Promotion Optimization in Retail

Maxime C. Cohen
NYU Stern School of Business, New York, NY 10012, maxcohen@nyu.edu

Georgia Perakis
MIT Sloan School of Management, Cambridge, MA 02139, georgiap@mit.edu

1. Introduction

This chapter presents some recent developments in retail promotions. In many retail settings such as supermarkets, promotions are a key driver for boosting profits. Promotions are often used on a daily basis in most retail environments including supermarkets, drugstores, fashion retailers, electronics stores, online retailers, convenience stores etc. For example, a typical supermarket sells several thousands of products, and needs to decide the price promotions for all the products at each time period. These decisions are of primary importance, as using the right promotions can significantly enhance the business' bottom line. In today's economy, retailers offer hundreds or even thousands of promotions simultaneously. Promotions aim to increase sales and traffic, enhance awareness when introducing new items, clear leftover inventory, bolster customer loyalty, and improve the retailer competitiveness. In addition, price promotions are often used as a tool for price discrimination among the different customers.

Surprisingly, many retailers still employ a manual process based on intuition and past experience in order to decide promotions. The unprecedented volume of data that is now available to retailers presents an opportunity to develop support decision tools that can help retailers improve promotion decisions. The promotion planning process typically involves a large number of decision variables, and needs to ensure that the relevant business constraints (called promotion *business rules*) are satisfied (more details can be found in Section 3.2). In this chapter, we discuss how analytics can help retailers decide the promotions for multiple items while accounting for many important modeling aspects observed in retail data. In particular, we consider practical models that are motivated from a collaboration between academia and industry. Most of the material discussed in this chapter is inspired by the results in Cohen et al. (2017) and in Cohen et al. (2018). For more details on the specifics of the algorithms, the proofs of the analytical results, and on the managerial insights, we refer the reader to the papers.

Several recent advances in operations management and marketing have focused on developing new methods to improve the process of deciding retail promotions. The ultimate goal is to increase the total profit by promoting the right items at the right time periods using the right price points. At a high level, retail promotions can be categorized as follows: (i) manufacturer versus retailer promotions, (ii) markdowns versus temporary price discounts, (iii) targeted versus mass campaigns, and (iv) price reductions versus alternative promotion vehicles. We next discuss these four categorizations.

Manufacturer versus retailer promotions: In retail settings, the brand manufacturer (e.g., Coca-Cola, Kellogg's) can directly offer a price discount either to the retailer or to the end-consumer. These incentives are often called trade funds, vendor funds or manufacturer coupons. This type of promotions usually come from long-term negotiations between the manufacturer and the retailer, and involve several contractual terms. For example, a manufacturer can offer a rebate to the retailer if the cumulative sales during the quarter exceed a certain target level. In exchange, the retailer will place the manufacturer's products in preferred locations (e.g., end-cap-displays). A second example is a shared promotion contract in which the manufacturer subsidizes some portion of the price discount offered to the consumers. A third example occurs when a manufacturer offers a coupon to the end-consumers who then need to claim the discount (at the store, on the Internet or toward future purchases). Typically, retailers have to decide when to accept such vendor funds and under what conditions. In many situations, manufacturers tend to be aggressive on the contractual terms by imposing long-term commitments, high volumes, and sometimes exclusivity restrictions (e.g., not allowing the promotions of competing brands).

Markdowns versus temporary price discounts: Markdowns typically refer to the practice of decreasing the price of an item at the end of the selling season. The regular price is decreased in order to clear the remaining inventory. Note that in such a case, the price may be reduced several times but cannot be increased back to the regular price. This is common practice in the fashion and tourism industries as well as in the business of selling tickets for media events (e.g., concerts). For example, an apparel from the summer collection may be discounted toward the end of the season if the remaining inventory is higher than anticipated. On the other hand, temporary price discounts are used in different contexts. A well-known such context is of the Fast-Moving Consumer Goods (FMCG). Examples of FMCG include processed foods and soft drinks, as well as household products (e.g., laundry detergent and toothpaste). Note that these products are usually non-perishable, and have a long shelf life. Such purchases are recurring, and retailers do not need to clear the remaining inventory. In order to increase the profit, it became common for most retailers to use temporary price reductions (e.g., 20% off the regular price during one week).

Targeted versus mass campaigns: Retailers can either decide to send promotions to a few targeted customers or to simply decrease the price of a particular product for all the potential buyers. Targeted marketing campaigns can be implemented via email redeemable coupons or by using advanced geo-localization techniques. Online retailers often use targeted promotions by tracking potential customers using cookies, and by sending promotional offers to selected sets of customers (e.g., active members that made a recent purchase). On the other hand, mass promotions are price discounts that apply to all customers. Brick-and-mortar retailers such as supermarkets mainly employ mass promotion campaigns.

Price reductions versus alternative promotion vehicles: Retailers can use different ways to promote a product. The most straightforward method is to use a price discount, in which the item is temporarily priced below its regular price. Other options include "buy-one-get-one", in-store flyers, coupons, tasting stands, placing products at the end of an aisle (end-cap-display), sending out flyers, broadcasting TV commercials, radio advertisements, etc. (these are often called promotion vehicles). Typically, a retailer can choose among 5-40 different promotion vehicles at each point in time.

In this chapter, we focus on the mass pricing promotion optimization problem faced by a retailer who sells FMCG products. Namely, we consider a retailer (e.g., a supermarket) who needs to decide which items to promote, at which price points, and when to schedule the promotions of the different items. The problems of setting the right manufacturer incentives, optimizing markdowns, designing targeted promotions, and optimizing promotion vehicles are also important retail questions, but are beyond the scope of this chapter. We will briefly refer to some of the relevant literature on these problems in Section 2.

The amount of money spent on promotions for FMCG products can be significant - it is estimated that FMCG manufacturers spend about \$1 trillion annually on promotions (Nielsen 2015). In addition, promotions play an important role in the FMCG industry as a large proportion of the sales is made during promotions. For example, retail data indicates that 12–25% of supermarket sales in five European countries were made during promotions (Gedenk et al. 2006). The market research group IRI found that more than half of all goods (54.6%) sold to UK shoppers in supermarkets and major retailers were on promotion.[1]

The promotion planning process faced by a medium to large size retailer is challenging for several reasons. First, one needs to carefully account for the cross-item effects in demand (cannibalization and complementarity). When promoting a particular item, the demand of some other products may also be affected by the promotion. Consequently, one needs to decide the promotions of all

[1] https://www.theguardian.com/business/2015/nov/02/majority-of-goods-sold-in-uk-stores-on-promotion-finds-study-multi-buys

the items in the category while accounting for those effects that can be directly estimated from data. Second, retail promotions are often constrained by a set of business rules specified by the retailer and/or the product manufacturers. Example of business rules include prices chosen from a set of discrete values, limited number of promotions (both per time period and for each item), and cross-item business rules that restrict the relationship between the prices of the different items (more details are provided in Section 3.2). Third, the demand usually exhibits a post-promotion-dip effect. This effect is induced by the promotion fatigue (i.e., repeating the same promotion may have a low marginal impact), and by the stockpiling behavior of consumers. More precisely, for certain categories of (non-perishable) products, customers tend to stockpile during promotions by purchasing larger quantities for future consumption. This ultimately leads to a reduced demand following the promotion period. Fourth, the problem is difficult due to its large scale. As we mentioned, an average supermarket offers several thousands of SKUs (Stock Keeping Units), and the number of items on promotion at any time can be very large. Consequently, this leads to a large number of decisions that need to be made by the retailer.

