM (*dotted line in gray*), PLCSM (*dotted line in black*), FHomSM (*long dashed line in gray*) and FHBSM (median scaling: *solid line*, maximum scaling: *dashed-dotted line in black*)

which benefitted most from accommodating heterogeneity and functional flexibility (compare Table 3). Estimated price response curves of PHomSM are represented by dotted lines in gray, those of PLCSM by short dotted lines in black (one curve for each segment), and those of FHomSM by long dashed lines in gray. In order to illustrate the amount of store heterogeneity for the retail chain, we further show estimated price effects of FHBSM for that store which exhibits the median of estimated scaling factors (solid line in black) as well as that store which exhibits the highest absolute value across estimated scaling factors (dashed-dotted line in black). Thus, the former illustrates an "average" store while the latter deviates most in price response from the "average" store. For a clear presentation, we abstain from plotting price response curves obtained from PHBSM since mean price elasticities are similar to those of PHomSM (compare Table 4) and store-specific price elasticities do not differ significantly across stores (see above). To investigate the functional forms of the estimated own-item price effects we show marginal price effects. Thus, we do not display odd price effects which would appear as sales spikes at prices ending in 9 or 99 cents. The tick marks along the base of the plots indicate the actual price levels set by DFF.

The parametric SUR models are exponentially decreasing price response functions by definition. Differences between the own-item price effects resulting from PHomSM and PLCSM are (if at all) rather small and only apparent in the low price range. The comparison of own-item price effects obtained from the two flexible SUR models reveals that they exhibit quite similar functional forms for each brand although FHomSM and FHBSM were estimated independently from each other. In particular, the price effect of FHomSM nearly coincides with that of FHBSM for the "average" store. The flexible SUR models yield L-shaped functions for Florida Gold (panel (a)) and a mixture of L- and reverse s-shaped functions for Tree Fresh (panel (b)). That way, threshold effects become visible indicating the potential for increasing sales below those price thresholds. For Tropicana Pure (panel (c)), the flexibly estimated price effects show a convex decreasing shape which is quite similar to the shape of the parametric price effects. In the low price range FHomSM as well as FHBSM (for the "average" store) suggest considerably larger sales than the parametric SUR models for all brands while differences are rather marginal for high prices. However, the price response curves resulting from FHBSM are heterogeneous across stores which becomes particularly evident from the price effects for the store corresponding to the maximum scaling factor (FHBSM (max.)) that in part considerably deviate from the price effects obtained for the "average" store (FHBSM (med.)), respectively.

Figure 2 exemplifies selected cross-item price effects. For Florida Gold, estimated price response functions with respect to the observed prices of competing brands in the national quality tier are shown (panel (a)). While the parametric SUR models provide almost linearly increasing cross-item price effects the flexible SUR models reveal a mixture of an s- and reverse L-shaped function. Especially in the upper price range sales of Florida Gold are considerably affected if one or more national brands reduce their prices. Moreover, threshold effects indicate that a certain amount of price reduction is necessary to further decrease sales of Florida Gold. The cross-item price effects of Tree Fresh with respect to observed prices of the other brands in the national quality tier exhibit s-shaped functions for the flexible SUR models (panel (b)) indicating a threshold effect as well. Accordingly, in the upper price interval above 2.70$ price cuts

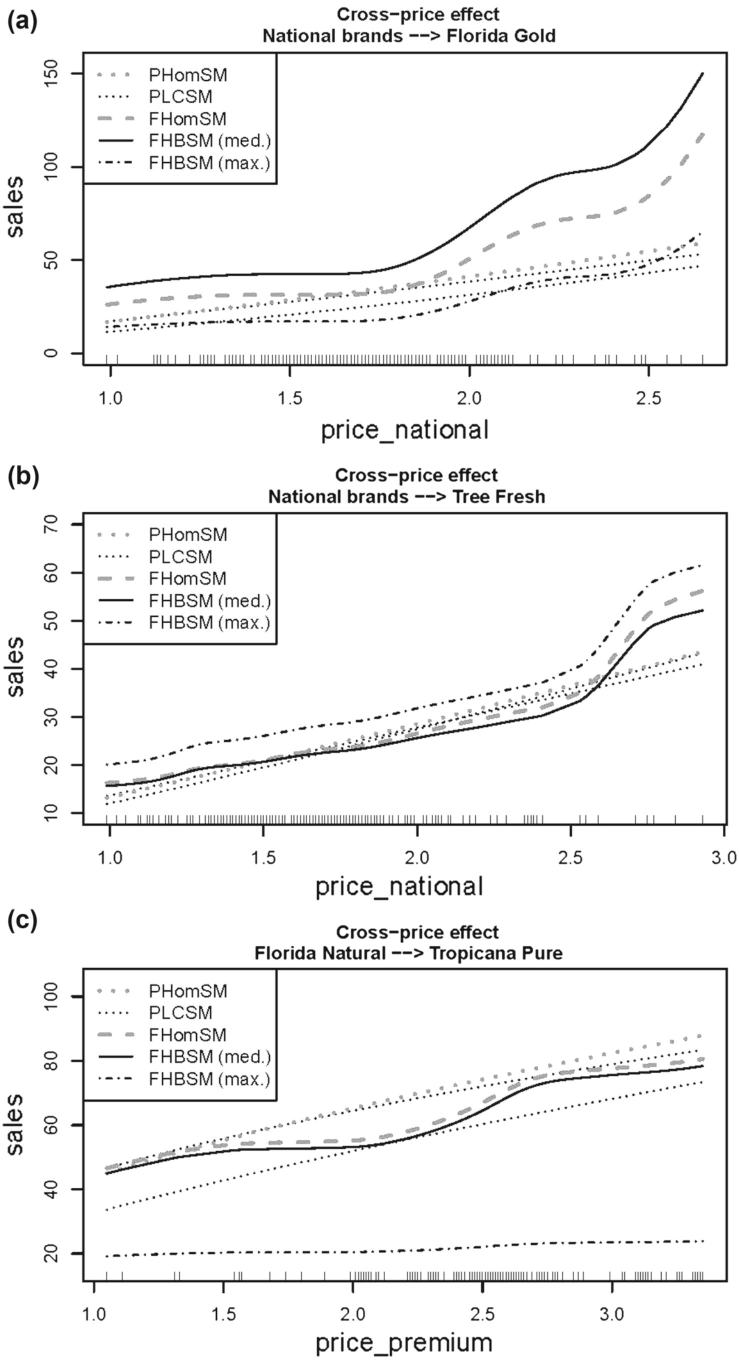

Fig. 2 Estimated cross-item price effects resulting from PHomSM (*dotted line in gray*), PLCSM (*dotted line in black*), FHomSM (*long dashed line in gray*) and FHBSM (median scaling: *solid line*, maximum scaling: *dashed-dotted line in black*)

Table 5 Optimization results based on the profit function of FHBSM

	Expected profits	Loss	Expected sales
Actual profit	2390363		4694096
$E\left[\Pi\left(p^*_{FHBSM}\right)\right]$	3564031		6304084
$E\left[\Pi\left(p^*_{PHomSM}\right)\right]$	3370934	−5.42	5892175
$E\left[\Pi\left(p^*_{PHBSM}\right)\right]$	3271465	−8.21	5729485
$E\left[\Pi\left(p^*_{PLCSM}\right)\right]$	3169675	−11.06	5560485
$E\left[\Pi\left(p^*_{FHomSM}\right)\right]$	3543383	−0.58	6271742

of the competing national brands have almost no impact on the sales of Tree Fresh. In contrast, if the price of a competing national brand falls below the threshold sales of Tree Fresh considerably decrease, especially within the price range from 2.50$ to 2.70$. The cross-item price effect of Florida Natural on Tropicana Pure (panel (c)) resembles that of the national brands on Tree Fresh (panel (b)). Again, we observe in part considerable differences in estimated cross-item price effects obtained from FHBSM at the median versus maximum scaling factor levels (see in particular panel (a)).

From the own- and cross-item price response curves considered here it becomes particularly apparent that complex nonlinearities in price response are existent which cannot be reproduced appropriately by parametric models on the one hand and that differences across stores are present on the other hand, which explains the superior predictive validity of the flexible (and heterogeneous) semiparametric SUR models once more (compare Tables 2 and 3).

