How big should my store be? On the interplay between shelf-space, demand learning and assortment decisions

Kevin Glazebrook[1], Joern Meissner[2], and Jochen Schurr[3]

[1] *Department of Management Science, Lancaster University Management School, United Kingdom.* k.glazebrook@lancaster.ac.uk,
[2] *Kuehne Logistics University, Hamburg, Germany.* joe@meiss.com
[3] *Department of Management Science, Lancaster University Management School, United Kingdom.* j.schurr@lancaster.ac.uk,

December 12, 2012

Abstract

A fundamental decision every merchant has to make is on is how large his stores should be. This is particularly true in light of the drastic changes retail concepts have seen in the last decade. There has been a noticeable tendency, particularly for food and convenience retailers, to open more and smaller stores. Also, there has been a well-documented recent shift in paradigm in apparel retailing with the so called fast-fashion business model. Short lead times have resulted in flexibility that allows retailers to adjust the assortment of products offered on sale at their stores quickly enough to adapt to popular fashion trends. Based on revised estimates of the merchandise's popularity, they then weed out unpopular items and re-stock demonstrably popular ones on a week-by-week basis. However, despite the obvious similarity of reliance on better demand learning, fashion-fashion retailers like Zara have opted to do exactly the opposite as groceries and opened sizable stores in premium locations. This paradox has not been explained in the literature so far. In this paper, we aim to calculate the profit of a retailer in such a complicated environment with demand learning and frequent assortment decisions in particular in dependence of the most valuable resource of a retailer: shelf-space. To be able to achieve this, we extend the recent approaches in the management literature to handle the sequential resource allocation problems that arises in this context with a concurrent need for learning. We investigate the use of multi-armed bandits to model the assortment decisions under demand learning, whereby this aspect is captured by a Bayesian Gamma-Poisson model. Our model enables us to characterize the marginal value of shelf-space and to calculate the optimal store size under learning and assortment decisions. An extensive numerical study confirms that the store size choices observed in real life can be explained by the varying length of selling seasons different retailers face.

Keywords: retailing, assortment planning, multi-armed bandit, store size

1 Introduction

The question of how big retail premises should be is so fundamental that every merchant has to face it. It is even more important under the strong competition most retailers face nowadays. As most of the industry works on razor thin margins, any firm deviating from the optimal store size for its particular business model will face the danger of reduced profits or even losses. The problems are even more pronounced in light of the fact that retail concepts have changed dramatically in the last decade, most notably by incorporating better strategies to learn about customer demand. On the one hand, there has been a tendency, particularly for food and convenience retailers, to open more and smaller stores. One of the earliest examples of this strategy was the now seminal case of 7-Eleven Japan. The company has opened many small stores in premium locations. This enables them to have location close to the customers convenience, aside from implementing innovative replenishment solutions. The ensuing demand is closely monitored and products on sale adjusted, up to the point that the assortment changes daily to perfectly fit the demand and to make the most use out of the scarce shelf-space. The strategy has been copied, for example, by Tesco in the United Kingdom.

A similar trend towards demand learning has taken place in the so called fast-fashion industry. Implementation of this new fast-fashion paradigm at Zara and others hinges on merchandize procurement strategies that permit lead times as short as two weeks. The resulting flexibility allows retailers to adjust the assortment of products offered on sale at their stores quickly enough to adapt to popular fashion trends. In particular, firms can choose from a large number of potential styles to produce and offer for sale, and then use weekly sales data to renew their estimates of specific items' popularity. Based on such revised estimates, they then weed out unpopular items, or else re-stock demonstrably popular ones on a week-by-week basis. In sharp contrast, traditional retailers such as Marks and Spencer face lead times on the order of several months. As such these retailers need to predict popular fashions months in advance and are allowed virtually no changes to their product assortments over the course of an entire sales season, which is typically several months in length. While providing cost benefits, this approach typically results in substantial unsold inventories at the end of a sales season while failing to identify other high selling styles. In view of the great deal of a-priori uncertainty in the popularity of a new fashion and the speed at which fashion trends evolve, the fast-fashion operations model is emerging as the de-facto standard operations model for fashion retailers.

However, one could observe that in contrast to the grocery stores mentioned above, fast-fashion retailers have tended to open bigger stores in high street locations. In this paper, we set out to explain this phenomenon and develop guidelines for managers having to make such decisions. A central aim of the paper is to determine the marginal value of shelf-space, a quantity of fundamental importance to any retailer operating in such an environment. We will develop a range of analytical and computational competences which will enable us to understand how this marginal value declines (as it does) as the store size grows. We implement numerical studies which lend support to the contention that sectors with more rapidly changing demand patterns, as evidenced by shorter selling seasons (i.e., fast fashion rather then groceries), will have larger stores, not least to facilitate rapid demand learning.

The paper is structured as follows: Section 2 contains a review of the related literature. We present a model for the problem concerning the dynamic assortment of retail products over a selling season which lies at the heart of the paper in section 3. In section 4 we use a Lagrangian relaxation of the dynamic assortment problem to develop a product index to serve as a means of calibrating the strength of a product's candidature for inclusion in an assortment. We propose a range of heuristic approaches to the use of these indices to produce admissible product assortments. In section 5 we report on an extensive numerical study which explores the closeness to optimality of our heuristic index policies and reports on how the marginal value of shelf-space (as a function of store size) relates to key problem features including the length of the selling season. This in turn yields the insights mentioned in the previous paragraph.

