

A Practical Price Optimization Approach for Omnichannel Retailing

Pavithra Harsha

IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, pharsha@us.ibm.com

Shivaram Subramanian

IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, subshiva@us.ibm.com

Markus Ettl

IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, msettl@us.ibm.com

Consumers are increasingly navigating across sales channels to maximize the value of their purchase. The existing retail practices of pricing channels independently at the one end, and matching channel and competitor prices at the other, are unable to achieve the desired profitable coordination required between channels. As part of a joint partnership agreement with IBM Commerce, we engaged with three major retailers over two years, and developed advanced omnichannel pricing (OCP) solutions that are used by several retail chains today. We develop an omnichannel framework to model location-specific cross-channel demand interactions. An integrated OCP optimization formulation profitably coordinates prices for non-perishable products across channels and store locations, while satisfying a variety of business rules, and taking into account the impact of competition, and sales goals. We present the OCP formulations with and without assortment effects. The resultant non-linear models are non-convex and NP-hard, and practically efficient optimization approaches are prescribed, along with computational results using real-world data. An OCP implementation for a large retail chain projected a 7% profit lift. IBM Commerce deployed proprietary versions of these models into production in 2014.

Key words: Omnichannel, pricing, assortment, attraction demand, nested model, regular pricing, cross-channel effects, elasticity

1. Introduction

Omnichannel retailing is a recent trend sweeping companies across the industry ([Brynjolfsson et al. 2013](#), [Bell et al. 2014](#)). It aims to revolutionize how companies engage with consumers by creating a seamless customer shopping experience across the retailer’s multiple sales channels. This is because today’s consumers navigate across the channels with ease to make purchases. Using smart phones, in-store shoppers can visit the mobile or web store of the same retailer or its competitors to find better deals and finalize a purchase. Omnichannel retailing also includes the use of advanced order fulfillment practices such as initiating ship-from-store fulfillment for e-commerce orders and offering a buy-online-pick-up-in store fulfillment option to enhance the convenience of receiving a product. Retailers require such capabilities to survive in a highly competitive marketplace that is

witnessing a fast pace of online sales growth ([US Census Bureau 2017](#) reported that online sales grew 14-16% compared to the previous year) and an ever-increasing market-share gain by e-tailers whose price-transparent product offerings eat into e-tail margins as well as store sales (e.g. due to 'showrooming').

Many of today's large retailers started as single channel retailers and their supply chain was designed to ensure maximum efficiency and scale in that channel. These retailers subsequently opened additional sales channels, and supported common and channel-specific assortments, to increase their customer base. However, these channels largely operated independently of each other in 'silos', with limited transparency and data integration even within the organization. From the perspective of operations research technologies, many retailers today maintain separate brick and online merchandizing divisions, and employ decision support tools for demand forecasting, pricing optimization, and inventory management that are channel specific, often procured from different vendors. Such tools largely ignore the multi-channel shopping path of today's customers, as well as the potential efficiencies of omnichannel retailing. The headline story from 'Black Friday' sales reports is that many large retail chains are beginning to see the value of an integrated approach to merchandizing ([Retail Dive 2015](#)).

This work is focused on developing a novel solution that overcomes some of these challenges by integrating key decisions, specifically pricing, across the different sales channels. Consider for example, a retail chain that operates two sales channels (say brick-and-mortar stores, and online). The retailer is faced with the question of how to price products across channels and locations for omnichannel consumers. Following the existing retail practices of pricing channels independently or matching channel and competitor prices cannot achieve the desired profitable coordination required between channels. First, the retailer has to model consumer channel switching behavior. Next, the store and online prices must be coordinated to be profitable, while also meeting global sales volume goals, brand-price image goals for products within an assortment, etc., and remain competitive with large e-tail giants. Given the increase in the number of digital channel offerings and the dynamic nature of the marketplace, the scale and speed of executing pricing decisions are equally critical. The integral omnichannel pricing solution that we have proposed and developed for the retail industry in this paper addresses these questions.

As part of a joint partnership agreement with [IBM Commerce](#), a leading provider of merchandizing solutions, we engaged with three major omnichannel retailers over a period of two years, who faced many of the challenges described above. Our goal was to develop advanced omnichannel retail analytics solutions for commercial use by current and future retail customers of IBM Commerce, keeping in mind the infrastructure and operational requirements of a deployable solution. The following are the contributions of the paper.

1. **Omnichannel demand modeling framework:** We develop a framework to model sales channels as substitutable purchase choices for consumers, while also capturing the heterogeneity of channel preferences across locations. With this framework, we can quantify own- and cross-channel demand interactions, especially, the impact of the innumerable number of store prices on online demand. This method overcomes the limitations of the legacy approach that treated the online channel as simply another store location. We use this framework to predict channel-specific demand by location.
2. **Omnichannel price optimization:** We study the omnichannel pricing (OCP) of nonperishable (basic) products in the context of developing a regular pricing solution (also referred to as base pricing). We formulate and solve an integrated price optimization problem across multiple channels and locations subject to certain practically important channel, volume, and price image goals.
 - (a) *Single product pricing:* We use attraction demand models to represent consumer preference across sales channels and observe that the resulting optimization model is a non-linear, non-convex NP-Hard problem due to certain complicating pricing constraints and the multi-location aspect of the problem. We employ specialized mathematical transformations to recover a computationally tractable and exact mixed-integer programming (MIP) reformulation that can be solved to (near) optimality using commercial off-the-shelf MIP solvers such as CPLEX. In certain special cases of two channel operations such the brick-and-mortar and online channels, we propose a pseudo polynomial decomposition method to OCP and derive an insightful price coordination result.
 - (b) *Assortment pricing:* Using nested attraction demand models, we additionally manage cross-product demand interactions within channel assortments. We exploit the concave structure of the nest attraction function and obtain a tractable MIP formulation that admits a variety of additional cross-product constraints. We show that this MIP yields an effective upper bound and can achieve (near) global optimal solutions.
3. **Implementation and business value assessment:** We performed a business value assessment as a part of our OCP implementation for one of the major omnichannel retailers in the United States. For 100 products in the two product categories that we analyzed, we found that the degree of cross-channel price sensitivity to demand can be up to 50% of the own channel price elasticity. We solved the resulting OCP formulation and obtained a projected profit lift of 7% using omnichannel pricing over their legacy pricing system, while also satisfying several other critical business and pricing goals. These results were presented at the retailer's site to a senior team that included their Vice-President for revenue management. Their response was overwhelmingly positive, and we had a similar experience with the other retailers. IBM

Commerce followed it up with an internal evaluation and deployed a proprietary version into production in 2014 as a cloud solution. This solution was showcased as one of the retail analytics success stories in the smarter-commerce global summit in 2014, and included a presentation on its capabilities by the retailer.

The remainder of the paper is structured as follows. In [Section 1.1](#) we review related literature. In [Section 2](#) we describe the omnichannel demand modeling framework. We discuss the omnichannel demand model without cross effects in [Section 3](#), and in [Section 4](#), we formulate the resultant price optimization problem and discuss its tractability. This is followed by the different solution methods in [Section 5](#). In [Section 6](#) we analyze the general omnichannel assortment pricing problem. In [Section 7](#) we discuss the OCP implementation and the business value assessment presented to the retail customer on their data. We conclude in [Section 8](#) with a brief discussion of data flow in the commercial deployment along with other practical use-cases of OCP, and some post-deployment highlights.

1.1. Literature Review

Studying and modeling the consumer preferences in an omnichannel environment is a first step that can pave the way for coordination of channel strategies ([Neslin et al. 2006](#)). Some recent papers in the marketing literature have explored consumer dynamics in a multi-channel environment, in particular, consumer migration across multiple channels such as web and catalog ([Ansari et al. 2008](#)) or online and brick channels ([Chintagunta et al. 2012](#)) respectively. [Goolsbee \(2001\)](#) finds significant cross-price elasticity between online sales and stores sales of computers. The decision to buy online is dependent on the store prices, and hence they conclude that channels cannot be treated as separate markets. These papers adopt logit models to calibrate the substitution behavior of consumers between two channels. Similar to the consideration in the above papers, we estimate consumer channel preferences using discrete choice models. We also estimate location-specific models, as price elasticities are not uniform across locations but vary depending on an area's household income, demography, and other factors ([Mulhern et al. 1998](#)).

From an operational perspective, there is substantial academic literature that focus on single and multiproduct pricing problems (for example, see the survey papers by [Bitran and Caldentey 2003](#), [Elmaghraby and Keskinocak 2003](#), [Chen and Simchi-Levi 2012](#)). To the best of our knowledge, the focus has been on single channel pricing and largely restricted to a single location. Commercially available pricing solutions employed by retailers today do not incorporate cross-channel interactions and often price the multiple channels and locations of a retailer separately or sequentially. In contrast, the focus of this paper is on an integrated multi-channel and multi-location pricing

problem in the presence of cross-channel and cross-product demand interactions and important operational considerations. [Montgomery \(1997\)](#) estimates the incremental gain achievable through location-specific pricing over uniform chain-level prices using data from supermarkets. [Zhang et al. \(2010\)](#) motivate that firms must strike a delicate balance between consumers' expectations of prices in different channels and the cost structure of each channel

From the perspective of price optimization using customer choice models, several papers in the literature have analyzed a variety of parametric and non-parametric approaches in the context of cross-product demand substitution. For the multinomial logit (MNL) demand model, [Hanson and Martin \(1996\)](#) show that the profit as a function of the prices is not *quasi-concave*. [Aydin and Porteus \(2008\)](#), [Akçay et al. \(2010\)](#) explored this problem further and show that the resultant profit function is unimodal in the price space. Meanwhile, [Song and Xue \(2007\)](#), [Dong et al. \(2009\)](#) proposed a market share variable transformation to demonstrate that the objective function is jointly concave in the space of the market share variables. This transformation idea for MNL demand models was later extended to a general class of attraction models by [Schön \(2010\)](#), and [Keller et al. \(2014\)](#).

Our work on single product pricing relates to the price optimization problem using mixture of attraction demand models. [Keller et al. \(2014\)](#) point that this is an open problem and develop a local optimal heuristic solution by employing an approximate demand model by assuming convexity. On the other hand, for the non-convex OCP problem we analyze, we develop tractable global optimization methods to solve large-scale problem instances that arise in practical omnichannel operations.

A few papers have explored the use of other demand models in the context of multi-item pricing problems. For example, pricing nested logit demand models has been studied by [Li and Huh \(2011\)](#), [Gallego and Wang \(2014\)](#), [Rayfield et al. \(2015\)](#) among others. In the presence of varying price elasticities across items, and unconstrained prices, [Gallego and Wang \(2014\)](#) show that the resultant problem can be computationally intractable as a transformed model is non-convex. [Rayfield et al. \(2015\)](#) provide an approximation method to this problem by discretizing the intrinsic value of a nest. [Davis et al. \(2016\)](#) study this problem under discrete prices using a specific type of price ordering and show that the number of feasible price vectors in a given nest is polynomial in the number of choices and price levels, and derive an exact solution. In the assortment OCP problem that we study, we use a nested attraction function similar to the above papers, but our problem includes certain required operational constraints as well as the multi-location setting (like a mixture of nested attraction demand models). [Subramanian and Sherali \(2010\)](#) studies a multi-item pricing using a hybrid MNL demand model and provide a piecewise linear approximation that manages a variety of practical business rules by simultaneously working in the price and market share space.

