Interaction Between Shelf Layout and Marketing Effectiveness and Its Impact On Optimizing Shelf Arrangements

Erjen van Nierop, Dennis Fok and Philip Hans Franses

ERIM REPORT SERIES *RESEARCH IN MANAGEMENT*	
ERIM Report Series reference number	ERS-2006-013-MKT
Publication	March 2006
Number of pages	35
Persistent paper URL	
Email address corresponding author	dfok@few.eur.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus Universiteit Rotterdam P.O.Box 1738 3000 DR Rotterdam, The Netherlands Phone: + 31 10 408 1182 Fax: + 31 10 408 9640 Email: info@erim.eur.nl Internet: www.erim.eur.nl

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:
www.erim.eur.nl

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES

RESEARCH IN MANAGEMENT

ABSTRACT AND KEYWORDS	
Abstract	Allocating the proper amount of shelf space to stock keeping units [SKUs] is an increasingly relevant and difficult topic for managers. Shelf space is a scarce resource and it has to be distributed across a larger and larger number of items. It is in particular important because the amount of space allocated to a specific item has a substantial impact on the sales level of that item. This relation between shelf space and sales has been widely documented in the literature. However, besides the amount of space, the exact location of the SKU on the shelf is also an important moderator of sales. At the same time, the effectiveness of marketing instruments of an SKU may also depend on the shelf layout. In practice, retailers recognize that these dependencies exist. However, they often revert to rules of thumb to actually arrange their shelf layout. We propose a new model to optimize shelf arrangements in which we use a complete set of shelf descriptors. The goal of the paper is twofold. First of all, we aim to gain insight into the dependencies of SKU sales and SKU marketing effectiveness on the shelf layout. Second, we use these insights to improve the shelf layout in a practical setting. The basis of our model is a standard sales equation that explains sales from item-specific marketing-effect parameters and intercepts. In a Hierarchical Bayes fashion, we augment this model with a second equation that relates the effect parameters to shelf and SKU descriptors. We estimate the parameters of the two-level model using Bayesian methodology, in particular Gibbs sampling. Next, we optimize the total profit over the shelf arrangement. Using the posterior draws from our Gibbs sampling algorithm, we can generate the probability distribution of sales and profit in the optimization period for any feasible shelf arrangement. To find the optimal shelf arrangement, we use simulated annealing. This heuristic approach has proven to be able to effectively search an enormous solution space. Our results indicate that our model is able to fit and forecast the sales levels quite accurately. Next, when applying the simulated annealing algorithm to the shelf layout, we appear to be able to increase profits for all the stores analyzed. We compare our approach to commonly used shelf optimization rules of thumb. Most sensible rules of thumb also increase expected profits (although not as much as our optimization algorithm). In particular, it is beneficial to put high-margin items close to the beginning of the aisle (or the “racetrack"). Finally, we provide managerial implications and directions for further research.
Free Keywords	Shelf Management, Sales Models, Hierarchical Bayes, Markov Chain Monte Carl, Simulated Annealing.
Availability	The ERIM Report Series is distributed through the following platforms: Academic Repository at Erasmus University (DEAR), DEAR ERIM Series Portal Social Science Research Network (SSRN), SSRN ERIM Series Webpage Research Papers in Economics (REPEC), REPEC ERIM Series Webpage
Classifications	The electronic versions of the papers in the ERIM report Series contain bibliographic metadata by the following classification systems: Library of Congress Classification, (LCC) LCC Webpage Journal of Economic Literature, (JEL), JEL Webpage ACM Computing Classification System CCS Webpage Inspec Classification scheme (ICS), ICS Webpage

Interaction between shelf layout and marketing effectiveness and its impact on optimizing shelf arrangements

Erjen van Nierop *
Department of Marketing, University of Groningen
Dennis Fok
Econometric Institute, Erasmus University Rotterdam
Philip Hans Franses
Econometric Institute and Department of Business Economics
Erasmus University Rotterdam

March 16, 2006

*Address for correspondence: University of Groningen, Faculty of Economics and Business, Room WSN 411, POBox 800, 9700 AV Groningen, email: j.e.m.van.nierop@rug.nl. We would like to thank Alan Montgomery for providing us with the data. We would also like to thank Paul Farris, Kusum Ailawadi and Richard Paap for useful comments. Finally, we thank participants at seminars at Carnegie Mellon University in Pittsburgh and Erasmus University Rotterdam.

Interaction between shelf layout and marketing effectiveness and its impact on optimizing shelf arrangements

Abstract

Allocating the proper amount of shelf space to stock keeping units [SKUs] is an increasingly relevant and difficult topic for managers. Shelf space is a scarce resource and it has to be distributed across a larger and larger number of items. It is in particular important because the amount of space allocated to a specific item has a substantial impact on the sales level of that item. This relation between shelf space and sales has been widely documented in the literature. However, besides the amount of space, the exact location of the SKU on the shelf is also an important moderator of sales. At the same time, the effectiveness of marketing instruments of an SKU may also depend on the shelf layout. In practice, retailers recognize that these dependencies exist. However, they often revert to rules of thumb to actually arrange their shelf layout.

We propose a new model to optimize shelf arrangements in which we use a complete set of shelf descriptors. The goal of the paper is twofold. First of all, we aim to gain insight into the dependencies of SKU sales and SKU marketing effectiveness on the shelf layout. Second, we use these insights to improve the shelf layout in a practical setting. The basis of our model is a standard sales equation that explains sales from item-specific marketing-effect parameters and intercepts. In a Hierarchical Bayes fashion, we augment this model with a second equation that relates the effect parameters to shelf and SKU descriptors. We estimate the parameters of the two-level model using Bayesian methodology, in particular Gibbs sampling. Next, we optimize the total profit over the shelf arrangement. Using the posterior draws from our Gibbs sampling algorithm, we can generate the probability distribution of sales and profit in the optimization period for any feasible shelf arrangement. To find the optimal shelf arrangement, we use simulated annealing. This heuristic approach has proven to be able to effectively search an enormous solution space.

Our results indicate that our model is able to fit and forecast the sales levels quite accurately. Next, when applying the simulated annealing algorithm to the shelf layout, we appear to be able to increase profits for all the stores analyzed. We compare our approach to commonly used shelf optimization rules of thumb. Most sensible rules of thumb also increase expected profits (although not as much as our optimization algorithm). In particular, it is beneficial to put high-margin items close to the beginning of the aisle (or the "racetrack"). Finally, we provide managerial implications and directions for further research.

Keywords: shelf management, sales models, Hierarchical Bayes, Markov Chain Monte Carlo, Simulated Annealing.

1 Introduction

Retailers have limited shelf space available. The choice of which items to stock and the allocation of scarce shelf space among the stocked items are relevant issues to the retailer. For individual SKUs these decisions are important determinants of the sales and the marketing effectiveness. At the aggregate level, shelf allocation is an important factor in the revenue, cost and eventual profit of a product category. Complementary to the amount of space to allocate to an item, there is the problem of the location of the item on the shelf. For example, items on the lower shelf usually get less consumer attention than items on upper shelves. The items on the lower shelves may therefore have lower sales and may also benefit less from promotions.

Finding the profit-maximizing shelf arrangement and, at the same time, meeting the requirements of manufacturers, is far from easy. Additional to the items currently in the assortment, there are also line extensions that are fighting for share of sales and share of shelf. This further complicates the retailer's optimization problem. A prerequisite to the actual shelf optimization, is a proper measurement of the effect of shelf layout on sales and marketing effectiveness. An adequate shelf management model would be a very useful aid to retailers to estimate these relations and to support their decisions and their negotiations with manufacturers.

In this paper, we propose such a shelf management model. The basis of our model is a standard sales equation that explains (the logarithm of) sales from item-specific marketing-effect parameters and intercepts. In a Hierarchical Bayes [HB] fashion, this model is augmented with a second equation that relates the marketing-effect parameters to shelf and SKU descriptors. This second equation provides the link between shelf allocation on the one hand and sales and marketing effectiveness on the other hand. We estimate the parameters of the two-level model using the Bayesian methodology, in particular Gibbs sampling. The estimated model parameters measure the effect of shelf layout on baseline sales and on the effectiveness of marketing instruments such as price and promotions. We use graphs to visualize these (non-linear) effects. To investigate how

the model performs in forecasting sales, we predict sales for a hold-out sample containing five weeks of data. Furthermore, and most importantly, the model is used to optimize shelf allocation. For this, we consider Simulated Annealing, for its ease of implementation and the ability to search across a large and complex solution space as well as for its ability to avoid getting stuck in a local optimum.

The remainder of this paper is organized as follows. In Section 2, we review the current literature on shelf management. We also indicate the added value of our approach with respect to the current literature. In Section 3, we discuss our approach in words. Next, we present the technicalities of our model in Section 4. Subsequently, we illustrate our shelf management approach using a database concerning the canned soup category. It contains a rich description of the shelf space and location on the shelf of a large number of products, where the shelf layouts were manipulated in an experimental setting. We conclude in Section 6.