3.4 Optimization results

Optimal prices were determined for all model versions and are denoted by p^*_{PHomSM}, p^*_{PHBSM}, p^*_{PLCSM}, p^*_{FHomSM} and p^*_{FHBSM}. To obtain comparable results, we then plugged those optimal prices into the profit function of the model with the highest predictive validity,[18], i.e., into the profit function of FHBSM (compare Table 2). As an example, consider that we plug p^*_{PHomSM} (computed under PHomSM) into Eq. (13) where the predicted unit sales are calculated using the parameter samples of FHBSM. Hence, we compute the total expected profit $E[\Pi(p^*_{PHomSM})]$ for the flexible FHBSM but use p^*_{PHomSM} instead of p^*_{FHBSM}. That way, we are able to measure the loss management incurs by not using the model with the best predictive performance (Hruschka 2006b). Table 5 reports the corresponding results for all model versions.

As a first result, we can observe that actual profits (computed by using actual sales) could be increased considerably by each model version. Optimizing prices led to expected profits which are substantially larger than actual profits obtained by DFF. In comparison with expected profits resulting from FHBSM, the parametric models

[18] According to van Heerde et al. (2002), one should choose the model with the best predictive performance.

reveal substantial losses ranging from about 6 to 11 %. The smallest loss is associated with PHomSM, while the more complex parametric heterogeneous models PLCSM and PHBSM suffer from still higher losses.[19] In contrast, expected profits resulting from FHomSM and FHBSM, respectively, do not differ to a noticable extent. Thus, the flexible models promise huge potential for increasing profits compared to each parametric model.

We further computed expected unit sales based on FHBSM (also shown in Table 5), which were obtained by plugging the optimal prices of each model (one at a time) into the sales equation of FHBSM. According to Natter et al. (2007), it is possible that increased expected profits are accompanied by decreased unit sales. Our results do not confirm the finding of Natter et al. (2007) since expected unit sales could be increased considerably by all model versions compared to the observed unit sales of DFF. Moreover, we find for each model version that the larger expected profits the larger expected unit sales are.

Brand-level optimization results averaged across weeks are summarized in Table 6 (with best values of expected profits indicated in **bold**). Comparison of observed prices and optimized prices reveals that prices on average should be increased for the national brands Florida Gold, Minute Maid, Tree Fresh and Tropicana (as well as for the premium brand Florida Natural according to the parametric SUR models) and decreased for the premium brand Florida Natural (according to the flexible SUR models), the national brand Citrus Hill as well as for the store brand Dominick's. Remember that the average price level of the product category is maintained in a given week (compare Eq. (19)). Thus, price increases of some brands had to be offset by price decreases of other brands in the product category. Average margins across weeks and stores could not be enhanced for Citrus Hill (and Florida Natural according to the flexible SUR models). Average weekly market shares could be increased considerably for DFFs own store brand, while they would clearly decrease for the premium brand Tropicana Pure and the national brand Minute Maid according to all model versions, respectively. Due to the higher margin, expected profits could however be improved for Minute Maid. Besides Tropicana, expected profits would also increase for all other national brands (except for Florida Gold according to PLCSM), and for the premium brand Florida Natural according to the flexible models. Furthermore, we can observe that optimization results are mixed across models at the brand-level. In particular, while FHBSM provides the largest expected profit at the category level this (on average) holds for only four brands at the brand-level. Nevertheless, expected profits obtained from FHBSM are rather similar to those obtained from FHomSM which provides the largest expected profits for two further brands. For Tropicana Pure and Tropicana, however, PLCSM yields the largest expected profits.[20] Altogether, the

[19] For the sake of completeness, we further determined *store-specific* optimal prices based on the latent class model PLCSM and the general heterogeneity model PHetSM (note that this model revealed one empty segment) and plugged them into the profit function of FHBSM. Relative to expected profits of FHBSM, optimal prices obtained from these two models result in losses of about 8 % in case of PLCSM and 10 % in case of PHetSM.

[20] Note that based on the profit function of FHBSM optimal prices of this model have to provide the largest expected profit for the whole product category in each week. Nevertheless, optimal prices of the other models can lead to better results at the brand-level for some brands.

Table 6 Brand-level optimization results

	Brand							
	FN	TP	CH	FG	MM	TF	T	D
Observed[a]	2.85	2.94	2.32	2.18	2.22	2.16	2.19	1.76
	0.32	0.28	0.24	0.32	0.24	0.29	0.29	0.30
	4	12	7	6	10	8	18	35
	1704	4051	1478	2259	1771	2127	4267	9506
PHomSM[a]	2.87	2.91	2.26	2.23	2.27	2.31	2.21	1.70
	0.33	0.29	0.23	0.35	0.28	0.34	0.31	0.30
	3	8	8	5	7	5	17	46
	1616	3966	2374	4053	2664	2489	6896	14250
PHBSM[a]	2.89	2.93	2.25	2.22	2.28	2.29	2.22	1.70
	0.33	0.29	0.22	0.34	0.28	0.33	0.31	0.30
	3	8	8	5	7	6	17	46
	1645	3979	2376	3087	2681	2464	6841	14105
PLCSM[a]	2.92	2.95	2.23	2.21	2.27	2.29	2.20	1.69
	0.34	0.30	0.22	0.34	0.27	0.33	0.30	0.30
	3	8	9	5	7	5	18	46
	1618	**4013**	2433	2024	2681	2440	**6898**	13911
FHomSM[a]	2.85	2.94	2.26	2.21	2.25	2.28	2.24	1.74
	0.32	0.29	0.23	0.34	0.26	0.33	0.32	0.31
	4	8	8	5	7	6	16	46
	1864	3803	2428	**4761**	2706	2689	6837	**15176**
FHBSM[a]	2.80	2.93	2.28	2.21	2.26	2.29	2.25	1.76
	0.31	0.28	0.23	0.34	0.27	0.33	0.32	0.32
	4	8	7	6	7	7	16	45
	2198	3801	**2442**	4753	**2765**	**2877**	6778	14885

[a] *1. Row* Average prices across weeks and stores, *2. Row* Average margins across weeks and stores, *3. Row* Average market shares across weeks for all stores of the chain, *4. Row* Average profits across weeks for all stores of the chain

brand-level results (on average) as well suggest the relative superiority of both flexible SUR model versions with respect to expected profits.

Pricing implications resulting from the different model versions are exemplified in Fig. 3 for the national brand Tree Fresh which benefitted considerably from accommodating heterogeneity and functional flexibility in terms of both predictive validity (compare Table 3) and expected profits (compare Table 6). For PHBSM and FHBSM, store-specific optimal prices are shown for each week while PLCSM yields segment-specific optimal prices. We further display (store-specific) actual prices of DFF in each week as indicated by the gray-shaded triangles. Optimal price levels resulting from PHomSM coincide with those for (at least) one segment obtained from PLCSM in about 50 % of the weeks. PHBSM provides up to 25 different price levels across stores. But taking a closer look leads to the conclusion that the most frequent price

