Analytics in Retail

Thanks to the abundant consumer transaction data collected and ample cheap computing power, retailers have begun to develop and employ analytical methods as decision support tools for their various operation and management tasks. They either hire analytically trained individuals or specialized retailing consulting firms to develop their proprietary software for that purpose. In either case, they turn to the wealth of academic models for ideas of developing their own practical models. Generally though, there are differences between the academic and practical models. Sometimes, academic models are overly simplified such that a lot of challenging issues are left unsolved; sometimes, academic models are too complicated such that it would be too hard or costly to implement them in practice. In this article, we focus on two types of commonly used analytics in retail practice, forecasting and optimization; and present two models, motivated by academic research, which are in production use or development in retail business settings.

1 Forecasting

Forecasting is one of the earliest and most common analytics carried out in retail practice. It provides vital input to almost all of the other analytics in retail functions such as marketing, merchandising and operation, and for management tasks such as planning, budgeting and controlling. Its methods can be roughly divided into judgmental and statistical methods. Judgmental methods prevailed in the early days of retailing, when there were only limited products as well as scarce recorded data. These methods are subjective and are nearly pure art with little science. They are still useful tools for forecasting sales in some cases such as innovative products (e.g. the iPad) or fashion goods, which have no direct relevant historical sales. And expert judgment can convey valuable knowledge and experience which can be used to improve the performance of statistical methods (Bunn & Wright, 1991). But statistical methods are the basic tools for modern forecasting system in retail. They make use of the large-scale data collected through an IT system. They are objective, scientific and can be automated. Statistical forecasting methods can be further divided into extrapolation and causal methods. Techniques of extrapolation methods include the simpler moving average and exponential smoothing family, and the more sophisticated Box-Jenkins approach (Box & Jenkins, 2008). They use only the time series data of the forecasting subject. Franses (Franses, 1998) discusses the application of extrapolation methods for business and economic forecasting. Causal methods, on the other hand, build statistical models using both the data of the forecasting subject and potential causal factors. Some causal factors are under the control of management, such as prices, promotions, advertising; others are not, such as competitor prices, weather, changes in competitive landscape and market demographics. Despite the simplicity of judgemental and extrapolation methods in retail practice, the more complicated causal forecasting methods play an important role in the retail industry and are the focus of this chapter.

In what follows, we use a promotion forecasting example to illustrate these causal forecasting methods. This example employs a multiple linear regression model for forecasting, which is most often used in causal methods. We highlight some of the issues commonly encountered in retail forecasting using causal statistical methods, and an innovative solution for one of them.

1.1 Promotion Forecasting

An important class of decisions made by retailers is related to planning a promotional strategy. These decisions include which products to promote, what promotional prices to offer, how to communicate the promotional offers to customers, either via various media channels or in-store, and how to execute the promotions. Promotion decisions may be made on a chain-wide basis, or may be tailored to specific markets, store types and stores. These decisions can be informed and supported by promotion forecasting, which aims to forecast the magnitude of the impact of a particular promotion on both the promoted products, and products complementary or substitute to the promoted ones. Promotions forecasting methods embedded in decision support have been shown to provide substantial increments in revenue and gross profit. The benefits accrue in three areas. First, retailers can make more informed decisions about promotional plans, including items, timing and targets, subject to a given level of support from vendors. By basing such decisions on better understanding of the profitability of a promotion, the retailer can lessen the risk of unprofitable promotions, and make the most effective use of promotional 'budgets'. Second, the retailer is able to provide higher service levels on the promoted items by improving the forecast accuracy, which translates into fewer lost sales and greater revenue. Finally, better forecasts for promotional items can also improve inventory management and hence reduce the ordering and holding costs.

The forecasting method described in this section features an integrated demand model which captures the sales effects of various promotional features, both in the item being promoted and other related items. Ordinary least square estimates of these effects, calibrated on historical sales and ancillary data, are used first in a promotional planning tool to support choices on promotional activities, and once plans are finalized, in the inventory management tools to plan inventory for items influenced by the planned promotions.

1.2 Forecast Constituents

1.2.1 Secular Effects

Secular effects are predictable effects of phenomena that are time based. While predicting these effects is not the chief goal of the promotion forecasting model, they are important for two reasons. First, in order to estimate the effects of promotional features, it is necessary to "untangle" the effects of the promotion on sales from the effects of these other phenomena. Second, when the retailer is considering the specific details of a promotion under consideration, the ultimate effects of the promotion on sales often depend on the level of sales when the promotion is absent, which in turn depends on these secular effects. Our model incorporates three secular effects: demand level, seasonal effects and holiday effects; it could be expanded to include a trend effect as well. The demand level is interpreted as the average (or deseasonalized) sales under a regular pricing regime. The seasonal effects indicate how sales vary over the time of a year. They are captured at the weekly level in the promotion forecasting model, which works well in practice. We include holiday effects in addition to the weekly seasonal effects because the same holiday may not happen at the same week each year.

1.2.2 Own Effects

Own promotional effects are the influence of the promotional features on the sales of the item being promoted. Any particular promotion is characterized by a vector of promotional features. These features can be categorized into four classes: the discount or temporary price reduction; the mechanism by which the offer is extended (e.g. straight discount, “buy one, get one free”); the way in which the offer is communicated or promoted in the store (e.g. on shelf, aisle end cap); and the way in which the offer is communicated outside the store (e.g. circular front page, brand advertising). For example, a particular promotion may be a 50% discount on shelf promotion with “buy one, get one free” mechanism and circular front page advertising. To forecast the own effects for such a promotion, we just need to combine effects of its feature vector. These effects are combined multiplicatively.

Own effects of a promotion can either be on sales which occur contemporaneously with the promotion, or on sales before or after the promotion. The effect on sales before or after the promotion is sometimes referred to as retiming, self-cannibalization or “pantry loading”. It occurs when consumers either defer purchases until the promotion is in effect, if it can be anticipated, or when consumers stock up on the item during the promotion and purchase less afterwards. Our model incorporates both the contemporaneous own effects and the retiming effects.