2 Literature Review

We review the literature in four four key areas, each of which plays a significant role in our models and analyses.

2.1 Assortment Planning

Central to the issues raised in the Introduction is the question of how retailers choose the assortment of products which they place on sale in their stores. Many publications in the field of (static) assortment planning formulate the problem as a mathematical program. A good example is Hariga et al. (2007). They introduced an optimization model for assortment planning with a focus on the joint consideration of inventory, assortment, shelf space and display area. This results in a mixed-integer program which can be solved with a commercial package. Their demand function is designed to incorporate the locational positioning as well as the shelf space allocated to the product and the cross-product-elasticity. It can be extended to incorporate the selling price.

A similar approach in terms of decision modeling was followed by Yücel et al. (2009). Their focus is on substitution effects. These are modeled by redirecting respective portions of the demand whenever a stock-out is encountered. In their objective function, the authors consider half a dozen cost types in order to tailor precisely the total profit, which is to be maximized. An exceptionally rich source giving an overview of assortment planning and adjacent topics is the book chapter written by Kök et al. (2008).

2.2 Dynamic Assortment Planning

Fisher and Raman (1996) studied the decisions arising in a two-period model consisting of an initial forecast of the demand and a refined forecast after observing some of the demand, thereby reducing the risk of stock-outs and obsolete inventory.

A more recent contribution to the field of dynamic assortment planning is the work of Caro and Gallien (2007). The authors focused on increasing revenue by dynamically optimizing the assortment, which is put onto the shelves in the show room, in the context of fast-fashion retailers, such as Zara, H&M or Mango. The model consists of a Bayesian Gamma-Poisson learning scheme for the demand and a multi-armed bandit for the decision process. A Lagrangian relaxation and further simplifications allow them to derive an index-based heuristic, which performs competitively against adapted versions of the work of Brezzi and Lai (2002) and Ginebra and Clayton (1995).

Recently, Saure and Zeevi (2009) chose another modeling path and considered a decision structure in which customers choose one or no product so as to maximize their utility. The asymptotic performance of their policy is within a quantity which is of the order of the logarithm of the time horizon below the theoretical full-knowledge optimum. This stream of work aims towards applications where a large number of measurements can be obtained.

2.3 Multi-Armed Bandits

Adaptive dynamic resource allocation problems (such as dynamic assortment planning) have long been modelled as multi-armed bandit problems. A celebrated contribution was that of Gittins (1979) who elucidated the optimality of policies of index form for an infinite horizon setting with discounted returns. Whittle (1980) subsequently developed a more transparent proof of the Gittins index theorem by demonstrating explicitly that the index policy's value function satisfies the Bellman equations. Some years later, Whittle (1988) introduced a class of so-called restless bandits which generalised the Gittins model by allowing state evolution in passive bandits. His mode of analysis via Lagrangian relaxation is now the tool of choice to analyse (generalisations of) such models and we shall use this approach in section 4. Weber and Weiss (1990) found conditions under which the index policies proposed by Whittle are asymptotically optimal in a regime in which the number of available bandits and the number chosen for processing at each epoch diverge to infinity in fixed proportion.

A focus of current research has been the analysis of a multitude of variations of the original multi-armed bandit and the restless bandit problem: Glazebrook et al. (2006) studied two practically motivated families of restless bandits. The spinning plates problem represents a situation in which assets tend to decline in profitability over time unless there is an impulse of either active management or investment. The second one is the so called squad system, which models the phenomenon of resource exploitation. Mahajan and Teneketzis (2007) presented various extensions of the classical bandit problem, such as arm-acquiring bandits, switching penalties, bandits with multiple plays and restless bandits. Farias and Madan (2008) considered so-called irrevocable multi-armed bandits. In this model, an arm once pulled and then discarded, can never be pulled again. Chakrabarti et al. (2008) considered so called mortal multi-armed bandits, where arms have a limited lifetime. The interested reader should consult Gittins et al. (2011), which provides an excellent overview of the index theory behind multi-armed bandits and related processes.

2.4 Shelf Space Optimization

Both academics and practitioners alike have identified shelf space optimization as one of the key levers for retailers. Fancher (1991) mentioned that with the emergence of computerized assortment software for retailers in the 1980s, a more analytical approach to optimizing the use of this valuable resource became possible. Corstjens and Doyle (1981) considered shelf space allocation as “a central problem in retailing” and proposed a model, which incorporates the elasticity of the demand function for each individual product as well as cross elasticity effects between different products. Similarly, some years later Bultez and Naert (1988) stated that “shelf space is the retailer’s scarcest resource.” and suggested an extension of the model Corstjens and Doyle to allow demand to interrelate within and across product-groups. The authors formulated a maximization problem with a shelf space constraint, developed a Lagrange relaxation and derived an optimal allocation by solving a simplified version of the relaxed problem. Citing a 1990s study, according to which shelf space optimization was among the top three purposes for the collection of scanner data, Wartenberg et al. (1997) stated that space allocation among concurrent products is a central and regularly recurring problem in retailing. Yang and Chen (1999) and Yang (2001) agreed that “shelf space is one of the most important resources” of a retail firm. Starting with a similar formulation to Corstjens and Doyle (1981), the authors introduced an alternate form that relies on a simpler, but computationally costlier integer programming approach. Yang (2001) extended the model to consider various constraints for shelf space usage. More recently, Nafari and Shahrabi (2010) took into account the price sensitivity of shelf space decisions. As prices affect the interrelations and substitutability between products, they affect shelf allocation decisions as well. Their model maximizes the total profit while taking account of the cost of inventory and transportation. An empirical study demonstrates the applicability of the approach.