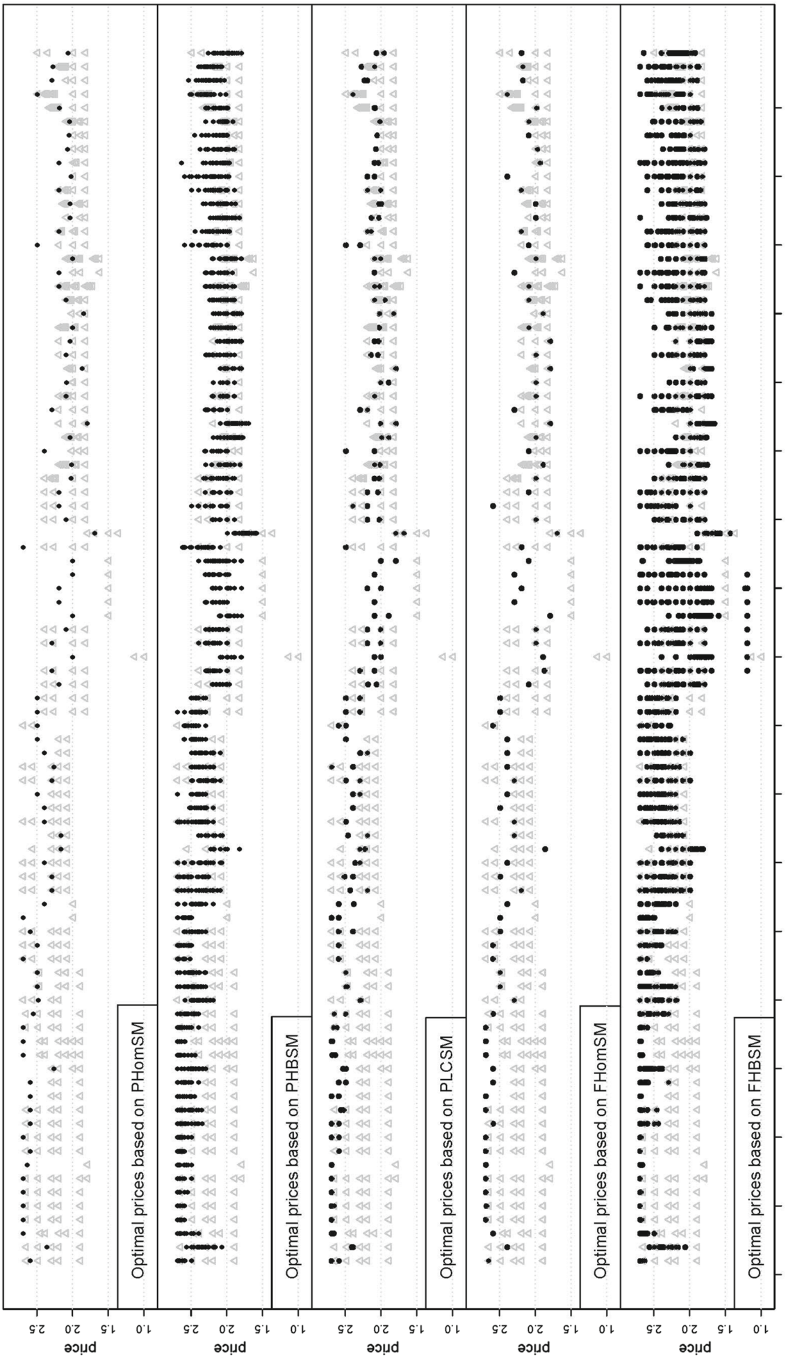

Fig. 3 Optimal prices for Tree Fresh based on PHomSM, PHBSM, PLCSM (with 2 segments), FHomSM and FHBSM

levels obtained from PHBSM coincide with those obtained from PHomSM in 42 % and with those obtained from (at least) one segment of PLCSM in about 57 % of the weeks. In contrast, the flexible homogeneous SUR model reveals the same price levels as PHomSM in only 27 % of the weeks, and the most frequent price levels resulting from FHBSM coincide with those of PHomSM only in about 29 % of the weeks. Interestingly, while DFF actually sets on average three different price levels across stores for Tree Fresh the heterogeneous model versions suggest a more differentiated pricing strategy with on average about 14 different price levels across stores.

Differences in optimal prices between the various model versions become particularly evident if we simultaneously consider optimal prices of all brands of the product category. Then, optimal category pricing resulting from PHomSM is different from that under FHomSM in all weeks and corresponds to that for at least one store under FHBSM in only 10 weeks. Optimal category pricing under PHomSM coincides with that for one segment under PLCSM in only 9 weeks, and PHBSM and FHBSM yield on average (across weeks) a store-specific category pricing different from PHomSM in about 71 and 65 % of the stores, respectively. In contrast, a similar optimal category pricing is obtained from FHomSM and for at least one store from FHBSM in 56 % of the weeks thus explaining the smaller difference in expected profits between the flexible model versions.

4 Summary and discussion

The present study deals with the comparison of different approaches for modeling price response from store sales. Several heterogeneous parametric response models are contrasted with a homogeneous and a heterogeneous flexible approach, and brand sales and profits are modeled simultaneously for all brands in a product category within a SUR regression framework. In particular, we compare a general heterogeneity SUR model (PHetSM) and its nested heterogeneous versions, i.e., the hierarchical Bayesian SUR model (PHBSM) and the latent class SUR model (PLCSM), as well as a hierarchical Bayesian semiparametric SUR model (FHBSM) and its nested homogeneous version (FHomSM) to a simple homogeneous parametric SUR model (PHomSM) used as benchmark. While the parametric heterogeneous models allow for differences in marketing mix effects across stores or groups of stores the FHomSM accounts for functional flexibility in price response, whereas the FHBSM accommodates both heterogeneity and functional flexibility and the PHomSM none of these features.

The main results of our empirical application with store-level scanner data can be summarized as follows: First, the less complex parametric SUR models clearly outperform the general heterogeneity SUR model with respect to model fit. This can probably be traced back to the fact that not even a PHetSM with only two segments is supported by our data (the two-segment PHetSM yielded one empty segment with no stores assigned to it and a corresponding segment size of almost zero). While PHBSM turned out to perform best, differences in model fit measured in terms of the DIC and the model likelihood between the nested versions of PHetSM (i.e., PHBSM, PLCSM, PHomSM) are rather marginal. The same applies to differences in predictive validity between those parametric models which turned out to be best for PHomSM and was highest for PLCSM with two segments as far as only the heterogenous model versions are considered. In contrast, the flexible SUR models could increase the predictive performance by about 6 and 10 %, respectively, compared to PHomSM. Thus, referring to the first question posed in the introductory section: “Does a category-level sales response model that allows for correlations between sales across brands benefit from accommodating either store heterogeneity or functional flexibility or both features in terms of fit and predictive validity? And if, which of those features pays off more?”, we can conclude that the benefits from incorporating heterogeneity into a parametric sales response model are rather limited whereas accommodating functional flexibility as well as allowing for heterogeneity and functional flexibility jointly really pays off with respect to predictive performance. This finding is also reflected by the results on estimated price elasticities. While differences in price elasticities between the parametric SUR model versions are rather small, the flexible SUR models lead to materially different price elasticities which vary across different price ranges and, in case of FHBSM, additionally across stores. And, we find that differences in price elasticities turned out much larger across different price ranges than across stores. Hence, this kind of heterogeneity of price elasticities seems to play a more important role when modeling price response than differences in price elasticities across stores or groups of stores. Estimated price effects of the semiparametric SUR models uncover complex nonlinearities and differences across stores in case of FHBSM which further helps to explain the superiority of the flexible SUR models with respect to predictive performance.

The computation of expected category profits revealed that the loss management incurs by using a parametric instead of the best performing model, i.e., FHBSM, ranges from about 6 % when optimal prices are set according to the most simple parametric model PHomSM up to about 8 and 11 % when optimal prices are set according to the heterogeneous models PHBSM and PLCSM, respectively. These results resemble our findings on predictive performance where the PHomSM also constituted the best parametric model. Interestingly, despite different underlying pricing strategies, expected category profits obtained from using FHomSM and FHBSM do not differ to a noticable extent. Thus, the findings on expected profits indicate the importance of a flexible modeling approach, too. Actually, the various models reveal materially different pricing implications where differences are not only due to the incorporation of heterogeneity into a homogeneous sales response model. Accounting for functional flexibility only also leads to a different uniform pricing strategy compared to PHomSM, and allowing for heterogeneity and functional flexibility yields a different micromarketing pricing strategy compared to PHBSM which is also reflected by the differences in expected profits. We therefore can respond to the second question raised in the introduction: “Do category-level store sales models with either different representations of heterogeneity and/or flexibly modeled price effects provide different implications with respect to expected profits and optimal prices?” with a clear ‘yes’. Certainly, expected profit calculations and related pricing implications are always dependent on the modeling approach chosen and may be different when extending our category management perspective to a cross-category management or shopper marketing point of view. On the other hand, we accounted for model uncertainty in our optimization step by using 200 Gibbs samples of coefficients instead of point estimates.