1.2.3 Cross-Effects

Cross effects are the influence of the promotion on items other than the one being promoted. Cross effects can happen either for complement or substitute items. Items which are substitutes for the promoted item may experience reduction in sales. These are sometimes referred to as promotional “victims” (e.g., different pack sizes of the same item, other brands of the same item). Other items, which are complements to the promoted item, may experience an increase in sales. These are sometimes referred to as “halo” items (e.g., snacks, when beer is promoted; accessories, when furniture is promoted).

Another type of cross effects arises when a promotion is undertaken to stimulate traffic in the whole store, from which sales of many items may benefit. Some of these items may be neither complement nor substitute to the promoted items. The promoted item is sometimes referred to as a “loss leader” or a “traffic-driver”. A typical example of a loss leader in grocery is milk. Retailers sometimes reduce the price of milk under its cost to attract shoppers to their stores. Our model incorporates both types of cross effects.

1.2.4 Full Model

We model the effects in the promotions forecast multiplicatively. For seasonal and holiday effects, this is merely a choice of convenience, and is consistent with most other work in the area of retail forecasting. For the own and cross-effects, the multiplicative formulation is motivated by the fact that the own and cross-effects clearly depend on the baseline sales of the promoted item and related items. Multiplicative models automatically capture this kind of dependence while additive models cannot.

The full forecast model for item i is given by:

$$
\begin{aligned}
Sales_{i,t} = {} & Level_i \cdot SeasonalEffects_{i,t} \cdot HolidayEffects_{i,t} \\
& \cdot \prod_{k=1}^{K} \beta_{i,k}^{own} \cdot PromoFactor_{k,t} \\
& \cdot \prod_{l=1}^{L} \prod_{j=1}^{J} \beta_{l,j}^{cross_i} \cdot RelatedItemPromoFactor_{j,t}
\end{aligned}
$$

where k indexes the factors representing the promotional features for item *i*, which are being offered at time *t*, and β^{own} are the own effects for each of the K factors; *l* indexes the items being promoted with cross effects on item *i*; *j* indexes the attributes of each of the L other items being promoted at time t and β^{cross} are the cross effects for each of the J factors.

Implementing such a model requires making choices about inclusion of effects and the way in which effects are modeled. We discuss some of these choices with an example.

In this promotions planning example, the promotion under consideration is a branded grocery product, peanut butter. The features of the promotion are a "buy one, get a second for half price" offer, and a large display at the end of the aisle. For simplicity we consider only two related items, a larger package of the same brand and a similar size package of a house brand. The promotion is planned to run for two weeks, corresponding to week indices 13 and 14. The estimated effects are presented in Table 1.

Promotions Forecast Effects	Premium Brand 14 oz	Premium Brand 32 oz	House Brand 15 oz
Demand Level	66	27.5	83.7
Seasonal Effect, Week 13	1.16	1.16	1.16
Seasonal Effect, Week 14	1.16	1.16	1.16
Seasonal Effect, Week 15	1.14	1.14	1.14
20-25% Discount	1.32		
Buy One, Get One Half Off	1.04		
Major End Aisle Display	1.23		
Promo prior week w/display	0.78		
Cross Effects of 14 oz Premium Brand Promo w/Display		0.41	0.68

Table 1: Promotions Forecasting Effects

The first column displays the own effects. The demand level and seasonal effects provide the baseline (i.e. non-promoted) forecast. This promotion offers an effective discount of 25%; we have an estimated effect (demand lift) of 1.32 for discounts in the range of 20-25%. An alternative way to model this effect is to treat the discount as a continuous variable and the effect as price elasticity. An advantage of the discrete discount buckets is that it is non-parametric, offering a more flexible way to model the discount effects. The promotional mechanism, "buy one, get one half off" is estimated to have a lift of 1.04. The fact that this is greater than 1.0 indicates that this form of the offer is more effective than simply offering a 25% discount. The way by which the promotion is communicated in the store, using a large end cap display, is estimated to have a lift effect of 1.23. Finally, we have a retiming effect of 0.78 on the sales of the promoted peanut butter the following week.

For many retailers, there can be a large set of promotional mechanisms and in-store communications, as well as promotional advertising possibilities. For the purposes of modeling and estimating the effects, these can be consolidated into a smaller number of alternatives of each of the three main categories.

The second and third columns in Table 1 display the baseline forecast effects and the cross promotional effects for the other two items in the subcategory. The cross effects predict the sales impacts of a promotion on the 14 oz Premium brand, on each of these other two items, given that the promotion includes a display. In the present case, the 0.41 effect for the 32 oz Premium brand indicates large reduction in sales volume, due to a high degree of substitutability across sizes of identical items. The effect on the sales of the like-sized House brand is less, indicating a lesser degree of substitutability with the promoted item.

As with the own effects, there are a variety of ways to model the cross effects. In this example, we strike a compromise between the simplest model which reflects only the presence or absence of a promotion of a related item; and a much more complex model which attempts to estimate the effects of all of the various promotional attributes on the sales of the related item. Note also that, in this promotions planning example, there are no complementarity cross effects.

1.3 Challenges

1.3.1 Data Preparation

The presented forecast model is based on information from the point-of-sales (POS) and the promotions planning and execution systems. Neither the POS nor the promotions execution systems were designed with a view towards supplying data to support a forecasting model and process, and thus the source data require inspection and cleansing prior to being used. The primary hygiene tasks are to ensure the plausibility of data values and consistency of data across systems. Once these checks are accomplished, the sales data are summed to weekly values, whose start and end dates correspond to the promotions calendar. These values are then further cleansed of outliers and are matched up to the historical promotions information, providing the essential ingredients to estimate the effects.

1.3.2 Typical Problem Sizes

In 2008, the average number of SKUs carried in a typical U.S. supermarket was 46,852 according to the Food Marketing Institute. The number of stores a large retail chain may have is in the order of thousands. For example, Wal-Mart has more than 8,000 retail stores under 53 different banners in 15 countries (Wal-Mart Stores, Inc.).