Retail promotions can have a significant impact on boosting sales, and on influencing customers. For example, a study from the International Council of Shopping Centers shows that 90% of adult consumers claim to be influenced by promotions in terms of the amount they spend, and the items they purchase.[2] Despite the complexity of the promotion planning process, it is still to this day performed manually in many supermarket chains. This motivates us to design and study promotion optimization models that can make promotion planning more efficient and automated. The goals of this line of research include the following:

- Formulate the promotion optimization problem for multiple items (labeled as Multi-POP). This formulation is directly motivated from practice, holds for general demand models (estimated from data), and can incorporate the relevant business rules.
- Discuss how the formulation captures several important economic factors which are present in retail environments. These factors include the post-promotion dip effect (due to the stockpiling behavior of consumers), the cross-item effects, and the demand seasonality.
- Develop an efficient approximation solution approach to solve the problem. We propose a discrete linearization method that allows the retailer to solve a large scale instance of the problem within seconds. We also convey that our solution approach provides a parametric worst-case bound on the quality of the approximation relative to the optimal (intractable) solution.

[2] https://retailleader.com/brick-and-mortar-makes-grade-back-school-shopping

- Present a beginning-to-end application of the entire process of optimizing retail promotions. We divide the process in five steps that the retailer needs to follow; from collecting and aggregating the data to computing the future promotion decisions.
- Discuss the potential impact of using data analytics and optimization for retail promotions. We convey that in our tested examples (calibrated with retail data), using the promotions suggested by our model can yield a 2-9% profit improvement. Such an increase is significant, as retail businesses typically operate under small margins.

This chapter is organized as follows. In Section 2, we review some of the related literature. In Section 3, we report the notation, assumptions, and problem formulation. In Section 4, we present a class of approximation methods to efficiently solve the promotion optimization problem. In Section 5, we use our model and solution approach to draw practical insights on promotion planning, and present a summary of how to apply our model to real-world retail environments. Finally, we report our conclusions in Section 6. As mentioned before, more details on the technical results and on the insights can be found in Cohen et al. (2017) and in Cohen et al. (2018).

2. Literature Review

The topic of retail promotions has been an active research area both in academia and industry. In particular, our problem is related to several streams of literature, including dynamic pricing, promotions in marketing, and retail operations.

Dynamic pricing: Dynamic pricing has been an extensive topic of research in the operations management community. Comprehensive reviews can be found in the books and review papers by Bitran and Caldentey (2003), Elmaghraby and Keskinocak (2003), Talluri and Van Ryzin (2006), Özer and Phillips (2012), as well as the references therein. A large number of recent papers study the problem of dynamic pricing under various contexts and modeling assumptions. Examples include Ahn et al. (2007), Su (2010) and Levin et al. (2010), just to name a few. In Ahn et al. (2007), the authors propose a demand model in which a proportion of customers strategically wait k periods, and purchase the product once the price falls below their willingness to pay. They then formulate a mathematical programming model, and develop solution techniques. In Su (2010), the author studies a model with multiple consumer types who may differ in their holding costs, consumption rates, and fixed shopping costs. The author solves the dynamic pricing model by computing the rational expectation equilibrium, and draws several managerial insights. In Levin et al. (2010), the authors consider a dynamic pricing model for a monopolist who sells a perishable product to strategic consumers. They model the problem as a stochastic dynamic game, and prove the existence of a unique subgame-perfect equilibrium pricing policy. A very prominent topic in the dynamic pricing literature is to study the setting in which consumers are strategic (or forward-looking) (see, e.g., Aviv and Pazgal 2008, Cachon and Swinney 2009, Levina et al. 2009, Besbes and

Lobel 2015, Liu and Cooper 2015, Chen and Farias 2015). The problem considered in this chapter is in the same spirit as the dynamic pricing problem. Nevertheless, we focus on a setting where the demand model is estimated from historical data, and the optimization formulation includes the simultaneous promotion decisions of several interconnected items. In addition, we require the dynamic pricing decisions to satisfy several business rules.

Promotions in marketing: Sales promotions are an important area of research in marketing (see Blattberg and Neslin (1990) and the references therein). However, the focus in the marketing community is typically on modeling and estimating dynamic sales models (econometric or choice models) that can be used to draw managerial insights (Cooper et al. 1999, Foekens et al. 1998). For example, Foekens et al. (1998) study econometrics models based on scanner data to examine the dynamic effects of sales promotions. It is widely recognized in the marketing community that for certain products, promotions may have a pantry-loading or a post-promotion dip effect, i.e., consumers tend to purchase larger quantities during promotions for future consumption (stockpiling behavior). This effect leads to a decrease in sales in the short term. In order to capture the post-promotion dip effect, many of the dynamic sales models in the marketing literature posit that the demand is not only a function of the current price, but also of the past prices (see, e.g., Ailawadi et al. 2007, Macé and Neslin 2004). Finally, note that several prescriptive works in the marketing community study the impact of retail coupons in the context of sales promotions (see, for example Heilman et al. 2002). The demand models used in this chapter also consider that the demand depends explicitly on the current and past prices as well as on the prices of other items.

Retail operations: Several academic papers study the topic of retail promotions from an empirical descriptive point of view. Van Heerde et al. (2003) and Martínez-Ruiz et al. (2006) use panel-data to empirically study how retail promotions induce consumers to switch brands. The recent work by Felgate and Fearne (2015) uses supermarket loyalty card data from a sample of over 1.4 million UK households to analyze the effect of promotions on the sales of specific products across different shopper segments. Another line of research discusses field experiments on pricing decisions implemented at retailers. A classical successful example is the implementation at the fashion retail chain Zara (see Caro and Gallien 2012). In their work, the authors report the results of a controlled field experiment conducted in all Belgian and Irish stores during the 2008 fall-winter season. They assess that the new process has increased clearance revenues by approximately 6%. An additional recent work can be found in Ferreira et al. (2015) in which the authors collaborated with Rue La La, a flash sales fashion online retailer. The authors propose a non-parametric prediction model to predict future demand of new products, and develop an efficient solution for the price optimization problem. They estimate a revenue increase for the test group by approximately 9.7%. Pekgün et al. (2013) describe a collaboration with the Carlson Rezidor Hotel Group. In this study, the authors show that demand forecasting and dynamic revenue optimization consistently increased revenue by 2-4% in participating hotels relative to non-participating hotels.

Other types of promotions: As mentioned before, retail promotions can be divided in several categories. While the models presented in this chapter focus on the mass pricing promotion optimization problem faced by a retailer who sells FMCG products, other studies have considered the alternative promotion types. Several papers consider the problem of vendor funds in the context of promotion planning (see, e.g., Silva-Risso et al. 1999, Nijs et al. 2010, Yuan et al. 2013, Baardman et al. 2017b). As mentioned before, an additional related topic is the one of markdown pricing, or markdown optimization. In this problem, the seller needs to decide when to decrease the price of the item(s) in order to clear the remaining inventory by the end of the season. There is a large number of academic papers that propose different models and methods to solve the markdown pricing problem. Examples include Yin et al. (2009), Mersereau and Zhang (2012), Zhang and Cooper (2008), Vakhutinsky et al. (2012), and Caro and Gallien (2012), just to name a few. As we explained before, the promotion optimization problem for FMCG products differs from the markdown optimization problem by the structure of the pricing policy and by the lack of inventory expiration. The topic of designing targeted promotions has recently attracted a lot of attention. Given that sending promotions to existing or new customers can be expensive and often results in low conversion rates, several firms aim to develop quantitative methods that exploit the large historical data sets in order to design targeted promotion campaigns. For example, retailers often need to decide which types of customers to target, and what are the most important features (e.g., geo-localization, demographics, and past behavior). Targeted marketing campaigns (email and mobile offers) have been extensively studied in the academic literature (see, e.g., Arora et al. 2008, Fong et al. 2015, Andrews et al. 2015, Jagabathula et al. 2018). Finally, in addition to price promotions, retailers typically need to decide how to assign the different vehicles (e.g., flyers and TV commercials). The recent work in Baardman et al. (2017a) addresses the problem of optimally scheduling promotion vehicles for a retailer.