Referring to the third question: “Should retailers adopt a store-level, segment-level or chain-level pricing policy?”, a uniform pricing strategy which is much less complex to implement than micromarketing for a retailer can be recommended for our data (as FHBSM and FHomSM perform nearly equally well concerning expected profits). However, this result implies that the observed pricing of DFF is suboptimal. We, therefore, investigated the pricing behavior of this retailer and found that he does not use a simple cost plus approach for setting prices (please compare Sect. 3.1). Following Chintagunta (2002), we further computed markups ($price - cost$) and margins ($[price - cost]/price$) as well as their correlations between brands, which turned out rather small (at most 0.35). Thus, it could be confirmed that DFF does not apply simple pricing rules like equal markups or equal margins (Blattberg and Neslin 1990). Since price zones of DFF are almost exclusively defined by the extent of local competition, it seems not unlikely at first glance that the pricing strategy of DFF is oriented toward prices of competitive retailers. We are not able to directly confirm this assumption due to the lack of (price) data of competing retailers. However, based on proxy variables relating to competitive characteristics of each stores trading area (distance in miles to nearest Omni warehouse, ratio of DFF store sales to nearest Omni warehouse, average distance in miles to nearest five Cubfoods supermarkets, ratio of DFF store sales to average of nearest five Cubfoods supermarkets), Hoch et al. (1995) already demonstrated that a relationship between price sensitivities and price zones does not exist. Therefore, even if DFF actually set prices based on competitive retailers, this pricing strategy might not be the optimal one from a profit-maximizing point of view. Alternatively, the retailer might use a different optimization approach for setting prices, such as maximizing store traffic or maximizing store profit taking into account cross-category effects, and/or the underlying sales response model might be a less complex model (e.g., a linear function). Nevertheless, the results of our study show that there seems to be huge potential for improving the retailer’s profits in the considered category, in particular when using a flexible modeling approach.

Of course, there are several limitations of our empirical study. First of all, results have to be replicated for numerous product categories to generalize our findings. Furthermore, our results are based on the underlying specification of our models. For instance, we did not account for price endogeneity constituting a recent and controversially discussed issue in the relevant marketing literature. Approaches to account for endogeneity in prices are the well-known instrument variable approach (e.g., Chintagunta 2000; Villas-Boas and Winer 1999), so-called control function approaches (e.g., Luan and Sudhir 2010; Petrin and Train 2010), or the explicit modeling of pricing behavior of firms/retailers in terms of a so-called supply-side model (e.g., Manchanda et al. 2004; Otter et al. 2011). Chintagunta et al. (2003) use wholesale prices as instruments for retail prices and demonstrate that controlling for endogeneity leads to a more elastic demand in the refrigerated orange juice product category of DFF. If this finding holds predicted unit sales as well as expected profits of our study would turn out smaller. However, according to Rossi (2014) manufacturers account for advertising and promotional events when setting wholesale prices leading to an invalid instrument that “can cause the estimates to differ even when there is no endogeneity bias” (Rossi

2014, p. 671).[21] Another important limitation of our study is that it does not consider cross-category effects (related studies based on household-level data are, e.g., Ainslie and Rossi 1998; Chib et al. 2002; Kim et al. 1999; Manchanda et al. 1999; Russell and Kamakura 1997; Russell and Petersen 2000 and studies based on store-level data are, e.g., Bezawada et al. 2009; Shankar and Kannan 2014; Wedel and Zhang 2004), which presumably would lead to different profit implications (as already mentioned above). Furthermore, we account for interdependencies between sales of brands only in terms of correlations of the error terms within a SUR framework. Accommodating the fact that sales of one brand are significantly affected by sales of other brands (e.g. Bezawada et al. 2009; Elrod et al. 2002; Shankar and Kannan 2014) would lead to a more general simultaneous equation model including the unit sales of competing brands as further covariates. However, the optimization problem of our study would then become much more complex. For this reason, we leave this issue to future research.

The optimization problem considered does not take into account every aspect which influences a retailer's pricing decision problem. For example, due to the lack of data we did not investigate the effects of competitive retailers. However, we preserved the retailer's current price image by imposing that the average price level in a given week before and after optimizing price levels must not change noticeably just in order to avoid competitive reactions (Montgomery 1997). Nevertheless, Shankar and Bolton (2004) showed that competitive factors are the most important determinants of a retailer's pricing decision. Thus, if possible, they should be taken into account. Moreover, retail prices of a brand are also influenced by the pricing history of the brand (e.g., Nijs et al. 2007), which could be accounted for, e.g., by the incorporation of lagged price variables into the sales response models, as well as by temporary deals of manufacturers (e.g., Levy et al. 2004) which could be passed on to consumers.

Another limitation of the present work is that optimal prices are based on a given promotional strategy, relating to the use of displays in our data. Of course, different scenarios of display usage could be considered in our optimization model but according to Ailawadi et al. (2009), pricing and promotional activities should be simultaneously investigated "in order to avoid sub-optimal decisions on both fronts" (Kopalle et al. 2009, p. 62). On the other hand, setting prices conditional on a fixed promotional calendar is common retailer practice (e.g., Kim et al. 1999; Chintagunta et al. 2003. The 'complete' optimization problem (if it exists) would further comprise aspects of space management (Hoch and Lodish 1998) like the planning of order quantities and shelf-space allocation (examples considering such aspects can be found in Hall et al. 2010; Murray et al. 2010; Tellis and Zufryden 1995). The research focus of the present study lies on the comparison of the statistical and managerial performance of the various models considered and of different pricing scenarios arising from these models "rather than on prescribing an approach for determining optimal prices" (Khan

[21] We repeated our optimization exercise for PHomSM over four weeks by simulating a more elastic demand. Similar to Chintagunta et al. (2003), we found that expected profits are lower when accounting for price endogeneity. However, this is reasonable since expected unit sales are in consequence lower, too. Optimal prices turn out somewhat different while the average price level of the product category does not change according to expression (19).

and Jain 2005, p. 523). As such, the computed expected profits must not be considered as numbers "set in stone", but should rather suggest the relative superiority of the flexible semiparametric models compared to parametric models. Further note that expected profits are larger than actual profits which leads to increased unit sales, as well. From the retailer's perspective, this fact can cause problems with respect to capacity, possible additional storage costs, shelf-space optimization, etc. Thus profit increases have to be critically reflected against these aspects.

Finally, optimal prices obtained here are based on the assumption that the retailer's objective is to maximize the chain-level profit of a product category. Other objectives could be the maximization of product category sales or store traffic (e.g., Basuroy et al. 2001; Klapper 2000; Reibstein and Gatignon 1984). For example, there are so-called 'loss leader' categories where prices are close to wholesale costs in order to generate store traffic rather than to maximize profits of these product categories (e.g., Kumar and Leone 1988; Wedel et al. 2004). The retailer could further optimize the overall profit of each store (e.g., Song and Chintagunta 2006) requiring the simultaneous determination of optimal prices (and other marketing activities) across different product categories. In the context of shopper marketing (Shankar et al. 2011) retailers are also maximizing customer lifetime value. This objective, however, would require the availability of individual-level data (panel data, loyalty card data). To address different optimization objectives would of course require to extend our modeling approach beyond category management taking into account cross-category effects, activities or reactions of competing retailers, spatially correlated effects of neighboring stores, or even effects of aisle management (Bezawada et al. 2009; Inman et al. 2009), for example.

For practitioners the question arises which approach of modeling price response should be used for prediction and optimization in which situation. In the refrigerated orange juice category of DFF considered in our empirical application we can observe some extremely deep price discounts favoring a flexible model version which is able to reproduce complex nonlinearities in price response based on such an extreme price distribution. Moreover, the extent of (store) heterogeneity in our data seems to be rather moderate. This may be explained by the fact that the data set comes from a supermarket chain located in one metropolitan area which is geographically limited. Thus, if in general the stores of a retailer are spread over a large area (e.g., the whole country) and price discounts are less extreme the general heterogeneity SUR model could perform similarly or even better than the other model versions, as well. In that case, one could further analyze mixed pricing strategies, as for example a uniform pricing policy in one segment and a micromarketing strategy in another segment. A simulation study with synthetic data could further help to get more insights how the underlying heterogeneity distribution of store-level data affects model estimates and managerial implications on pricing strategies. Moreover, an interesting and challenging aspect for future research would be the development of a flexible general heterogeneity SUR model.

Acknowledgements The data for our empirical study was provided by the James M. Kilts Center, GSB, University of Chicago. We further thank three anonymous referees for their critical comments and valuable recommendations on our way toward a publishable manuscript.