2 Literature

In the 1960s and 1970s, a number of experiments were conducted to measure the effect of shelf space on sales, see for example Brown and Tucker (1961), Cox (1970) and Curhan (1972). These authors only considered the problem of measuring this effect. Models to (partly) solve the shelf management problem have been proposed in the past decades. Corstjens and Doyle (1981, 1983) were the first to optimize store profitability with respect to space allocation. They consider both the main and the cross-space elasticities in their multiplicative demand function, and specify a cost function that moderates the profitability of the allocation. This shelf-space optimization problem is solved within a geometrical programming framework. In a comparison of their approach with alternative procedures they find that their general model leads to significantly different allocation rules and better profit performance.

Bultez and Naert (1988) build on the work of Corstjens and Doyle (1981, 1983) in their SH.A.R.P. (Shelf Allocation for Retailer's Profit) model. The authors derive an

expression for the optimal shelf space to be allocated to an SKU. This expression depends on the cross-space elasticities between the items. Commonly used rules of thumb for space allocation are compared and shown to be special, though inferior, cases of the optimal rule derived. The authors apply the model to experimental data with six brands and find that the proposed model improves upon current profit levels and that it is better than the rules of thumb. The optimization only focuses on the shelf space devoted to an item, and does not include other shelf layout descriptors such as shelf height and the horizontal position of an item on the shelf, nor does it include marketing instruments such as feature and price.

Drèze et al. (1994) conduct a series of field experiments in which they measure the effectiveness of two shelf management techniques: "space-to-movement", where the shelf is customized based on historic store-specific movement patterns, and "product reorganization", where product placement is manipulated to facilitate cross-category merchandizing or ease-of-shopping. The authors find sales gains of about 4% with the first manipulation and 5-6% with the second. The impact of shelf positioning and facing allocations on sales of individual items is also analyzed. In particular, location appears to have a large impact on sales. For example, in most categories, products perform best when placed at eye level.

Borin et al. (1994) develop a category management model formulated as a constrained optimization problem, with assortment and allocation of space as the decision variables. The parameters of the model are based on judgmental estimates, that is, they are not based on an econometric model. In the next step, the authors use simulated annealing to improve the shelf layout for two data sets. The two data-sets analyzed contain 6 SKUs and 18 SKUs, respectively. In a follow-up study (Borin and Farris, 1995), the authors examine the sensitivity of the analysis to errors in the judgements. More specifically, they find the maximum degree of error that may be introduced before the model yields assortments and shelf allocation that are inferior compared to those produced by the merchandizing rule of thumb to set share-of-shelf equal to share-of-sales. Their results show that as much as 50% variation in the estimates of parameters is allowed before the

model appears unusable.

In a more or less separate stream of research, optimization routines for shelf allocation have been investigated. Several routines have been proposed to optimize the shelf layouts. Yang and Chen (1999) use a simplified version of the integer programming model of Corstjens and Doyle (1981), whereas Yang (2001) uses the knapsack algorithm. Lim et al. (2004) build on this work by optimizing profits with two metaheuristic approaches, that is, Tabu Search and Squeaky-Wheel Optimization. Their method appears to outperform Yang's heuristic. However, by using simulated sales data and by using fixed and known parameters, these approaches assume that the effect of shelf layout on sales is given. In a real-life situation, this is of course not true and one needs to estimate the relation between sales and shelf layout for a particular situation.

In this paper, we propose a Hierarchical Bayes [HB] model to estimate the interaction between shelf layout and sales and the interaction between marketing instrument effectiveness. In an HB model, the parameters for individual items are assumed to be samples from a common distribution, with possibly different means. In this way, the parameter estimates for the separate SKUs will be "shrunk" towards reasonable values, thereby dampening some of the undesirable variation that separate, independent, estimators could have. The (marketing) literature contains many papers using hierarchical models, see for example, Blattberg and George (1991), Montgomery (1997) and Boatwright et al. (1999). In these papers it is documented that the hierarchical model reduces the problem of coefficient instability across equations and that it improves predictive power.

Based on this model we develop an optimization procedure for shelf management using simulated annealing. In contrast to the existing literature, we explicitly account for a moderating effect of shelf layout on marketing-mix elasticities. Furthermore, instead of restricting the analysis to shelf space, we also consider other shelf descriptors such as the number of items stacked on top of each other and the horizontal and vertical position of an item on the shelf. Moreover, we develop our model for a large number of items. Instead of considering the market at the brand level, we consider the individual SKUs.

Finally, we appropriately take into account uncertainty in sales and uncertainty in the model parameters. We believe that this situation comes closer to actual practice.

To summarize, our modeling and optimization approach is in various ways related to previous papers in the literature. We extend most previous shelf optimization approaches in at least one out of four important ways (i) we account for dependencies between shelf characteristics and marketing-mix elasticities, (ii) we use a rich description of the shelf layout instead of just focusing on shelf space, (iii) we optimize the layout while taking into account the uncertainty in sales and model parameters, and (iv) our model can easily be considered for a large number of items. In Table 1 we give an overview of the present literature and their most important features.

3 Our approach

The model we propose in this paper aims to accurately measure the effect of shelf space and shelf placement on sales levels and on marketing instrument effectiveness. Furthermore, we avoid unnecessary simplifications in the shelf optimization.

First of all, we describe the set of shelf descriptors we use in our model. Following most papers cited above, we have the number of facings as an important determinant of demand. In addition, we use the shelf number, expecting that products that are higher on the shelf have higher visibility. At the same time, some decreasing returns of shelf height may also appear. To capture this, we also use the distance to the middle shelf (usually the third shelf) as a moderating variable. Thirdly, we use the distance of an item to the end of the shelf as a shelf descriptor. Products that are closer to the beginning of the shelf may benefit from people reaching the item quicker coming from the back isle, or the "racetrack" (Larson et al., 2005). On the other hand, we may find the opposite effect in that items that are in the middle, get more attention from consumers, who may often end up in the middle of the shelf for the category. To capture these potential non-linear position effects, we also add the distance to the middle of the shelf to our set of shelf descriptors. Next to characteristics of the shelf layout, characteristics of the product

Table 1: Papers on shelf layout optimization

Paper	**Shelf descriptors**	**Moderated variables**	**Estimated shelf effects?**[a]	**Optimization method**	**Max. number of items in application**
Corstjens and Doyle (1981)	Shelf space	Baseline sales	yes	geometric programming	5 (product groups)
Corstjens and Doyle (1983)	Shelf space	Baseline sales	yes	none	4 (fictitious products)
Bultez and Naert (1988)	Shelf space	Baseline sales	yes	non-linear programming heuristic	20
Drèze et al. (1994)	Shelf space & shelf location	Baseline sales	yes	integer programming	235
Borin et al. (1994) Borin and Farris (1995)	Shelf space	Baseline sales	yes	Simulated Annealing	18
Yang and Chen (1999)	Shelf space	Baseline sales	no	non-linear integer programming	6
Yang (2001)	Shelf space	Baseline sales	no	Heuristics + knapsack	10
Lim et al. (2004)	Shelf space	Baseline sales	no	Tabu search, Squeaky-Wheel	100
This study	Shelf space & shelf location	Baseline sales & marketing elasticities	yes	Simulated Annealing	407

[a] Indicates whether the effects of shelf layout on sales are estimated, as opposed to assumed known as is often done in the optimization stream of the shelf optimization literature.

itself, such as item width and brand name, are also incorporated as we expect they will also influence demand and elasticities.

The basis of our model is a standard sales equation that explains (log)sales from item- and time-specific intercepts and marketing-effect parameters. We augment the sales equation with a second equation that relates the intercepts and effect parameters to shelf and SKU descriptors. To estimate the parameters of this two-level HB model, we use MCMC methodology, in particular Gibbs sampling. We use graphs to display the potentially complicated non-linear effects captured by our model. To investigate how well the model performs in describing and forecasting the sales (and thus, profits in the shelf optimization), we predict sales for a hold-out sample of five weeks of data. A comparison of our forecasts to the actual sales and to forecasted sales using SKU-level regressions provides information on the absolute and relative performance of our model.

Next, we optimize the total profit for the final week by changing the shelf arrangement. Using the posterior draws of the model parameters from our Gibbs sampling algorithm, we can obtain the probability distribution of sales and profit in the optimization period for any feasible shelf arrangement. We evaluate the shelf allocations using the mean of the posterior gross profit distribution. Note that this measure gives the expected profit over all sources of uncertainty, that is, uncertainty in sales as well as uncertainty in the (estimated) parameters. The minimum number of facings for each item equals one. We do not allow items to have zero facings, that is, we do not consider assortment decisions. We believe the model and the data are not suited for these decisions, since it is very likely that moving from two to one facings implies a different elasticity than moving from one to zero facings. Since item deletions are not observed in the data, it is not possible to go this far in the optimization.

There are many ways to search for the optimal shelf arrangement. Given the complexity of the problem, an algorithm that yields a guaranteed optimal solution is hard, if not impossible, to obtain. Therefore, we opt for simulated annealing, which is a heuristic approach, to search for the optimal shelf arrangement. The algorithm starts with a

random shelf arrangement and then searches the neighborhood of the current solution for better ones. To avoid getting stuck in local maxima, an inferior solution may also be temporarily accepted, but the probability of this decreases as the algorithm proceeds, see Johnson et al. (1989).

4 A new model for shelf management

In this section we describe our modeling and optimization approach in detail. In Section 4.1, we discuss the model. Technical details concerning the estimation of the model parameters are presented in the Appendix. Next, in Section 4.2 we describe how we optimize the shelf layout.