Methodology: From a methodological perspective, the tools used in this chapter are related to the literature on nonlinear and integer optimization. We formulate the promotion optimization problem as a nonlinear mixed integer program (NMIP). Due to the general classes of demand functions we consider, the objective function is typically non-concave, and such NMIPs are generally difficult from a computational complexity standpoint. Under certain special structural conditions (see, e.g., Hemmecke et al. (2010) and the references therein), there exist polynomial time algorithms for solving NMIPs. However, many NMIPs do not satisfy these conditions and are solved using techniques such as Branch and Bound, Outer-Approximation, Generalized Benders and Extended Cutting Plane methods (Grossmann 2002).

In the special instance of the Multi-POP with linear demand and continuous prices, one can formulate our problem as a Cardinality-Constrained Quadratic Optimization (CCQO) problem. It has been shown in Bienstock (1996) that such a problem is NP-hard. Thus, tailored heuristics have been developed in order to solve this type of problems (see, for example, Bienstock 1996, Bertsimas and Shioda 2009). The general instance of our problem has discrete variables, and considers a general demand function. Note that our problem was also shown to be NP-hard (Cohen et al. 2016). Our solution approach is based on approximating the objective function by exploiting the discrete nature of the problem. Given that we consider general demand functions, it is not possible to use linearization approaches such as in Sherali and Adams (1998). Our main approximation method results in a formulation which is related to the field of Quadratic Programming. Such problems were extensively studied in the literature (see, e.g., Frank and Wolfe 1956, Balinski 1970, Rhys 1970, Padberg 1989, Nocedal and Wright 2006).

3. Problem Formulation

In this section, we formulate the promotion optimization problem (labeled as Multi-POP). We first introduce the notation and our assumptions. We then discuss the various business rules that the retailer needs to satisfy when deciding price promotions. Finally, we present the resulting optimization formulation.

Consider a retailer who sells several FMCG products. Very often, retailers decide the price promotions of their products for each category separately. Consequently, we focus our presentation on a single category (e.g., ground coffee, soft drinks) composed of N items (or SKUs). The goal of the category manager is to maximize the total profit over a selling horizon composed of T periods (for example, one quarter of 13 weeks). We denote by p_t^i the price of item i at time t.[3] We also denote by c_t^i the (exogenous) cost of a single unit of item i at time t. In other words, we assume that the cost of each item at each time is known, and that the retailer needs to decide the prices of all N items during all T time periods. A summary of our notation can be found at the end of this section.

3.1. Assumptions

To gain tractability, we impose the following assumptions.

Assumption 1. 1. *The retailer decides all the price promotions at the beginning of the season.*
2. *The retailer carries enough inventory to meet the demand for each item in each time period.*[4]
3. *The demand is expressed as a deterministic time-dependent nonlinear function of the prices.*

[3] Throughout this chapter, the subscript (resp. superscript) index corresponds to the time (resp. item).

[4] We therefore use the words *demand* and *sales* interchangeably.

4. *The demand function depends explicitly on self past and current prices, and on cross current prices.*

We next briefly discuss the validity of the above assumptions. Assumption 1.1 applies to a setting where the retailer needs to commit upfront for the entire selling season. For example, such restrictions can emerge from vendor funds or can be imposed by sending out flyers through different advertising channels.

Note that Assumption 1.2. does not apply to all products and retail settings (e.g., very often in the fashion industry, limited inventory is produced to induce scarcity). Unlike fashion items which may be seasonal, FMCG products are typically available all year round. These products have a long shelf life, and customers have been conditioned to always find these products in stock at retail stores. Since FMCG products are usually easy to store and have a high degree of availability, FMCG retailers typically do not stock out. In Cohen et al. (2017), the authors analyze two years of supermarket data for FMCG products, and convey that (i) the demand forecast accuracy for this type of products is often high (good out-of-sample R^2 and MAPE), and (ii) the inventory is not issue as very few stock-outs occurred over a two-year period. This can be justified by the fact that supermarkets have a long experience with inventory decisions, and collected large data sets allowing them to develop sophisticated forecasting demand tools to support ordering decisions (see, e.g., Cooper et al. 1999, Van Donselaar et al. 2006). Finally, grocery retailers are aware of the negative effects of being out of stock for promoted products (see, e.g., Corsten and Gruen 2004, Campo et al. 2000). However, for settings where inventory is limited, one needs to consider a different formulation than the one presented in this chapter.

Assumption 1.3 translates to denoting the demand of item i at time t by $d_t^i(\mathbf{p})$, where $\mathbf{p}$ is a vector of current and past prices (see more details below). We assume that the demand is a deterministic function as we observed a high out-of-sample prediction accuracy using our data. Extending our model when the demand is a stochastic function is an interesting direction for future research (e.g., by using learning algorithms).

Assumption 1.4 implies that the demand does not explicitly depend on cross past prices. In other words, the demand of item i does not depend on the past prices of the other items in the category. This assumption was validated by running demand prediction models using retail datasets (more details can be found in Cohen et al. 2018). Consequently, the demand of item i at time t can be any nonlinear and time dependent function of the form: $d_t^i\big(p_t^i, p_{t-1}^i, \ldots, p_{t-M^i}^i, \mathbf{p_t^{-i}}\big)$, where M^i represents the memory parameter of item i (i.e., the number of past prices that affect the current demand), and $\mathbf{p_t^{-i}}$ denotes the vector of prices of all the items except i at time t. Note that in practice M^i is estimated from the historical data, and can be different across items.

Note that the demand of item i at time t depends on several factors:

- The self current price p_t^i – This captures the price sensitivity of the consumers toward the item.
- The self past prices $\left(p_{t-1}^i, \ldots, p_{t-M^i}^i\right)$ – This captures the post promotion dip effect (induced by the stockpiling behavior of consumers).
- The cross current prices $\mathbf{p_t^{-i}}$ – This captures the cross-item effects on demand (substitution and complementarity).
- Other potential features such as demand seasonality (weekly, monthly or quarterly), trend factor, store effect, holiday boosts, etc.

Concrete demand models such as the log-log demand function can be found in Cohen et al. (2017).

In most product categories, a promotion for a particular item affects its own sales, but also the sales of other items in the category. As mentioned, we capture these cross-item effects by assuming that the demand of item i depends on the prices of the other items (at the same time period). The standard example of substitutable items are competing brands such as Coke and Pepsi. In this case, it is clear that promoting a Coke product potentially increases Coke's sales but it may also decrease Pepsi's sales. Mathematically, one can assume that if items i and $j \neq i$ are substitutes, then $\frac{\partial d_t^i}{\partial p_t^j} \geq 0$ and $\frac{\partial d_t^j}{\partial p_t^i} \geq 0$ for some t. Two products i and j are complements if the consumption of i induces customers to purchase item j (and vice versa), e.g., shampoo and conditioner. Mathematically, one can assume that if items i and $j \neq i$ are complements, then $\frac{\partial d_t^i}{\partial p_t^j} \leq 0$ and $\frac{\partial d_t^j}{\partial p_t^i} \leq 0$ for some t.

3.2. Business Rules

In the retail setting we consider, there are typically two classes of business rules: (i) business rules on each item separately (called *self business rules*); and (ii) business rules that impose joint pricing constraints on several items (called *cross-item business rules*). The self business rules are identical to the ones presented in Cohen et al. (2017), while the cross-item business rules are similar to Cohen et al. (2018).

Self business rules

1. *Prices are chosen from a discrete price ladder.* For each product, there is a finite set of permissible prices. In particular, we consider that each item $i = 1, \ldots, N$ can take several prices: the regular price denoted by q^{i0}, and $K^i = |Q_i| - 1$ promotion prices denoted by q^{ik}. The total number of price points for item i is called the size of the price ladder (denoted by $|Q_i|$).[5] Consequently, the price of item i at time t can be written as $p_t^i = \sum_{k=0}^{K^i} q^{ik} \gamma_t^{ik}$, where the binary decision variable γ_t^{ik} is equal to 1 if the price of item i at time t is selected to be q^{ik}, and 0 otherwise.