Non-parametric approaches to multi-item pricing have been explored by [Rusmevichientong et al. \(2006\)](#) and [Aggarwal et al. \(2004\)](#) using heuristic approaches and approximation algorithms.

2. Demand modeling framework to quantify cross-channel substitution

In the omnichannel environment customers navigate across channels and retailers to finalize a purchase that maximizes their own benefit. Therefore, a fundamental aspect that omnichannel demand models should aim to capture is the channel switching behavior of consumers, in other words, the cross-channel substitution effects. From a demand modeling stand point, an omnichannel demand model should include cross-channel causals in addition to the conventional same-channel causals. For illustration, assuming price as the only driver of demand and the set J denoting the brick store locations,

$$D_{B_j} := D_{B_j}(p_{B_j}, p_O) \quad \forall j \in J, \quad D_O := D_O(p_O, p_{B_1}, p_{B_2}, \dots) \quad (2.1)$$

where D_{B_j}, D_O and p_{B_j}, p_O are the demands and prices for brick-store location B_j and the online channel O respectively. Here, store demand is a function of its own physical store price and the online price, while the online channel demand is a function of the online price and *all* the brick-and-mortar store prices. The latter is because the online channel virtually connects all the physical stores.

Legacy forecasting systems currently do not model such cross-channel effects. This allows the legacy pricing system to conveniently treat the online channel just as an additional independent store location. The legacy forecasting systems can estimate the impact of online causals on any store demand by including the online causals as modeling features (see [Eq. \(2.1\)](#)), but they fail to accurately quantify the interactions in the reverse direction. This is because of the sheer number of physical store locations (ranging from several hundreds to a few thousand), and the usage of location-specific pricing makes this task impractical. It also precludes an accurate quantification of location-specific impact of store causals on the online demand. Note that besides price, there are a variety of other location-specific demand influencers such as store promotions, competitors and local events, whose cross effects are additionally important to capture in an omnichannel demand model.

We overcome this challenge by employing a geographical partitioning of the online store (transactions of which originate from a continuum of customer zip-codes) into discrete virtual online stores (virtual stores, for brevity). We follow this with the assumption that the customers within the zip-codes associated with any virtual store, choose to purchase from this virtual store or the

physical store(s) in that location and are not influenced by the prices (and other causals) in other physical store locations, i.e.,

$$D_O(p_O, p_{B_1}, p_{B_2}, \dots) = \sum_{j \in J} D_{O_j}(p_O, p_{B_1}, p_{B_2}, \dots) = \sum_{j \in J} D_{O_j}(p_O, p_{B_j}). \quad (2.2)$$

By gainfully localizing online demand using geographical partitions, we reduce the dimensionality of interactions terms from $O(|J|^2)$ to $O(|J|)$, enabling us to accurately and tractably estimate location-specific cross-channel effects. The virtual stores that we propose can be created using appropriate geographical clustering methods (e.g., retail trade analysis).

While partitioning online transactions to their respective virtual stores, it is important to track the final fulfillment destination of a sale because omnichannel retailers offer buy-online-pickup-instore options and execute ship-from-store fulfillments, and often, the point-of-sales data does not encapsulate this difference.

In Fig. 1 we provide an example where we geo-spatially clustered a omnichannel retailer's brick-and-mortar stores (more than 1500) into 50 zones using a k-means ($k=50$) algorithm based on the latitude-longitude coordinates of the stores. We geo-tag and aggregate all the historical transactions in the TLOG data using zones based on the purchase channel and the final fulfillment destination of the sale. The figure also shows the zonal distribution of the sales (the volume is proportional to the pie size) and channel share between brick (red) and online (blue) for one product category. Observe the heterogeneity of the online channel share across zones (e.g., 4% to 11%). This zone-tagged data by item is used to calibrate zone level (bi-directional) cross-channel demand models described in the following section.

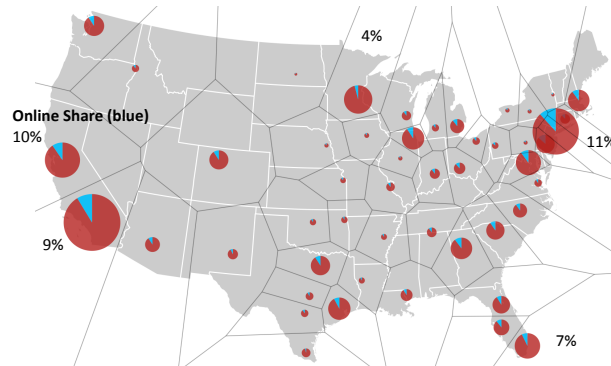


Figure 1 Distribution of sales over 50 zones for a product category. Sales volume is proportional to the pie size. The pie in each zone shows the relative frequency of brick-and-mortar sales and online sales.

3. Omnichannel demand model for a non-perishable product

Consider an omnichannel retailer selling a single non-perishable product using M sales channels to customers in J locations. Let $V \subset M$ be the set of virtual channels like website, mobile, social, which are partitioned into virtual stores by location $j \in J$. Let p_{jm} be the price for the product sold in channel $m \in M$ and location $j \in J$ and \mathbf{p}_j be the corresponding vector of prices in all channels at location j . Note that p_{jm} is often the same across $j \in J$ for virtual channels $m \in V$. Let $\mathbf{D}_j(\mathbf{p}_j)$ be the vector of demands originating from location $j \in J$ in all the channels. As motivated in [Section 2](#) we assume that the demand for a product in a specific channel and location depends on the attributes of all channels at that location. We refer to this representation as the omnichannel demand model.

Discrete choice demand models are one of the commonly used demand functions to model consumer choice in marketing, economics, and more recently, in the revenue management literature. They generalize the well-known *multinomial logit* (MNL) and the *multiplicative competitive interaction* (MCI) demand models, and have their foundations in the random utility theory in economics ([McFadden 1974](#), [Urban 1969](#)). We use these demand functions to model consumer channel demand in an omnichannel environment using market shares as follows:

$$D_{mj}(\mathbf{p}_j) = \text{Market Size of location } j * \frac{\text{Market Share of channel } m \text{ in location } j}{1} \quad (3.1)$$

$$= \tau_j \frac{f_{mj}(p_{mj})}{1 + \sum_{m' \in M} f_{m'j}(p_{m'j})} \quad (3.2)$$

where τ_j is the market size of location j and $f_{mj}(p_{mj})$ is the attraction function of customers in location j to channel m . The market size represents the measure of consumers interested in the product and the market share represents the relative attractiveness of a choice over all choices that includes the no-purchase option, whose attractiveness without loss of generality is normalized to 1. In the omnichannel environment, we model consumer choice to be the option of buying in one of the M channels, or the option of no-purchase. If the attractiveness of a channel drops (for example, due to a channel price increase) then that channel share of the product reduces and it gets proportionally distributed among the other channels. This models the cross-channel substitution (i.e., switching) behavior of consumers.

Examples of the attraction function for demand models include the MNL demand model where $f_{mj}(p_{mj}) = e^{a_{mj} + b_{mj}p_{mj}}$, the MCI demand model where $f_{mj}(p_{mj}) = a_{mj}p_{mj}^{b_{mj}}$, and the linear attraction demand model where $f_{mj}(p_{mj}) = a_{mj} + b_{mj}p_{mj}$. Here, a_{mj}, b_{mj} are constants that ensure the negative price elasticity of demand. In general, we assume that attraction functions are strictly decreasing and continuous, in that their inverse exists. We focus only on the price causal as it is our decision variable but other demand drivers such as promotions, seasonal variations, holidays

and even competitor prices, if available, are included in the demand functions during calibration and we highlight this in the computational experiments in [Section 7.1](#).

A discrete choice function is operationally convenient because of its parsimony in the number of coefficients to be estimated and maintained. In particular, the number of coefficients in the discrete choice demand model is $O(|M|)$ for purchasing choices in set M (this models $O(|M|^2)$ cross-channel interactions).

The standard methods to estimate discrete choice models require historical information about every choice, which in our setting, would include the no purchase data ([Domencich and McFadden 1975](#), [Berkson 1953](#)). Omnichannel retailers rarely have complete information about lost sales and must calibrate their demand models using incomplete data. We employ an integrated mixed-integer programming (MIP) approach that jointly estimates market size and the market share parameters in the presence of censored lost sales data proposed by [Subramanian and Harsha \(2017\)](#). Their method performs imputations endogenously in the MIP by estimating optimal values for the probabilities of the unobserved censored choice. Under mild assumptions, they show the method is asymptotically consistent. Besides being a computationally fast single step method, this estimation approach is capable of jointly calibrating market-size covariates (e.g., with temporal causals), a critical feature with real data. We incorporated model enhancements such as regularization using lasso and ridge penalties, and sign constraints on price coefficients to enable an automated demand estimation environment that is required for operational deployment.

4. Omnichannel price optimization (OCP) for a non-perishable product

In this section, we formulate the omnichannel price optimization model for a non-perishable product in order to identify the most profitable prices in all channels and locations, subject to various retailer's product category goals, channel strategy, sales targets and practical business rules.

We assume that there are well established replenishment policies, and that out-of-stock inventory effects are negligible. This is a reasonable assumption for non-perishable goods (e.g., basic items such as office stationery, printer supplies, etc). Mathematically, it allows one to view the integrated pricing problem across the retail chain as a single period pricing problem without inventory effects.

Using the notation introduced earlier in [Section 3](#), we formulate the general non-linear omnichannel price optimization problem denoted by OCP as follows:

$$\text{OCP: } \max_{\mathbf{p}_j} \sum_{j \in J} (\mathbf{p}_j - \mathbf{c}_j)^T \mathbf{D}_j(\mathbf{p}_j) \quad (4.1)$$

$$\sum_j \mathbf{A}_{kj} \mathbf{D}_j(\mathbf{p}_j) \leq u_k \quad \forall k = 1, \dots, K \quad (4.2)$$

$$\sum_j \mathbf{B}_{lj} \mathbf{p}_j \leq v_l \quad \forall l = 1, \dots, L \quad (4.3)$$

$$p_{m,j} = p_{m,j'} \quad \forall m \in V, j, j' \in J \quad (4.4)$$

$$p_{mj} \in \Omega_{mj} \quad \forall m \in M, j \in J. \quad (4.5)$$

The decision variables in the above OCP formulation are the prices in all locations and channels, and the objective is to maximize the total profitability of the retailer across the retail chain. Constraints (4.2–4.3) are generic polyhedral constraints on demands and prices defined with known matrices $\mathbf{A}_k, \mathbf{B}_l \in \mathbb{R}^M \times \mathbb{R}^J$ and vectors $\mathbf{u} \in \mathbb{R}^K, \mathbf{v} \in \mathbb{R}^L$. These generic constraints encapsulate the retailer's goals and critical pricing business rules that are required for operations. We provide several examples of these constraints in this section below. Constraint (4.4) ensures that the retailer offers the same price across all the virtual stores. This constraint is particularly relevant within our omnichannel framework because we explicitly partitioned the virtual channels by location in order to model bi-directional cross-channel effects, and this constraint binds them back together from the view of the customer. Discrete pricing constraints, which are typical in retail operations, are encapsulated in constraint (4.5).