To our knowledge, the fusion of ideas in the current paper is quite novel. We formulate an issue of shelf space usage as a dynamic assortment problem in a manner which permits the determination of *inter alia* the marginal value of shelf space. This in turn yields insights concerning optimal store size.

3 A Model for Dynamic Assortment Planning over a Selling Season

In this section we develop a stylized model for dynamic assortment planning in a context similar to fast fashion retailing. However, we do not see the application of our model limited to this domain only. The focus is set on central questions of operations management, namely which products to offer at which time, while other aspects may be incorporated in the developed model in hindsight. Our model makes the following assumptions: The retailer is managing a single store and this store has a show room with a fixed amount of available shelf space.

The core question within the given problem is, how the retailer should choose a limited number of items out of a huge set of products, such that he achieves maximal profit. Each decision is valid for one period (week), that is during each week the assortment is only replenished, not replaced by other products. Many consecutive periods form a season and the retailer needs to make a decision at the beginning of each period concerning how he wants to make use of his shelf space. The information (sales data) gained in past periods provide the basis for future decisions.

The presence of a limited selling season and the fact that the set of products is changed almost completely after each season, raises the need to gather information about the demand of each item to make an informed decision and on the other hand to exploit this knowledge to achieve the primary goal, which is profit maximization. It is therefore necessary to find a balance between testing the market for new products and selling as many as possible of the highest demand items found so far. It is rather obvious that exploration will have more importance in the beginning and exploitation is stressed towards the end of the season. More specific results on that will be discussed at a later point.

The constraining quantity is shelf space and the main uncertainty the demand for each of the products. In order to fully concentrate on the assortment aspects, inventory considerations are postponed and thus perfect replenishment is assumed. The demand and profit per item, from here on mostly called the reward, are assumed to be constant, justified by the fact that markdowns are rare in fast fashion stores. Moreover in the basic model there is no time lag between decision making and execution and no switchover cost associated with changing the assortment, i.e. putting a different item on the shelf than in the previous period.

The selling season is divided into T equally sized periods (usually weeks) and the time index t gives the remaining sales periods, i. e. time steps are counted backwards. Shelf space is expressed in units, which can relate to area or a custom measure of the retailer. The maximum capacity of shelf space is $C_{\max}$ units and before each period, the retailer makes a decision concerning how to make use of it. There is a number of S items in the set of products and each one, denoted by s, comes with a fixed reward r_s and a shelf space requirement c_s of integer value. There is an unknown, constant mean demand rate γ_s for product s and in a given period a random number of n_s of them are sold.

The model formulates decision making as a multi-armed bandit with $\sum_{s=1}^{S} c_s$ arms, where each arm stands for one unit of shelf space. In each period $C_{\max}$ arms can be pulled, where for each item s the number of arms to be pulled is either 0 or c_s. This refers to the wish of a retailer to have at least a certain amount of a product on the shelf, if any. In case the product is included, a reward of r_s per unit sold is earned. The retailer's objective is to maximize the expected sum of rewards over all pulls and all periods.

Whenever an arm or a cluster of arms belonging to item s is pulled, the number of sales n_s, which is computed as the realization of a Poisson distribution, gives further information about the unknown rate, i. e. the mean weekly demand γ_s. We adopt a Bayesian approach to parametric uncertainty in which the initial beliefs about the demand rates γ_s are summarised by independent Gamma distributions with shape parameters m_s and scale parameters α_s. This implies that the prior mean is $\mathbb{E}[\gamma_s] = m_s/\alpha_s$ and the prior variance $\mathbb{V}[\gamma_s] = m_s/\alpha_s^2$. This set up yields an easy to handle Gamma-Poisson Bayesian learning mechanism as in Aviv and Pazgal (2005). For a Poisson likelihood, the Gamma distribution is said to be a conjugate prior, which means that the posterior for γ_s remains in the class of Gamma distributions for an arbitrary number of Bayesian updates. Doing the algebra leads to the insight that the parameter transition from one period to the next depending on the observation n_s is $(m_s, \alpha_s) \to (m_s + n_s, \alpha_s + 1)$ when the arms corresponding to item s are pulled. If item s is not selected for the assortment, there is no transition in the demand parameter state. Another quantity of interest is the predictive distribution for n_s, which turns out to be a negative binomial distribution with parameters m_s and $\alpha_s(\alpha_s + 1)^{-1}$. Therefore the expectation of n_s in advance of sampling is $\mathbb{E}[n_s] = m_s/\alpha_s$ and its variance is $\mathbb{V}[n_s] = m_s(\alpha_s + 1)/\alpha_s^2$.