[5] For simplicity, we assume that the elements of the price ladder are time independent, but our results still hold when this assumption is relaxed. In addition, we assume without loss of generality that the regular non-promotion price $q^{i0} = q^0$ is the same across all items $i = 1, \ldots, n$ and all time periods (this assumption can be relaxed at the expense of a more cumbersome notation).

2. *Limited number of promotions.* The retailer may want to limit the promotions frequency for a product in order to preserve the image of their store, and not train customers to be deal-seekers. For example, it may be required to promote item i at most $L^i = 3$ times during the quarter. This requirement for item i is captured by the following constraint: $\sum_{t=1}^{T} \sum_{k=1}^{K^i} \gamma_t^{ik} \leq L^i$.

3. *Separating periods between successive promotions (no-touch constraint).* A common additional requirement is to space out two successive promotions by a minimal number of separating periods, denoted by S^i. This constraint also helps retailers preserve their store image and discourage consumers to be deal-seekers. In addition, this type of requirement may be dictated by the manufacturer that sometimes restricts the frequency of promotions in order to preserve the brand image. Such a requirement for item i translates to adding the following constraint: $\sum_{\tau=t}^{t+S^i} \sum_{k=1}^{K^i} \gamma_\tau^{ik} \leq 1 \ \forall t$.

Cross-item business rules

1. *Total limited number of promotions.* The retailer may want to limit the total number of promotions throughout the selling season. For example, at most $L_T = 20$ promotions may be allowed during the season. Mathematically, one can impose the following constraint:

$$\sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{k=1}^{K^i} \gamma_t^{ki} \leq L_T. \tag{1}$$

Note that L_T should satisfy $L_T < \sum_{i=1}^{N} L^i$ for this constraint to be relevant.

2. *Inter-item ordinal constraints.* Several price relations can be dictated by business rules. For example, smaller size items should have a lower price relative to similar larger-sized products, and national brands must be more expensive when compared to private labels. These constraints can be captured by linear inequalities among the prices (e.g., if item i should be priced no higher than item j, one can add the constraint: $p_t^i \leq p_t^j \ \forall t$).

3. *Simultaneous promotions.* Sometimes, retailers require particular items to be promoted simultaneously as part of a manufacturer incentive or a special promotional event. If items i and j should be promoted simultaneously, one can impose: $\gamma_t^{0i} = \gamma_t^{0j} \ \forall t$, where γ_t^{0i} (resp. γ_t^{0j}) is a binary variable that is equal to 1 if item i (resp. item j) is not promoted at time t.

4. *Limited number of promotions in each period.* One can impose a limitation on the number of promotions in each time period. For example, at most $C^t = \frac{N}{10}$ promotions may be allowed i.e., only at most 10% of the items. Mathematically, we have:

$$\sum_{i=1}^{N} \sum_{k=1}^{K^i} \gamma_t^{ki} \leq C^t \ \forall t. \tag{2}$$

5. *Cross no-touch constraints.* An additional requirement can be to space out the promotions of a set of similar items by a minimal number of separating periods, denoted by S^c. As before, this is

motivated by the wish to preserve the store image and to mitigate the incentives for consumers to be deal-seekers. In this case, we need to separate successive promotions for two (or more) products. Mathematically, one can impose: $\sum_i \sum_{\tau=t}^{t+S^c} \sum_{k=1}^{K^i} \gamma_\tau^{ki} \leq 1 \ \forall t$, where the sum on i can be over any given subset of items in the category. Note that when $S^c = 0$, this corresponds to never promoting the items simultaneously in order to impose an exclusive offer (very common in practice).

3.3. Problem Formulation

In what follows, we present the promotion optimization problem for multiple items:

$$
\begin{aligned}
\max_{\gamma_t^{ik}} \ & \sum_{i=1}^{N} \sum_{t=1}^{T} (p_t^i - c_t^i) d_t^i \big(p_t^i, p_{t-1}^i, \ldots, p_{t-M^i}^i, \mathbf{p_t^{-i}}\big) \\
\text{s.t.} \ & p_t^i = \sum_{k=0}^{K^i} q^{ik} \gamma_t^{ik} \qquad \forall i \\
& \sum_{t=1}^{T} \sum_{k=1}^{K^i} \gamma_t^{ik} \leq L^i \qquad \forall i \\
& \sum_{\tau=t}^{t+S^i} \sum_{k=1}^{K^i} \gamma_\tau^{ik} \leq 1 \qquad \forall i, t \\
& \sum_{k=0}^{K^i} \gamma_t^{ik} = 1 \qquad \forall i, t \\
& \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{k=1}^{K^i} \gamma_t^{ki} \leq L_T \\
& \sum_{i=1}^{N} \sum_{k=1}^{K^i} \gamma_t^{ki} \leq C^t \qquad \forall t \\
& \gamma_t^{ik} \in \{0, 1\} \qquad \forall i, t, k
\end{aligned}
\tag{Multi-POP}
$$

In this problem, the objective is to maximize the total profit from all the N items during the selling season. Note that in the formulation above, we include all the self business rules, as well as the constraints on the total limited number of promotions from (1), and on the limited number of promotions in each period from (2). One can naturally include additional cross-item business rules into the formulation, depending on the requirements. It is worth mentioning that even in the absence of cross-item business rules, the N items are linked through the cross-item effects present in the demand functions.

Summary of Notation:

T - Length of the selling season.

N - Number of different items in the category.

c_t^i - Cost of item i at time t (assumed to be known).

p_t^i - Price of item i at time t (decision variable).

$\mathbf{p_t^{-i}}$ - Vector of prices of all items but i at time t.

$d_t^i\big(p_t^i, p_{t-1}^i, \ldots, p_{t-M^i}^i, \mathbf{p_t^{-i}}\big)$ - Demand of item i at time t, which is assumed to be a function of the self current and past prices as well as of the cross current prices (estimated from data).

M^i - Memory parameter of item i, i.e., the number of past prices that affect the current demand (estimated from data).

L^i - Limitation of the number of promotions for item i.

S^i - No-touch period for item i, i.e., the minimal number of time between two successive promotions.

K^i - Number of promotion prices in the price ladder of item i.

q^0 - Regular price (assumed to be the same across the different items).

$|Q_i| = K^i + 1$ - Total number of possible prices for item i.

q^{ik} - Price point k for item i $(k = 1, \ldots, K^i)$.

γ_t^{ik} - Binary decision variable to indicate if the price of item i at time t is equal to q^{ik}.

$MPOP$ - Objective function of the (Multi-POP) problem, i.e., the total profit generated by all items at all times.

$SPOP$ - Objective function of the problem for a single item.

4. Solution Approach

Our goal is to solve the optimization problem (Multi-POP). Since the problem is a nonlinear Integer Program, solving the formulation efficiently is not straightforward. Consequently, we develop an approximation solution approach. The requirements are twofold: (i) the solution method needs to be efficient and to run fast, and (ii) the approximation solution needs to be near optimal. In retail settings, retailers typically solve the (Multi-POP) problem for a large number of items. In addition, retailers often solve several instances of the problem in order to test the robustness of the solution before implementing it. More precisely, these routine tests are called *what-if scenarios*. They consist of solving perturbed versions of the nominal optimization problem, where some of the demand parameters and some of the business are rules are slightly modified (more details are discussed in Section 5.2). In what follows, we describe the solution approaches developed in Cohen et al. (2017) and in Cohen et al. (2018).

4.1. Single Item Setting

We first present an efficient solution approach to solve the single item problem. While the most interesting and relevant case is the problem with multiple items, the single item setting is used as a starting point for the presentation, and is interesting in its own right. In some retail categories, the different items can be independent, i.e., the demand of each item depends solely on its prices, and not on the prices of the other items. In this case, the (Multi-POP) problem decomposes in N

independent single item problems (assuming that there are no cross-item business rules), and one can solve each problem separately.