Promotions planning systems will typically support dozens to hundreds of different promotion features. Localization of promotional planning and execution of some large retailers can increase the large size retail forecasting problem even further. For example, a grocery chain with approximately 1000 stores may have 25-30 different price zones in which there are different promotional effects to be forecast. Although costs of computing are not much of a constraint, especially with the advent of cloud computing, the large size of the problem requires careful design of the automation of the estimation process, which can take many hours to execute. Further, with so many intersections of item, location, and period, the majority of histories used to calibrate the model are sparse. Aggregation and pooling are strategies for facilitating estimate of the effects in the presence of sparsity.

1.3.3 Aggregation and Pooling

In forecasting problems amenable to the extrapolation or time series methods, sparsity is often addressed through aggregation and disaggregation methods (Zotter, Kalchschmidt, & Caniato, 2005), (Kahn, 1998). In promotional forecasting models (and causal models generally), aggregation is problematic, since SKUs which might be aggregated are likely to have histories with different promotion timing and features.

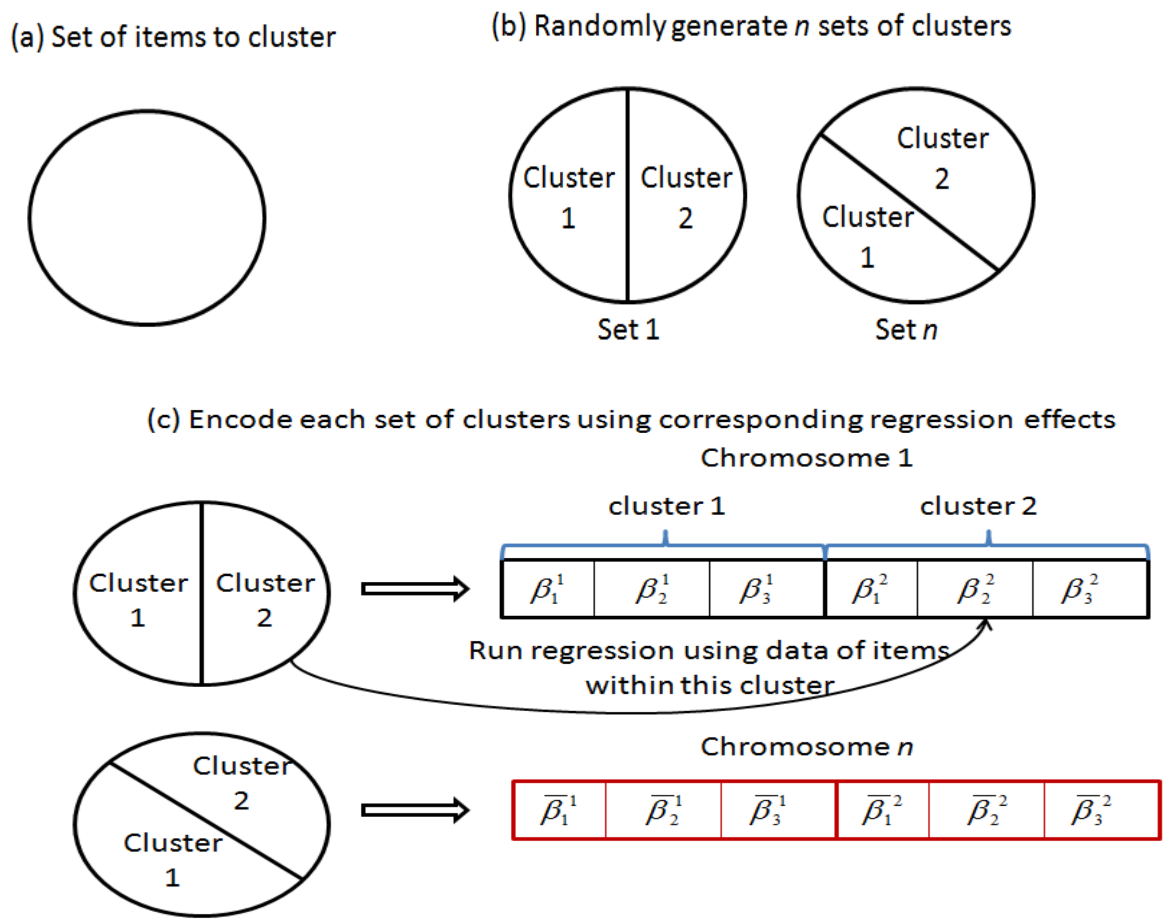
In causal forecasting models, pooling data for different items and different locations can improve the estimability and reliability of effects estimates. The challenge is to devise a means of defining pools or clusters of items that are homogeneous in the effects.

One approach to find item-level clusters is a two-stage method. We first assign each item an attribute vector. After that, we can group items together according to the assigned attribute vector by some clustering algorithm like K-means or hierarchical clustering. For example, Zotter et al. (Zotter, Kalchschmidt, & Caniato, 2005) used normalized store sales to group stores with similar seasonal effects. Products are often grouped according to certain physical or usage attributes, such as product category or performance grade. The problem with this approach is that it is hard, *a priori*, to align the attribute vector with the regression effects across which we want to pool.

1.3.4 KMeans – GA algorithm

An alternative approach not relying on these artificially assigned attribute vectors is cluster-wise linear regression. The basic idea of this modeling approach is to consider clustering and regression concurrently. In this way, the difficulty of aligning the attribute vector and the regression effects is bypassed.