The first constraint mentioned above concerns shelf space availability. We now introduce a second one, which reflects a retailer's wish for a well-balanced mix of products in the store. Every item within the set of products is assigned a unique affiliation with a product category, such as shoes, shirts, accessories, etc. Each category needs to be allocated a preset minimum amount (potentially 0) of shelf space coverage. This is to ensure that marketing related aspects like cross-selling or customer perception as to what kind of retailer the store belongs to are fulfilled.

Every product belongs to exactly one of k categories. The resulting partition of $\{1,\ldots,S\}$ into k subsets is denoted $K_1,\ldots,K_k$. An admissible assortment needs to allocate at least L_i units of shelf space to category i products. Let u_s be a logical decision variable indicating whether or not to include product s in the assortment. Then the second constraint reads $\sum_{s\in K_i} c_s u_s \ \geq \ L_i, \ i=1,\ldots,k$.

The set of admissible assortments is therefore given by

$$\mathcal{U} := \{\mathbf{u} \in \{0,1\}^S : \ \sum_{s=1}^{S} c_s u_s \leq C_{\max} \ \wedge \ \sum_{s\in K_i} c_s u_s \geq L_i, \ i=1,\ldots,k\}. \tag{1}$$

Note that the existence of an admissible assortment is subject to the choice of parameters $\boldsymbol{c}$, $(L_i)_{i=1,\ldots,k}$, $(K_i)_{i=1,\ldots,k}$, and $C_{\max}$. Accordingly, there is a need to check the well-posedness of the problem.

4 The Development of Index Heuristics for the use of shelf-space

Our analysis of the shelf-space problem outlined in the previous section will proceed in three steps. In step 1 we shall use a Lagrangian relaxation of the dynamic programming (DP) problem to develop a computable upper bound on the optimal return achievable over a selling season. This upper bound will be used in the following section as a means of evaluating competing heuristic solutions to the shelf-space problem. In step 2 we give an account of a product index for evaluating how competitive a product is for inclusion in the selected assortment. Finally in step 3 we describe heuristics which use index values to construct close-to-optimal assortments.

4.1 A computable upper bound on the optimal return

Due to the sequential character of the problem, we can formulate an analytical solution via dynamic programming. The Bellman equations for the optimal profit-to-go function J_t^* are given by

$$
\begin{aligned}
J_t^*(\mathbf{m},\boldsymbol{\alpha}) &= \max_{\mathbf{u}\in\mathcal{U}} \sum_{s=1}^{S} r_s \tfrac{m_s}{\alpha_s} u_s + \mathbb{E}_{\mathbf{n}}[J_{t-1}^*(\mathbf{m}+\mathbf{n}\cdot\mathbf{u},\boldsymbol{\alpha}+\mathbf{u})] \\
J_0^*(\mathbf{m},\boldsymbol{\alpha}) &= 0
\end{aligned}
\tag{2}
$$

where the expectation is with respect to the predictive distribution mentioned above. The operator "·" is to be understood as the component wise product of two vectors, resulting in a vector of the same size as both input vectors.

As the above Bellman equations reveal, the only dependence between choices is induced by the constraints. Therefore a natural next step is to relax the constraints taking a Lagrangian approach. The Lagrangian H, the sum of the original function and the weighted deviation from meeting the constraints, satisfies

$$
\begin{aligned}
H_t^{\boldsymbol{\lambda},\boldsymbol{\mu}}(\mathbf{m},\boldsymbol{\alpha}) &= C_{\max}\lambda_t(\mathbf{m},\boldsymbol{\alpha}) \; - \textstyle\sum_{i=1}^{k} L_i\, \mu_{i,t}(\mathbf{m},\boldsymbol{\alpha}) \\
&\quad + \max_{\mathbf{u}\in\{0,1\}^S} \sum_{s=1}^{S} \Big(\tfrac{r_s}{c_s}\tfrac{m_s}{\alpha_s} - \lambda_t(\mathbf{m},\boldsymbol{\alpha}) + \mu_{\kappa(s),t}(\mathbf{m},\boldsymbol{\alpha}) \Big) c_s u_s \\
&\quad + \mathbb{E}_{\mathbf{n}}[H_{t-1}^{\boldsymbol{\lambda},\boldsymbol{\mu}}(\mathbf{m}+\mathbf{n}\cdot\mathbf{u},\boldsymbol{\alpha}+\mathbf{u})] \\
H_0^{\boldsymbol{\lambda},\boldsymbol{\mu}}(\mathbf{m},\boldsymbol{\alpha}) &= 0.
\end{aligned}
$$

The Lagrange multipliers $\lambda_t(\mathbf{m},\boldsymbol{\alpha})$ and $\mu_{i,t}(\mathbf{m},\boldsymbol{\alpha})$ depend on the period t and the state $(\mathbf{m},\boldsymbol{\alpha})$. We introduce the shorthand notation $\boldsymbol{\lambda} = (\lambda_t)_{t=1..T}$ and similarly $\boldsymbol{\mu_1},\ldots,\boldsymbol{\mu_k}$. The vectors $\boldsymbol{\mu_1},\ldots,\boldsymbol{\mu_k}$ will notationally be further condensed to a single quantity $\boldsymbol{\mu}$. For the purpose of referencing from product index to category we use a category affiliation vector $\boldsymbol{\kappa} \in \{1,\ldots,k\}^S$.