Appendix 1: General heterogeneity SUR model

The marginal model

We use the partly marginalized Gibbs sampler suggested by Frühwirth-Schnatter et al. (2004) where the random effects are integrated out when sampling S and $\gamma = (\alpha, \beta_1^G, \ldots, \beta_K^G)$. If segment membership is known the random effects β_i can be rewritten as

$$\beta_i = \beta_k^G + b_i, \quad b_i \sim N(0, Q_k^G). \tag{22}$$

Under the assumption that b_i and ϵ_i are independent (Frühwirth-Schnatter 2006) the marginal model is obtained by substituting (22) into model (4):

$$y_i = X_i\alpha + W_i\beta_k^G + \tilde{\epsilon}_i, \quad \tilde{\epsilon}_i \sim N(0, W_i Q_k^G (W_i)' + \Sigma \otimes I_{T_i}). \tag{23}$$

Introducing indicator variables

$$H_i^{(k)} = \begin{cases} 1, & \text{if } S_i = k, \\ 0, & \text{otherwise}, \end{cases} \qquad k = 1, \ldots, K, \tag{24}$$

leads to the following representation of the marginal model:

$$y_i = Z_i\gamma + \epsilon_i^*, \quad \epsilon_i^* \sim N(0, V_i) \tag{25}$$

with the design matrix

$$Z_i = \begin{pmatrix} X_i & W_i H_i^{(1)} & \cdots & W_i H_i^{(K)} \end{pmatrix} \tag{26}$$

and the covariance matrix

$$V_i = W_i H_i^{(1)} Q_1^G (W_i)' + \cdots + W_i H_i^{(K)} Q_K^G (W_i)' + \Sigma \otimes I_{T_i}. \tag{27}$$

Prior specifications

The parameter vectors α and β_k^G $(k = 1, \ldots, K)$ are assumed to be a priori independent leading to a joint prior $N(\mu_{0\gamma}, \Sigma_{0\gamma})$ for $\gamma = (\alpha, \beta_1^G, \ldots, \beta_K^G)$ with $\mu_{0\gamma} = (\mu_{0\alpha}, \mu_{0\beta_1^G}, \ldots, \mu_{0\beta_K^G})'$. $\Sigma_{0\gamma}$ is a block diagonal matrix with diagonal elements $(\Sigma_{0\alpha}, \Sigma_{0\beta_1^G}, \ldots, \Sigma_{0\beta_K^G})$. We determined the prior means $\mu_{0\alpha}$ for the fixed effects α and $\mu_{0\beta_k^G}$ for the group-specific effects β_k^G $(k = 1, \ldots, K)$ by estimating a homogeneous SUR model via Gibbs sampling implemented in the R function *rsurGibbs()* contained in the R package *bayesm* (see Rossi et al. 2005 for details).

About the remaining parameters we stay nearly noninformative choosing $\Sigma_{0\alpha} = \Sigma_{0\beta_1^G} = \cdots = \Sigma_{0\beta_K^G} = 50I$, $a_{0\Sigma} = 1$ and $B_{0\Sigma} = 0.005I$ and $e_{0k} = 1$ for

$k = 1, \ldots, K$. The covariance matrices Q_k^G $(k = 1, \ldots, K)$ are defined as diagonal matrices with diagonal elements $(\tau_{k1}^2, \ldots, \tau_{kr}^2)$ (Gelfand et al. 1995, p. 483) where $r = \sum_{m=1}^{M} r_m$ corresponds to the dimension of β_i. Thus, inverse gamma priors $IG(a_{0h}^k, b_{0h}^k)$ are placed on the variance parameters τ_{kh}^2 $(h = 1, \ldots, r)$ with hyperparameters $a_{0h}^k = b_{0h}^k = 0.001$ $(h = 1, \ldots, r, k = 1, \ldots, K)$. For our data used in the empirical application, we found that applying diagonal covariance matrices Q_k^G $(k = 1, \ldots, K)$ considerably improved the model fit of the hierarchical Bayesian SUR model and the general heterogeneity SUR model (measured by the model likelihood). In the context of a single (HB) regression model, Andrews et al. (2008) have shown that the predictive performance of the model can be improved (and besides model complexity will be reduced) by not estimating covariances between store-specific effects. The respective assumption that, for example, the price sensitivity of a store i is independent of the display activity in another store i' is reasonable in the context of sales data belonging to *one* retail chain since customers usually shop in the store which is closest to their home (cf. Andrews et al. 2008, p. 25). If the covariance matrices Q_k^G $(k = 1, \ldots, K)$ were non-diagonal matrices we could choose $a_{0Q_K^G} = r + 1$ and $B_{0Q_K^G} = rI$ for $k = 1, \ldots, K$ as hyperparameters for the inverse Wishart prior.

Parameter estimates are based on the last 2000 iterations of the MCMC algorithm. Due to high computing times the burn-in period was limited to 1000 iterations for the one-segment models and the unidentified versions of the two-segment models, and to 8000 iterations when the latent class SUR model had to be identified via constrained permutation sampling. Trace plots of sampled coefficients indicated that these numbers of burn-in iterations are sufficient to ensure convergence of the MCMC algorithm.

Bayesian inference

The joint posterior of the latent random effects $\beta^I = (\beta_1, \ldots, \beta_N)$, the latent segment indicator $S = (S_1, \ldots, S_N)$ and all unknown parameters $\phi = (\alpha, \beta_1^G, \ldots, \beta_K^G, Q_1^G, \ldots, Q_K^G, \eta_1, \ldots, \eta_K, \Sigma)$ given the data $y = (y_1, \ldots, y_N)$ is proportional to

$$\pi(\beta^I, S, \phi | y) \propto f(y|\beta^I, S, \phi)\pi(\beta^I | S, \phi)\pi(S|\phi)\pi(\phi). \tag{28}$$

For a fixed number K of segments the following sampling scheme results (compare Frühwirth-Schnatter et al. 2004, 2005, and the derivation of steps 4 (a2) and (b) in Weber (2015)):

1. Sample S_i, $i = 1, \ldots, N$, from

$$P(S_i = k|\cdot) \propto \eta_k \cdot f(y_i|\alpha, \beta_k^G, Q_k^G, \Sigma), \quad k = 1, \ldots, K, \tag{29}$$

   where the likelihood $f(y_i|\alpha, \beta_k^G, Q_k^G, \Sigma)$ is represented by the density of the normal distribution

$$N(X_i\alpha + W_i\beta_k^G, W_i Q_k^G (W_i)' + \Sigma \otimes I_{T_i}).$$

2. Sample η from the Dirichlet distribution $D(e_{0,1} + N_1, \ldots, e_{0,K} + N_K)$ with $N_k = \#\{S_i = k\}$.
3. $\alpha, \beta_1^G, \ldots, \beta_K^G$ and β^I are conditionally independent and can be sampled within two blocks.
   (a) Sample γ from the normal distribution $N(\mu_\gamma, \Sigma_\gamma)$ with

$$\Sigma_\gamma = \left(\sum_{i=1}^{N} (Z_i)' (V_i)^{-1} Z_i + \left(\Sigma_{0\gamma}\right)^{-1} \right)^{-1},$$

$$\mu_\gamma = \Sigma_\gamma \left(\sum_{i=1}^{N} (Z_i)' (V_i)^{-1} y_i + \left(\Sigma_{0\gamma}\right)^{-1} \mu_{0\gamma} \right).$$

   (b) Sample $\beta_i, i = 1, \ldots, N$, from the normal distribution $N\left(\mu_{\beta_i}, \Sigma_{\beta_i}\right)$ with

$$\Sigma_{\beta_i} = \left((W_i)' \left(\Sigma \otimes I_{T_i}\right)^{-1} W_i + \left(Q_{S_i}^G\right)^{-1} \right)^{-1},$$

$$\mu_{\beta_i} = \Sigma_{\beta_i} \left((W_i)' \left(\Sigma \otimes I_{T_i}\right)^{-1} (y_i - X_i \alpha) + \left(Q_{S_i}^G\right)^{-1} \beta_{S_i}^G \right).$$

4. The covariance matrices $Q_1^G, \ldots, Q_K^G$ and Σ are conditionally independent so that sampling can be done within two blocks.
   (a1) If the covariance matrices $Q_k^G, k = 1, \ldots, K$, are non-diagonal matrices they can be sampled from the inverse Wishart distribution $IW(a_{Q_k}, B_{Q_k})$ with

$$a_{Q_k} = a_{0Q_k^G} + N_k/2, \quad B_{Q_k} = B_{0Q_k^G} + \frac{1}{2} \sum_{i=1}^{N} H_i^{(k)} (\beta_i - \beta_k^G)'(\beta_i - \beta_k^G).$$