Some examples of the generic business rules used in practice are as follows:

Volume (or sales goal) constraints

$$\sum_{m \in M_k, j \in J_k} D_{mj}(\mathbf{p}_j) \geq u_k, \quad (4.6)$$

where $M_k \subset M$ and $J_k \subset J$ and depending on the choice of M_k, J_k these constraints can be employed to support a retailer's global or channel and location-specific sales goals. For example, constraint (4.6) can ensure that the total sales volume by channel does not drop below a user-specified threshold, u_k , thereby balancing profitability and market share objectives. Such constraints also act as a practical guard that prevent the drastic price increases that can occur while optimizing pricing for weakly elastic products.

General price monotonicity constraints

$$p_{mj} \leq \gamma_{mm'} p_{m'j} + \delta_{mm'j} \quad \forall j \in J \text{ and for some } m, m' \in M. \quad (4.7)$$

The goal of constraint (4.7) is to enforce that prices in certain channels are cheaper than others by a specified percentage $\gamma_{mm'}$ and/or a constant $\delta_{mm'j}$. This constraint can also account for the variation in unit-cost across channels, i.e., the overhead cost of operating a physical store. An extension of constraint (4.7) is the *price-matching* constraint across the retail chain where the inequality is replaced by an equality and setting $\gamma_{mm'j} = 1, \delta_{mm'j} = 0$. Here, consumers can buy the same product anywhere in the retail chain at the same price. One can view constraint (4.7) also as a *volume measure* constraint. Sometimes a channel exclusively sells a larger volume measure or pack of the same product (for example, a 12-pack case of white board markers sold online versus

a 6-pack case of markers sold in-store). Here, $\gamma_{mm'}$ is a scaling factor between channels that is employed to achieve price parity per unit measure.

Price bounds

$$\underline{\mu}_{mj} \leq p_{mj} \leq \bar{\mu}_{mj} \quad \forall j \in J, m \in M. \quad (4.8)$$

Here, $\underline{\mu}_{mj}$ and $\bar{\mu}_{mj}$ are upper and lower bounds that are often imposed as a percentage of historically offered prices or as a percentage of competitor prices to ensure the competitiveness of the retailer.

Discrete prices

$$p_{mj} \in \Omega_{mj} \quad \forall j \in J, m \in M. \quad (4.9)$$

Ticket prices are naturally discrete (e.g., dollars and cents). Often ‘magic number’ endings (e.g., those ending with 9 or other odd pricing strategies) are important to a retailer and are required to be encoded as a business rule. Furthermore, when a retailer re-optimizes prices, [constraint \(4.9\)](#) can be employed to generate a price ladder that proactively excludes trivial price changes to avoid the substantial labor cost incurred in physically changing the sticker prices in stores.

In practice, the choice of the business rules differs by product category and by the retailer. But in general, we classify all the business rules as either *inter-channel* or *inter-location* constraints. From a computational complexity perspective, inter-channel constraints are relatively easier to satisfy (see [Claim 1](#) below for a counter example), compared to inter-location constraints. We discuss this issue in the following subsection.

4.1. Computational complexity of the OCP problem

Market share transformations are commonly used for discrete choice models to achieve convexity in the pricing problem. The market share variables are defined as follows for each $j \in J$:

$$\theta_{mj} = \frac{f_{mj}(p_{mj})}{1 + \sum_{m' \in M} f_{m'j}(p_{m'j})} \quad \forall m \in M, \text{ and} \quad (4.10)$$

$$\bar{\theta}_j = 1 - \sum_m \theta_{mj}, \quad (4.11)$$

with a one-to-one transformation to the price variables, given by $p_{mj} = f_{mj}^{-1}\left(\frac{\theta_{mj}}{\bar{\theta}_j}\right)$.

CLAIM 1. *Under the market share transformations, the resultant price monotonicity (or a volume measure) [constraints \(4.7\)](#) are non-linear and non-convex.*

We provide an example to prove the above claim in [Appendix A](#). The OCP objective function for a single location is well-known to be uni-modal in the price space and convex in the market-share space. These results do not extend in the multi-location setting. In [Fig. 2](#) we plot the values of the objective function [Eq. \(4.1\)](#) for an OCP instance having a single virtual channel, say online,

and two locations, with [constraint \(4.4\)](#) ensuring that the online price across locations is the same. We observe from the figure that the objective function in this example is non-convex and has multiple peaks. [Constraint \(4.4\)](#) is similar to the price monotonicity [constraint \(4.7\)](#), except that it is across locations and not choices (see [Claim 1](#)). Because the lost sales probabilities, $\bar{\theta}_j$'s, vary across locations, the resulting constraint in the market share space is likely to inject a higher degree of non-linearity and non-convexity into the problem, when compared to the price monotonicity constraint.

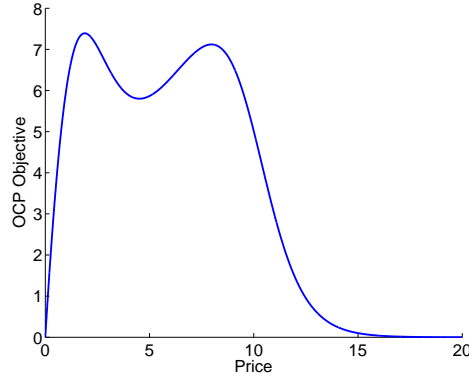


Figure 2 Example of a OCP objective for a single virtual channel and two locations as a function of the virtual channel price for an MNL demand model with $a_{11} = 10, a_{12} = 1, b_{11} = 1, b_{12} = 1, \tau_1 = 1$ and $\tau_2 = 10$.

CLAIM 2. *The OCP problem with multiple virtual channels and at least two locations is NP-hard.*

The proof can be found in [Appendix B](#). A relevant question for this paper is the complexity of the OCP problem having multiple locations but a limited number of channels (with one or more virtual channels). While this remains an open question, the non-convexity and non-linearity of the OCP problem is apparent.

Typical solution approaches in industry involve obtaining solutions using gradient-based non-linear programming techniques. For non-convex problems such as OCP, the solutions can be stuck in local optima which can be far away from the global optimum, resulting in poor quality pricing recommendations. Sometimes, a post-facto rounding of the prices to the nearest feasible value can result in infeasibility. The practical value of global optimization goes beyond just suboptimal prices because such approaches can induce an inconsistent pricing response from the application (e.g., profit increases after a constraint is added) resulting in unsatisfactory user experience and loss in credibility (we provide an example in [Section 8.1](#)).

5. Exact methods to solve the OCP problem

In the following section, we provide two tractable methods that can achieve the global optimum to the OCP problem and compare their computational performance. The first method is an empirically tractable MIP for the general multi-channel case, while the second method is an efficient pseudo-polynomial algorithm for the two-channel case. These methods gainfully operate both in the discrete price and the market share space to address the non-linearity concerns discussed in [Section 4.1](#).

5.1. A mixed-integer programming approach for OCP

We present a MIP re-formulation of the OCP problem. Let the feasible discrete prices for each channel $m \in M$ and location $j \in J$ be denoted by \bar{p}_{mji} for $i \in I_{mj}$. Here, the set I_{mj} denotes the index set of feasible prices. Let z_{mji} be a binary variable which is nonzero only if the price in channel $m \in M$ at location $j \in J$ is \bar{p}_{mji} . Note that for a virtual channel $v \in V$ the prices across all locations are the same. Therefore, the corresponding prices \bar{p}_{vi} , the price index set I_v and binary variable z_{vi} are location independent, and $\bar{p}_{vji} = \bar{p}_{vi}$, $I_{vj} = I_v$ and $z_{vji} = z_{vi}$.

Using these definitions, the key term to linearize in the OCP problem is the demand function $D_{mj}(\mathbf{p}_j)$ which depends on prices in all channels. This is proportional to $\prod_{m \in M} z_{mji_m}$ with the discrete price ladder, where Π refer to the product function. A naive way of linearizing this function requires the introduction of $|I|^{|M|}$ additional binary variables when $I_{mj} = I \forall m \in M$ (note that there are already $|I||M|$ binary variables due to discrete prices). This results in a MIP that explodes in size very quickly making it impractical to solve. We present an exact alternative linearization that exploits the special structure of the discrete choice model and does not require any additional binary variables, resulting in a computational tractable MIP. Assuming $q_{mji} = \tau_j(\bar{p}_{mji} - c_{mj})f_{mj}(\bar{p}_{mji})$, $r_{mji} = f_{mj}(\bar{p}_{mji})$, $\alpha_{kmji} = A_{kmj}r_{mji}$ and $\beta_{kmji} = B_{lmj}\bar{p}_{mji}$, the objective and [constraint \(4.2\)](#) of the OCP problem can be rewritten as

$$\max_{z_{mji}, z_{vi}} \sum_{j \in J} \frac{\sum_{m \in M} \sum_{i \in I_{mj}} q_{mji} z_{mji}}{1 + \sum_{m \in M} \sum_{i \in I_{mj}} r_{mji} z_{mji}}, \text{ and} \quad (5.1)$$

$$\sum_{j \in J} \sum_{m \in M} \sum_{i \in I_{mj}} \frac{\alpha_{kmji} z_{mji}}{1 + \sum_{m \in M} \sum_{i \in I_{mj}} r_{mji} z_{mji}} \leq u_k \quad \forall k \in K. \quad (5.2)$$

We use the fractional programming transformations proposed by [Charnes and Cooper \(1962\)](#) to overcome the non-linearity arising from the ratio terms. Let

$$y_j = \frac{1}{1 + \sum_{m \in M} \sum_{i \in I_{mj}} r_{mji} z_{mji}} \quad \forall j \in J, \text{ and} \quad (5.3)$$