We have developed a heuristic algorithm, termed the KMeans-GA algorithm, which is able to find clusters of similar items efficiently and effectively. This KMeans-GA algorithm hinges on the work by Maulik and Bandyopadhyay (Maulik & Bandyopadhyay, 2000) where the authors embed a K-means procedure into the genetic algorithm framework for the standard clustering problem. Along the same lines, our KMeans-GA algorithm also embeds a structure exploring K-means like procedure into the genetic operations. For our cluster-wise linear regression problem, each cluster of items is encoded by a vector that stores the regression effects. These regression effects define similarity among items, according to which the K-means part assigns items to different clusters. The vector of regression effects is updated by running regression over new clusters generated through the K-means part. The genetic algorithm part performs genetic operations on the population of vectors of regression effects. The genetic operations include the standard selection, crossover, mutation, and elitism. We refer to Figure 1 for a simple example that illustrates the KMeans-GA algorithm. In this example, each cluster i has the same regression model $y = \sum_{j=1}^{3} x_j \beta_j^i$ with different estimates for the regression effects $\beta_j^i{}'s$. We are trying to divide items into 2 groups.



(d) Select 2 chromosomes for genetic operations

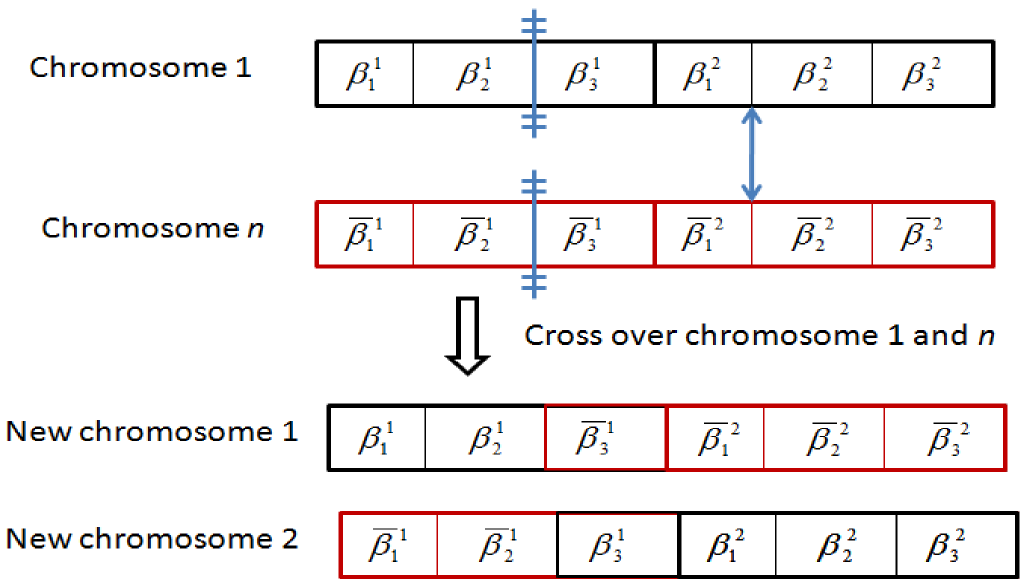
Chromosome 1	β_1^1	β_2^1	β_3^1	β_1^2	β_2^2	β_3^2

Select chromosome 1 with probability $\dfrac{\text{fitness}_1}{\sum_{i=1}^{n} \text{fitness}_i}$

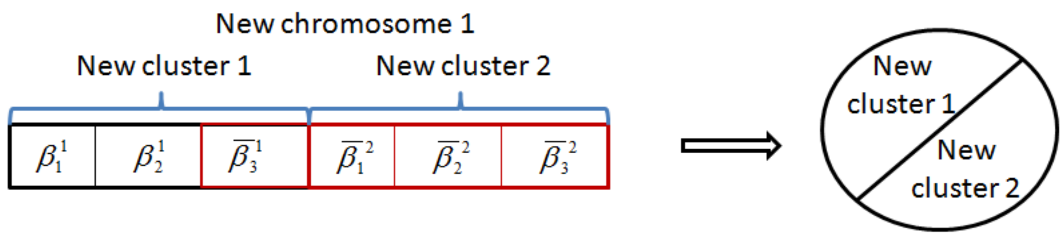
Here fitness_1 is the fitness measure for chromosome 1

$$\text{fitness}_1 = \frac{1}{\sum_{j=1}^{2} \sum_{k=1}^{K} (y_k - \beta_1^j x_{k1} - \beta_2^j x_{k2} - \beta_3^j x_{k3})^2}$$

(e) Cross over in genetic operation



(f) Generate new set of clusters through K-means like procedure



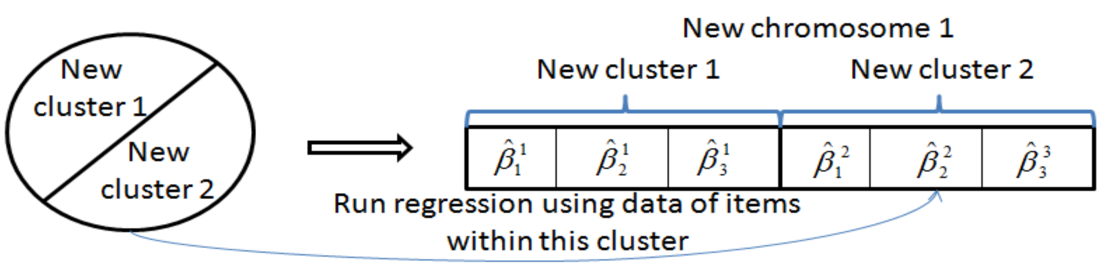
K-means like procedure:
Item A is assigned to new cluster 1 if and only if

$$\sum_{k=1}^{K}(y_k - \beta_1^1 x_{k1} - \beta_2^1 x_{k2} - \bar{\beta}_3^1 x_{k3})^2 \leq \sum_{k=1}^{K}(y_k - \bar{\beta}_1^2 x_{k1} - \bar{\beta}_2^2 x_{k2} - \bar{\beta}_3^2 x_{k3})^2$$

Otherwise, Item A is assigned to new cluster 2.

Here $(y_k, x_{k1}, x_{k2}, x_{k3})_{k=1}^{K}$ is the data for item A.

(g) Update coding for new chromosome



(h) Update population by elitism

Replace chromosome 1 with new chromosome 1 if the latter has larger fitness.