4.1 Representation of the sales model

First we introduce some notation. We denote the number of SKUs in the market by I, the number of observations for SKU i by T_i, and the number of item attributes by L. Among the L attributes, there are $C \leq L$ shelf characteristics such as the number of facings, the shelf number and the distance to the end of the shelf. To explain sales we have K explanatory variables, such as price and promotion. Let $\ln S_{i,t}$ be the natural log of sales of SKU i at time $t = 1, ..., T_i$. We model the log sales by a standard log-linear model (see for example Wittink et al., 1988), that is,

$$\ln S_{i,t} = X'_{i,t}\beta_{i,t} + \varepsilon_{i,t}, \quad i = 1, \ldots, I, \;\; t = 1, \ldots, T_i \tag{1}$$

where $X_{i,t}$ denotes a $(K+1) \times 1$ dimensional vector containing an intercept and the K explanatory variables for SKU i at time t and where $\beta_{i,t} = (\beta_{0,i,t}, \ldots, \beta_{K,i,t})'$ with $\beta_{k,i,t}$ the coefficient measuring the effect of the k-th explanatory variable for SKU i at time t. The vector of explanatory variables will in general contain (log-transformed) marketing instruments such as price, feature and 0/1 dummy variables such as promotion. We let the error term $\varepsilon_{i,t}$ be independently distributed $N(0, \sigma_i^2)$.

All β parameters in (1) potentially differ across SKUs and across time. To describe how these parameters vary over these two dimensions, we add a second layer to the model. In this layer we specify a model for the marketing-effect parameters. As explanatory variables in this second-level model, we use item-specific attribute data and shelf-layout descriptives, both of which may, in general, vary over time. Of these attributes, the shelf characteristics are most likely to change. This can for example happen due to a relocation of items during the observational period. We denote the observed attributes of item i by an $(L+1) \times 1$ vector $Z_{i,t}$. This vector contains an intercept, the SKU characteristics, and the shelf layout characteristics at time t. We introduce the following linear relation between the item-specific parameters and the attribute space, that is,

$$\beta_{i,t} = \gamma Z_{i,t} + \eta_i, \quad \eta_i \sim N(0, \Sigma_\eta) \tag{2}$$

where γ is a $(K+1) \times (L+1)$ matrix of parameters. The coefficients $\gamma_{k,l}$ represent the effect of attribute l on the effect size of marketing instrument k. For $k = 0$ the coefficients represent the effects of the SKU characteristics on the intercept of the sales equation. In other words, these coefficients give the direct effect of the shelf layout on sales. Note that the current literature usually restricts the analysis to only these effects (Corstjens and Doyle, 1981; Yang and Chen, 1999; Yang, 2001; Lim et al., 2004). Also, most papers only use facings as a shelf layout descriptor.

Of course, there may be relevant attributes that we do not observe, or there may be intangible attributes such as brand equity that also influence the baseline sales and the marketing-instrument effectiveness. We represent the joint effect of such attributes by a normally distributed disturbance term in (2), that is, $\eta_i = (\eta_{i,0}, \eta_{i,1}, \ldots, \eta_{i,K})' \sim N(0, \Sigma_\eta)$. Note that we assume that these intangible characteristics are fixed over time. This implies that we assume that a relocation of the products will not affect η_i. The degree of uncertainty may differ across instruments, and we therefore allow the variance of $\eta_{i,k}$ to depend on k. Furthermore, we may expect that some unobserved attributes simultaneously affect multiple marketing instruments. For example, if an item has a high feature effectiveness it may also be very effective with display. Such relations will lead to

positive correlations between $\eta_{i,k}$ and $\eta_{i,h}$. To capture such correlations we allow Σ_η to be non-diagonal.

An alternative view on (2) is that an SKU can be represented by a specific point in an attribute space. The second layer of our model then specifies a (linear) mapping from the attribute space to the model parameters in (1). Furthermore, by explicitly recognizing that items that are close in attribute space will have similar parameters, we efficiently make use of the data to estimate marketing-effectiveness parameters.

In sum, the combination of (1) and (2) gives our attribute-based sales model. The joint estimation of these two equations gives more precise estimates of the attribute mapping than a two-step approach, in which (1) would be estimated separately per SKU and where the resulting estimates of β_i would then be regressed on SKU and shelf characteristics. Our HB approach yields more accurate estimates as it combines all the available information and accounts for uncertainty in estimates of the marketing instrument effectiveness. Furthermore, in a two-step approach it would be difficult to deal with changes in characteristics of the shelf allocation. In the Appendix, we discuss an MCMC algorithm that can be used to estimate the model parameters and which, as a by-product, gives draws from the distribution of all parameters conditional on the data.

4.2 Shelf optimization

The output of the Gibbs sampling algorithm allows us to draw inference on the posterior distribution of any function of the parameters. The total profit of the category for a particular week and store is one example of such a function. Our model contains a complete set of shelf arrangement descriptors in the number of facings, shelf height and distance to the end of the shelf. We can therefore obtain the posterior distribution of the profit based on the current shelf layout as well as based on any feasible alternative layout, which is key to our approach. Note that the posterior profit distributions are conditional on the (in-sample) data and they represent both the uncertainty in the sales levels themselves as the uncertainty in the parameters. In turn, we can use these distributions to optimize

the total profit for the final week (out of sample), conditional on the data. In this paper we use the mean of the posterior profit distribution to measure the quality of the associated shelf arrangement. As an alternative, one could also consider the mode, or even the 5%-percentile of the profit distribution. The latter corresponds to maximizing the profit under the "worst-case scenario".

4.2.1 The shelf optimization problem

In our representation of the layout, items are allocated a number of facings on a specific shelf with a specified distance to the end of the shelf. The number of products stacked on top of each other will be determined based on the available shelf height and the dimensions of the item, that is,

$$Z_{i,\text{stack}} = \left\lfloor \frac{\text{ShelfHeight}(Z_{i,\text{shelf}})}{Z_{i,\text{height}}} \right\rfloor, \forall i, \tag{3}$$

where $\lfloor x \rfloor$ gives the floor of x.

Let $\mathrm{E}[S_i(Z_i)]$ be the expected sales for item i, given its shelf allocation and item characteristics Z_i. Let m_i denote the per unit contribution for item i. Also, let $c_i(Z_i)$ be the replenishment cost for carrying item i for a given layout Z_i [1]. Furthermore, define Π as the total profit for the category.

$$\Pi = \sum_{i=1}^{I} (m_i \mathrm{E}[S_i(Z_i)] - c_i(Z_i)). \tag{4}$$

The issue of interest is to maximize Π given several restrictions. The main restrictions concern the logical consistency of the shelf layout. Formal mathematical restrictions that correspond to these consistency requirements are difficult to formulate. Previous papers that did specify formal mathematical restrictions only consider the number of facings as a decision variable, while the exact location on the shelf is not taken into account. In this case, the restrictions are much easier to specify in a mathematical programming format. However, even with more shelf descriptors, all restrictions are easy to check in practice, for

[1] This in turn also depends on the expected sales given layout Z_i, as more sales means more replenishment activity.

example, (i) each SKU must be assigned to a shelf, and (ii) the total width of shelf space used by items may not exceed the total shelf space available; (iii) the shelf space allocated to a particular SKU may not (partly) overlap with another SKU. If one would optimize the layout using, for example, linear programming all these restrictions would have to be translated into formal mathematical equations. We choose to approach these restrictions in a different way. In the search for the optimal shelf layout we only consider feasible layouts. Thereby we make sure that the layout always satisfies the given constraints.

If needed, additional restrictions can easily be added for the particular retailer's situation at hand. For example, it may be interesting to add restrictions on the capacity of the shelf space allocated to SKUs. In some cases the capacity of the allocated shelf space must at least be equal to the minimum packout. That is, in case of restocking of the item one full packout has to fit on the shelf. Incorporating such a restriction in our optimization strategy is very simple. We again just have to make sure that we do not consider layouts that violate these restrictions.

Given the enormous number of possible combinations of facings, shelf numbers and the other decision variables, it is impossible to find a closed-form solution for this optimization problem, in particular if the number of SKUs is large. The geometrical programming framework, or branch-and-bound procedure, as employed by Corstjens and Doyle (1981), would also have a hard time finding an optimal solution in the high-dimensional space. Therefore, a heuristic optimization technique as Simulated Annealing is necessary to search for the profit-maximizing shelf layout in a practical retailer situation.

4.2.2 Simulated Annealing applied to shelf optimization

Simulated Annealing [SA] was proposed by Kirkpatrick et al. (1983). One of the advantages of this algorithm is that in each step a feasible solution is guaranteed. In our setting this means that the layout in each iteration will comply with all logical consistency restrictions. In each iteration of the algorithm new layouts in the neighborhood of the current solution are considered. If a candidate solution performs better than the current

one, the current solution is discarded in favor of the candidate. With SA, an inferior candidate solution may also be accepted, but this happens with a certain probability. This probability is decreasing in the difference in profit between the two solutions and it also decreases as the algorithm proceeds. In terms of the SA algorithm, this probability depends on the so-called temperature of the system, which decreases as the algorithm progresses. By allowing for the acceptance of inferior solutions, the algorithm lowers the probability of becoming trapped at local minima. At the end the final solution is the best candidate solution found during the progress of the algorithm.