Because z_{mji} are binary variables and r_{mji} are non-negative constants, $0 \leq y_j \leq 1 \forall j \in J$. Now we define

$$x_{mji} = y_j z_{mji}. \quad (5.4)$$

It is easy to see that $0 \leq x_{mji} \leq 1 \forall m \in M, i \in I_{mj}$. Eq. (5.3) introduces a computationally beneficial convexification via a market-share transformation, where the y-variables represent the lost-share values at locations, and the x-variables represent the corresponding channel-shares. However, we retain the original z-variables to manage the multitude of business rules but lose convexity due to the bilinear terms. We use the reformulation and linearization technique (RLT) proposed by Sherali and Adams (1999) to eliminate those non-linearities. The RLT transformations exploit the discrete nature of the binary variables, allowing us to recover an exact reformulation of the OCP problem (reason discussed below). Substituting these transformations and linearizing, the resulting reformulated OCP problem is as follows:

$$\max_{z,y,x} \sum_{j \in J} \sum_{m \in M} \sum_{i \in I_{mj}} q_{mji} x_{mji} \quad (5.5)$$

$$\sum_{j \in J} \sum_{m \in M} \sum_{i \in I_{mj}} \alpha_{kmji} x_{mji} \leq u_k \quad \forall k \in K \quad (5.6)$$

$$\sum_{j \in J} \sum_{m \in M} \sum_{i \in I_{mj}} \beta_{lmji} z_{mji} \leq v_l \quad \forall l \in L \quad (5.7)$$

$$y_j + \sum_{m \in M} \sum_{i \in I_{mj}} r_{mji} x_{mji} = 1 \quad \forall j \in J \quad (5.8)$$

$$x_{mji} \leq y_j \quad \forall i \in I_{mj}, m \in M, j \in J \quad (5.9)$$

$$x_{mji} \leq z_{mji} \quad \forall i \in I_{mj}, m \in M, j \in J \quad (5.10)$$

$$\sum_{i \in I_{mj}} x_{mji} = y_j \quad \forall m \in M, j \in J \quad (5.11)$$

$$\sum_{i \in I_{mj}} z_{mji} = 1 \quad \forall j \in J, m \in M \quad (5.12)$$

$$z_{vi} = z_{vji} \quad \forall j \in J, v \in V \quad (5.13)$$

$$z_{mji} \in \{0, 1\} \quad \forall i \in I_{mj}, m \in M, j \in J \quad (5.14)$$

$$y_j, x_{mji} \geq 0 \quad \forall i \in I_{mj}, m \in M, j \in J \quad (5.15)$$

In the above formulation, constraints (5.12) and (5.14) model the discrete nature of channel prices. The objective function, the general business rules on volumes and prices and the uniform virtual price constraint of the OCP problem given by (4.1) and constraints (4.2–4.4) are encapsulated in the above formulation in (5.5) and constraints (5.6–5.7) and (5.13) respectively. Constraint (5.8) linearizes Eq. (5.3). Constraints (5.9–5.10) along with the objective linearize the product term $x_{mji} = y_j z_{mji}$. Note that the linearization is exact because if $z_{mji} = 0$, constraint (5.10) ensures $x_{mji} = 0$, and if $z_{mji} = 1$, constraint (5.9) together with the objective ensure $x_{mji} = y_j$ (because q_{mji} are positive). The RLT constraints (5.11) are implied by constraint (5.9) in the integer sense and are valid linear inequalities that serve to tighten the underlying LP relaxation. In our numerical

computations, we observed that the addition of [cuts \(5.11\)](#) yielded a considerable improvement in the computational performance.

Observe that the transformed OCP formulation allows for any number of channels and can incorporate a variety of important and complex business rules that are employed in practice. Furthermore, the above formulation is now a linear MIP and a commercial optimization software package like IBM ILOG CPLEX can be used to solve this problem to optimality. In [Section 7](#) we report that it takes no more than 3 seconds for practical sized instances. The formulation is computationally efficient and this stems from its compact reformulation that gainfully navigates both in the price and market share space.

5.2. A decomposition method for OCP in the case of the brick and online channel

In the two channel (brick and online) setting, a simple decomposition algorithm for the OCP problem is as follows: *fix the online price and solve the corresponding single channel multi-location brick problem to optimality, and repeat this search over all online prices.* This decomposition method described in this section is especially beneficial when there are several inter-channel constraints because they can all be managed locally within the brick subproblem.

We first consider the OCP problem without inter-location constraints and provide extensions of the decomposition method in their presence towards the end of the section. Given an online price p_o , the resultant OCP problem decomposes into separable univariate location-specific subproblems $OCP_j(p_o)$ at every location j , wherein the inter-channel constraints translate into a bound-restricted store price p_{bj} as shown below:

$$OCP_j(p_o): \quad \Pi_j(p_o) = \max_{p_{bj} \in \Omega_{bj}} \sum_{m \in \{o,b\}} (p_{mj} - c_{mj}) \frac{f_{mj}(p_{mj})}{1 + f_{bj}(p_{bj}) + f_{oj}(p_o)} \quad (5.16)$$

$$s.t. \quad \underline{h}_j(p_o) \leq p_{bj} \leq \bar{h}_j(p_o) \quad (5.17)$$

PROPOSITION 1. *Suppose for all $j \in J$, the function $f_{bj}(\cdot)$ is strictly decreasing and twice differentiable, with $\lim_{x \rightarrow -\infty} f_{bj}(x) = \infty$, $\lim_{x \rightarrow \infty} x f_{bj}(x) = 0$ and that the function $g_{bj}(\cdot) = f_{bj}^{-1}(\cdot)$ satisfies the following condition:*

$$2g'_{bj}(y) + yg''_{bj}(y) \leq 0 \quad \forall y > 0. \quad (5.18)$$

Then the optimal solution to $OCP_j(p_o)$ can be obtained by rounding the optimal solution of its underlying continuous relaxation, which in turn, reduces to solving the following differential equation that always admits a solution, say z^ :*

$$g_{bj}(z) + z \left(1 + \frac{z}{1 + f_{oj}(p_o)} \right) g'_{bj}(z) \stackrel{def}{=} \sigma_j(p_o) = (p_o - c_o) \frac{f_{oj}(p_o)}{1 + f_{oj}(p_o)} + c_{bj}, \quad (5.19)$$

The proof of the proposition is in [Appendix C](#). The well-know discrete choice models like the MNL, MCI and linear attraction choice models satisfy the assumptions on the attraction function and [Eq. \(5.18\)](#). The proposition shows that the optimal solution to $\text{OCP}_j(p_o)$ can be obtained in near closed form without having to evaluate every discrete store price. The solution to [Eq. \(5.19\)](#) can be derived using root finding algorithms (e.g., Newton Raphson method).

The example in [Fig. 2](#) depicts a function that is non-convex and can have multiple peaks. Therefore, in order to find the optimal online price that solves the OCP problem, we search over every discretized online price point (using [Proposition 1](#) to determine the corresponding store prices) and report the best omnichannel price combination.

REMARK 1. *The runtime complexity of the decomposition algorithm in the absence of inter-location constraints and using an attraction demand model is pseudolinear, and is given by $O(IJR)$ where I is the size of the price ladder in the online channel, J is the number of locations, and R is the complexity of a root finding algorithm.*

If multiple inter-location constraints are active, the location-level subproblems for a given online price are no longer independent and we recommend using the MIP approach. In the presence of a single inter-location constraint such a global volume goal, the location-level subproblems can be jointly solved as a multiple-choice knapsack problem (MCKNP) problem in pseudo-polynomial time using dynamic programming ([Pisinger 1994](#)).

An interesting question in the context of the OCP problem is the structure of the optimal channel prices, and how the brick and online prices coordinate with each other. This is a non-trivial question in the presence of discrete prices and business rules. In the absence of both, the optimal channel prices satisfy a simple and insightful channel price coordination equation. We provide the equation in the special case of MNL demand model and defer the general case and its proof to [Appendix D](#).

PROPOSITION 2. *For an MNL demand model where $f_{mj}(p_{mj}) = e^{a_{mj} - b_{mj}p_{mj}}$, the optimal prices of the OCP problem in the absence of business rules satisfies the following equation*

$$p_o^* = c_o + \sum_{j \in J} w_j(\mathbf{p}) \left[p_{bj}^* - c_{bj} + \frac{1}{b_{oj}} - \frac{1}{b_{bj}} \right] \quad \forall o, b \in M, \quad (5.20)$$

where $w_j(\mathbf{p})$ are the normalized values corresponding to the (weighted) online demand i.e., $b_{oj}D_{oj}(\mathbf{p}^*)$, over all the locations $j \in J$ such that $\sum_{j \in J} w_j(\mathbf{p}) = 1$.

[Eq. \(5.20\)](#) implies that the optimal online margin is a weighted linear combination of the optimal brick margins and certain constant terms across all locations, generalizing the result for multiple products (and single location) that the optimal margins are identical when $b_o = b_b$. In contrast to the single location result, [Eq. \(5.20\)](#) is just a necessary condition for optimality and not sufficient

(see Fig. 2). One way this result can be used in practice is to identify products for which the business rules (e.g., with discrete prices or volume goals or price rules) are binding. We provide examples in Section 7.2.

5.3. Comparison of computational performance between the MIP and the decomposition approach to solve OCP

Fig. 3 compares the average running time of the decomposition method and the MIP approach for two channels (brick and online) as a function of the number of locations using simulated demand models that were motivated from real data. To enable comparison, we ignore inter-location constraints, and report run times for the decomposition methods under the assumption of full parallelization of subproblem solutions across locations. An MNL attraction model was used in the simulations, and the resulting optimization problem was solved using the decomposition method based on Proposition 1. The simulations incorporated the business rules related to the price bounds and 20 discrete prices per product per channel. Not surprisingly, Fig. 3 shows that the runtime for *parallelized* decomposition method marginally increases when the number of locations is no more than 128, and then remains constant thereafter. In contrast, the MIP approach was faster when the number of locations was less than 200, but increased non-linearly with the number of locations thereafter. In practice, we employed no more than 100 locations/zones to model retail chains having thousand or more physical stores. The run time trend of the single online channel MIP is like that exhibited by the two-channel MIP, but grows at a slower rate as the number of locations increase.

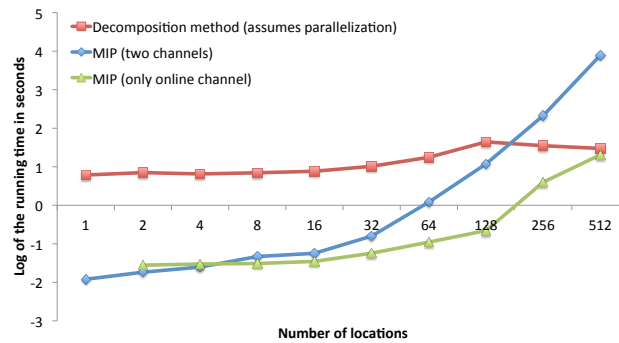


Figure 3 Average run times of the decomposition model and the MIP over 25 simulated instances as a function of the number of locations

6. Omnichannel assortment pricing problem

*** Section not included in the public version at the moment. ***

7. OCP implementation for a major US retailer

In this section, we report results of an OCP implementation for a major U.S. omnichannel retailer. We worked with IBM Commerce and engaged with the retailer over the course of 8 months to demonstrate the business value of the integrated omnichannel regular pricing over their existing channel independent regular pricing method on representative product categories based on historical data. The first few weeks were spent on working with all stakeholders to define the business problem, success criteria, select the categories to be analyzed, and collect and process the historical data. Thereafter, we performed the demand model calibration and the value assessment. The details of the different steps are provided below but first, we describe the retailer and their business process (while maintaining their anonymity).