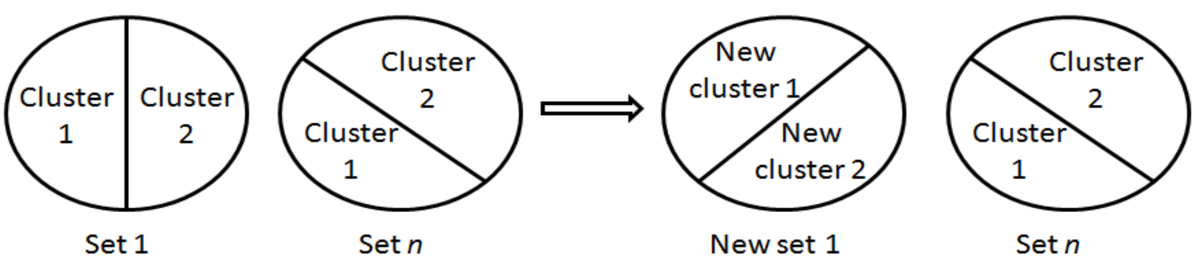


Figure 1 Illustration of the KMeans-GA algorithm

We tested our KMeans-GA algorithm on the promotion forecasting problem to find SKUs with similar seasonality patterns. We found that the heuristic approach performs very well; the difference in total sum of squared regression errors between our heuristic and optimal solutions is below 2%. In addition, the improvement of our KMeans-GA algorithm over the attribute based two-stage method is significant; reductions in SSE of greater than 30%. Finally, in a number of test cases, the procedure revealed evidently better and more granular predictions about seasonal effects than a two-stage method. Figure 2 shows the seasonal effects estimated for a group of SKUs pooled on the basis of an attribute, the subcategory of the product. The only significant seasonal pattern was observed in weeks 51 and 52. For the same set of SKUs, the method revealed pools with appreciably different and more pronounced seasonal effects, as shown in Figure 3. These seasonal patterns were confirmed by retail experts for the SKUs under study.

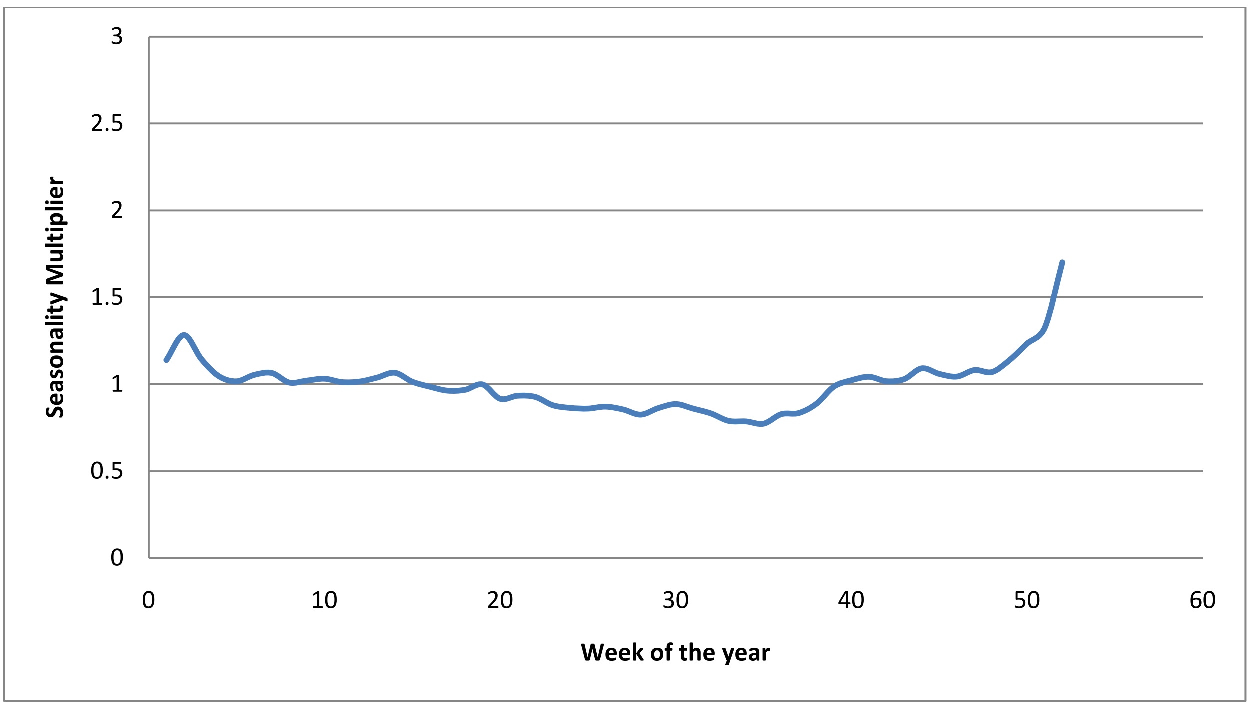


Figure 2: Seasonal pattern for an attribute-based cluster of approximately 350 SKUs