Even in the case of a single item, the problem is hard to solve (the problem is shown to be NP-hard in Cohen et al. 2016). We observe that the constraints in the (Multi-POP) formulation are linear. However, the objective function is nonlinear, and usually neither concave nor convex, as we do not want to impose restrictions on the form of the demand functions. This motivates us to propose a way to approximate the objective function by using a linear approximation, and by exploiting the discrete nature of the problem. In particular, we approximate the objective function by the sum of the marginal contributions of having a single promotion at a time. For example, if the item is on promotion at times 2, 3, and 7, we approximate the objective by the sum of the marginal deviations of having a single promotion at time 2, a single promotion at time 3, and a single promotion at time 7. We next present this approach, called *App*(1), in more detail.

The *App*(1) approximation method ignores the second-order interactions between promotions and captures only the direct effect of each promotion. Since we consider the same set of constraints as in the original problem, the solution remains feasible. We next introduce some additional notation. We consider a particular item, and hence we drop the superscript i in the remaining of this subsection. For a given price vector $\mathbf{p} = (p_1, \ldots, p_T)$, we define the corresponding total profit (of the item under consideration) throughout the season:

$$SPOP(\mathbf{p}) = \sum_{t=1}^{T} (p_t - c_t) d_t(\mathbf{p_t}).$$

Next, we define the price vector $\mathbf{p_t^k}$ such that the promotion price q^k is used at time t, and the regular price q^0 (no promotion) is used at all the remaining periods. We denote the regular price vector by $\mathbf{p^0} = (q^0, \ldots, q^0)$, for which the regular price is set at all the time periods. We define the coefficients b_t^k as follows:

$$b_t^k = SPOP(\mathbf{p_t^k}) - SPOP(\mathbf{p^0}). \tag{3}$$

The coefficients in (3) represent the unilateral deviations in the total profit by using a single promotion. One can compute these TK coefficients before starting the optimization procedure so that it does not affect the complexity of the method. The approximated objective function is then given by:

$$SPOP(\mathbf{p^0}) + \max_{\gamma_t^k} \sum_{t=1}^{T} \sum_{k=1}^{K} b_t^k \gamma_t^k, \tag{4}$$

while the set of constraints is the same as in the original problem. Consequently, the approximation optimization problem is linear, and can be solved using a solver. As mentioned before, two important requirements for our solution approach are (i) a low running time, and (ii) a close to optimal solution. We next summarize the properties (both theoretical and practical) for the single item setting.

Summary for the single item setting: We solve the promotion optimization problem for a single item by using the $App(1)$ approximation. This approximation linearizes the objective solution by computing the sum of the marginal contributions of each promotion separately. The following properties hold:

- The formulation is integral, i.e., one can solve the problem by considering the Linear Programming (LP) relaxation.
- Under two general demand models which are discussed below (multiplicative and additive), we derive a parametric worst-case bound on the quality of the approximation relative to the optimal profit.
- In many tested instances (calibrated with retail data), the approximation yields a solution which is optimal or very close to optimal.

We next discuss the implications of the above summary. Since one can get a solution by solving an LP, the approach is efficient (we can solve large instances in milliseconds). Consequently, the retailer can use this approach in practical settings. The approach works for general demand function, and for any objective function. If we further impose some structure on the demand function, we can derive a parametric bound on the quality of the approximation. We do so by considering two general classes of demand functions:

1. **Multiplicative demand:**

$$d_t\big(p_t, p_{t-1}, \ldots, p_{t-M}\big) = f_t(p_t) \cdot g_1(p_{t-1}) \cdot g_2(p_{t-2}) \cdots g_M(p_{t-M}), \tag{5}$$

   where the demand (of the item under consideration) can be written as the product of M functions that each depends on a single price. Note that since we consider a single item setting, the demand does not depend on the prices of the other items. The class of demand functions in (5) includes the log-log and the log-linear functions, which are commonly used in retail.

2. **Additive demand:**

$$d_t\big(p_t, p_{t-1}, \ldots, p_{t-M}\big) = f_t(p_t) + g_1(p_{t-1}) + g_2(p_{t-2}) + \ldots + g_M(p_{t-M}), \tag{6}$$

   where the demand (of the item under consideration) can be written as the sum of M functions that each depends on a single price. Note that the class of demand functions in (6) includes the linear function as a special case.

For the two classes of demand functions presented above, one can derive bounds on the quality of the $App(1)$ approximation. These bounds explicitly depend on the problem parameters, and depict a very high performance on all the instances we tested based on retail data. More details can be found in Cohen et al. (2017).

4.2. Multiple Items

In this section, we consider the more general setting where the retailer needs to decide the prices of N interconnected items by solving the (Multi-POP) problem. Recall that in this case, a promotion in item i may have an effect on the demand of item $j \neq i$. The cross-item effects on demand can be directly estimated from data. A potential simple approach can be the following: Solve the (Multi-POP) problem by applying the *App*(1) solution approach, i.e., approximate the objective by the sum of the marginal contributions of each item at each time period (as discussed in Section 4.1). We tested this approach, and observed a poor performance (especially in cases where the cross-item effects are significant). In particular, it fails to accurately capture the cross-item effects, and may find a promotion strategy far from optimal. For example, it may suggest to promote two items simultaneously, whereas this pair of items highly cannibalize each other. As a result, one needs to develop an alternative solution approach that can capture the cross-item effects, and at the same time remains efficient. We introduce the following sequence of methods, *App*(κ), for any given $\kappa = 1, 2, \ldots, N$.

- *App*(1) is the approximation applied to (Multi-POP) in a similar fashion as in the single item setting discussed in Section 4.1. In particular, it approximates the objective function by the sum of the marginal contributions of a single promotion for each item and time period separately. As we previously discussed, in the case of multiple items, it will generally provide a poor performance guarantee relative to the optimal solution.
- *App*(2) is an alternative approximation applied to (Multi-POP) that includes the marginal contributions (same as *App*(1)), as well as the pairwise contributions (i.e., having two items promoted at the same time). *App*(2) is described in full details below.
- *App*(N) is an alternative approximation that includes the marginal contributions, the pairwise contributions, and so on, up to all the possible combinations of having the N items promoted simultaneously.

One can also naturally consider any intermediate method for $2 < \kappa < N$. Note that there is a clear trade-off between simplicity (as well as speed) and performance (in terms of accuracy of the approximation relative to the optimal solution). On one extreme, *App*(1) is a simple approach that only requires computing the marginal contributions of having a single promotion at a time, but can perform poorly as it does not capture the cross-item effects at all. On the other extreme, *App*(N) is clearly more accurate, as it successfully captures all the cross-item effects. But this benefit comes at the expense of being more complex, as one needs to compute the marginal contribution of each possible combination of items that could be promoted simultaneously. In particular, it requires us to compute an exponential number of coefficients, and to solve an Integer Program (IP) that grows exponentially with the number of items. Note that when $T = 1$ or $M^i = 0 \ \forall i$, *App*(N) is exact as it captures accurately all the cross-item effects. Nevertheless, for a general dynamic problem with $T>1$ periods and non-zero memory parameters, *App*(N) is still not an exact algorithm, as it approximates the time effects induced by the past prices. We next describe *App*(2) in more details as it will be used in the sequel.