Retailer and their current business process: The omnichannel retailer we worked with sells a variety of products, including office supply product categories. They operate a brick-and-mortar channel with a network of well over 1500 stores across the United States. The online channel is used to complete sales transactions that are routed through their website, as well as mobile and paper-catalog orders. The organizational structure of the retailer results in two different divisions separately managing the planning and operations of the two channels. Both divisions use a regular price optimization (RPO) solution to manage prices for many non-perishable products, referred to as UPCs (universal product code). The prices for the remaining products are controlled partly by the manufacturer, or are price-matched with certain competitors. The incumbent RPO solution produces demand forecasts that are independent of the other channel or competitors, and identifies regular or base price for the (non-perishable) products that maximizes the retailer’s profitability over a specified finite horizon subject to some business constraints. Small and infrequent price changes (e.g., less than 30%) are typical. One price is found for every geographical cluster of brick stores identified by the retailer as a ‘price zone’. The entire online channel is treated as a separate price zone. The regular prices identified are treated as ticket prices but can be overlaid with various promotions. The retailer can re-optimize prices using the RPO solution as needed (e.g., weekly). The pricing solution requires weekly sales and promotion data to calibrate its models.

Business Problem: More than 20% of UPCs that are priced by the retailer are sold in both channels. They contribute to a significant portion of the retailer’s category revenue (details provided below). Due to the rapid growth of the online channel, the retailer was primarily concerned about how to optimally coordinate prices between the two channels while accounting for competitor effects. To remain competitive, growing their online presence was important to the retailer, but because the vast majority of their customers purchase the product in-store, ignoring channel switching effects can adversely impact their KPIs such as gross profit and sales volume. Note that for

weakly-elastic items, one can achieve short-term margin gain by simply raising prices. The incumbent business process that optimized channel prices independently was likely to increase prices in both channels, additionally motivating the need for integrated decision making. In the future, OCP would be used to frequently re-optimize online prices to respond to competitor price changes. For the business value assessment, they suggested ignoring cross-product assortment effects and that in general, they were not keen on enforcing a price match between the brick and online channels.

Data Summary: To support the business value assessment, we were provided with 52 weeks of U.S. sales transaction and promotion data (date range of July 2012 to July 2013) for the brick and online channels for two categories: (1) inkjet cartridges and (2) markers and highlighters. The top 50 UPCs in terms of historical volume that were sold in both the channels (channel volume share of at least 1%) were selected for the business value assessment. Some statistics about the data are summarized in [Table 7](#).

Category	No. of UPCs	Avg. Final Price		% of category revenue	Online volume share
		Brick	Online		
Inkjet Cartridges	50	\$36.4	\$32.1	30%	12%
Markers and Highlighters	50	\$8.7	\$8.3	42%	12%

Table 1 Summary of the 2012-13 data provided.

In the inkjet cartridges and the markers and highlighters category, the 50 UPCs that were selected contributed to about 30%, and 42% of the category revenues respectively, and have a 12% online volume share in each case. Although the online share of the retailer sales in 2012-13 was relatively low, this number has been steadily increasing year to year.

We also note that the inkjet cartridges category consists of products that are more expensive than the markers and highlighters category and that the retailer sells products in the brick channel at a slightly higher price than the online channel, even though a vast majority of customers buy in-store. Although one would attribute it to the higher holding costs in-store, this information was not provided to us. We were only provided with the wholesale cost information for the different UPCs and this was the same across the sales channels.

There were approximately 40 distinct geographical price zones in the brick channel. The continuum of online sales was disaggregated by the brick channel's geographical price zones using the omnichannel framework described in [Section 2](#). The total sales rates across the 40 zones were not evenly distributed across locations, and we found that the top 10 zones contributed to 54% of the total sales, and the top 20 zones accounted for 83% of the total sales.

For each of the UPC-zone pairs, we obtained by channel, the weekly aggregated sales, volume weighted weekly average ticket price, discounts and promotions, weekly holidays and seasonality

factors. For a small sample of UPCs, we were also provided a time series of three online competitor's prices. We observed that the products in both these categories exhibited a relatively steady sales rate, which is typical of non-perishable basic products.

7.1. Demand Estimation

We use the zone-tagged data to estimate the omnichannel demand model described in Eq. (3.2) based on the MNL attraction function at the UPC-zone level. Model selection and cross-validation on a variety of training instances yielded the following market-size model for a zone that predicts the customer arrivals for any week t in the selling season:

$$\log(\text{Market Size}_t) = \gamma_0 + \sum_{1,k} \gamma_{1,k} \text{TEMPORAL-VARIABLES}_{k,t}. \quad (7.1)$$

and the following market share model to predict channel shares in week t :

$$\begin{aligned} \log(\text{Channel Attraction}_t) = & \beta_0 + \beta_1 \text{PRICE}_t + \sum_k \beta_{2,k} \text{PROMOTION-VARIABLES}_{k,t} \\ & + \sum_j \beta_{3,j} \text{COMPETITOR PRICES (optional)}_{j,t}. \end{aligned} \quad (7.2)$$

The promotion variables included discounts and other promotional indicators and the temporal variables included seasonality, holiday effects and trend. Competitor prices were introduced as attributes in the channel specific utilities, whenever they were available. Since lost sales data was unavailable, the coefficients in Eq. (7.1) and Eq. (7.2) were jointly estimated using the approach discussed in Section 3.

The preferred method for the retailer to track the forecast accuracy was using the weighted mean absolute percentage error (WMAPE) metric where t represents the week index:

$$WMAPE = \frac{\sum_t |\text{predicted sales}(t) - \text{actual sales}(t)|}{\sum_t \text{actual sales}(t)} * 100 \quad (7.3)$$

Table 2 reports the achieved out-of sample WMAPE metric for eight-week ahead predictions of weekly sales at the UPC-zone level. The WMAPE value compares well with the estimates of Fisher and Vaidyanathan (2014) who report an out-of-sample sales-weighted MAPE of 40.1% at the store-UPC level and 25.8% at the chain-UPC level for store-sales of automobile tires. Overall, the achieved forecast accuracy at this fine level of aggregation satisfied the retailer's expectation.

Estimated own and cross-channel price elasticities: We present the average same-channel and cross-channel price elasticity values evaluated at the average channel price in Table 3. These elasticities range between -2.0 to 0. The relatively low elasticity values are typical of essential consumer products. For example, Krugman and Wells (2008) report price elasticities for various

Inkjet Cartridges		Markers and highlighters	
Brick	Online	Brick	Online
29	36	36	24
(23,31)	(24,45)	(26,44)	(20,27)

Table 2 Eight week average and range (25th and 75th quantile) of the out-of-sample WMAPE fit over the UPC-zone pairs in each category.

essential and luxury product types and a value of -0.5 for stationery goods. It can also be seen from Table 3 that the cross-channel price elasticities are significant for these categories, and as high as 50% of the own channel price elasticity. Note also that the cross-elasticities are asymmetric in that the impact of brick prices on the online sales tends to be higher than the impact of the online prices on brick sales. It is indicative of the heterogeneity of the customers shopping in the different channels as well as the volume share of these channels (the absolute change in volume of brick sales is much higher than that for the online channel). As the online share rises in the future, we can expect the online price to exert an increasing influence on the brick channel sales.

Channel \ Price	Inkjet Cartridges		Markers and highlighters	
	Brick price	Online price	Brick price	Online price
Brick sales	-0.66	0.06	-1.19	0.04
	(-1.84,-0.4)	(0,0.22)	(-1.87,-0.01)	(0, 0.05)
Online sales	0.31	-1.04	0.19	-0.78
	(0,1.46)	(-1.99,-0.01)	(0,0.39)	(-1.96,-0.01)

Table 3 Average and the range (10th and 90th quantile) of the own and cross-channel price elasticities.

It must be noted that the cross-channel elasticities (off diagonal entries in Table 3), and in particular the impact of brick prices on online sales at a location specific level (lower off-diagonal entry), cannot be computed using traditional pricing models employed in the industry. The value assessment presented in the following section are predictions based on this calibrated demand model.

7.2. Business value assessment

For each of the 50 UPCs in the two product categories we jointly optimize for the UPC prices in all the physical stores and the online channel using the estimated zone-level demand models through a MIP formulation of the OCP model. For the business case, we explicitly focused on the multi-period regular pricing problem (a typical RPO setting) wherein the goal is to find the optimized price for a product over the last 8 consecutive out-of-sample prediction weeks in 2013.

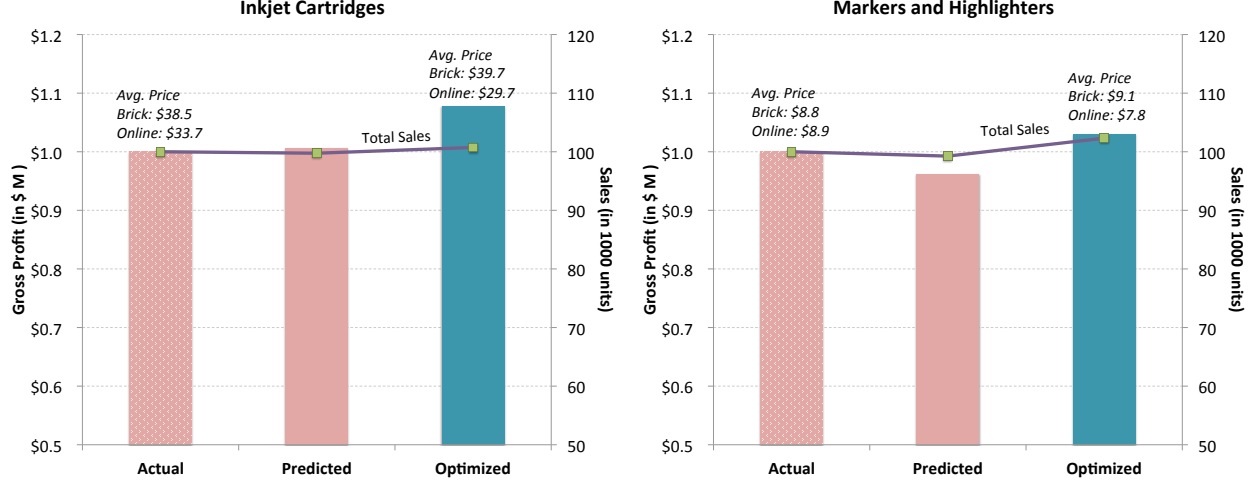


Figure 4 Re-normalized gross-profit and sales during a 8 week period for actuals from historical data, predicted using omnichannel demand models at actual prices and the optimization

The OCP model was extended to include multiple time periods and its demand variations, via a constraint similar to (4.4) that runs across time.

The retailer specified using the following goals and business rules for the value assessment: (1) price bounds to ensure that recommended prices were within historical values allowing at least a 10% variation (20 price points per ladder with magic number endings); (2) a global volume goal to preserve sales volume that required the predicted volume at the optimal prices to be no less than the predicted volume at the baseline (actual) prices; and (3) a price balancing constraint that required the average channel price to be no higher than the corresponding baseline value. The motivation for the last constraint is to prevent a myopic response to raising prices in both channels that can result in lost customers in the long run. We used 20 price points per ladder. We included the constraints in the OCP model to simulate the profit that OCP would achieve when implemented using the business rules applied by the retailer.

The OCP optimization models developed as a JAVA API were evaluated on a Windows-7 computer having 8GB RAM, and an Intel Core i7 processor. CPLEX 12.6.2 with its out-of-box parameter settings was used to solve the resultant MIPs to global optimality. The average problem size after preprocessing was about 30K rows, 20K columns, and 550 SOS1 variables. The average run-time per UPC was 1.7 seconds, and no more than 3 seconds in the worst case. Thus, the solution response is fast enough not just for a weekly refresh but also multiple price updates within a day (we describe the benefit of this in Section 8.1).