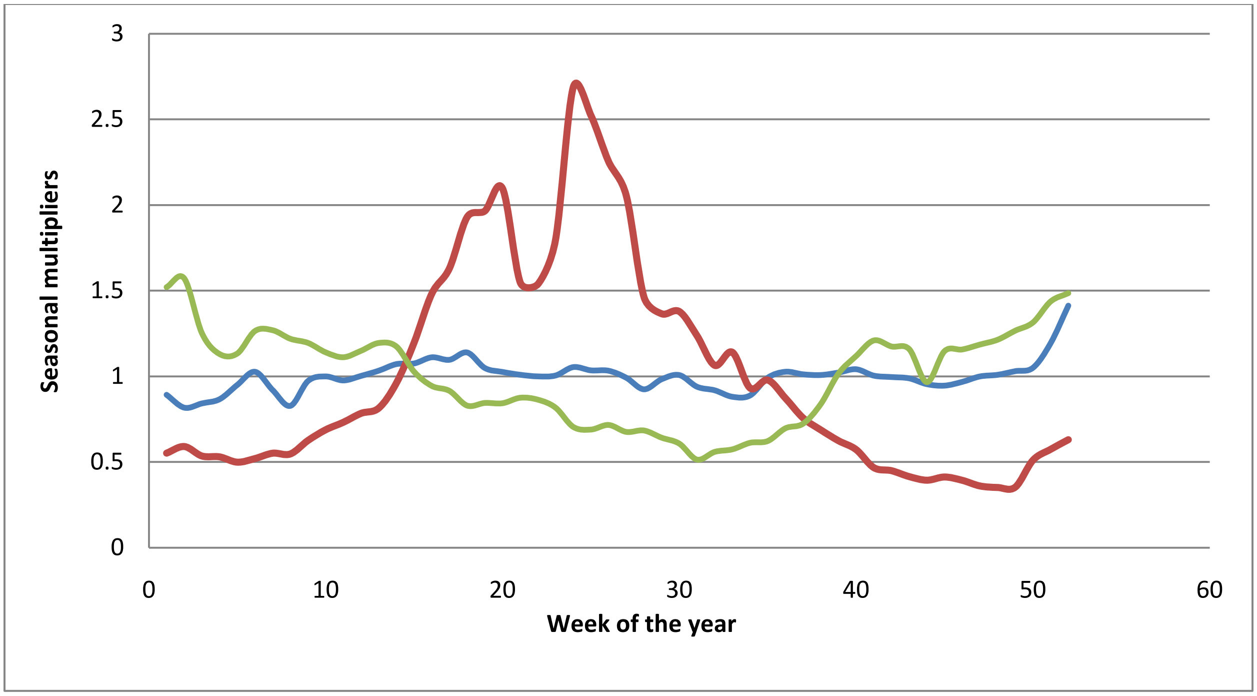


Figure 3: Three distinct seasonal patterns revealed for same SKUs

2 Optimization

Optimization is not as a commonly used tool in retailing as forecasting. Retailers are more comfortable making decisions by asking "what-if" questions than resorting to optimization models. The reason is the lack of the underlying intuition of an optimization model in absence of technical background and confidence. But thanks to recent collaborations between researchers in academia and some leading retailers, selected optimization models have made their headway into retail practice. Among them, assortment planning and price optimization models are the two prominent classes. Nevertheless, there are gaps between academic and industry models. Please refer to Section 4.2.1 *Retail Applications* for academic models for assortment planning and price optimization. In what follows, we introduce a novel optimization model jointly optimizing price, assortment, and presentation for a group of substitutable products. The model strikes a balance between complexity, relevance, and adoptability.

2.1 Optimizing Price, Assortment and Presentation

In a typical retailer, macro-level assortment decisions, e.g., how much space to allocate to each category or department, are strategic decisions driven by considerations of the image or position of the chain, an appreciation of what items have been winners and losers, relationships with vendors and the size and layout of the store. As a matter of operational practice, these macro assortment decisions either ignore heterogeneity in tastes and preferences across markets, or accommodate it coarsely by differentiating assortment plans by geographic store groups. More micro assortment decisions are typically made by buyers or category managers and presentation decisions are typically undertaken by space planners, especially in grocery retailing. They are to decide, in each subcategory or finer categorization of items, precisely what size package of which brand, and how many facings of each to display. These decisions are made on the basis of the particular display space available, visual appeal, negotiations with vendors, and received wisdom about what has 'worked' in the past. Finally, item-level pricing decisions are the responsibility, in some cases, of a pricing department, and in other cases, by the merchants or category managers. For most retailers, these decisions are made based on judgment, informed of high-level policy rules governing markups, price image, and promotional positioning. In some more sophisticated retailers, analytic-based tools are used to support these decisions, sometimes including price optimization tools. The latter rely on models of demand which predict sales changes due to price changes, and may even include substitution and complementarity effects. But even these sophisticated price optimization models are insensitive to assortment and presentation decisions.

Studies of these assortment, pricing, and presentation decisions suggest that category profits may be improved by upwards of 50% by optimizing these decisions independently, (McIntyre & Miller, 1999), (Green & Savitz, 1994), (Phillips, 2005), (Talluri, 2004), (Fisher, 2009).

2.2 Demand Effects

2.2.1 Substitutable Item Groups and Price Effects

A "substitutable item group," or SIG, is a set of candidate items to offer and present which are strong, but imperfect substitutes for each other. They would typically comprise items of the same brand in different packages and flavors, as well as comparable items in different brands. A typical grocery retailer, for example, will sell 30,000 to 50,000 items in 50 to 150 categories. Each category would comprise 15 to 100 SIG's. Thus, a SIG may comprise from as few as 2 to 20 or more items. In practice, on average, there are 7 to 10 items per SIG.

We first model demand for the SIG as a whole, and then use a demand share model to obtain the demand for each individual item. Let M be the set of all items in the SIG and M^* be the subset of M comprising the items actually assorted. Quantity Q_{SIG}, the total demand of the entire SIG, is assumed to be responsive to $\boldsymbol{p} = (p_i)_{i \in M^*}$, the selling prices of the items offered:

$$Q_{SIG}(\boldsymbol{p}, M^*) = Q_{SIG}^r \cdot \left(p_{SIG}(\boldsymbol{p}, M^*) \Big/ p_{SIG}^r(M^*) \right)^{e_{SIG}}, \tag{1}$$

where Q_{SIG}^r is the reference sales for the SIG under the reference prices , e_{SIG} is the own-price elasticity of demand for the SIG, and p_{SIG} and p_{SIG}^r are the reference share-weighted selling prices and reference prices, respectively, of the items in the SIG that are offered . More specifically,

$$p_{SIG}(\boldsymbol{p}, M^*) = \sum_{i \in M^*} S_i^r p_i \Big/ \sum_{i \in M^*} S_i^r$$

$$p_{SIG}^r(M^*) = \sum_{i \in M^*} S_i^r p_i^r \Big/ \sum_{i \in M^*} S_i^r$$

where S_i^r is the demand share for item i under the reference prices p_i^r.