The Simulated Annealing algorithm amounts to a pair of nested loops. The outer loop controls the acceptance probability of inferior candidate solutions and the inner loop considers a fixed number of candidate solutions. The way in which the temperature (acceptance probability) is decreased is known as the cooling schedule. A commonly used cooling schedule is the proportional cooling schedule ($T_{\text{new}} = rT_{\text{old}}$) where $r < 1$. For a maximization problem, the Simulated Annealing algorithm in pseudo-code is displayed in Figure 1. We refer to Johnson et al. (1989) for a more detailed description of SA.

1. Get an initial shelf layout W. Set $W^{\text{best}} = W$.
2. Get an initial temperature $T > 0$
3. While not yet frozen do
   (a) Perform the following loop Q times
      i. Pick a random neighbor W' of W.
      ii. If Profit(W') > Profit(W^{best}) then set $W^{\text{best}} = W$.
      iii. Let Δ = Profit(W') - Profit(W).
      iv. if $\Delta \geq 0$ then $W = W'$.
      v. if $\Delta < 0$ then set $W = W'$ with probability $\exp(\Delta/T)$.
   (b) Set $T = rT$ (reduce temperature).
4. Return W^{best}.

Figure 1: Simulated annealing for profit maximization problem

Generally, a difficult aspect with Simulated Annealing is to determine how many candidate solutions to consider at each temperature (Q). In theory one could reduce Q as the temperature drops. In practice, the balance between the maximum step size and the number of Monte Carlo steps is important. Both depend very much on the characteristics of the search space. In our application, we will choose a small step size and keep Q constant, as described below.

For our shelf optimization problem, we let the SA algorithm start at the best of many randomly generated layouts. We generate a preset number of layouts at random and choose the one that has the highest predicted profit as the starting point. One can generate as many layouts as desired. This makes it (even) less likely for the algorithm to get trapped in local optima. To gain insight in good starting temperatures, cooling schedules and a good search length Q, we experiment with various settings. As long as the settings are not such that the optimization terminates very quickly, there appear to be only small differences among settings. Therefore, we choose to report those that generate the best profit in a reasonable computation time. We use a value of 0.8 for r and 50 for Q.

In the search for a neighborhood solution W', we employ two methods. The first method generates a new layout by interchanging two randomly chosen SKUs as far as their shelf height and position on that shelf are concerned. The number of facings for each SKU is then adapted upwards or downwards according to the space available in the new location. The second method randomly selects a shelf and on this shelf it randomly selects two SKUs. If feasible, the first SKU loses one facing while the other gains one. If this does not work, for instance when the first SKU is already at the minimum number of facings allowed, the other way around is tried, that is, the first SKU gains one facing while the other loses one. If this is also not feasible, a new shelf and a new set of items are randomly drawn. As items may have different package widths, an extra check here is needed to make sure the items still fit on the shelf. If not, the gaining item loses its extra facing again. By searching the space in this way, we use the smallest step size available.

Larger steps would involve interchanging several items at once, or using larger facing increases and decreases. Although computation time increases, we prefer small steps, as it prevents missing out on potentially promising solutions.

5 Illustration

To illustrate our method we present a detailed analysis of an interesting and extensive data set. In Section 5.1 we briefly describe the data. Section 5.2 concerns the estimation results and the forecasting performance of our model. In Section 5.3, we illustrate how our model can be used to optimize the shelf layout in each of the stores in our data set.

5.1 Data description

The data analyzed in this paper is a scanner data set with the sales levels of canned soup. The data concerns one of the categories studied by Drèze et al. (1994). The experiments in this study were carried out at Dominick's Finer Foods, a leading supermarket chain in Chicago. Sixty stores participated in the tests, where each store was randomly assigned to a control or test condition. There were two test conditions – "space-to-movement", where the shelf sets are customized based on store-specific movement patterns, and "product reorganization", where product placement is manipulated to facilitate cross-category merchandizing or ease-of-shopping. We choose to analyze the canned soup category as this category has a large number of items and shows relatively frequent price changes. Furthermore, this category has large variation in shelf layout since one of the test conditions was to alphabetize the items on the shelf.

In our analysis, we only look at stores that have data in the test condition. We have 36,044 observations for 407 canned soup SKUs, for five randomly selected test stores. Three stores carry 81 each and two carry 82 each of these items. There may be overlap between these items, but we treat them separately, as items in different stores will differ in their position on the shelf and even if an item would have the exact same location, it is unlikely that it will have the same demand and elasticity parameters. For each SKU we

have around 100 weeks of observations. As explanatory variables in the sales equation, we use an intercept, price, and a promotion variable, which is a combination of the variables bonus-buy and display available in the database.

The unique feature of this dataset is that we have information on a number of shelf characteristics and item attributes. Several of these variables appear to correlate strongly with each other. After iteratively removing halves of the pairs that correlate most, ten attributes remain. We list these attributes in Table 2. We include both facings and ln(facings) to model the potentially diminishing effects of the number of facings on marketing instrument effectiveness. If the available data would allow this, one can extend this list with additional item characteristics, such as flavor, type (condensed or not) and package type (for example, Easy Open lid or not). These variables could contribute to the explanatory power of our model and yield additional insights.

Table 2: Available variables in attributes equation. More variables were available in the dataset, but removed due to too much correlation with shown variables.

Variable	**Description**
Facings	
facings	Number of facings on shelf in units
ln(facings)	Log of number of facings on shelf in units [a]
Vertical measures	
Shelf number	Shelf number (1 being the bottom shelf)
Vertical distance to middle	Distance of shelf variable to the middle shelf
Horizontal measures	
Distance to shelf end (racetrack)	Distance to the beginning of the shelf, measured in inches
Horizontal distance to middle	Distance of item to the middle of the shelf, measured in inches
Capacity	
Depth	Depth of shelf in units
Stack	The number of items that are stacked on top of each other
Item characteristics	
Item width	Width of item in inches
Campbell	Item is from the Campbell brand (1=yes, 0=no)

[a] Included to model diminishing effects

5.2 Estimation and forecast results

In this subsection, we discuss the estimation results for our model, and report the out-of-sample forecast performance.

5.2.1 Estimation

As explanatory variables in the sales equation we use an intercept, log price and promotion. The parameters associated with these variables are each item- and time-specific. In the model, changes in the parameters across time for a specific item are completely attributed to changes in its characteristics. Obviously, differences across items can only partly be explained by differences in characteristics. The random component in (2) allows for unexplained differences in the parameters across items. Shrinkage estimation in the hierarchical structure allows us to estimate these parameters with sufficient accuracy. For the estimation of the parameters, we generate 20,000 iterations of the Gibbs sampler for burn in and 20,000 iterations for analysis, where we retain every tenth draw to reduce the effects of autocorrelation between consecutive draws. The (unreported) iteration plots are inspected to see whether the sampler has converged.

The marketing effectiveness parameter $\beta_{i,t}$ varies across items and time. Even though the values do not change each and every period, there are obviously too many values to display in a table. A histogram per marketing instrument, as given in Figure 2, summarizes the dimensions 'item' and 'time' in an insightful way. The number of observations that constitute the histogram, is equal to $\sum_{i=1}^{I} T_i = 36,044$. We see the expected signs for each of the three explanatory variables. The intercept is positive for all observations. The price effect is negative for most periods and items. Finally, the promotion variable has the expected positive effect for most items and periods. From this figure, it may be hard to see what the actual expected β parameter values are for all items in the data set. We list the posterior mean of the average β over all items and time periods ($\frac{1}{I}\sum_i \frac{1}{T_i}\sum_t \beta_{it}$) in Table 3. From this table it is clear that over all items and time periods the parameters have the expected signs.

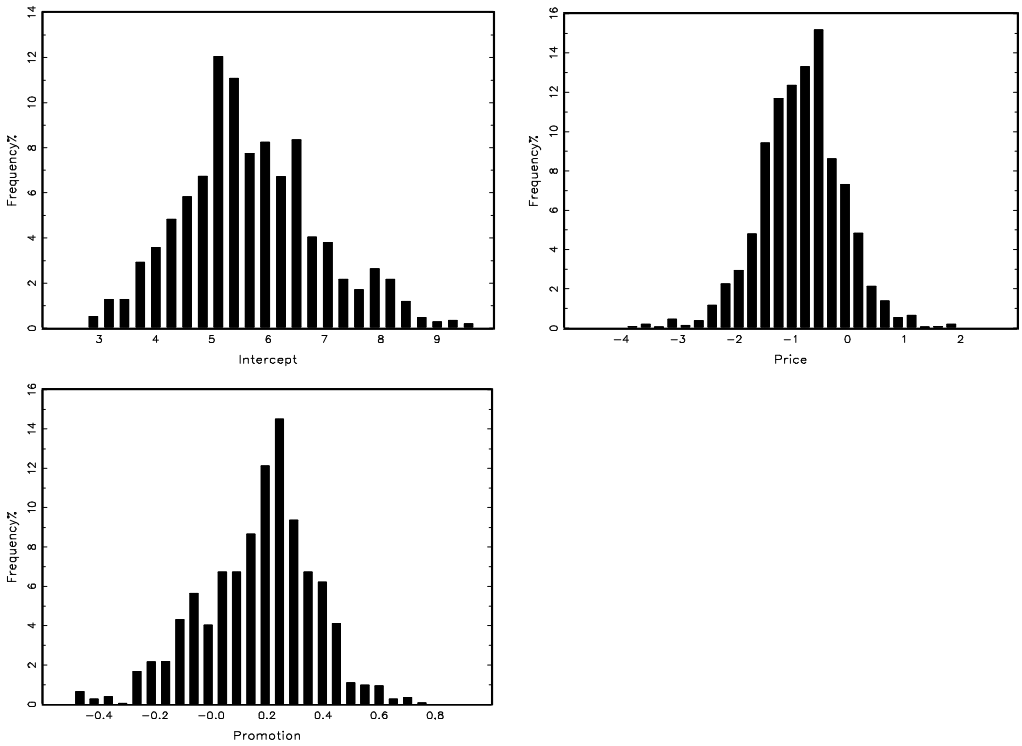

Figure 2: Histogram per marketing instrument across all $\beta_{i,t}$.