The results of this optimization are presented in Fig. 4 along with two baselines. The first baseline actual represents the KPIs achieved by the retailer’s incumbent single channel forecasting and pricing system. The second baseline (predicted) represents the KPIs using the actual prices

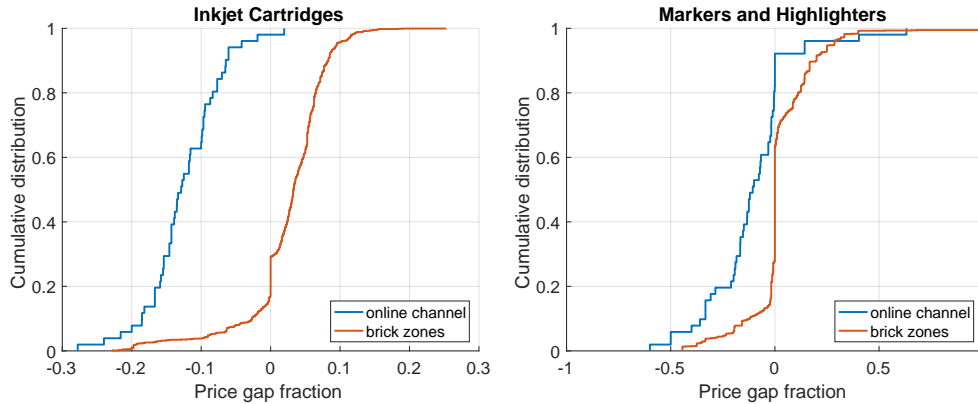


Figure 5 Cumulative distribution of relative price change of the optimized compared to the actuals.

and the omnichannel demand model. To protect the retailer’s data privacy, the actual gross profits are normalized to \$1M and the sales volume to 100K units and hence the results for predicted and optimized are relative to these normalized actuals. We can observe that the predicted and realized (or actual) metrics are relatively close to each other for both categories, i.e., within 1% in terms of sales volume and gross profit each for inkjet cartridges category, and within 1% in terms of sales volume and 4% in terms of gross profit for markers and highlighters category. The optimized metrics, on the other hand, indicate a 7% gross profit lift in the inkjet cartridges and the markers and highlighters categories each with respect to the predicted metrics, while also achieving a 1% and 3% increase in aggregate sales volume respectively.

Fig. 5 displays the cumulative distribution of the relative change in brick store prices and the online channel prices compared to their respective baselines (note that there are more observations in the former than the latter). The average optimized online price is about 12% lower than the actuals, while the average optimized store prices are 2.5% higher in each of the product categories. In the inkjet category, we observe that the optimization increased brick prices in 70% of the zones, while retaining or lowering brick prices in 30% of the locations. On the other hand, for the markers, only 37% of the brick locations witness a price rise (due to relatively higher own brick price elasticity). In both categories, the optimization predominantly lowered the online channel price for all UPCs compared to the brick zones, because of a relatively high elasticity of the online channel in comparison. Such a pricing solution boosts online sales, allowing the retailer to be more competitive with online retailers in the marketplace without trying to aggressively match the low price of e-tail giants. Furthermore, by accounting for the presence of cross-channel price elasticity, a portion of the brick customers at more price sensitive zones can switch channels and purchase the product at the retailer’s online website at a lower price.

Fig. 6 is a scatter plot of the constrained optimal online price against online prices predicted by Eq. (5.20) (using the constrained optimal prices on the right-hand side of the equation). From

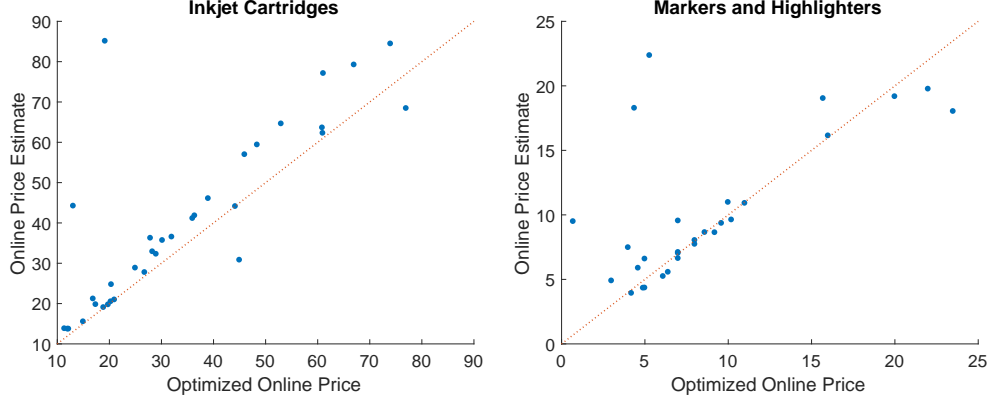


Figure 6 ScatterPlot of the online price predicted by Eq. (5.20) and the optimized online price

Section 5.2 we recall that if the constrained optimal prices do not satisfy Eq. (5.20) then it implies that certain constraints are binding. For a majority of the UPCs this is the case, which is not surprising as the products are weakly elastic. We observe that at higher prices, constraints tend to be increasingly binding. Recall that the size of the problem is significant even after the presolve step in CPLEX (which focuses on logical reduction of the problem inputted and is an industry standard), which indicates a large feasible space and hence, the value of the models proposed in the paper.

In summary, our business value assessment projected an incremental annual profit gain of 7% for the retailer for the categories analyzed while also satisfying a variety of important business rules and balancing different goals. The above results were presented at the retailer site to a team of pricing analysts and senior executives, including their Vice-President for revenue management. OCP's demonstrated ability to accurately model omnichannel customer behavior and integrate competitive channel pricing strategies was highly valued by IBM Commerce and their retail customers. The overall feedback was positive, and with similar experiences with other retailers in a variety of product categories, IBM Commerce followed it up with their own internal evaluation within their system. With equally positive results, a proprietary version of the solution was approved for commercial deployment.

8. Deployment and Commercialization

Fig. 7 provides a high level view of the data flow of the OCP implementation. A big-data platform was employed to manage the challenging data-integration tasks across channels, and implement the specific data extraction algorithms required to create the omnichannel modeling framework. This platform is scalable to the enterprise level and allows for extraction, transformation, loading (ETL), and can process large volumes of diverse transaction data. A data visualization tool was employed to view a variety of results derived from the omnichannel framework. The data preprocessing is

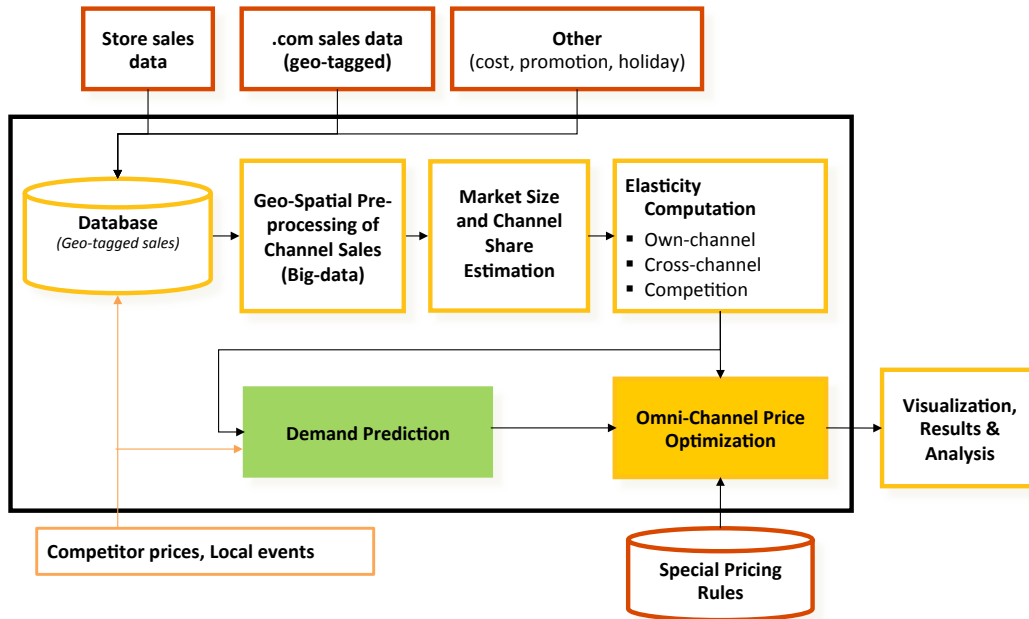


Figure 7 Data flow diagram for OCP

followed by the omnichannel models that includes demand prediction and the optimization engine which were delivered as a JAVA API. The API is configurable with a variety of business rules and goals that can vary across UPCs and allows for an in-memory processing of inputs and outputs.

Proprietary versions of these models were developed and handed to IBM Commerce following multiple sessions of knowledge transfer. These assets were deployed into production by IBM Commerce in 2014 and reside in the IBM cloud. The methods and the systems described in this paper are patent pending.

8.1. Impact of operations research in the deployment of the OCP models

We now delineate two additional applications of the OCP optimization model and discuss the positive impact of operations research on the resulting pricing decisions and prior business practices.

Profitability threshold based price matching and the practical value of global optimality: Consider a retailer who would rather price-match the channels if the profitability gain (by not imposing the constraint) is insufficient. The retailer can conveniently specify this tradeoff via a profitability threshold, where in, a price match constraint is accepted only if the resultant drop in profitability is within the threshold limit.

A natural way of implementing this feature is to solve the OCP problem with, and without the price-matching constraint and then choose the preferred solution. Employing an exact solution approach turns out to be critical in this context. The use of a local-optimum based heuristic

approach to solve the OCP problem with and without the price matching constraint, can result in incorrect profitability gaps, producing ‘false-positive’ price-matching recommendations. Note that for such heuristic methods, a price-matching constraint can operate like a cutting-plane that deletes a local (but not global) optimal solution to the unconstrained problem (i.e., without the price-matching constraint), potentially yielding an improved profitability objective. In such cases, the heuristic approach is likely to approve price-matching, whereas, the unconstrained profit value it achieves may have been far from optimality. Numerical testing demonstrated that such false-positives were not uncommon. On the other hand, an optimal approach always generates the correct price-coordination recommendation, and the application produces stable and predictable responses from a user perspective.

Asynchronous channel-specific dynamic price optimization to support the existing business process: It is often required to frequently change the online prices to respond to rapidly changing competitor prices compared to brick prices, which typically incurs additional labor cost. Therefore, asynchronous channel-specific optimization becomes necessary along with the ability to execute rapid data refreshes. We demonstrate our ability to solve the integrated OCP on a weekly basis, and our algorithms were fast enough to support frequent re-optimization of online prices (one or more times a day or near real-time) while keeping the brick prices at all locations fixed at their most recent optimized values, thereby, accounting for their cross-channel impact. Analyzing and implementing a retailer’s competitive price response strategy in the omnichannel era is an important topic of study with research underway.