As we previously mentioned, *App*(2) approximates the objective of (Multi-POP) by the sum of unilateral deviations (i.e., having a single promotion at a time) and the pairwise contributions (i.e., having two items promoted simultaneously). More precisely, the approximated objective is:

$$MPOP(\mathbf{p^0})+\ \max_{\gamma}\Big\{\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{k=1}^{K^i}b_t^{ki}\gamma_t^{ki}+\sum_{i,j:i>j}^{N}\sum_{t=1}^{T}\sum_{k=1}^{K^i}\sum_{\ell=1}^{K^j}b_t^{k\ell ij}\gamma_t^{k\ell ij}\Big\}, \tag{7}$$

where the coefficients b_t^{ki} and $b_t^{k\ell ij}$ are formally defined in equations (8) and (9) respectively. We denote the regular price vector by $\mathbf{p^0}=(q^0,\dots,q^0)$, which means that the regular price is set for all items at all times. The first term, denoted by $MPOP(\mathbf{p^0})$, represents the total profit generated by all the items throughout the selling season, without any promotion. The second term captures all the marginal contributions of having a single promotion, i.e., for one item at one time period. More precisely, we define the price vector $\mathbf{p_t^{kj}}$ as follows:

$$(\mathbf{p_t^{kj}})_\tau=\begin{cases}q^{kj}; & \text{if } \tau=t \text{ and } i=j,\\ q^0; & \text{otherwise.}\end{cases}$$

In other words, the vector $\mathbf{p_t^{kj}}$ has the promotion price q^{kj} for item j at time t, and the regular price q^0 (no promotion) is used at all the remaining periods for item j, and for all the other items at all times. The coefficient b_t^{kj} is then given by:

$$b_t^{kj}=MPOP(\mathbf{p_t^{kj}})-MPOP(\mathbf{p^0}), \tag{8}$$

and represents the marginal contribution in the total profit by having a single promotion for item j at time t, using price q^{kj}.

The third term in equation (7) represents all the pairwise contributions of having two items on promotion at the same time. More precisely, we define the price vector $\mathbf{p_t^{k\ell ju}}$ for any pair of items $j>u$ as follows:

$$(\mathbf{p_t^{k\ell ju}})_\tau=\begin{cases}q^{kj}; & \text{if } \tau=t \text{ and } i=j,\\ q^{\ell u}; & \text{if } \tau=t \text{ and } i=u,\\ q^0; & \text{otherwise.}\end{cases}$$

In other words, the vector $\mathbf{p_t^{k\ell ju}}$ uses the promotion price q^{kj} for item j at time t, the promotion price $q^{\ell u}$ for item u at time t, and the regular price q^0 for items j and u in all the remaining periods, and for all the other items at all times. The coefficient $b_t^{k\ell ju}$ is given by:

$$b_t^{k\ell ju}=MPOP(\mathbf{p_t^{k\ell ju}})-MPOP(\mathbf{p_t^{kj}})-MPOP(\mathbf{p_t^{\ell u}})+MPOP(\mathbf{p^0}), \tag{9}$$

and represents the marginal pairwise contribution in the total profit by having two simultaneous promotions. Finally, in order to make the formulation consistent, we should ensure that when both items i and j are on promotion, we count the pairwise contribution but also both unilateral deviations, i.e., for each pair of items i and $j < i$, $\gamma_t^{ki} = \gamma_t^{\ell j} = 1$ if and only if $\gamma_t^{k\ell ij} = 1$ for each t and k, ℓ. One can encode this set of conditions by incorporating the following constraints to the formulation for each pair of items i, $j < i$, each t, and each promotion prices q^{ki} and $q^{\ell j}$:

$$\begin{aligned}
&\gamma_t^{k\ell ij} \leq \gamma_t^{ki}, \\
&\gamma_t^{k\ell ij} \leq \gamma_t^{\ell j}, \\
&\gamma_t^{k\ell ij} \geq 0, \\
&\gamma_t^{k\ell ij} \geq \gamma_t^{ki} + \gamma_t^{\ell j} - 1.
\end{aligned}$$

When maximizing the objective of the approximated problem in equation (7), the decisions are the binary variables γ. In particular, there is one such variable for each item/time/price (i.e., $NT(K+1)$, assuming for simplicity that $K^i = K \ \forall i$), and one such variable for any pair of items $i > j$ at each time/price (i.e., $\frac{N(N-1)}{2} TK^2$). As we previously mentioned, for *App*(N), this number grows exponentially with N and K and hence, it may not be practical to go beyond *App*(3) or *App*(4). We next summarize the main results for the multiple item setting.

Summary for the multiple item setting: We solve the promotion optimization problem for multiple items by using the *App*(2) approximation. The following properties hold:

- Assuming that the cross-item effects for each item are additively separable, i.e.,

$$d_t^i\big(p_t^i, p_{t-1}^i, \ldots, p_{t-M^i}^i, \mathbf{p_t^{-i}}\big) = h_t^i\big(p_t^i, p_{t-1}^i, \ldots, p_{t-M^i}^i\big) + \sum_{j \neq i} H_t^{ji}(p_t^j), \tag{10}$$

  then *App*$(2) =$ *App*$(3) = \ldots =$ *App*(N).

- If we further assume that the function $h_t^i\big(p_t^i, p_{t-1}^i, \ldots, p_{t-M^i}^i\big)$ is additively separable for each item, i.e.,

$$h_t^i\big(p_t^i, p_{t-1}^i, \ldots, p_{t-M^i}^i\big) = f_t^i(p_t^i) + g_1^i(p_{t-1}^i) + \ldots + g_{M_i}^i(p_{t-M^i}^i), \tag{11}$$

  then the *App*(2) solution is optimal.

- Consider the class of demand models in (10) and $K^i = 1$ (i.e., the regular price and a single promotion price). For substitutable items, the *App*(2) formulation can be solved efficiently in the absence of business rules.

- Under two general demand models (multiplicative and additive price dependence), we derive a parametric bound on the quality of the approximation relative to the optimal profit.

- In many tested instances, the approximation yields a solution which is optimal or very close to optimal.

We next discuss the implications of the above summary. Interestingly, for demand functions with additively separable cross-item effects (in practice, several demand models satisfy this property), it is sufficient to consider $App(2)$ as opposed to include higher order terms. In the special case where each item can take two prices, the $App(2)$ approximation can be solved efficiently when all the items are substitutable. Having two prices is common in practice as the promotion price is often negotiated upfront with the manufacturer. In the more general case, where the retailer can choose among several promotion prices, we observed computationally that one can still solve the IP within low runtimes for realistic size instances. It is worth mentioning that for most categories of supermarket items, the products within a category are either independent (i.e., no cross-item effects) or substitutable. In particular, for categories such as coffee, tea and chocolate, we could not find any complementarity effects in the data we analyzed. Note also that even if some of the products are complement, we observed by extensive testing that solving the relaxation of the $App(2)$ formulation yields an optimal integer solution very often. More details on such computational tests are presented in Cohen et al. (2018).

5. Insights and Practical Impact

In this section, we summarize the insights we have been able to draw by solving the (Multi-POP) using our solution approach. We then describe how to concretely apply our model to a real-world retail setting.

5.1. Insights

We briefly discuss several insights that were drawn by using our promotion optimization model. Very often, retailers want to infer the impact of promoting the different items at the different time periods. Our solution approach can easily be used to test various promotion strategies in order to reach a better understanding on the impact of retail promotions. As we mentioned before, several economic factors are present in the context of our problem: the cross-item effects on demand, the post-promotion dip effect, the seasonality, and the presence of business rules. It is definitely valuable for the retailer to learn the tradeoffs between these different effects, and to understand how they impact the promotion decisions. Our model can help retailers to deepen their knowledge on the following topics:

- **Understanding the structure of the cross-item effects:** In a given category of items, the retailer needs to decide the price promotions by accounting for the cross-item effects on demand. Using our model, the retailer can infer the impact of promoting a specific item on the demand of each item in the category. This can ultimately allow the retailer to carefully decide which set of items should be promoted simultaneously, and which should not. For example, when two (or more) items have strong substitution effects (i.e., promoting an item increases its own sales, but also significantly decreases the sales of the other items), the retailer should not promote those items. More details on such insights can be found in Cohen et al. (2018).