   (a2) If the covariance matrices $Q_k^G, k = 1, \ldots, K$, are diagonal matrices of the form $Q_k^G = \text{diag}(\tau_{k1}^2, \ldots, \tau_{kr}^2)$ $(k = 1, \ldots, K)$ the variance parameters τ_{kh}^2, $k = 1, \ldots, K$ and $h = 1, \ldots, r$, can be sampled from the inverse gamma distribution $IG(a_{\tau_{kh}}, b_{\tau_{kh}})$ with

$$a_{\tau_{kh}} = a_{0h}^k + N_k/2, \quad b_{\tau_{kh}} = b_{0h}^k + \frac{1}{2} \sum_{i=1}^{N} H_i^{(k)} (\beta_{ih} - \beta_{kh}^G)^2.$$

   (b) Sample Σ from the inverse Wishart distribution $IW(a_\Sigma, B_\Sigma)$ with

$$a_\Sigma = a_{0\Sigma} + \frac{1}{2} \sum_{i=1}^{N} T_i, \quad B_\Sigma = B_{0\Sigma} + \frac{1}{2} \sum_{i=1}^{N} A^{(i)}$$

   with $A^{(i)} = (Y_i - P_i^* B_i)' (Y_i - P_i^* B_i)$, $Y_i = (y_{i1}, \ldots, y_{iM})$, $P_i^* = (P_{i1}, \ldots, P_{iM})$, $P_{im} = (X_{im}, W_{im})$, $B_i = diag(\gamma_{i1}, \ldots, \gamma_{iM})$ and $\gamma_{im} = (\alpha_m, \beta_{im})'$.

The R code of the estimation procedure is available from the authors upon request.

Appendix 2: Hierarchical Bayesian semiparametric SUR model

Inference uses MCMC simulation, drawing from full conditionals of single parameters or blocks of parameters given the rest and the data. Let $y_m = (y_{m11}, \ldots, y_{mNT})'$ and $\eta_m = (\eta_{m11}, \ldots, \eta_{mNT})'$ denote the vector on the m-th response variable and the corresponding vector of predictors. Then the additive predictors in (7) can be written as

$$\eta_m = \sum_{j=1}^{M} A_{mj} X_{mj} \beta_{mj} + \sum_{l=0}^{L} W_{ml} Z \gamma_{ml} + V_m \delta_m, \tag{30}$$

where A_{mj} is a $n \times n$ diagonal matrix with possible entries $1 + \alpha_{m1j}, \ldots, 1 + \alpha_{mNj}$ depending on the store $i = 1, \ldots, N$ a particular observation pertains to (with n being the total number of observations of brand m), X_{mj} corresponds to the design matrix for the jth price effect of brand m with elements given by the B-spline basis functions evaluated at the observed prices P_{jit}, W_{ml} is the (diagonal) design matrix for the lth parametrically and store-specifically modeled variable D^l_{mit}, Z is a 0/1 incidence matrix indicating if a particular observation belongs to store i, V_m is the design vector for the homogeneous parametric effect of E_t, $\beta_{mj} = (\beta_{mj1}, \ldots, \beta_{mjO_{mj}})'$ is the vector of regression parameters for function f_{mj}.

An alternative formulation in terms of the random coefficients is

$$\eta_m = \sum_{j=1}^{M} (f_{mj} + \tilde{X}_{mj} Z \alpha_{mj}) + \sum_{l=0}^{L} W_{ml} Z \gamma_{ml} + V_m \delta_m, \tag{31}$$

where $\tilde{X}_{mj} = \mathrm{diag}(f_{mj}(P_{m11}), \ldots, f_{mj}(P_{mNT}))$ and $\alpha_{mj} = (\alpha_{m1j}, \ldots, \alpha_{mNj})'$ the vector of random coefficients for function f_{mj}.

Let $\beta = (\ldots, \beta'_{mj}, \ldots)'$ and $\alpha = (\ldots, \alpha'_{mj}, \ldots)'$ denote the stacked vector of all regression parameters, $\tau^2 = (\ldots, \tau^2_{mj}, \ldots)'$, $\phi^2 = (\ldots, \phi^2_{mj}, \ldots)'$, $\psi^2 = (\ldots, \phi^2_{ml}, \ldots)'$ the vectors of corresponding variances τ^2_{mj}, ϕ^2_{mj}, ψ^2_{ml} and $\delta = (\delta'_1, \ldots, \delta'_M)'$ the stacked vector of all fixed effects parameters.

Posterior analysis is then based on

$$\pi\left(\beta, \alpha, \tau^2, \phi^2, \psi^2, \delta, \Sigma \middle| y\right) \propto f\left(y \middle| \beta, \alpha, \delta, \Sigma\right) \prod_{m=1}^{M} \prod_{j=1}^{M} \left[\pi\left(\beta_{mj} \middle| \tau^2_{mj}\right) \pi\left(\alpha_{mj} \middle| \phi^2_{mj}\right) \pi\left(\tau^2_{mj}\right) \pi\left(\phi^2_{mj}\right)\right] \tag{32}$$

$$\times \prod_{m=1}^{M} \prod_{l=0}^{L} \left[\pi\left(\gamma_{ml} \middle| \psi^2_{ml}\right) \pi(\psi^2_{ml})\right] \pi(\delta)\, \pi(\Sigma) \tag{33}$$

with $f(y|\cdot)$ denoting the likelihood of the data.

Parameters are estimated via Gibbs sampling within the following blocks (Lang et al. 2003):

1. Sample β_{mj}, $m = 1, \ldots, M$, $j = 1, \ldots, M$. The full conditional for β_{mj} is Gaussian, $\beta_{mj} | \cdot \sim N(\mu_{mj}, P_{mj}^{-1})$, with precision matrix

$$P_{mj} = \frac{X'_{mj} A^2_{mj} X_{mj}}{\sigma^2_{m|-m}} + \frac{K_{mj}}{\tau^2_{mj}}$$

and mean

$$\mu_{mj} = P_{mj}^{-1} \left(\frac{1}{\sigma^2_{m|-m}} X'_{mj} A_{mj} (y_m - o_m) \right).$$

Here, $\sigma^2_{m|-m}$ is the (conditional) variance

$$\sigma^2_{m|-m} = \sigma^2_m - \Sigma_{m,-m} \Sigma_m^{-1} \Sigma'_{m,-m},$$

derived from partitioning Σ into

$$\Sigma = \begin{pmatrix} \sigma^2_m & \Sigma_{m,-m} \\ \Sigma'_{m,-m} & \Sigma_m \end{pmatrix}$$

(after reordering for the mth component of the error variable). The vector o_m is an offset vector defined in Lang et al. (2003).

2. Sample α_{mj}, $m = 1, \ldots, M$, $j = 1, \ldots, M$. The full conditionals for the scaling factors are derived from (31) with $\alpha_{mj} | \cdot \sim N(\mu_{mj}, P_{mj}^{-1})$ and

$$P_{mj} = \frac{Z' \tilde{X}^2_{mj} Z}{\sigma^2_{m|-m}} + \frac{I}{\phi^2_{mj}}$$

and mean

$$\mu_{mj} = P_{mj}^{-1} \left(\frac{1}{\sigma^2_{m|-m}} Z' \tilde{X}_{mj} (y_m - o_m) \right).$$

3. Sample γ_{ml}, $m = 1, \ldots, M$, $l = 0, \ldots, L$. Full conditionals are standard and omitted here.
4. Sample δ_m, $m = 1, \ldots, M$. Full conditionals are standard and omitted here.
5. Sample τ^2_{mj}, $m = 1, \ldots, M$, $j = 1, \ldots, M$. Full conditionals for the variance parameters τ^2_{mj} are inverse Gamma distributions with parameters

$$a_{mj} = 0.001 + \frac{rank(K_{mj})}{2}, \quad b_{mj} = 0.001 + \frac{1}{2} \beta'_{mj} K_{mj} \beta_{mj}. \qquad (34)$$

6. Sample ϕ^2_{mj}, $m = 1, \ldots, M$, $j = 1, \ldots, M$. Full conditionals for the variance parameters ϕ^2_{mj} are inverse Gamma distributions with parameters

$$a_{mj} = 0.001 + \frac{N}{2}, \quad b_{mj} = 0.001 + \frac{1}{2}\alpha'_{mj}\alpha_{mj}. \tag{35}$$

7. Sample Σ. The full conditional for Σ is an inverse Wishart distribution with parameters

$$a = 1 + \frac{n}{2}, \quad B = 0.005I + \frac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it} - \eta_{it})(y_{it} - \eta_{it})'. \tag{36}$$

where $Y_{it} = (y_{1it}, \ldots, y_{Mit})'$ and $\eta_{it} = (\eta_{1it}, \ldots, \eta_{Mit})'$

The complete sampling scheme can be found in Lang et al. (2003, pp. 270–271) in combination with Lang et al. (2015). The estimation procedure of the semiparametric SUR model is implemented in the software BayesX and the code is available from the authors.