8.2. Post-deployment highlights

We conclude the paper with some important highlights. Our analytical solution was showcased as one of the retail analytics success stories in the smarter-commerce global summit in 2014. It included a presentation on its capabilities by a partner retail chain. Today, several large global retail chains are regular users of the commercial offering including those with whom we engaged. In November 2015, this work was formally recognized by IBM as one of the major accomplishments in 2015 by the research division.

Appendix A: Example to prove Claim 1

Consider a price monotonicity (or a volume measure) [constraint \(4.7\)](#) of the form $p_{mj} \leq \gamma p_{m'j}$, for example, where γ is a constant. The constraint in the attraction space translates to $f_{mj}(p_{mj}) \leq f_{mj}(\gamma p_{m'j})$ which in the market share space translates to $\theta_{mj} \leq \frac{f_{mj}(\gamma p_{m'j})}{f_{m'j}(p_{m'j})} \theta_{m'j}$. Even for simple attraction models, the ratio of the attractions is not a constant and the resulting non-linearity cannot be linearized. In the special case when $\gamma = 1$ and the b_{mj} are identical for all $m \in M$ in an MNL or linear attraction demand model, the right

hand side can be transformed into an affine function in the market share space. We observe that b_{mj} are not identical across $m \in M$ for any of the product categories that we analyzed across retailers in our customer engagements. This can be attributed to the heterogeneity in people's shopping preferences at a location across different channels. Therefore, a market share transformation fails to recover convex constraints in the market share space.

Appendix B: Proof of Claim 2

We now show that the OCP problem having two or more virtual channels and just two locations is an NP-hard problem. This result is achieved by performing a reduction from the *2-class logit assortment optimization problem* (2CL). The goal of 2CL is to identify an optimal assortment of items in a set V to offer to customers who can potentially belong to one of two segment classes that are unknown to the seller. The inputs to this problem include the item profits, the relative weight of the classes and the preference weight of each item in each class. [Rusmevichientong et al. \(2010\)](#) showed that the 2CL is NP-hard. The reduction is as follows. Consider an arbitrary instance of 2CL and map every item in set V to a distinct virtual channels in the OCP problem. If an item $v \in V$ is part of an assortment then offer a product in the virtual channel v at some finite price that results in the attraction value of the channel-class being equal to the preference weight of the item in that class. If an item $v \in V$ is not part of an assortment then we offer the product in channel v at a sufficiently high price that reduces the attraction value of channel v to zero. The item profits in 2CL correspond to the margin of the channels, and the relative weight of the classes correspond to the market size of each location. As a reduction from this 2CL assortment optimization problem for an MMNL demand model, we deduce that the OCP problem is NP hard. \square

Appendix C: Proof of Proposition 1

We drop subscript j in this proof as we are working with a specific location. Let θ_b denote the market share of brick as follows:

$$\theta_b = \frac{f_b(p_b)}{1 + f_b(p_b) + f_o(p_o)}. \quad (\text{C.1})$$

The lost market share and the online market share in terms of θ_b are then $\frac{1-\theta_b}{1+f_o(p_o)}$ and $\frac{f_o(p_o)(1-\theta_b)}{1+f_o(p_o)}$ respectively. We define the inverse attraction function as $g_b(y) = f_b^{-1}(y)$, $y > 0$. Therefore, we can write $p_b = g_b\left(\frac{B\theta_b}{1-\theta_b}\right)$ where $B = 1 + f_o(p_o)$. Substituting this for p_b in continuous relaxation of the OCP(p_o) problem, we get

$$\theta_b^* = \arg \max_{\theta_b \in [\Theta_L, \Theta_U]} A(1 - \theta_b) + \theta_b \left(g_b \left(\frac{B\theta_b}{1 - \theta_b} \right) - c_b \right) \quad (\text{C.2})$$

where $\Theta_U = \theta_b|_{p_b=\underline{h}(p_o)}$, $\Theta_L = \theta_b|_{p_b=\bar{h}(p_o)}$ and $A = (p_o - c_o) \frac{f_o(p_o)}{1+f_o(p_o)}$. Under [Eq. \(5.18\)](#), it is easy to check that the second derivative of the objective function is positive which means this relaxed problem has a concave objective. Therefore, a solution that sets the first derivative of the objective to zero exists and it is an optimal solution to the unconstrained problem. We use this in deriving an optimal solution to the constrained problem as well. Now taking first derivative and setting it to zero, we get:

$$g_b \left(\frac{B\theta_b^*}{1 - \theta_b^*} \right) + \theta_b^* g_b' \left(\frac{B\theta_b^*}{1 - \theta_b^*} \right) B \frac{1}{(1 - \theta_b^*)^2} = A + c_b. \quad (\text{C.3})$$

Substituting $z = \frac{B\theta_b^*}{1-\theta_b^*}$, we get

$$g_b(z) + z \left(1 + \frac{z}{B}\right) g'_b(z) = A + c_b$$

which is the same as Eq. (5.19).

The constrained optimal solution to the problem takes one of the following three values because the objective in the problem (C.2) is concave in θ_b : (a) optimal value, θ_b^* if $\Theta_L \leq \theta_b^* \leq \Theta_U$; (b) Θ_U if $\theta_b^* > \Theta_U$; or (c) Θ_L if $\theta_b^* < \Theta_L$. In the price space, because θ_b and p_b have a one-to-one correspondence and have an inverse relationship, this optimal solution simplifies to $p_b^*(p_o) = \max \{ \underline{h}(p_o), \min \{ g_b(z^*), \bar{h}(p_o) \} \}$.

Because problem (C.2) is a concave maximization problem, with a one-to-one correspondence to the price space, a simple rounding algorithm around the continuous optimal $p_b^*(p_o)$ that checks for the maximum objective at the ceiling and floor of $p_b^*(p_o)$ with respect to the discretization Ω_b within the feasible region can be employed to get the optimal solution to OCP(p_o).

Let $\psi(z, p_o)$ denote the left hand side of Eq. (5.19). We now show that Eq. (5.19) always has a solution. Consider the derivative of $\psi(z, p_o)$ w.r.t. z :

$$\frac{\partial \psi(z, p_o)}{\partial z} = (2g'_b(z) + zg''_b(z)) \left(1 + \frac{z}{1 + f_o(p_o)}\right) \leq 0. \quad (\text{C.4})$$

The last inequality is because the first product term is always non-negative because of Eq. (5.18) and second product term is positive because z represents market share ratios and is always positive. This negative derivative implies that $\psi(z, p_o)$ is a non-increasing function in z .

From the assumption on the properties of the attraction function, It is easy to gather that the inverse function $g_b(z) = f_b^{-1}(p_b)$ will satisfy the following:

$$g'_b(z) \leq 0, \quad \lim_{z \rightarrow 0} g_b(z) = \infty, \quad \text{and} \quad \lim_{z \rightarrow \infty} g_b(z) = 0.$$

This implies, $\lim_{z \rightarrow 0} \psi(z, p_o) = \infty$, and $\lim_{z \rightarrow \infty} \psi(z, p_o) \leq 0$. In turn, this implies the differential Eq. (5.19) always has a solution because the right hand side is a positive constant for any given p_o . \square

Appendix D: Proof of Proposition 2

Consider the OCP problem in the absence of business rules:

$$Obj = \sum_{j \in J} \tau_j \sum_{m \in M} (p_{mj} - c_{mj}) \frac{f_{mj}(p_{mj})}{1 + \sum_{m' \in M} f_{m'j}(p_{m'j})} \Big|_{p_{mj}=p_m \quad \forall m \in V} \quad (\text{D.1})$$

We take the first derivative and set it equal to zero and obtain the following two conditions where $p_{mj} = p_m \quad \forall m \in V$:

$$\sum_{j \in J} \tau_j \left[\frac{f_{oj}(p_o) + (p_o - c_o) f'_{oj}(p_o)}{1 + \sum_{m' \in M} f_{m'j}(p_{m'j})} - \sum_{m \in M} (p_{mj} - c_{mj}) \frac{f_{mj}(p_{mj}) f'_{oj}(p_o)}{[1 + \sum_{m' \in M} f_{m'j}(p_{m'j})]^2} \right] = 0 \quad \forall o \in V \quad (\text{D.2})$$

$$\frac{f_{bj}(p_{bj}) + (p_{bj} - c_{bj}) f'_{bj}(p_{bj})}{1 + \sum_{m' \in M} f_{m'j}(p_{m'j})} - \sum_{m \in M} (p_{mj} - c_{mj}) \frac{f_{mj}(p_{mj}) f'_{bj}(p_{bj})}{[1 + \sum_{m' \in M} f_{m'j}(p_{m'j})]^2} = 0 \quad \forall b \in M \setminus V, j \in J \quad (\text{D.3})$$

Because the first order conditions are necessary conditions for optimality, all the optimal prices satisfy these conditions, maybe in addition to other prices.

Simplifying by substituting $\epsilon_{mj}(p_{mj}) = \frac{-f'_{mj}(p_{mj})p_{mj}}{f_{mj}(p_{mj})}$ in the conditions we get

$$\sum_{j \in J} \tau_j \frac{f'_{oj}(p_o)}{1 + \sum_{m' \in M} f_{m'j}(p_{m'j})} \left[\frac{p_o}{\epsilon_{oj}(p_o)} + (p_o - c_o) - \sum_{m \in M} (p_{mj} - c_{mj}) \frac{f_{mj}(p_{mj})}{1 + \sum_{m' \in M} f_{m'j}(p_{m'j})} \right] = 0 \quad \forall o \in V \quad (\text{D.4})$$

$$\frac{p_{bj}}{\epsilon_{bj}(p_{bj})} + (p_{bj} - c_{bj}) - \sum_{m \in M} (p_{mj} - c_{mj}) \frac{f_{mj}(p_{mj})}{1 + \sum_{m' \in M} f_{m'j}(p_{m'j})} = 0 \quad \forall b \in M \setminus V, j \in J \quad (\text{D.5})$$

Multiplying each of the latter location specific condition by $\tau_j \frac{f'_{oj}(p_o)}{1 + \sum_{m' \in M} f_{m'j}(p_{m'j})}$ and subtracting it from the first condition, we get,

$$\sum_{j \in J} \tau_j \frac{f'_{oj}(p_o)}{1 + \sum_{m' \in M} f_{m'j}(p_{m'j})} \left[(p_o - c_o) - (p_{bj} - c_{bj}) - \frac{p_o}{\epsilon_{oj}(p_o)} + \frac{p_{bj}}{\epsilon_{bj}(p_{bj})} \right] = 0 \quad (\text{D.6})$$

We refer to this as the general attraction price coordination equation.

Now consider an MNL demand model where $f_{mj}(p_{mj}) = e^{a_{mj} - b_{mj}p_{mj}}$, Eq. (D.6) reduces to:

$$p_o^* = c_o + \sum_{j \in J} w_j(\mathbf{p}) \left[p_{bj}^* - c_{bj} + \frac{1}{b_{oj}} - \frac{1}{b_{bj}} \right] \quad \forall o, b \in M, \quad (\text{D.7})$$

where $w_j(\mathbf{p})$ are the normalized values corresponding to the (weighted) online demand i.e., $b_{oj}D_{oj}(\mathbf{p}^*)$, over all the locations $j \in J$ such that $\sum_{j \in J} w_j(\mathbf{p}) = 1$. \square

Appendix E: Proof of Claim ??