Table 4 shows the posterior means for γ, that is, the parameters linking the attributes to the effectiveness of own marketing instruments. We will investigate these estimates both with the numbers in the table and with graphs for the attributes that appear in a non-linear fashion in the model, i.e., facings, shelf and distance to shelf end. From the numerical estimates, it can be seen that the logarithm of the number of facings has a positive influence on the intercept in the sales equation (0.513). This indicates that items that have many facings, have a higher expected sales level when there is no promotion, feature activity, or otherwise. This is the effect that was studied in the previous literature on shelf management.[2] The number of facings appears to make the price effect stronger.

[2]In practice, there could be some feedback effects working here as well. An item with a high sales level may be granted more shelf space in the store, and thus result in more facings for the item. This feedback effect is not analyzed in our model, since with the experimental data, it is expected to be less of a concern, see Corstjens and Doyle (1981) and Bultez and Naert (1988), where the same assumption is made.

Table 3: Posterior means (and standard deviations) for $\frac{1}{I}\sum_i \frac{1}{T_i}\sum_t \beta_{it}$

Variable	**Mean**	**St. Dev.**
intercept	5.625	0.079
price	-0.813	0.063
promotion	0.164	0.028

Table 4: Posterior means (and standard deviations)[a] for γ. The numbers in the cells reflect the effect of a layout characteristic (left) on a marketing mix instrument (top).

Variable	**Intercept**		**Price**		**Promotion**	
Intercept	9.480***	(1.276)	-1.443	(1.446)	3.265***	(1.002)
Number of facings	-0.038**	(0.017)	0.114***	(0.02)	-0.035*	(0.019)
Log(number of facings)	0.513***	(0.114)	-0.608***	(0.11)	0.080	(0.117)
Shelf number (1,...,5)	0.168***	(0.047)	-0.252***	(0.05)	-0.127***	(0.053)
Vertical distance to middle shelf	-0.092***	(0.030)	0.041	(0.03)	0.054	(0.036)
Horizontal distance to racetrack	-0.003***	(0.001)	0.003***	(0.00)	-0.002***	(0.001)
Horizontal distance to shelf middle	0.009***	(0.001)	-0.009***	(0.00)	-0.002***	(0.001)
Shelf depth in units	-0.230***	(0.062)	0.246***	(0.06)	-0.127**	(0.062)
Stack in units	0.146	(0.091)	-0.222**	(0.10)	-0.061	(0.095)
Item width in inches	-1.892***	(0.418)	0.542	(0.47)	-0.558**	(0.216)
Campbell dummy	1.457***	(0.216)	-0.785***	(0.20)	0.150	(0.095)

[a] *, **, *** Zero not contained in 90%, 95% or 99% highest posterior density region, respectively.

To further investigate the effects of the two facing variables available in the model, we calculate the effects of varying values for the number of facings on the posterior mean of β. The results of these effects are depicted in Figure 3. The first graph shows the effect facings have on β_i's intercept, that is, the direct effect of facings on sales. As discussed above, a higher number of facings causes a higher intercept, which in turn results in more sales. However, this effect levels off as the number of facings increases. Apparently, the

effect of the number of facings on sales exhibits diminishing returns. Next, in the second graph (top right), the effect of facings on the price-parameter is displayed. For a low number of facings, the price elasticity appears to be higher than -0.5. This could be caused by the fact that items that have only 1 or 2 facings may be niche brands, which are purchased by only a few customers who really search for an SKU and are generally speaking relatively price insensitive. However, the price sensitivity increases as the number of facings goes up. Apparently, having more facings makes the SKU more visible, causing an increased awareness of the price level and thus creating more price sensitivity among customers. Again this effect levels off as the number of facings increases further. Finally, the third panel of Figure 3 shows the impact of facings on the promotion effectiveness. Interestingly, items that have more facings, have a slightly weaker promotion effect. For items that are already visible on the shelf, a promotion does not generate much additional attention.

The shelf height also seems to correlate positively with sales. This can be seen from the value of 0.168 for the effect of shelf on the intercept (see Table 4). Just as with facings, we see that a higher value (i.e. a higher shelf location) makes consumers more price sensitive. This is no surprise, as consumers see the prices for higher located products more easily than for those at the bottom shelf. This can also be seen from Figure 4. Even though our variable "distance to middle shelf" allows for a nonlinear impact of shelf height, the effect appears to be pretty much linear. This is in particular true for price where the variable "vertical distance to the middle shelf" was non-significant in Table 4.

The distance to the shelf end has a negative effect on sales, i.e., the further away an item is from the racetrack, the lower the expected sales. The horizontal distance to the middle of the shelf has a small positive effect, so being further from the middle may increase sales. The combined expected effect can be seen in Figure 5. Obviously, being close to the racetrack is optimal. Note however that the price sensitivity is highest for these items, as is promotional sensitivity. Being in the middle may hurt sales, although it makes consumers less price sensitive. Promotion effectiveness decreases as items move

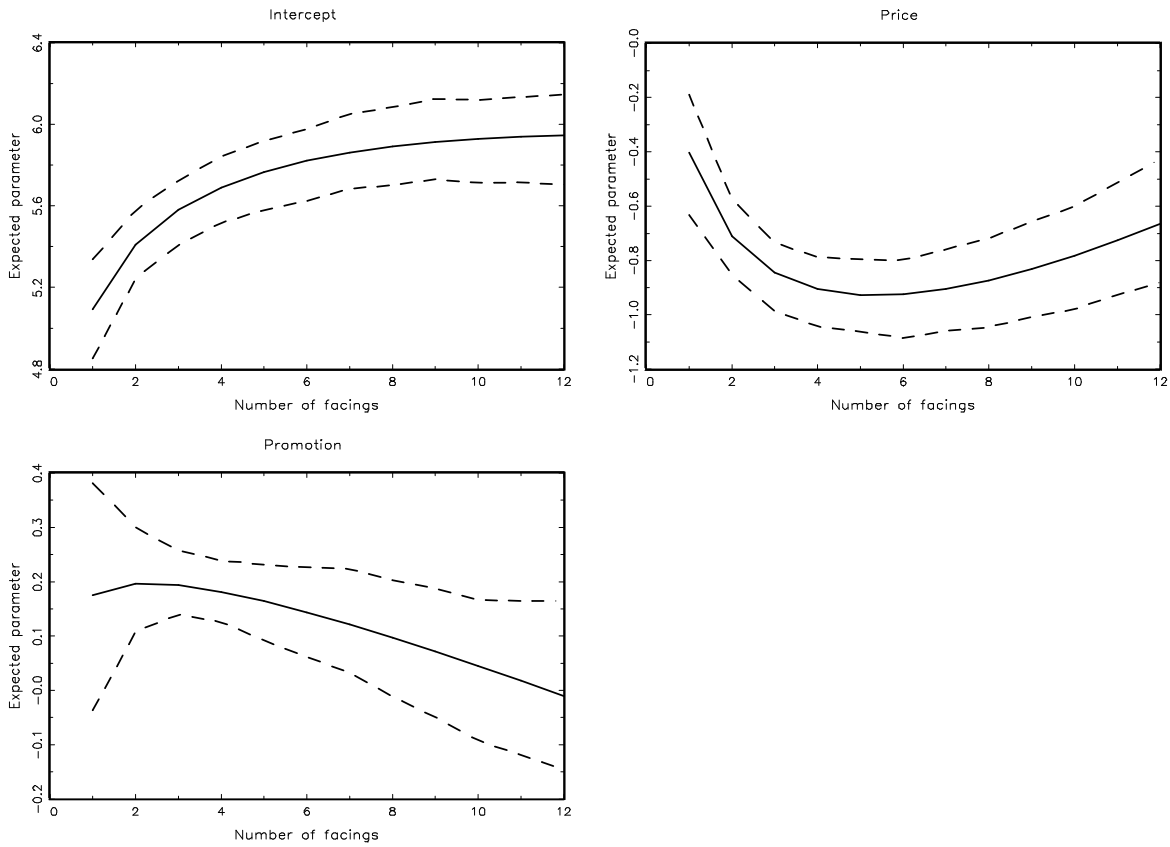

Figure 3: Impact of different number of facings on marketing-effectiveness parameters (dashed lines show 95% HPD).

beyond the middle of the shelf. In Table 4 we see that the more products are stacked on top of each other, the stronger the price effect. The brand Campbell has higher expected sales and consumers appear to be more price sensitive for this brand.