To model demand shares for individual items assorted, we include an "outside" good to account for no-purchase decisions. Both the own items' shares, and the outside good's share respond to their prices:

$$ST(\boldsymbol{p}, M^*) = \sum_{i \in M^*} \left(S_i^r \cdot \left({p_i}/{p_i^r} \right)^{e_i} \right) + S_{OG}^r \cdot \left({p_{OG}}/{p_{SIG}} \right)^{e_{OG}}$$

$$S_i(\boldsymbol{p}, M^*) = S_i^r \cdot \left({p_i}/{p_i^r} \right)^{e_i} / ST(\boldsymbol{p}, M^*) \qquad i \in M^*$$

$$S_i(\boldsymbol{p}, M^*) = 0 \qquad i \notin M^*$$

$$S_{OG}(\boldsymbol{p}, M^*) = S_{OG}^r \cdot \left({p_{OG}}/{p_{SIG}} \right)^{e_{OG}} \Big/ ST(\boldsymbol{p}, M^*)$$

where e_i is the own-price share elasticity of item *i*, and e_{OG} analogously for the outside good. S_i^r and S_{OG}^r are the reference shares of item *i* and the outside good respectively; they can be thought of as the intrinsic preference weights for the item. Quantity $ST(\boldsymbol{p}, M^*)$ normalizes the shares to ensure they sum to unity. We take p_{OG} as fixed, since it is outside the control of the decision maker; note however, that S_{OG}^r does depend on $\boldsymbol{p}$. This share model is equivalent to the widely used Multiplicative Competitive Interaction (MCI) model due to Nakanishi and Cooper ((Nakanishi & Cooper, 1974), (Nakanishi & Cooper, 1982), (Cooper & Nakanishi, Market-Share Analysis, 1988), (Cooper, Market Share Models, 1993)).

In its most basic form, demand Q_i of item i and demand Q_{OG} for the "outside" good is given by:

$$Q_i(\boldsymbol{p}, M^*) = Q_{SIG}(\boldsymbol{p}, M^*) \cdot S_i(\boldsymbol{p}, M^*) \,, \tag{2}$$

and

$$Q_{OG}(\boldsymbol{p}, M^*) = Q_{SIG}(\boldsymbol{p}, M^*) \cdot S_{OG}(\boldsymbol{p}, M^*).$$

Note that by definition we have $\sum_{i \in M^*} S_i(\boldsymbol{p}, M^*) + S_{OG}(\boldsymbol{p}, M^*) = 1$.

2.2.2 Assortment or Variety Effects

Retailer experience suggests that having a larger assortment, with greater variety, can lead to larger sales. This can be understood either as accommodating consumer heterogeneity, or as a consumer preference for variety (Kim, Allenby, &

Rossi, 2002). We use a simple extension of the basic SIG demand model in Eq. (1) to accommodate these effects:

$$\bar{Q}_{SIG}(\boldsymbol{p}, M^*) = \left[\left(Q_{SIG}^{\,r} \cdot \left({p_{SIG}(\boldsymbol{p}, M^*)}\middle/{p_{SIG}^{r}(M^*)}\right)^{e_{SIG}}\right) - Q_{OG}(\boldsymbol{p}, M^*)\right] \cdot |M^*|^{\delta_{SIG}},$$

where $0 < \delta_{SIG} < 1$ is the "assortment elasticity." While this is not strictly elasticity, we adopt that label for expository convenience and restrict parameter values to model diminishing returns to assortment breadth. Note that here, the effects of assortment on demand of a SIG depend only on the size of the assortment.

2.2.3 Presentation Effects

As is the case with the assortment or variety effects, retailers acknowledge that the extent of a presentation of an item is also correlated with sales for that item. A greater number of facings of an item will make it more prominent and attract more demand; and the more facings allocated to an item, the less likely that the display quantity will be depleted prior to restocking. We employ a simple extension of the basic item quantity model in Eq. (2) to accommodate these effects:

$$Q_i(\boldsymbol{p}, M^*, f_i) = (\bar{Q}_{SIG}(\boldsymbol{p}, M^*) \cdot S_i(\boldsymbol{p}, M^*)) \cdot f_i^{\gamma_i},$$

where f_i is the number of "facings" of item *i* and $0 < \gamma_i < 1$ is the "presentation elasticity" of item *i*. Again, this is not strictly elasticity, but we adopt that label for expository convenience and restrict parameter values to model diminishing returns to presentation intensity. Finally, we note that if $f_i = 0$, there are no sales, which corresponds to the case of *i* being excluded from the assortment. The set of assorted items, M^*, is determined by $\boldsymbol{f} = (f_i)_{i \in M^*}$ as following:

$$M^*(\boldsymbol{f}) = \{i | f_i > 0\}.$$

2.2.4 Full Model and Parameters

Combining all of these elements, we have an item-level sales model that reflects the effects of prices on all items in the SIG, as well as the items that are assorted and the presentation, reflected in the number of facings of each item:

$$Q_i(\boldsymbol{p},\boldsymbol{f}) = \left(\left[Q_{SIG}(\boldsymbol{p},M^*(\boldsymbol{f}))(1 - S_{OG}(\boldsymbol{p},M^*(\boldsymbol{f}))) \cdot |M^*(\boldsymbol{f})|^{\delta_{SIG}}\right] \cdot S_i(\boldsymbol{p},M^*(\boldsymbol{f}))\right) \cdot f_i^{\gamma_i}$$

We are interested in the choices that confront the decision-maker, to set values for p_i and f_i (and thus for M^*, as well). We assume that p_{OG} and p_i^r are fixed and given, and that Q_{SIG}^r, S_i^r, S_{OG}^r, $\delta_{SIG}, \gamma_i, e_i, e_{SIG}$ and e_{OG} are parameters to be set by expert judgment or to be calibrated from historical data.