Appendix 3: Overview of model specifications

Model	Specification	Modeled effects			
		Intercept	Price effects	Display and price ending effects	Holiday effect
PHomSM	Parametric, homogeneous	Store-specific	Equal across stores	Equal across stores	Equal across stores
PLCSM	Parametric, heterogeneous	Store-specific	Segment-specific	Segment-specific	Equal across stores
PHBSM	Parametric, heterogeneous	Store-specific	Store-specific	Store-specific	Equal across stores
PHetSM	Parametric, heterogeneous	Store-specific	Store-specific within segments	Store-specific within segments	Equal across stores
FHomSM	Semiparametric, homogeneous	Store-specific	Equal across stores and nonparametric	Equal across stores	Equal across stores
FHBSM	Semiparametric, heterogeneous	Store-specific	Store-specific and nonparametric	Store-specific	Equal across stores

Appendix 4: Optimization details

The optimization algorithm implemented in the R function *genoud()* is able to solve problems for objective functions that are nonlinear or even discontinuous in parameters and provides a high probability of finding the global optimum (Sekhon and Mebane 1998). The evolutionary algorithm starts with a population of trial solutions (in our application one trial solution corresponds to a vector of prices of the 8 brands). Then, a set of heuristic rules or operators (basically reproduction, mutation, crossover) is used to modify the trial solutions in order to increase their fitness values (i.e., the values of the function which is optimized). The selection of trial solutions for reproduction depends on their value of the objective function. The best trial solution is reproduced in each generation. The remaining trial solutions are recombined or mutated by applying the other operators (for details compare Sekhon and Mebane 1998, p. 192). That way, a new population (generation) results that "tends to be, on average, better than its predecessor" (Mebane and Sekhon 2011, p. 3). The following pseudo-code (adapted from Eiben and Smith 2003, p. 16) presents the general procedure:

```
BEGIN
   INITIALIZE population with random trial solutions
   EVALUATE each trial solution
   REPEAT UNTIL (termination criterion is satisfied) DO
      1 SELECT trial solutions
      2 RECOMBINE pairs of trial solutions
      3 MUTATE trial solutions
      4 EVALUATE each new trial solution
      5 SELECT new trial solutions for the next generation
   OD
END
```

Since the evolutionary algorithm is basically a genetic one where the code-strings are floating-point vectors (see, e.g., Ali and Törn 2004 for a description of genetic algorithms) its performance strongly depends on the population size (option $pop.size$) which must to be sufficiently large (Mebane and Sekhon 2011). In our application we chose $pop.size = 3000$ which leads to relatively high computing times but increases the probability of finding the global optimum. The maximum number of generations was set to $max.generations = 300$. The algorithm stops if the objective function is not improved anymore in a fixed number ($wait.generations$) of generations. We set $wait.generations = 10$. On average 25 generations were required to solve the optimization problems. The option $Domains$ comprises lower and upper bounds for each variable which correspond to the observed price range for each price variable in our applications. Finally, $boundary.enforcement = 2$ ensures candidates that lie within the bounds specified in $Domains$. The R code is available from the authors upon request.

In our empirical application, optimization is done separately for each of the 89 weeks as well as separately for each store/segment in case of the heterogeneous models.

The following table summarizes the size of the respective optimization problems for each model:

Model	Number of problems to solve
PHomSM, FHomSM	89 weeks = 89 problems to solve
	1 problem = 1 run for all brands and all stores
PLCSM (2)	89 weeks × 2 segments = 178 problems to solve
	1 problem = 1 run for all brands and all stores of 1 segment
PHBSM, FHBSM	89 weeks × 81 stores = 7209 problems to solve
	1 problem = 1 run for all brands (in 1 store)

References

Ailawadi KL, Beauchamp JP, Donthu N, Gauri D, Shankar V (2009) Communication and promotion decisions in retailing: a review and directions for future research. J Retail 85:42–55

Ainslie A, Rossi PE (1998) Similarities in choice behavior across product categories. Mark Sci 17:91–106

Ali MM, Törn A (2004) Population set-based global optimization algorithms: some modifications and numerical studies. Comput Oper Res 31:1703–1725

Allenby GM, Arora N, Ginter JL (1998) On the heterogeneity of demand. J Mark Res 35:384–389

Andrews RL, Currim IS, Leeflang PSH, Lim J (2008) Estimating the SCAN*PRO model of store sales: HB, FM or just OLS? Int J Res Mark 25:22–33

Basuroy S, Mantrala MK, Walters RG (2001) The impact of category management on retailer prices and performance: theory and evidence. J Mark 65:16–32

Bezawada R, Balachander S, Kannan PK, Shankar V (2009) Cross-category effects of aisle and display placements: a spatial modeling approach and insights. J Mark 73:99–117

Blattberg RC, Neslin SA (1990) Sales promotion: concepts. methods and strategies. Prentice Hall, Englewood Cliffs

Bolton RN, Shankar V (2003) An empirically derived taxonomy of retailer pricing and promotion strategies. J Retail 79:213–224

Brezger A, Steiner WJ (2008) Monotonic regression based on Bayesian P-splines: an application to estimating price response functions from store-level scanner data. J Bus Econ Stat 26:90–104

Celeux G, Hurn M, Robert CP (2000) Computational and inferential difficulties with mixture posterior distributions. J Am Stat Assoc 95:957–970

Chen Y, Hess JD, Wilcox RT (1999) Accounting profits versus marketing profits: a relevant metric for category management. Mark Sci 18:208–229

Chib S, Seetharaman PB, Strijnev A (2002) Analysis of multi-category purchase incidence decisions using IRI market basket data. Adv Econom 16:57–92

Chintagunta PK (2000) A flexible aggregate logit demand model. Working paper, University of Chicago

Chintagunta PK (2002) Investigating category pricing behavior at a retail chain. J Mark Res 39:141–154

Chintagunta PK, Dubé JP, Singh V (2003) Balancing profitability and customer welfare in a supermarket chain. Quant Mark Econ 1:111–147

Dobson PW, Waterson M (2008) Chain-store competition: customized vs. uniform pricing. Working paper, University of Warwick

Eiben AE, Smith JE (2003) Introduction to evolutionary computing. Springer, Berlin

Elrod T, Russell G, Shocker A, Rao V, Bayus B, Caroll D, Kamakura W, Shankar V (2002) Inferring market structure from customer response to competing and complementary products. Mark Lett 13:221–232

Fahrmeir L, Kneib T, Lang S (2007) Regression - Modelle, Methoden und Anwendungen. Springer, Berlin

Frühwirth-Schnatter S (2001) Markov chain Monte Carlo estimation of classical and dynamic switching and mixture models. J Am Stat Assoc 96:194–209