We transformed the objective (??) into objective (??) with constraints (??) and (??-??) and it follows the same steps as in Section 5.1 and hence exact in the integer sense. Because we use SOS2 variables (??-??) transformation is also exact. What remains to show is that $\sum_{s \in S} \bar{U}_{ms} \rho_{ms}$ converges to term $\sum_{s \in S} \bar{U}_{ms} \bar{\theta} w_{ms}$ in the limit, i.e., when the knots increase which happens when $|\bar{U}_{ms1} - \bar{U}_{ms2}| \rightarrow 0$ for any $s1, s2$ that are adjacent indicies. Because $w_{ms} \forall s \in S$ are SOS2 variables, without loss of generality, say $w_{ms1}, w_{ms2} \geq 0$ and the rest are all zero. Therefore, $w_{ms1} + w_{ms2} = 1$ and $\rho_{ms1} + \rho_{ms2} = \bar{\theta}$. Now substituting for ρ_{ms2} and w_{ms2} in the respective terms, they reduce to $(\bar{U}_{ms1} - \bar{U}_{ms2})\rho_{ms1} + \bar{U}_{ms2}\bar{\theta}$ and $(\bar{U}_{ms1} - \bar{U}_{ms2})\bar{\theta}w_{ms1} + \bar{U}_{ms2}\bar{\theta}$, which converge in the limit as $|\bar{U}_{ms1} - \bar{U}_{ms2}| \rightarrow 0$. \square

Appendix F: Proof of Claim ??

We first show that the new formulation is always an upper bound. Because all the steps are exact (in the integer sense) except the lower and upper bounds on U_m , it suffices to show that even with the RLT relaxation and introduction of the variable ρ_{ms} , the lower and upper bounds to the true value of U_m given by Eq. (??) hold. This means we have to show that the following hold:

$$\sum_{s \in S} \bar{U}_{ms} \rho_{ms} \leq \bar{\theta} (R_m^{\max} \eta)^{\gamma_m} \quad (\text{F.1})$$

$$\bar{\theta} (R_m^{\max} \eta)^{\gamma_m} \leq \bar{U}_{mt} \left[(1 - \lambda_m) \bar{\theta} + \lambda_m \sum_{s \in S} \frac{\bar{\eta}_{ms}}{\bar{\eta}_{mt}} \rho_{ms} \right] \quad \forall t \in T \quad (\text{F.2})$$

Because $w_{ms} \forall s \in S$ are SOS2 variables, without loss of generality, say $w_{ms1}, w_{ms2} \geq 0$ and the rest are all zero. Therefore, $w_{ms1} + w_{ms2} = 1$ and $\rho_{ms1} + \rho_{ms2} = \bar{\theta}$. Note that here $\eta = \bar{\eta}_{ms1}w_{ms1} + \bar{\eta}_{ms2}w_{ms2}$.

Now the left hand side (LHS) of [constraint \(F.1\)](#) can be simplified to $\bar{\theta} [\bar{U}_{ms1} \frac{\rho_{ms1}}{\bar{\theta}} + \bar{U}_{ms2} \frac{\rho_{ms2}}{\bar{\theta}}]$. Now substituting the value of \bar{U}_{ms} as a function of $\bar{\eta}_{ms}$, we can deduce that it is always less than the right hand side (RHS) of [constraint \(F.1\)](#) as the RHS is a concave function and LHS is a piecewise linear lower bound.

Now consider the RHS of [constraint \(F.2\)](#). This can be simplified as follows: $\bar{U}_{mt} \bar{\theta} \left[(1 - \lambda_m) + \lambda_m \frac{\bar{\eta}_{ms1} \rho_{ms1} + \bar{\eta}_{ms2} \rho_{ms2}}{\bar{\theta} \bar{\eta}_{mt}} \right]$. This simplifies to a tangent to the concave function on the LHS and hence is always an upper bound. Thus this proves that even with the RLT relaxation of the SOS2 variables, [constraints \(F.1–F.2\)](#) relax the function at its true value and hence the optimal solution of this formulation is an upper bound on OCPN.

The proof of asymptotic optimality follows the same steps as Claim ??, except that here we use it both on the lower and the upper bound. \square

Acknowledgments

First we thank the anonymous omnichannel retailers who engaged with us for this project. Next we thank all the people from IBM Commerce who guided us and helped commercialize OCP with a special thanks to Charles Tze Ng, Emrah Zarifoglu and Suzanne Valentine.

References

- Aggarwal, Gagan, Tomás Feder, Rajeev Motwani, An Zhu. 2004. Algorithms for multi-product pricing. *Automata, Languages and Programming*. Springer Berlin Heidelberg, 72–83.
- Akçay, Yalçın, Harihara Prasad Natarajan, Susan H Xu. 2010. Joint dynamic pricing of multiple perishable products under consumer choice. *Management Science* **56**(8) 1345–1361.
- Ansari, Asim, Carl F Mela, Scott A Neslin. 2008. Customer channel migration. *Journal of Marketing Research* **45**(1) 60–76.
- Aydin, Goker, Evan L Porteus. 2008. Joint inventory and pricing decisions for an assortment. *Operations Research* **56**(5) 1247–1255.
- Bell, D. R., S. Gallino, A. Moreno. 2014. How to win in an omnichannel world. *MIT Sloan Management Review*.
- Berkson, Joseph. 1953. A statistically precise and relatively simple method of estimating the bio-assay with quantal response, based on the logistic function. *Journal of the American Statistical Association* **48**(263) 565–599.
- Bitran, Gabriel, René Caldentey. 2003. An overview of pricing models for revenue management. *Manufacturing & Service Operations Management* **5**(3) 203–229.
- Brynjolfsson, E., Y. J. Hu, M. S. Rahman. 2013. Competing in the age of omnichannel retailing. *MIT Sloan Management Review*.
- Charnes, Abraham, William W Cooper. 1962. Programming with linear fractional functionals. *Naval Research logistics quarterly* **9**(3-4) 181–186.

- Chen, Xin, David Simchi-Levi. 2012. *The Oxford handbook of pricing management*, chap. Pricing and inventory management. Oxford University Press, 784–822.
- Chintagunta, Pradeep K, Junhong Chu, Javier Cebollada. 2012. Quantifying transaction costs in online/off-line grocery channel choice. *Marketing Science* **31**(1) 96–114.
- Davis, James M, Huseyin Topaloglu, David P Williamson. 2016. Pricing problems under the nested logit model with a quality consistency constraint. *INFORMS Journal on Computing* **29**(1) 54–76.
- Domencich, Thomas A, Daniel McFadden. 1975. *Urban travel demand-a behavioral analysis*. North-Holland.
- Dong, Lingxiu, Panos Kouvelis, Zhongjun Tian. 2009. Dynamic pricing and inventory control of substitute products. *Manufacturing & Service Operations Management* **11**(2) 317–339.
- Elmaghraby, Wedad, Pinar Keskinocak. 2003. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science* **49**(10) 1287–1309.
- Fisher, Marshall, Ramnath Vaidyanathan. 2014. A demand estimation procedure for retail assortment optimization with results from implementations. *Management Science* **60**(10) 2401–2415.
- Gallego, Guillermo, Ruxian Wang. 2014. Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities. *Operations Research* **62**(2) 450–461.
- Goolsbee, Austan. 2001. Competition in the computer industry: Online versus retail. *The Journal of Industrial Economics* **49**(4) 487–499.
- Hanson, Ward, Kipp Martin. 1996. Optimizing multinomial logit profit functions. *Management Science* **42**(7) 992–1003.
- Keller, Philipp W, Retsef Levi, Georgia Perakis. 2014. Efficient formulations for pricing under attraction demand models. *Mathematical Programming* **145**(1-2) 223–261.
- Krugman, Paul, Robin Wells. 2008. *Microeconomics*. 2nd ed. New York: Worth Publishers.
- Li, Hongmin, Woonghee Tim Huh. 2011. Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications. *Manufacturing & Service Operations Management* **13**(4) 549–563.
- McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics* 105–142.
- Montgomery, Alan L. 1997. Creating micro-marketing pricing strategies using supermarket scanner data. *Marketing science* **16**(4) 315–337.
- Mulhern, Francis J, Jerome D Williams, Robert P Leone. 1998. Variability of brand price elasticities across retail stores: Ethnic, income, and brand determinants. *Journal of Retailing* **74**(3) 427–446.
- Neslin, Scott A, Dhruv Grewal, Robert Leghorn, Venkatesh Shankar, Marije L Teerling, Jacquelyn S Thomas, Peter C Verhoef. 2006. Challenges and opportunities in multichannel customer management. *Journal of Service Research* **9**(2) 95–112.

- Pisinger, David. 1994. A minimal algorithm for the multiple-choice knapsack problem. *European Journal of Operational Research* **83** 394–410.
- Rayfield, W Zachary, Paat Rusmevichientong, Huseyin Topaloglu. 2015. Approximation methods for pricing problems under the nested logit model with price bounds. *INFORMS Journal on Computing* **27**(2) 335–357.
- Retail Dive. 2015. A chilly black friday for some brick-and-mortar retailers online sales appear strong as line blurs between shopping experiences. URL <http://www.retaildive.com/news/a-chilly-black-friday-for-some-brick-and-mortar-retailers/409884/>.
- Rusmevichientong, Paat, David Shmoys, Huseyin Topaloglu. 2010. Assortment optimization with mixtures of logits. Tech. rep., Tech. rep., School of IEOR, Cornell University.
- Rusmevichientong, Paat, Benjamin Van Roy, Peter W Glynn. 2006. A nonparametric approach to multi-product pricing. *Operations Research* **54**(1) 82–98.
- Schön, Cornelia. 2010. On the product line selection problem under attraction choice models of consumer behavior. *European Journal of Operational Research* **206**(1) 260–264.
- Sherali, Hanif D, Warren P Adams. 1999. Reformulation-linearization techniques for discrete optimization problems. *Handbook of combinatorial optimization*. Springer, 479–532.
- Song, Jing-Sheng, Zhengliang Xue. 2007. Demand management and inventory control for substitutable products. *Working paper*.
- Subramanian, Shivaram, Pavithra Harsha. 2017. Demand modeling in the presence of unobserved lost sales. Submitted, IBM Research.
- Subramanian, Shivaram, Hanif D. Sherali. 2010. A fractional programming approach for retail category price optimization. *Journal of Global Optimization* **48**(2) 263–277.
- Train, Kenneth E. 2009. *Discrete choice methods with simulation*. 2nd ed. Cambridge university press.
- Urban, Glen L. 1969. A mathematical modeling approach to product line decisions. *Journal of Marketing Research* 40–47.
- US Census Bureau. 2017. Quarterly retail e-commerce sales. URL https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf.
- Zhang, Jie, Paul W Farris, John W Irvin, Tarun Kushwaha, Thomas J Steenburgh, Barton A Weitz. 2010. Crafting integrated multichannel retailing strategies. *Journal of Interactive Marketing* **24**(2) 168–180.