Finally, Table 5 shows the posterior means for Σ_η. From these estimates we conclude that there is quite a large proportion of the differences in baseline sales and price elasticities across the items, that we cannot explain using item and shelf characteristics.

5.2.2 Forecasting

In each run of the Gibbs sampler, we simulate sales forecasts for the last five periods in the data set. These data are not used for parameter estimation. We use the posterior

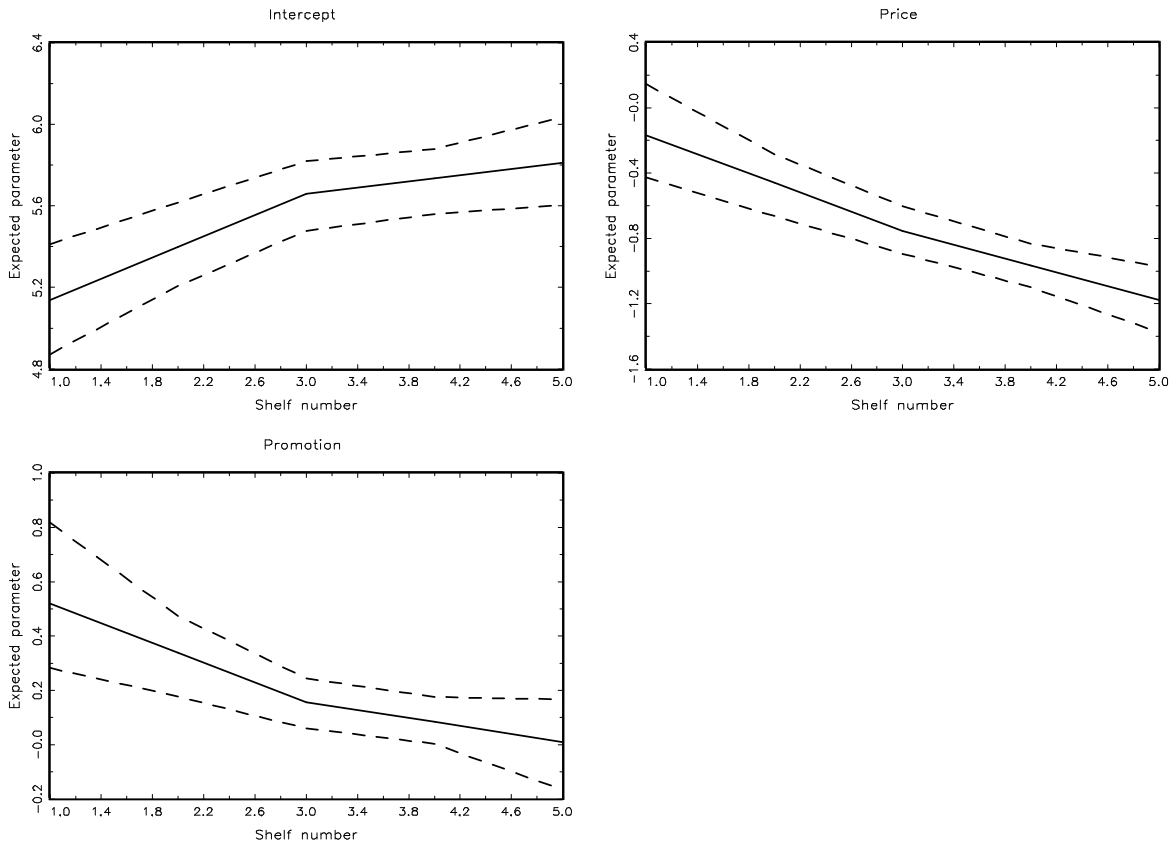

Figure 4: Impact of shelf number (varying from 1 to 5) on marketing-effectiveness parameters (dashed lines show 95% HPD).

mode of these forecasts as the out-of-sample prediction. The correlation between actual and predicted sales equals 88%. When we compare our HB model with a model which concerns a regression per item, we see that our model performs about 4%-points better, both in-sample and out-of-sample. This is most likely due to the extra information used in the attributes and shelf characteristics. The real power of this extra information however amounts to our ability to optimize the shelf arrangement, as we will see in the next subsection.

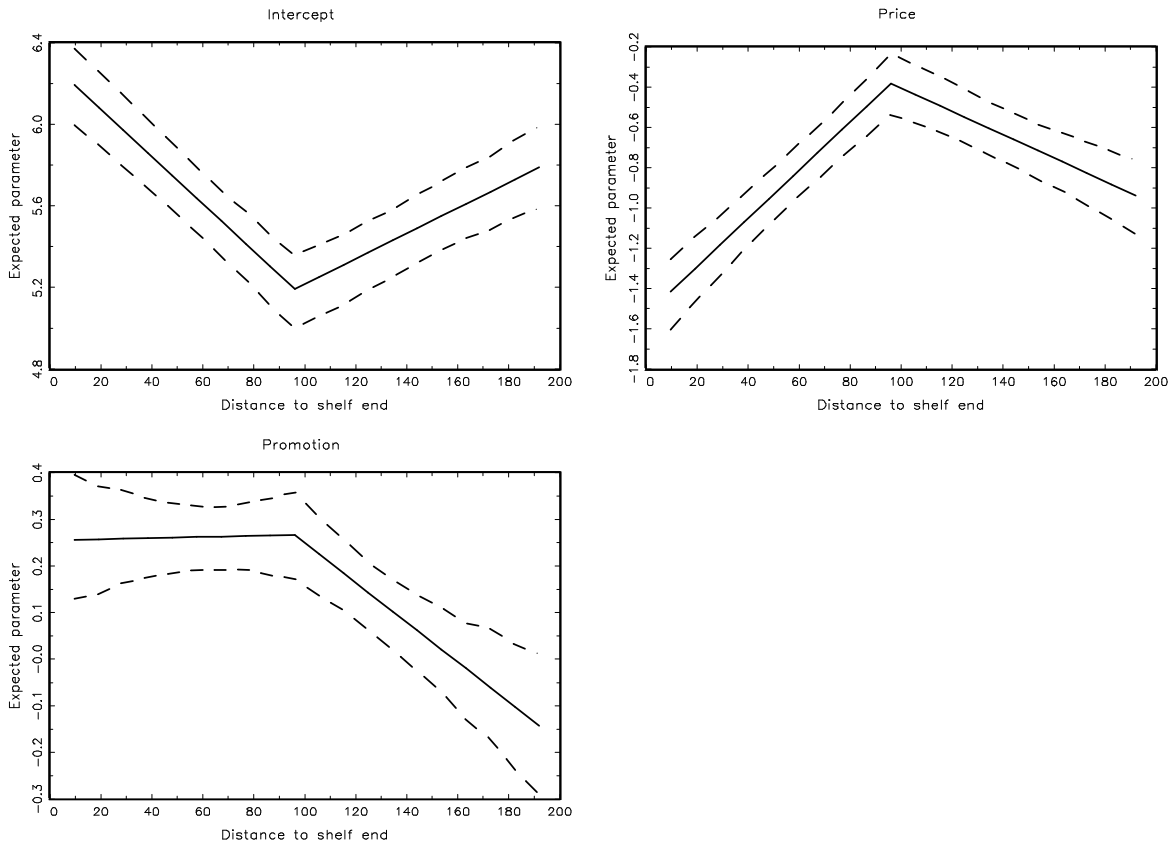

Figure 5: Impact of distance to shelf end on marketing-effectiveness parameters (dashed lines show 95% HPD).

5.3 Shelf optimization

We perform the optimization of the shelf layout for each of the five stores separately. The replenishment costs $c_i(Z_i)$ in equation (4) are currently assumed to be equal to 0.[3] To start up the SA search process, we generate 10,000 random shelf arrangements and have the algorithm start at the arrangement that has the highest mean profit.

Besides our profit-optimization routine, we also compute profits for commonly applied rules of thumb. Since these rules do not completely prescribe the shelf layout, we generate

[3]This setting can easily be changed. Experiments with different settings in the c_i-function showed substantively identical results.

Table 5: Posterior means (and standard deviations)[a] for Σ_η

Mean	**intercept**		**price**		**promotion**	
intercept	1.720***	(0.216)	-0.957***	(0.163)	-0.044	(0.054)
price	-0.957***	(0.163)	0.865***	(0.132)	0.047	(0.032)
promotion	-0.044	(0.054)	0.047	(0.032)	0.076***	(0.012)

[a] *, **, *** Zero not contained in 90%, 95% or 99% highest posterior density region, respectively.

10,000 random layouts based on each rule at hand. The profit for the best of each of these 10,000 layouts is displayed in Table 6. The first is "share-of-shelf equals share-of-log-sales". We have chosen the version of this rule with log-sales, since in practice the large items hardly ever get their share of sales in shelf space, and smaller items usually get more than their sales share justifies. The second rule of thumb is "share-of-shelf equals share-of-margin". Retailers often devote more shelf space to products that have high margins, rather than sell much. This rule does not appear to work very well, as can be seen in Table 6. Varying the position of high-margin items also has its consequences. As can be seen in the table, putting high-margin items close to the beginning of the shelf results in higher profits. The reason for this, as shown by Larson et al. (2005), is that shoppers do not always travel the entire aisle. In fact, once they enter an aisle, shoppers rarely make it to the other end. Instead, they travel by short excursions into and out of the aisle rather than traversing its entire length. This may lead them to purchase more from the beginning of the aisle than from the middle.