Frühwirth-Schnatter S (2006) Finite mixture and Markov switching models. Springer, New York
Frühwirth-Schnatter S, Tüchler R, Otter T (2004) Bayesian analysis of the heterogeneity model. J Bus Econ Stat 22:2–15
Frühwirth-Schnatter S, Tüchler R, Otter T (2005) Capturing unobserved consumer heterogeneity using the Bayesian heterogeneity model. In: Taudes A (ed) Adaptive information systems and modelling in economics and management science. Springer, Wien, pp 57–70
Gelfand AE, Sahu SK, Carlin BP (1995) Efficient parametrisations for normal linear mixed models. Biometrika 82:479–488
Greene WH (2008) Econometric analysis, 6th edn. Prentice Hall, New Jersey
Hall JM, Kopalle PK, Krishna A (2010) Retailer dynamic pricing and ordering decisions: category management versus brand-by-brand approaches. J Retail 86:172–183
Hoch SJ, Lodish LM (1998) Store brands and category management. Working paper, The Wharton School, University of Pennsylvania
Hoch SJ, Kim BD, Montgomery AL, Rossi PE (1995) Determinants of store-level price elasticity. J Mark Res 32:17–29
Hruschka H (2006a) Relevance of functional flexibility for heterogeneous sales response models: a comparison of parametric and semi-nonparametric models. Eur J Oper Res 174:1009–1020
Hruschka H (2006) Statistical and managerial relevance of aggregation level and heterogeneity in sales response models. Mark J Res Manage 2(2006):94–102
Hruschka H (2007) Clusterwise pricing in stores of a retail chain. OR Spectr 29:579–595
Inman JJ, Winer RS, Ferraro R (2009) The interplay between category characteristics and shopping trip factors on in-store decision making. J Mark 73:19–29
Kadiyali V, Chintagunta P, Vilcassim N (2000) Manufacturer-retailer channel interactions and implications for channel power: an empirical investigation of pricing in a local market. Mark Sci 19:127–148
Kamakura WA, Kang W (2007) Chain-wide and store-level analysis for cross-category management. J Retail 83:159–170
Khan RJ, Jain DC (2005) An empirical analysis of price discrimination mechanisms and retailer profitability. J Mark Res 42:516–524
Kim B, Blattberg RC, Rossi PE (1995) Modeling the distribution of price sensitivity and implications for optimal retail pricing. J Bus Econ Stat 13:291–303
Kim B, Srinivasan K, Wilcox R (1999) Identifying price sensitive customers: the relative merits of demographic vs. purchase pattern information. J Retail 75:1–21
Klapper D (2000) Einflußgrößen von regulären Preiselastizitäten, Preisaktionselastizitäten und Kreuzpreiselastizitäten - Determinants of regular, promotional and cross price elasticities. OR Spectr 22:135–157
Koop G (2003) Bayesian econometrics. Wiley, Chichester
Kopalle PK, Biswas D, Chintagunta PK, Fan J, Pauwels K, Ratchford BT, Sills JA (2009) Retailer pricing and competitive effects. J Retail 85:56–70
Kumar V, Leone RP (1988) Measuring the effect of retail store promotions on brand and store substitution. J Mark Res 25:178–185
Lang S, Brezger A (2004) Bayesian P-splines. J Comput Graph Stat 13:183–212
Lang S, Adebayo SB, Fahrmeir L, Steiner WJ (2003) Bayesian geoadditive seemingly unrelated regression. Comput Stat 18:263–292
Lang S, Steiner W, Weber A, Wechselberger P (2015) Accommodating heterogeneity and nonlinearity in price effects for predicting brand sales and profits. Eur J Oper Res 246:232–241
Lenk PJ, DeSarbo WS (2000) Bayesian inference for finite mixtures of generalized linear models with random effects. Psychometrika 65:93–119
Levy M, Chen H, Ray S, Bergen M (2004) Asymmetric price adjustment in the small: an implication of rational inattention. Discussion Paper Series 04-23, Utrecht School of Economics
Luan YJ, Sudhir K (2010) Forecasting marketing-mix responsiveness for new products. J Mark Res 47:444–457
Manchanda P, Ansari A, Gupta S (1999) The "shopping basket": a model for multicategory purchase incidence decisions. Mark Sci 18:95–114
Manchanda P, Rossi PE, Chintagunta P (2004) Response modeling with nonrandom marketing-mix variables. J Mark Res 41:467–478
Mebane WR Jr, Sekhon JS (2011) Genetic Optimization using derivatives: the rgenoud package for R. J Stat Softw 42:1–26

Montgomery AL (1997) Creating micro-marketing pricing strategies using supermarket scanner data. Mark Sci 16:315–337
Montgomery AL, Bradlow ET (1999) Why analyst overconfidence about the functional form of demand models can lead to overpricing. Mark Sci 18:569–583
Montgomery AL, Rossi PE (1999) Estimating price elasticities with theory-based priors. J Mark Res 36:413–423
Mulhern FJ, Leone RP (1991) Implicit price bundling of retail products: a multiproduct approach to maximizing store profitability. J Mark 55:63–76
Murray CC, Talukdar D, Gosavi A (2010) Joint optimization of product price, display orientation and shelf-space allocation in retail category management. J Retail 86:125–136
Natter M, Reutterer T, Taudes A (2007) An assortmentwide decision-support system for dynamic pricing and promotion planning in DIY retailing. Mark Sci 26:576–583
Nijs VR, Srinivasan S, Pauwels K (2007) Retail-price drivers and retailer profits. Mark Sci 26:473–487
Otter T, Gilbride TJ, Allenby GM (2011) Testing models of strategic behavior characterized by conditional likelihoods. Mark Sci 30:686–701
Petrin A, Train K (2010) A control function approach to endogeneity in consumer choice models. J Mark Res 47:3–13
Reibstein DJ, Gatignon H (1984) Optimal product line pricing: the influence of elasticities and cross-elasticities. J Mark Res 21:259–267
Rossi PE (2014) Even the rich can make themselves poor: a critical examination of IV methods in marketing applications. Mark Sci 33:655–672
Rossi PE, Allenby GM, McCulloch R (2005) Bayesian statistics and marketing. Wiley, Chichester
Russell GJ, Kamakura WA (1997) Modeling multiple category brand preference with household basket data. J Retail 73:439–461
Russell GJ, Petersen A (2000) Analysis of cross category dependence in market basket selection. J Retail 76:367–392
Sekhon JS, Mebane WR Jr (1998) Genetic optimization using derivatives. Political Anal. 7:187–210
Shankar V, Bolton RN (2004) An empirical analysis of determinants of retailer pricing strategy. Mark Sci 23:28–49
Shankar V, Kannan PK (2014) An across store analysis of intrinsic and extrinsic cross-category effects. Cust Needs Solut 1:143–164
Shankar V, Krishnamurthi L (1996) Relating price sensitivity to retailer promotional variables and pricing policy: an empirical analysis. J Retail 72:249–272
Shankar V, Inman JJ, Mantrala M, Kelley E, Rizley R (2011) Innovations in shopper marketing: current insights and future research issues. J Retail 87S:S29–S42
Silva-Risso JM, Bucklin RE, Morrison DG (1999) A decision support system for planning manufacturers' sales promotion calendars. Mark Sci 18:274–300
Smith M, Kohn R (2000) Nonparametric seemingly unrelated regression. J Econ 98:257–281
Song I, Chintagunta PK (2006) Measuring cross-category price effects with aggregate store data. Manag Sci 52:1594–1609
Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A (2002) Bayesian measures of model complexity and fit. J Roy Stat Soc: Ser B 64:583–639
Steiner WJ, Brezger A, Belitz C (2007) Flexible estimation of price response function using retail scanner data. J Retail Consum Serv 14:383–393
Sudhir K (2001) Structural analysis of manufacturer pricing in the presence of a strategic retailer. Mark Sci 20:244–264
Tellis GJ, Zufryden FS (1995) Tackling the retailer decision maze: which brands to discount, how much, when and why? Mark Sci 14:271–299
van Heerde HJ, Leeflang PSH, Wittink DR (2002) How promotions work: SCAN*PRO-based evolutionary model building. Schmalenbach Bus Rev 54:198–220
Verbeke G, Lesaffre E (1996) A linear mixed-effects model with heterogeneity in the random-effects population. J Am Stat Assoc 91:217–221
Vilcassim NJ, Chintagunta PK (1995) Investigating retailer product category pricing from household scanner panel data. J Retail 71:103–128
Villas-Boas JM, Winer RS (1999) Endogeneity in brand choice models. Manag Sci 45:1324–1338
Weber A (2015) Modeling price response from store sales: The roles of store heterogeneity and functional flexibility. Shaker Verlag, Aachen

Wedel M, Zhang J (2004) Analyzing brand competition across subcategories. J Mark Res 41:448–456
Wedel M, Zhang J, Feinberg F (2004) A model-based approach to setting optimal retail markups. Working paper, University of Michigan
Weicker K (2007) Evolutionäre Algorithmen, 2nd edn. Teubner Verlag, Wiesbaden
Zellner A (1962) An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. J Am Stat Assoc 57:348–368
Zenor MJ (1994) The profit benefits of category management. J Mark Res 31:202–213