In the practical context of product choice the multipliers $\boldsymbol{\lambda}$ can be thought of as a price for one shelf space unit per period. Note that $\lambda_t \geq 0$ due to the nature of the constraint equation. The multipliers $\boldsymbol{\mu}$ can be thought of as a price incurred when introducing the marketing related demands of the second constraint. As the differing direction of the inequality in the second constraint resulted in a negative sign in the Lagrangian, we also have $\mu_{i,t} \geq 0$.

For the sake of computational feasibility, a slight simplification now delivers a product-wise decomposition within the equations. Let the multipliers $\boldsymbol{\lambda},\boldsymbol{\mu}$ be constant over the states $(\mathbf{m},\boldsymbol{\alpha})$ and hence dependent only on the time t. Straightforward algebraic manipulations yield

$$\begin{aligned} H_t^{\boldsymbol{\lambda},\boldsymbol{\mu}}(\mathbf{m},\boldsymbol{\alpha}) &= \sum_{\tau=1}^{t}\left(C_{\max}\lambda_\tau - \sum_{i=1}^{k} L_i\,\mu_{i,\tau}\right) + \sum_{s=1}^{S} H_{t,s}^{\boldsymbol{\lambda},\boldsymbol{\mu}}(m_s,\alpha_s) \qquad \text{with} \\ H_{t,s}^{\boldsymbol{\lambda},\boldsymbol{\mu}}(m_s,\alpha_s) &= \max\{r_s\tfrac{m_s}{\alpha_s} - \lambda_t c_s + \mu_{\kappa(s),t}c_s + \mathbb{E}_{n_s}[H_{t-1,s}^{\boldsymbol{\lambda},\boldsymbol{\mu}}(m_s+n_s,\alpha_s+1)], \\ &\qquad H_{t-1,s}^{\boldsymbol{\lambda},\boldsymbol{\mu}}(m_s,\alpha_s)\}\ . \end{aligned} \tag{3}$$

Following standard Lagrangian theory, the sought maximum of the profit-to-go function can now be found by minimizing $H_t^{\boldsymbol{\lambda},\boldsymbol{\mu}}$ with respect to $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$. Denote that minimum by H_t^*. It can be proven quite easily, see for example Caro and Gallien (2007), that the solution H^* of the dual problem never underestimates the maximum, but serves as an upper bound (Weak DP Duality). We have

$$J_t^*(\mathbf{m},\boldsymbol{\alpha}) \;\le\; H_t^*(\mathbf{m},\boldsymbol{\alpha}) := \min_{\substack{\boldsymbol{\lambda},\boldsymbol{\mu_i}\ge 0,\\ i=1,\ldots,k}} H_t^{\boldsymbol{\lambda},\boldsymbol{\mu}}(\mathbf{m},\boldsymbol{\alpha}) \tag{4}$$

for all periods t, states $(\mathbf{m},\boldsymbol{\alpha})$, and Lagrange multipliers $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$.

Note that, in (3) we have achieved a decomposition into single product sub-problems and hence a considerable saving in computational effort. The resulting upper bound H^* is now computable for problems involving several hundred products. It will emerge that it is also sufficiently tight that it can be effectively used in section 5 as a means of assessing the closeness-to-optimality of our proposed heuristic solutions to the shelf-space problem.

A primary focus of the paper concerns how the optimal return from the assortment problem depends on the amount of available shelf-space $C_{\max}$. Given the tightness of the upper bound H^* and its consequential key role in the numerical exploration of this issue in section 5, we record in Proposition 1 some important facts about how it depends upon $C_{\max}$. In preparation for this, simply observe that (3) and (4) together imply, for suitably chosen $\Phi : \mathbb{R}^{(1+k)T} \to \mathbb{R}$, that $H_T^*(\mathbf{m},\boldsymbol{\alpha})$ has a representative form

$$H_T^*(\mathbf{m},\alpha) = \min_{\lambda,\mu_i\ge 0, i=1,2,\ldots,k}\left\{\Phi(\mathbf{m},\alpha,\lambda;\mu_i,1\le i\le k) + C_{\max}\left(\sum_{t=1}^{T}\lambda_\tau\right)\right\}. \tag{5}$$

Observe that, for all values of $T,\mathbf{m},\alpha,\lambda;\mu_i,1\le i\le k$, the bracketed quantity on the right hand side of (5) is linear and non-decreasing in $C_{\max}$. Proposition 1 follows from standard results.

Proposition 1. *For all $T,\mathbf{m},\alpha$, the upper bound $H_T^*(\mathbf{m},\alpha)$ is concave and non-decreasing in $C_{\max}$.*

It will turn out to be very important to achieve the purposes set out in the Introduction to have access to the rate of change of the upper bound on achievable returns $H_T^*(\mathbf{m}, \alpha)$ with respect to the available shelf-space $C_{\max}$. For reasons which will emerge more fully in the numerical study of section we call this *the marginal value of shelf-space.* The reader will find more information about this key quantity in subsection 5.3.