For the five stores in our data set, the SA algorithm manages to find profit increases relative to the current situation, that is, profit increases vary from 10% to 15%. Furthermore, the SA algorithm performs better than the rules of thumb described above, with increases ranging from 6% to 20%. Note that while the rules-of-thumb did not lead to an increase in profit for stores 1 and 4, our optimization method does succeed in finding better shelf layouts. When inspecting the optimized shelves for all stores, we find that

items that gain profit do not necessarily have an increased number of facings. It may also happen that it is put on a different shelf, closer to the racetrack, or a combination of these things. In Figure 6 we display the number of facings an item had before and after optimization.

We expect the reported profit increases to decrease when more retailer-specific restrictions are built into the model. Possible restrictions would be to put all Campbell soup cans in the same area, or to have the private label at eye-height. Our optimization algorithm can easily cope with these restrictions, by not considering neighboring layouts that violate these restrictions.

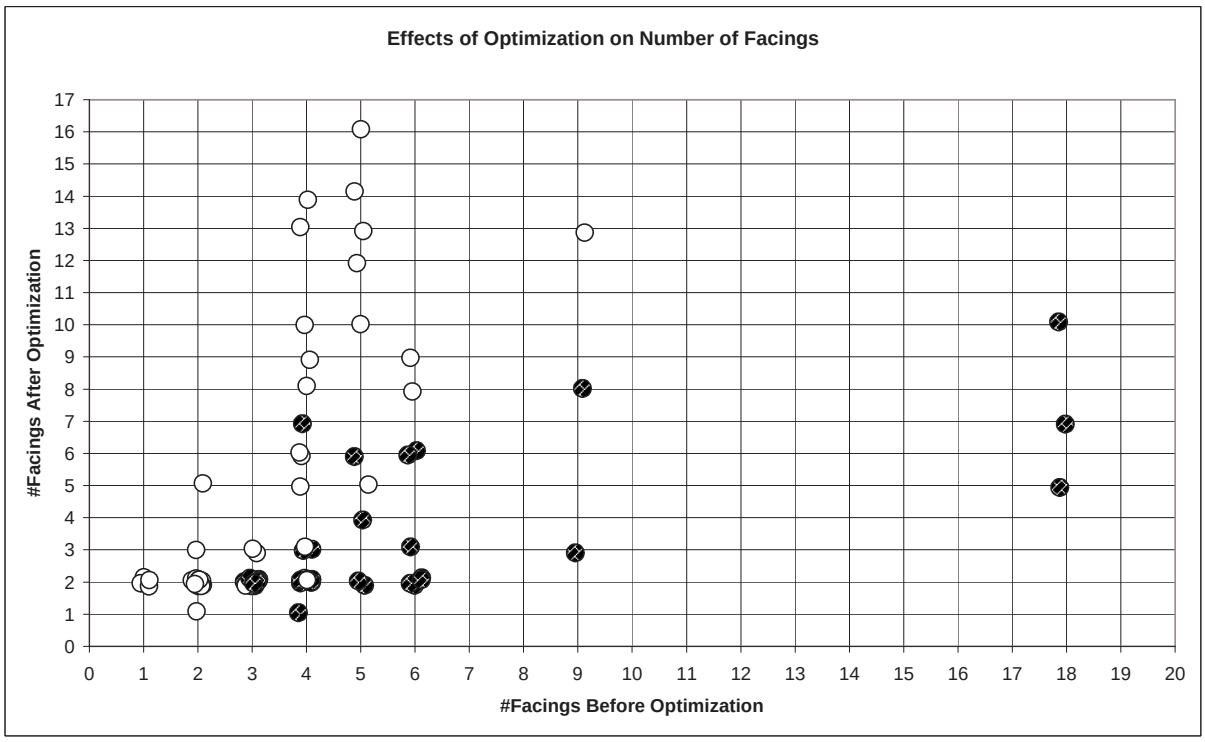

Figure 6: Number of facings before and after optimization for store 5. Black bubbles reflect items that have reduced profit after optimization, white bubbles are used for items that have increased profit after optimization.

Table 6: Profit results for current layout, various rules of thumb and optimization algorithm.

Layout	**Resulting maximum profit**				
	Store 1	**Store 2**	**Store 3**	**Store 4**	**Store 5**
Current layout[a]	\$ 3,410	\$ 2,340	\$ 1,713	\$ 3,114	\$ 2,442
Rule of thumb					
Share of shelf = share of sales[b]	\$ 3,281	\$ 2,488	\$ 1,726	\$ 2,954	\$ 2,512
Share of shelf = share of margin[b]	\$ 3,202	\$ 2,414	\$ 1,624	\$ 2,922	\$ 2,342
Put high margin items close to racetrack[c]	\$ 3,347	\$ 2,548	\$ 1,763	\$ 2,874	\$ 2,449
Put high margin items far from racetrack[c]	\$ 2,980	\$ 2,228	\$ 1,677	\$ 2,772	\$ 2,330
Optimization					
Optimized layout[d]	\$ 3,742	\$ 2,696	\$ 1,943	\$ 3,557	\$ 2,792
Improvement over current layout	10%	15%	13%	14%	14%
Improvement over best rule of thumb	12%	6%	10%	20%	11%

[a] Profit based on predicted sales (not actual).
[b] Achieved by generating 10,000 random layouts, where an item gets devoted the share of shelf space based on sales or margins. The profit for the best of each of these 10,000 layouts is displayed.
[c] Same as "Share of shelf = share of margin", in addition, high-margin items are put close to beginning or end of shelf (where the beginning is the back-isle or "racetrack").
[d] The profit for the optimized layout results after running our simulated annealing algorithm.

6 Conclusion and further research

In this paper we have presented a new approach to optimize shelf arrangements. By introducing shelf characteristics in an Hierarchical Bayes fashion into a sales model, we were able to model the direct effect of the shelf layout on sales as well as the moderating effect of the layout on the marketing instrument effects. After estimating the model parameters on experimental data, we found that the shelf layout has significant effects on baseline sales and marketing effectiveness. This not only holds for the number of facings allocated to an item, but also other shelf descriptors such as shelf height and distance to the end of the aisle. Our HB-setup allowed for interesting (graphical) insights into the effects of shelf-layout on often-used marketing instruments such as price and promotion. Our managerial implications derived from these graphs are the following. As expected,

an SKU with more facings has higher sales. However, the additional benefits of one extra facing does decrease. Also, price-effects appear to be weaker for niche SKUs, i.e. items with few facings. Finally, the results show that promotion effects are weak when products have more facings or are located further away from the racetrack. These are implications that could not be derived from previously proposed sales and shelf management models.

Furthermore, our approach allowed us to optimize the shelf arrangement by cleverly searching through the huge dimensions of the search space that a reasonably large sized category provides. The Simulated Annealing algorithm managed to find increases in profits for all stores in our data set. Optimized profits were also higher when compared to several rules of thumb. Most sensible rules of thumb also increase profits when compared to the current layout. It helps in particular to give high-margin items more shelf space and stock them closer to the beginning of the aisle (or the racetrack). Our optimization technique allows for the identification of high-potential SKUs that could give more profit to the retailer when put on the proper location on the shelf.

We provide several directions for further research. It would be interesting to analyze more stores for the current category and also other categories. Furthermore, if there is sufficient variation in the observed data, one could combine shelf optimization with price optimization. At any rate, we like to see our model as a useful tool in analyzing the effects of shelf layout on marketing instrument effectiveness, optimizing the shelf layout, and determining the value of SKUs to the retailer.

Appendix: Parameter estimation

This appendix describes the algorithm for sampling from a Markov Chain that has the posterior distribution of the model parameters as its stationary distribution, see (Tierney, 1994; Casella and George, 1992). In particular we use the Gibbs sampling technique of Geman and Geman (1984) with data augmentation, see Tanner and Wong (1987). The latent variables $\eta_i,\ i = 1, \ldots, I$ are sampled alongside with the model parameters. In our model we define η_i as the latent variable of interest instead of $\beta_{i,t}$ as the changes in $\beta_{i,t}$ over time are deterministic.

The likelihood function corresponding to the model in (1) and (2) equals

$$L(\text{data}|\theta) = \prod_{i=1}^{I} \int_{\eta_i} \prod_{t=1}^{T} \phi(\varepsilon_{i,t}(\gamma, \eta_i); 0, \sigma_i^2)\phi(\eta_i; 0, \Sigma_\eta) d\eta_i, \tag{A.1}$$

where $\theta = (\text{vec}(\gamma)', \sigma_1^2, \ldots, \sigma_I^2, \text{vec}(\Sigma_\eta)')$ is the vector of all model parameters and

$$\varepsilon_{i,t}(\gamma, \eta_i) = \ln S_{i,t} - X'_{i,t}(\gamma Z_{i,t} + \eta_i). \tag{A.2}$$

We impose flat priors on all parameters but the covariance of η_i. For this covariance we use an inverted Wishart prior. The full prior distribution equals

$$p(\gamma, \sigma_1^2, \ldots, \sigma_I^2, \Sigma_\eta) \propto \prod_{i=1}^{I} \sigma_i^{-2} \times f(\Sigma_\eta; \lambda, S), \tag{A.3}$$

where $f(\Sigma; \lambda, S)$ is the density function of an inverted Wishart distribution with λ degrees of freedom and scale parameter S evaluated at Σ. Although the influence of this prior on the posterior distribution is only marginal, the performance of the MCMC chain is significantly improved by imposing the inverted Wishart prior, see Hobert and Casella (1996).