- **Inferring the strength of the post-promotion dip effect:** It is well known that promoting a FMCG product induces a boost in its current demand as well as a potential decrease in its future demand, due to the stockpiling behavior of consumers and the promotion fatigue effect. The strength of the post-promotion dip effect can vary significantly depending on the category under consideration. For example, in Cohen et al. (2017), the authors found that the number of past prices that affect the current demand (which is one possible way to measure the post-promotion dip effect) highly depends on the item and on the category. For example, the post-promotion dip effect tends to be weaker for perishable products and for luxury/expensive brands, as expected.
- **Identifying the presence of a loss leader effect:** The loss leader is a common phenomenon in which one item is priced below its cost in order to extract significant profits on complementary items (see, e.g., Hess and Gerstner 1987). It is reported in Cohen et al. (2018) that the model considered in this chapter can identify the presence of a loss-leader effect. This can be a very important information for the retailer that can use one (or more) items in order to profitably trigger a loss-leader strategy.
- **Learning the impact of the business rules:** As discussed in Section 4, the retailer can easily solve several instances of the problem, with and without the presence of some of the business rules. Consequently, this allows the retailer to quantify the profit impact of relaxing some of the business requirements. This can ultimately help retailers decide which vendor funds to accept and under what terms.

In practice, retailers often solve the (Multi-POP) for large scale instances that involve a large number of different factors. It does not seem possible for managers, as experienced as they are, to understand and anticipate the impact of all the conflicting tradeoffs. Using an optimization tool calibrated with actual data can take into account all the different tradeoffs, and compute a close to optimal solution for the promotion planning problem.

5.2. Practical Impact

We next consider a concrete application of the (Multi-POP) optimization problem. We propose a generic process that can be used by any retailer who seeks to improve its promotion planning decisions. This process consists of the five following steps:

1. Data collection, cleaning, and aggregation,

2. Store and product clustering,
3. Demand estimation,
4. Optimization and sensitivity analysis,
5. Quantifying the impact.

We next describe each step in more details.

Data collection, cleaning, and aggregation: The first step is to collect and store the relevant data. In our context, retailers need to simply collect the data from the past transactions. Each observation typically includes: the store, the date/time, the items purchased, the prices, the promotion vehicles that were used, as well as various features of the item (brand, size, flavor, etc.). After gathering a large enough dataset, one needs to carefully clean the data, and perform the appropriate aggregations. Various techniques exist for cleaning and aggregating data but this is beyond the scope of this chapter (see, e.g., the book by Han et al. 2011). At a high level, one wants to deal with the missing data, remove the outliers, and perform some basic statistical tests. Once the data is cleaned, one needs to decide the level of aggregation. Depending on the context, one can perform the analysis at the brand, item, or category level. Similarly, one can aggregate the data at the week, day, or hour level. Once the data is cleaned and aggregated at the right level, one can start using it for estimation and prediction purposes. For example, in Cohen et al. (2017), the authors decided to aggregate the data at the brand-week level.

Store and product clustering: In many retail settings, the available historical data can be sparse. As a result, one needs to combine the data from multiple sources in order to obtain more reliable forecasts. Two common techniques widely used in retail consist of merging several stores together or clustering similar products. The idea is to use the data from several stores that share similar features (e.g., geographical location, size, management team). Similarly, items from the same brand (e.g., different sizes or flavors) can often be clustered together so as to improve the prediction accuracy of the models.

Demand estimation: This step is the actual first stage of using our promotion optimization model. As an input to the optimization, one first needs to estimate the demand models. The modeler has several degrees of freedom: choice of the demand function (e.g., log-log, log-linear), selection of the dependent variables, choice of the instrumental variables (if any), and choice of the estimation procedure. In many applications, one can simply run a linear regression (e.g., ordinary or weighted least squares, ridge regression, lasso). The typical process also includes splitting the data for out-of-sample testing. The demand estimation step is completed once the prediction models yields a high accuracy out-of-sample. In practice, one needs to test different models and assumptions in order to reach a good and robust prediction model. In Cohen et al. (2017), the authors present a prediction model for two coffee brands based on using ordinary least squares to predict a log-log model that includes past prices, weekly seasonality, and the trend effect. The resulting out-of-sample R^2 (resp. MAPE) was around 0.90 (resp. 0.11).

Optimization and sensitivity analysis: Once the demand models are accurately estimated from data, one can use them as an input to the (Multi-POP). The retailer also needs to specify the various business rules that need to be satisfied, the number of time periods in the selling season, and the cost of each item at each time period. At this point, one can use the $App(2)$ approximation method presented in Section 4.2 to solve the problem. As discussed before, this yields a near optimal solution by computing the price promotions of all the items during each period of the selling season. Usually, retailers want to check the robustness of the solution prior to a potential implementation. To this end, one can resolve the (Multi-POP) by perturbing several input parameters (e.g., estimated demand coefficients, business rules parameters). If the suggested solution appears to be robust with respect to variations in the problem input, this provides a higher confidence on the validity of the solution.

Quantifying the impact: The last step is to assess the potential impact of the entire process. For example, one can compare the simulation results obtained by using the optimized promotion prices relative to the profit generated by using the original promotion prices set by the retailer. In our experience, by applying our model to several retailers, we observed a profit improvement of 2-9%, depending on the product category and the store under consideration.

6. Conclusions

Retail promotions are important decisions faced by most retailers. Promoting the right set of items at the right time using the right price points can have a significant impact on the retailer's bottom line. In settings such as supermarkets, the retailer needs to simultaneously decide the price promotions for multiple items throughout the selling season. Historically, many retailers were designing their promotion strategy based on past experience and on trial-and-error attempts. The unprecedented volume of available data has now changed the picture. Using past transactions data, retailers can improve the demand forecasting accuracy, and exploit the data to develop quantitative tools for promotion planning. In particular, the combination of data analytics and optimization allows retailers to decide promotions in a more systematic and profitable fashion. In this chapter, we considered a retailer selling FMCG products who needs to decide the (mass) price promotions for all the items in a category. We first formulated the problem as a nonlinear integer optimization program. This formulation holds under general demand functions estimated from data, and includes several practical business rules on price promotions. Given that the resulting formulation is hard to solve, we presented an approximation solution approach. This approach can solve the problem efficiently in short timeframes, and admits a parametric worst-case bound on the quality of the approximation. We first considered the single item setting, and then extended the presentation to the more general instance with multiple items. In each case, we presented the model, the approximation solution approach, and its analytical and practical properties. Overall, the methods presented in this chapter run fast, and provide a near optimal solution for many tested instances (calibrated with actual retail data).

We then summarized an application of this model to a real-world setting. In particular, we proposed a beginning-to-end process for retailers that consists of five steps: (1) Data collection, cleaning, and aggregation, (2) Store and product clustering, (3) Demand estimation, (4) Optimization and sensitivity analysis, and (5) Quantifying the impact. By following these steps, retailers can potentially improve their promotion planning process. In our own experience, we observed a profit improvement of 2-9%, which is a significant impact in the retail industry.

While most of the results presented in this chapter are borrowed from previous publications (mainly from Cohen et al. 2017, 2018), it provides a summary of this line of research by presenting the different complementary studies in a single report. This chapter has focused on the mass pricing promotion optimization problem faced by a retailer who sells FMCG products. As we mentioned in Section 1, several alternative promotion problems are also important in retail. Interesting research directions can be the development of new decision tools based on data analytics and optimization for those practical retail problems.

References

Ahn Hs, Gümüş M, Kaminsky P (2007) Pricing and manufacturing decisions when demand is a function of prices in multiple periods. *Operations Research* 55(6):1039–1057.

Ailawadi KL, Gedenk K, Lutzky C, Neslin SA (2007) Decomposition of the sales impact of promotion-induced stockpiling. *Journal of Marketing Research* 450–467.

Andrews M, Luo X, Fang Z, Ghose A (2015) Mobile ad effectiveness: Hyper-contextual targeting with crowdedness. *Marketing Science* 35(2):218–233.

Arora N, Dreze X, Ghose A, Hess JD, Iyengar R, Jing B, Joshi Y, Kumar V, Lurie N, Neslin S, et al. (2008) Putting one-to-one marketing to work: Personalization, customization, and choice. *Marketing Letters* 19(3-4):305.

Aviv Y, Pazgal A (2008) Optimal pricing of seasonal products in the presence of forward-looking consumers. *Manufacturing & Service Operations Management* 10(3):339–359.