4.2 Product indices for shelf-space use

We develop the analysis further by utilizing the Lagrangian relaxation in (3) to produce an index for each product which calibrates its status for inclusion in any selected assortment. To achieve this we take the single product problem in (3) and introduce a single combined multiplier $G_{s,t} := \lambda_t - \mu_{\kappa(s),t}$ which can be understood as a net charge per unit of shelf-space used by product s at time t from the end of the selling season. Hence the single product DP in (3) has a simple interpretation which concerns whether or not product s should be included in the assortment when inclusion is subject to such a charge. As we shall now focus on a single product we drop the identifier s from the notation and thus have a collection of multipliers $\{G_t, t \geq 1\}$ which, for tractability we assume to take the form $G_t = G\theta_t$, where $G \in \mathbb{R}$ and $\{\theta_t, t \geq 1\}$ is a decreasing sequence of positive real numbers. Crucial simplifications now follow concerning the nature of solutions to the DP. We thus consider the single product problem given by

$$H_t^G(m, \alpha) = \max\left\{r\frac{m}{\alpha} - G\theta_t c + \mathbb{E}_n\left[H_{t-1}^G(m+n, \alpha+1)\right], H_{t-1}^G(m, \alpha)\right\} \tag{6}$$

We say that the DP in (6) is a *stopping problem* if it is true that whenever it is optimal to exclude the product from the assortment at t then it must also be optimal to exclude the product at all times $t' \leq t$. This will be the case if $\forall m, \alpha, t \geq 2$

$$H_t^G(m, \alpha) = H_{t-1}^G(m, \alpha) \quad \Rightarrow \quad H_{t-1}^G(m, \alpha) = H_{t-2}^G(m, \alpha) \tag{7}$$

We further say that the single product problem is *indexable* if $H_t^G(m, \alpha)$ is decreasing in G $\forall m, \alpha, t \geq 1$. The reason for this terminology will become clear later.

Theorem 1. *The single product problem in (6) is* (a) *a stopping problem and* (b) *is indexable.*

Proof. We first prove (a). First note that we may restrict to $G \geq 0$ without loss of generality since if $G < 0$ it is trivially always beneficial to include the product in the assortment. We establish (7) using a proof by contradiction. We first note from (6) that

$$H_t^G(m, \alpha) \geq H_{t-1}^G(m, \alpha) \quad \forall m, \alpha, t \geq 1 \tag{8}$$

We suppose now that $\exists m, \alpha, t \geq 1$ for which it is strictly optimal to exclude the product from the assortment at t but optimal to include it at time $t-1$. It must then follow that

$$r\frac{m}{\alpha} - G\theta_t c + \mathbb{E}_n \left[H_{t-1}^G (m+n, \alpha+1) \right] < H_{t-1}^G (m, \alpha) = r\frac{m}{\alpha} - G\theta_{t-1} c + \mathbb{E}_n \left[H_{t-2}^G (m+n, \alpha+1) \right] \tag{9}$$

But using the facts that $G \geq 0, \theta_t \leq \theta_{t-1}$ within (9) we conclude that

$$\mathbb{E}_n \left[H_{t-1}^G (m+n, \alpha+1) \right] < \mathbb{E}_n \left[H_{t-2}^G (m+n, \alpha+1) \right]$$

which contradicts (8). This concludes the proof of (a). To establish (b), we first note that it follows from the fact that our single product problem is a stopping problem that the DP equations may be recast as

$$H_t^G (m, \alpha) = \max_{\lambda \geq 0} \left\{ r\frac{m}{\alpha} - G\theta_t c + \mathbb{E}_n \left[H_{t-1}^G (m+n, \alpha+1) \right], 0 \right\}, \tag{10}$$

where $H_0^G = 0$. An argument which uses a simple induction on t now establishes that $H_t^G (m, \alpha)$ is decreasing in G $\forall m, \alpha, t \geq 1$ as required. This completes the proof. □

We now introduce the *product index* $G_t (m, \alpha)$, where

$$G_t (m, \alpha) := \inf \left\{ G : H_t^G (m, \alpha) = 0 \right\}. \tag{11}$$

Note that since trivially we have that $H_t^G (m, \alpha) > 0$ whenever $G < 0$, it is clear that all product indices must be positive. With this in place, the solution to the single product problem is as described in the following result which we give without proof.

Proposition 2. *If the single product is in state (m, α) at time t to the end of the season then it is optimal to include the product in the assortment if and only of $G \leq G_t (m, \alpha)$.*

Please note that it must follow that the solution to the multi-product Lagrangian relaxation in (3) must have an index form when we set $\lambda_t - \mu_{\kappa(s),t} = G\theta_{s,t}$ $\forall s, t$ with $\{\theta_{s,t}, t \geq 1\}$ decreasing and positive $\forall s$. By the above account, we are guaranteed the existence of indices $G_{s,t} (m_s, \alpha_s)$ such that if the system state is $(\mathbf{m}, \alpha)$ at time t to the end of the season then in the relaxed problem in (3) precisely those products are included in the optimal assortment whose indices exceed G.