Sampling of γ

After combining (1) and (2) and stacking the equations over t we obtain

$$\ln S_i = X_i^{*\prime} \begin{pmatrix} \gamma Z_{i1} \\ \gamma Z_{i2} \\ \vdots \\ \gamma Z_{i,T_i} \end{pmatrix} + X'_i \eta_i + \varepsilon_i, \tag{A.4}$$

where $\ln S_i = (\ln S_{i1}, \ldots, \ln S_{i,T_i})'$, $X_i = (X_{i1}, \ldots, X_{i,T_i})$, $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{i,T_i})'$ and

$$X_i^* = \begin{pmatrix} X_{i1} & & & \\ & X_{i2} & & \\ & & \ddots & \\ & & & X_{i,T_i} \end{pmatrix}. \tag{A.5}$$

After some manipulations (A.4) becomes

$$\ln S_i - X_i'\eta_i = X_i^{*\prime}(Z_i' \otimes \mathbf{I}_{K+1})\text{vec}(\gamma) + \varepsilon_i, \tag{A.6}$$

where $Z_i = (Z_{i1}, \ldots, Z_{i,T_i})$, and where $\mathbf{I}_m$ denotes a m-dimensional identity matrix. This equation can compactly be written as

$$W_i = V_i\text{vec}(\gamma) + \varepsilon_i, \tag{A.7}$$

where $\varepsilon_i \sim N(0, \sigma_i^2\mathbf{I}_{T_i})$. From (A.7) it is easy to derive that the full conditional posterior distribution of vec(γ) is normal with mean

$$\left(\sum_{i=1}^{I} \frac{1}{\sigma_i^2} V_i'V_i\right)^{-1} \left(\sum_{i=1}^{I} \frac{1}{\sigma_i^2} V_i'W_i\right), \tag{A.8}$$

and variance

$$\left(\sum_{i=1}^{I} \frac{1}{\sigma_i^2} V_i'V_i\right)^{-1}, \tag{A.9}$$

see, for example, Zellner (1971, Chapter III).

Sampling of $\boldsymbol{\eta_i}$

The relevant equations for sampling η_i for $i = 1, \ldots, I$ are

$$\begin{aligned} \frac{1}{\sigma_i}\left[\ln S_i - X_i^{*\prime}\text{vec}(\gamma Z_i)\right] &= \frac{1}{\sigma_i} X_i'\eta_i + \frac{1}{\sigma_i}\varepsilon_i \\ 0 &= \Sigma_\eta^{-1/2}\eta_i + \nu_i, \end{aligned} \tag{A.10}$$

where $\nu_i \sim N(0, \mathbf{I}_{K+1})$. The second line in (A.10) represents the second layer of our model. We see that the full conditional posterior distribution of η_i conditional on γ and σ_i^2 is normal. Denoting the first equation of (A.10) by $\tilde{Y}_i = \tilde{X}_i\eta_i + \tilde{\varepsilon}_i$, the mean of this posterior distribution is $(\tilde{X}_i'\tilde{X}_i + \Sigma_\eta^{-1})^{-1}\tilde{X}_i'\tilde{Y}_i$ and the variance equals $(\tilde{X}_i'\tilde{X}_i + \Sigma_\eta^{-1})^{-1}$.

Sampling of $\boldsymbol{\sigma}_i^2$

Conditional on the data and the other parameters, σ_i^2 has an inverted Gamma-2 distribution with scale parameter $\sum_{t=1}^{T_i} \varepsilon_{it}(\gamma, \eta_i)^2$ and degrees of freedom T_i. To sample σ_i^2 we use that

$$\frac{\sum_{t=1}^{T_i} \varepsilon_{it}(\gamma, \eta_i)^2}{\sigma_i^2} \sim \chi^2(T_i), \tag{A.11}$$

where $\varepsilon_{i,t}$ is given in (A.2).

Sampling of $\boldsymbol{\Sigma_\eta}$

Conditional on the other parameters, the covariance matrix Σ_η can be sampled from an inverted Wishart distribution with scale parameter $\sum_{i=1}^{I} \eta_i \eta_i' + S$ and degrees of freedom $I + \lambda$.

References

Blattberg, Robert C., Edward I. George, 1991. Shrinkage Estimation of Price and Promotional Elasticities: Seemingly Unrelated Equations. *Journal of the American Statistical Association* **86** 304–315.

Boatwright, Peter, Robert E. McCulloch, Peter E. Rossi, 1999. Account-Level Modeling for Trade Promotion: An Application of a Constrained Parameter Hierarchical Model. *Journal of the American Statistical Association* **94** 1063–1073.

Borin, Norm, Paul W. Farris, 1995. A Sensitivity Analysis of Retailer Shelf Management Models. *Journal of Retailing* **71**(2) 153–171.

Borin, Norm, Paul W. Farris, James R. Freeland, 1994. A Model for Determining Retail Product Category Assortment and Shelf Space Allocation. *Decision Sciences* **25**(3) 359–384.

Brown, W., W. T. Tucker, 1961. The Marketing Center: Vanishing Shelf Space. *Atlanta Economic Review* **11**(10) 9–13.

Bultez, Alain, Philippe Naert, 1988. SH.A.R.P.: Shelf Allocation for Retailers' Profit. *Marketing Science* **7**(3) 211–231.

Casella, George, Edward I. George, 1992. Explaining the Gibbs Sampler. *American Statistician* **46** 167–174.

Corstjens, M., P. Doyle, 1981. A Model for Optimizing Retail Space Allocations. *Management Science* **27** 822–833.

——, 1983. A Dynamic Model for Strategically Allocating Retail Space. *Journal of the Operational Research Society* **34** 943–951.

Cox, K. K., 1970. The Effect of Shelf Space Upon Sales of Branded Products. *Journal of Marketing Research* **7**(1) 55–58.

Curhan, R., 1972. The Relationship between Shelf Space and Unit Sales. *Journal of Marketing Research* **9** 406–412.

Drèze, Xavier, Stephen J. Hoch, Mary E. Purk, 1994. Shelf Management and Space Elasticity. *Journal of Retailing* **70**(4) 301–326.

Geman, S., D. Geman, 1984. Stochastic Relaxations, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transaction on Pattern Analysis and Machine Intelligence* **6** 721–741.

Hobert, James P., George Casella, 1996. The Effect of Improper Priors on Gibbs Sampling in Hierarchical Linear Mixed Models. *Journal of the American Statistical Association* **91**(436) 1461–1473.

Johnson, David S., Cecilia R. Aragon, Lyle A. McGeoch, Catherine Schevon, 1989. Optimization By Simulated Annealing - An Experimental Evaluation. *Operations Research* **37**(6) 865–892.

Kirkpatrick, S., C.D. Gelatt, M.P. Vecchi, 1983. Optimization by Simulated Annealing. *Science* **220** 671–680.

Larson, Jeffrey S., Eric Bradlow, Peter Fader, 2005. An Exploratory Look at Supermarket Shopping Paths. *International Journal of Research in Marketing, forthcoming, http://ssrn.com/abstract=723821* .

Lim, Andrew, Brian Rodrigues, Xingwen Zhang, 2004. Metaheuristics with Local Search Techniques for Retail Shelf-Space Optimization. *Management Science* **50**(1) 117–131.

Montgomery, Alan L., 1997. Creating Micro-Marketing Pricing Strategies Using Supermarket Scanner Data. *Marketing Science* **16** 315–337.

Tanner, Martin A., Wing Hung Wong, 1987. The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association* **82** 528–550.

Tierney, Luke, 1994. Markov Chains for Exploring Posterior Distributions. *Annals of Statistics* **22** 1701–1762.

Wittink, Dick R., Michael J. Addona, William J. Hawkes, John C. Porter, 1988. SCAN-PRO: The Estimation, Validation and Use of Promotional Effects Based on Scanner Data. Technical report, Cornell University Johnson Graduate School of Management.

Yang, Ming-Hsien, 2001. An Efficient Algorithm to Allocate Shelf Space. *European Journal of Operational Research* **131** 107–118.

Yang, Ming-Hsien, Wen-Cher Chen, 1999. A Study on Shelf Space Allocation and Management. *International Journal of Production Economics* **60-61** 309–317.

Zellner, Arnold, 1971. *An Introduction to Bayesian Inference in Econometrics*. Wiley, New York.

Publications in the Report Series Research* in Management

ERIM Research Program: "Marketing"

2006

The Effect of Feedback and Learning on DSS Evaluations
Ujwal Kayande, Arnaud de Bruyn, Gary Lilien, Arvind Rangaswamy and Gerrit Van Bruggen
ERS-2006-001-MKT
http://hdl.handle.net/1765/7302

Interaction Between Shelf Layout and Marketing Effectiveness and Its Impact On Optimizing Shelf Arrangements
Erjen van Nierop, Dennis Fok and Philip Hans Franses
ERS-2006-013-MKT

---

* A complete overview of the ERIM Report Series Research in Management:
https://ep.eur.nl/handle/1765/1

ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Marketing
F&A Finance and Accounting
STR Strategy and Entrepreneurship