Baardman L, Cohen MC, Panchamgam K, Perakis G, Segev D (2017a) Scheduling promotion vehicles to boost profits, forthcoming in Management Science.

Baardman L, Panchamgam K, Perakis G (2017b) Pass-through constrained vendor funds for promotion planning, working paper.

Balinski ML (1970) On a selection problem. *Management Science* 17(3):230–231.

Bertsimas D, Shioda R (2009) Algorithm for cardinality-constrained quadratic optimization. *Computational Optimization and Applications* 43(1):1–22.

Besbes O, Lobel I (2015) Intertemporal price discrimination: Structure and computation of optimal policies. *Management Science* 61(1):92–110.

Bienstock D (1996) Computational study of a family of mixed-integer quadratic programming problems. *Mathematical programming* 74(2):121–140.

Bitran G, Caldentey R (2003) An overview of pricing models for revenue management. *Manufacturing & Service Operations Management* 5(3):203–229.

Blattberg RC, Neslin SA (1990) *Sales promotion: Concepts, Methods, and Strategies* (Prentice Hall).

Cachon GP, Swinney R (2009) Purchasing, pricing, and quick response in the presence of strategic consumers. *Management Science* 55(3):497–511.

Campo K, Gijsbrechts E, Nisol P (2000) Towards understanding consumer response to stock-outs. *Journal of Retailing* 76(2):219–242.

Caro F, Gallien J (2012) Clearance pricing optimization for a fast-fashion retailer. *Operations Research* 60(6):1404–1422.

Chen Y, Farias VF (2015) Robust dynamic pricing with strategic customers. *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, 777–777.

Cohen MC, Gupta S, Jeremy, Perakis G (2016) An efficient algorithm for dynamic pricing using a graphical representation, working paper.

Cohen MC, Kalas J, Perakis G (2018) Optimizing promotions for multiple items in supermarkets, working paper.

Cohen MC, Leung NHZ, Panchamgam K, Perakis G, Smith A (2017) The impact of linear optimization on promotion planning. *Operations Research* 65(2):446–468.

Cooper LG, Baron P, Levy W, Swisher M, Gogos P (1999) Promocast: A new forecasting method for promotion planning. *Marketing Science* 18(3):301–316.

Corsten D, Gruen T (2004) Stock-outs cause walkouts. *Harvard Business Review* 82(5):26–28.

Elmaghraby W, Keskinocak P (2003) Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management science* 49(10):1287–1309.

Felgate M, Fearne A (2015) Analyzing the impact of supermarket promotions: A case study using tesco clubcard data in the uk. *The Sustainable Global Marketplace*, 471–475 (Springer).

Ferreira KJ, Lee BHA, Simchi-Levi D (2015) Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing & Service Operations Management* 18(1):69–88.

Foekens EW, SH Leeflang P, Wittink DR (1998) Varying parameter models to accommodate dynamic promotion effects. *Journal of Econometrics* 89(1):249–268.

Fong NM, Fang Z, Luo X (2015) Geo-conquesting: Competitive locational targeting of mobile promotions. *Journal of Marketing Research* 52(5):726–735.

Frank M, Wolfe P (1956) An algorithm for quadratic programming. *Naval Research Logistics (NRL)* 3(1-2):95–110.

Gedenk K, Neslin SA, Ailawadi KL (2006) Sales promotion. *Retailing in the 21st Century*, 345–359 (Springer).

Grossmann IE (2002) Review of nonlinear mixed-integer and disjunctive programming techniques. *Optimization and Engineering* 3(3):227–252.

Han J, Pei J, Kamber M (2011) *Data mining: concepts and techniques* (Elsevier).

Heilman CM, Nakamoto K, Rao AG (2002) Pleasant surprises: Consumer response to unexpected in-store coupons. *Journal of Marketing Research* 39(2):242–252.

Hemmecke R, Köppe M, Lee J, Weismantel R (2010) Nonlinear integer programming. *50 Years of Integer Programming 1958-2008*, 561–618 (Springer).

Hess JD, Gerstner E (1987) Loss leader pricing and rain check policy. *Marketing Science* 6(4):358–374.

Jagabathula S, Mitrofanov D, Vulcano G (2018) Customized individual promotions: Model, optimization, and prediction, working paper.

Levin Y, McGill J, Nediak M (2010) Optimal dynamic pricing of perishable items by a monopolist facing strategic consumers. *Production and Operations Management* 19(1):40–60.

Levina T, Levin Y, McGill J, Nediak M (2009) Dynamic pricing with online learning and strategic consumers: an application of the aggregating algorithm. *Operations Research* 57(2):327–341.

Liu Y, Cooper WL (2015) Optimal dynamic pricing with patient customers. *Operations Research* 63(6):1307–1319.

Macé S, Neslin SA (2004) The determinants of pre- and postpromotion dips in sales of frequently purchased goods. *Journal of Marketing Research* 41:339–350.

Martínez-Ruiz MP, Mollá-Descals A, Gómez-Borja MA, Rojo-Álvarez JL (2006) Assessing the impact of temporary retail price discounts intervals using svm semiparametric regression. *International review of retail, distribution and consumer research* 16(02):181–197.

Mersereau AJ, Zhang D (2012) Markdown pricing with unknown fraction of strategic customers. *Manufacturing & Service Operations Management* 14(3):355–370.

Nielsen (2015) The path to efficient trade promotions. URL `http://www.nielsen.com/us/en/insights/reports/2015/the-path-to-efficient-trade-promotions.html`.

Nijs V, Misra K, Anderson ET, Hansen K, Krishnamurthi L (2010) Channel pass-through of trade promotions. *Marketing Science* 29(2):250–267.

Nocedal J, Wright SJ (2006) *Sequential quadratic programming* (Springer).

Özer Ö, Phillips R (2012) *The Oxford handbook of pricing management* (Oxford University Press).

Padberg M (1989) The boolean quadric polytope: some characteristics, facets and relatives. *Mathematical programming* 45(1):139–172.

Pekgün P, Menich RP, Acharya S, Finch PG, Deschamps F, Mallery K, Van Sistine J, Christianson K, Fuller J (2013) Carlson rezidor hotel group maximizes revenue through improved demand management and price optimization. *Interfaces* 43(1):21–36.

Rhys J (1970) A selection problem of shared fixed costs and network flows. *Management Science* 17(3):200–207.

Sherali HD, Adams WP (1998) *A reformulation-linearization technique for solving discrete and continuous nonconvex problems*, volume 31 (Springer).

Silva-Risso JM, Bucklin RE, Morrison DG (1999) A decision support system for planning manufacturers' sales promotion calendars. *Marketing Science* 18(3):274–300.

Su X (2010) Intertemporal pricing and consumer stockpiling. *Operations research* 58(4-part-2):1133–1147.

Talluri KT, Van Ryzin GJ (2006) *The theory and practice of revenue management*, volume 68 (Springer Science & Business Media).

Vakhutinsky A, Kushkuley A, Gupte M (2012) Markdown optimization with an inventory-depletion effect. *Journal of Revenue and Pricing Management* 11(6):632–644.

Van Donselaar K, Van Woensel T, Broekmeulen R, Fransoo J (2006) Inventory control of perishables in supermarkets. *International Journal of Production Economics* 104(2):462–472.

Van Heerde HJ, Gupta S, Wittink DR (2003) Is 75% of the sales promotion bump due to brand switching? no, only 33% is. *Journal of Marketing Research* 40(4):481–491.

Yin R, Aviv Y, Pazgal A, Tang CS (2009) Optimal markdown pricing: Implications of inventory display formats in the presence of strategic customers. *Management Science* 55(8):1391–1408.

Yuan H, Gómez MI, Rao VR (2013) Trade promotion decisions under demand uncertainty: A market experiment approach. *Management Science* 59(7):1709–1724.

Zhang D, Cooper WL (2008) Managing clearance sales in the presence of strategic customers. *Production and Operations Management* 17(4):416–431.