Reverting again to the single product, in what follows we shall make the choice $\theta_t = 1, t \geq 1$, and defer to future work an exploration of the benefits which exploitation of the $\theta-$sequence may confer. For example, the above analysis makes clear that we can produce index solutions for cases where the per unit charge for shelf-space increases through the season. In any event, from the above proposition when $\theta_t = 1$ the quantity $G_t(m, \alpha)$ has an interpretation as a *fair charge* for a unit of shelf space to be allocated to the product when in state (m, α) at time t. If a charge for shelf space is levied which is in excess of $G_t(m, \alpha)$ then it will be optimal *not* to field the product. This decision is reversed for charges below the value of the product index.

We can infer important structural properties of the indices $G_t(m, \alpha)$ from properties of the value $H_t^G(m, \alpha)$. The proof details for Lemma 1 and Lemma 2 are available from the authors. Lemma 3 follows directly from the previous two using (11).

Lemma 1. *$H_t^G(m, \alpha)$ is increasing in m.*

Lemma 2. *$H_t^G(m, \alpha)$ is decreasing in α.*

Lemma 3. *$G_t(m, \alpha)$ is increasing in m and decreasing in α.*

Further, from (3) we see that $H_t^G(\mathbf{m}, \alpha)$ and hence $G_t(m, \alpha)$ are increasing in t. This is expressed in Lemma 4.

Lemma 4. *$G_t(m, \alpha)$ is increasing in t.*

The above results all accord with intuition. Since the mean demand per unit of time in state (m, α) is m/α, increasing m and decreasing α are alternative ways of making the current state more attractive. That the product index should increase in line with such changes is thus unsurprising. It is the result in the final lemma which in many ways is the most insightful. For larger t there is greater time and opportunity to exploit whatever is learnt about the product's true demand rate as the product is fielded. The product index increases with t as a result.

Even with the simplification offered by taking $\theta_t = 1, t \geq 1$, the task of index computation is an exacting one and we shall exploit the availability of index approximations in the literature deployed in simpler contexts than the current one. We shall use three approximations to the index $G_t(m, \alpha)$ in what follows. Firstly, the *greedy index* is given $\forall m, \alpha, t \geq 1$ by

$$G_t^{GDY}(m, \alpha) = G_1(m, a) = \frac{rm}{c\alpha} \tag{12}$$

and assesses products on the sole basis of expected return in the next period. Heuristics based on the greedy index are myopic and ignore longer term consequences of decisions including the opportunity of learning about demands for products. The other two approximations both exploit a normal approximation to the predictive distribution for demand in state (m, α) whose exact form is negative binomial. They both supplement the greedy index by a second term which reflects the benefit derived by the opportunity to learn about the product's true demand rate. The approximating index derived from the work of Brezzi and Lai (2002) takes the form

$$G_t^{BL}(m,\alpha) = \frac{rm}{c\alpha} + \frac{r\sqrt{m}}{c\alpha}\psi\left(\frac{1}{\alpha\log\left(\frac{t}{t-1}\right)}\right), \tag{13}$$

where $\psi(s)$ is increasing over the range $s \in \mathbb{R}^+$ and is given by

$$\psi(t) = \begin{cases} \sqrt{s/2} & \text{if } s \leq 0.2, \\ 0.49 - 0.11\sqrt{s} & \text{if } 0.2 < s \leq 1, \\ 0.63 - 0.26\sqrt{s} & \text{if } 1 < s \leq 5, \\ 0.77 - 0.58\sqrt{s} & \text{if } 5 < s \leq 15, \\ (2\log s - \log\log s - \log 16\pi)^{1/2} & \text{if } s > 15. \end{cases}$$

Please note that it is straightforward to show that the index G_t^{BL} shares the properties of the true index described in Lemmas 3 and 4 above and reduces to the greedy index (and hence also the true index) when $t = 1$. A further approximating index derived from the work of Caro and Gallien (2007) is given by

$$G_t^{CG}(m,\alpha) = \frac{rm}{c\alpha} + \frac{z_t r\sqrt{m}}{c\sqrt{\alpha^2+\alpha^3}}, \tag{14}$$

where z_t is the unique solution in z to the equation $(t-1)\Psi(z) = z$, with Ψ the normal error function. The sequence $\{z_t, t \geq 1\}$ is increasing and concave, with $z_1 = 0$ from which it follows that the index G_t^{CG} again shares the properties of the true index described in Lemmas 3 and 4 above and again reduces to the greedy index (and hence also the true index) when $t = 1$. In the next subsection we shall use G^{app} to denote a generic approximating index.