2.3 Optimization Model

The decision problem is to choose the prices p_i and the number of facings f_i for each item in the SIG (which subsumes the problem of choosing the set of items to include in the assortment $M^*(\boldsymbol{f})$). In practice, a retailer may wish to maximize revenue, but in concept, the optimization problem is to maximize some measure of gross profit. Thus, the model in Section 3.2.4 yields the following optimization problem:

$$\max_{\boldsymbol{p},\boldsymbol{f}} \sum_{i \in M} \left(\left[Q_{SIG}(\boldsymbol{p},M^*(\boldsymbol{f}))(1 - S_{OG}(\boldsymbol{p},M^*(\boldsymbol{f}))) \cdot |M^*(\boldsymbol{f})|^{\delta_{SIG}}\right] \cdot S_i(\boldsymbol{p},M^*(\boldsymbol{f}))\right) \cdot f_i^{\gamma} \cdot (p_i - c_i) \ ,$$

where $p_i - c_i$ is the gross profit for item *i*.

The dominant constraint in this problem is the space constraint. We assume that each item occupies a certain amount of linear shelf space, per facing. If we have K_{SIG} linear space allocated to the SIG, then we have:

$$\sum_{i \in M} w_i \cdot f_i \le K_{SIG}$$

where w_i is the width of item *i*. Typically, there are other business constraints, such as an item being required to be assorted, or the maximum space that can be allocated to an item. These can be condensed into a set of range constraints for each item:

$$l_i \le f_i \le u_i,$$

where l_i and u_i are minimum and maximum allowable number of facings for item i. It is also typical that there are limits on the magnitude of price changes. For simplicity, we assume that the current prices are the reference prices, which yields for each item constraints:

$$p_i^r \cdot (1 - \theta_i) \le p_i \le p_i^r \cdot (1 + \theta_i),$$

where θ_i is the maximum allowable price change of item i .

This model is hard to solve, however we resort to the fact that the number of items in the SIG is relatively low. To this end, the algorithm enumerates all possible assortments M and for each of them we solve the resulting non-linear continuous optimization problem.

3 Conclusion

Models much like the promotions forecasting model in Section 1 are presently used by managers in some retail chains, both for the purposes of planning promotions and improving supply chain decisions. The assortment, space and price optimization model in Section 2 is the subject of active research, but to the authors' knowledge, is not employed at present in a production system. Kök and Fisher (2007) present estimation results and management implications from implementing a similar model at a European grocery retailer.

Retailing has been imagined to be "a paradise for operations researchers," (Fisher, 2009). It is rich with interesting problems and an increasing appetite for adapting to some of the methods and tools of OR. We have described a few areas where the fertile interplay between academic research, analytic solution vendors and consultants, and retail decision-makers is yielding new and useful OR applications.

4 Bibliography

Alon, I., Qi, M., & Sadowski, R. J. (2001). Forecasting aggregate retail sales: A comparison of artificial neural networks and traditional methods. *Journal of Retailing and Consumer Services,* , 147–156.

Bell, D. R., & Chaing, J. P. (1999). The Decomposition of Promotional Response: An Empirical Generalization. *Marketing Science* , 504-526.

Box, G. E., & Jenkins, G. M. (2008). *Time Series Analysis: Forecasting and Control.* Hoboken: John Wiley & Sons, Inc.

Bunn, D., & Wright, G. (1991). Interaction of Judgemental and Statistical Forecasting Methods: Issues and Analysis. *Management Science , 37* (5), 501-518.

Cooper, L. G. (1993). Market Share Models. In J. Eliashberg, & G. L. Lilien, *Handbooks in Operations Research and Management Science: Marketing* (Vol. 5, pp. 259-354). Amsterdam, The Netherlands: Elsevier Science Publishers.

Cooper, L. G., & Nakanishi, M. (1988). *Market-Share Analysis.* Norwell, Massachusetts: Kluwer Academic Publishers.

Fisher, M. (2009). Rocket Science Retailing:The 2006 Philip McCord Morse Lecture. *Operations Research* , 527-540.

Franses, P. H. (1998). *Time series models for business and economic forecasting.* Cambridge: Cambridge University Press.

Green, P. E., & Savitz, J. (1994). Applying Conjoint Analysis to Product Assortment and Pricing in Retailing Research. *Pricing Strategy and Practice* , 4-19.

Kahn, K. B. (1998). Revisiting top-down Versus bottom-up forecasting. *The journal of business forecasting* , 14-19.

Kim, J., Allenby, G. M., & Rossi, P. E. (2002). Modeling Consumer Demand for Variety. *Marketing Science* , 229-250.

Kök, A. G., & Fisher, M. L. (2007). Demand Estimation and Assortment Optimization under Substitution: Methodology and Application. *Operations Research* , 1001-1021.

Mantrala, M. K., Kaul, R., Gopalakrishna, S., & Stam, A. (2006). Optimal Pricing Strategies for an Automotive Aftermarket Retailer. *Journal of Marketing Research* , 588-604.

Maulik, U., & Bandyopadhyay, S. (2000). Genetic algorithm-based clustering technique. *Pattern Recognition* , 1455-1465.

Mclntyre, S. H., & Miller, C. (1999). The Selection and Pricing of Retailing Assortments: An Empirical Approach. *Journal of Retailing* , 588-604.

Nakanishi, M., & Cooper, L. G. (1974). Parameter Estimation for a Multiplicative Competitive Interaction Model: Least Square Approach. *Journal of Marketing Research* , 303-311.

Nakanishi, M., & Cooper, L. G. (1982). Simplified Estimation Procedures for MCI Models. *Marketing Science , 1*, 313-322.

Phillips, R. L. (2005). *Pricing and Revenue Optimization.* Stanford: Stanford University Press.

Talluri, K. a. (2004). *The Theory and Practice of Revenue Management.* Boston: Kluwer.

Wal-Mart Stores, Inc. (n.d.). Retrieved Feburary 4, 2010, from http://investors.walmartstores.com/phoenix.zhtml?c=112761&p=irol-irhome

Zotter, G., Kalchschmidt, M., & Caniato, F. (2005). The impact of aggregation level on forecasting performance. *International Journal of production economics* , 479-491.