4.3 The construction of index heuristics for shelf-space use

A natural approach to the construction of an admissible assortment in any system state is first within each product category to allow products into the assortment in descending order of the appropriate indices (highest first) until the category constraints are met. Second, produce a common list of remaining products and allow these into the assortment in decreasing index order as long as shelf capacity allows. If necessary, in this phase we skip items and continue with lower index products which can utilize unused capacity. We call this approach *top down.* More formally, in each system state $(\mathbf{m}, \boldsymbol{\alpha})$ and time to go t, an assortment is constructed as follows:

Top-down policy

I. Meeting the product category constraints.

For $i = 1, \ldots, k$ do the following:

1. Let $\nu \in \{1, \ldots, S\}^{|K_i|}$ reference the items in category i. Let $\xi_i \in \{1, \ldots, |K_i|\}^{|K_i|}$ be a permutation such that $G^{app}_{\nu(\xi_i)}$ yields a list of indices in descending order, namely $G^{app}_{\nu(\xi_i(1))} \geq \cdots \geq G^{app}_{\nu(\xi_i(|K_i|))}$.
2. The number of category i items needed to meet the category constraint is $s_i := \min\left\{s; 1 \leq s \leq |K_i| \text{ and } \sum_{j=1}^{s} c_{\nu(\xi_i(j))} \geq L_i\right\}$ and the category i items included in step **I** is $I_i := \{\nu(\xi_i(1)), \ldots, \nu(\xi_i(s_i))\}$.

II. Utilizing the remaining common area.

1. Following step **I**, the remaining products are $R := \{1, \ldots, S\} \setminus \cup_{i=1}^{k} I_i$ and the remaining capacity is $c_R := C_{\max} - \sum_{i=1}^{k} \sum_{j=1}^{s_i} c_{\nu(\xi_i(j))}$, assumed non–negative.
2. Let $\pi \in \{1, \ldots, S\}^{|R|}$ reference the items in R and let $\xi_0 \in \{1, \ldots, |R|\}^{|R|}$ be a permutation such that $G^{app}_{\pi(\xi_0)}$ yields a list of indices in descending order, namely $G^{app}_{\pi(\xi_0(1))} \geq \cdots \geq G^{app}_{\pi(\xi_0(|R|))}$.
3. Introduce an auxiliary variable $\rho \in \{0, 1\}^{|R|}$ which is developed inductively to indicate inclusion as follows:

$$\rho(j) = 1 \Leftrightarrow \sum_{l=1}^{j-1} c_{\pi(\xi_0(l))}\rho(l) + c_{\pi(\xi_0(j))} \leq c_R.$$

   The set $I_0 := \{\pi(\xi_0(j)) : 1 \leq j \leq |R| \text{ and } \rho(j) = 1\}$ is the corresponding collection of items to be included to utilize the common area.

The set of items to be included in the arrangement by this heuristic is $I := \cup_{i=0}^{k} I_i$.

Note that it is possible, though unlikely, that the values of $\sum_{i=1}^{k} L_i$ may be very close to $C_{\max}$. If this is the case then the problem is close to having an additive decomposition by category. In this event, overshooting of the category thresholds L_i during step **I** of the top-down approach may mean that $c_R < 0$ and that the total capacity $C_{\max}$ is exceeded in this phase of the algorithm. In such a case the heuristic would fail to deliver a feasible solution on occasion even when the problem is well-posed. Note that this cannot happen under the condition $\sum_{i=1}^{k}(L_i - 1 + \max_{s\in K_i} c_s) \leq C_{\max}$ which we would expect to be satisfied in practice.

A slightly more sophisticated version again uses step **I** of the top-down approach for all individual categories. The residual capacity is then filled optimally with respect to index values by solving a 0–1 knapsack problem for the remaining products. We shall call this the *mixed method*. It constructs assortments as follows:

Mixed Method

I. Meeting the product category constraints. As for the top-down policy.
II. Utilizing the remaining common area.

1. Following step **I**, the remaining products are $R := \{1,\ldots,S\}\setminus\cup_{i=1}^{k} I_i$ and the remaining capacity is $c_R := C_{\max} - \sum\limits_{i=1}^{k}\sum\limits_{j=1}^{s_i} c_{\nu(\xi_i(j))}$, assumed non–negative.

2. Let $\pi \in \{1,\ldots,S\}^{|R|}$ reference the items in R. Solve the optimization problem:

$$\max \sum_{j=1}^{|R|} G^{app}_{\pi(j)} c_{\pi(j)} x_j,$$

subject to

$$\sum_{j=1}^{|R|} c_{\pi(j)} x_j \leq c_R, x_j \in \{0,1\}$$

and write $x^* \in \{0,1\}^{|R|}$ for an optimal solution. The set $I_0 := \left\{\pi(j) : 1 \leq j \leq |R| \text{ and } x_j^* = 1\right\}$ is the corresponding set of items to be included to utilize the common area.

The set of items to be included in the arrangement by this heuristic is $I := \cup_{i=0}^{k} I_i$.

The above approaches have the advantages of simplicity and will produce strongly performing feasible solutions when the stated sufficient condition is satisfied. However we conclude with a so-called *knapsack method* which is guaranteed to produce a feasible solution for any well posed problem. Before outlining the algorithm we give a brief rationale.

The challenge we have concerns the fact that, in the absence of the individual product category constraints, a knapsack solution may wish to leave capacity idle. For each product category i, we therefore meet these constraints by considering a range of knapsack-like problems aimed at using capacity $L_i + h$ where h is no greater than some overlap o_i which is determined by the capacity used in the first stage of the top-down approach. This phase leaves us with $o_i + 1$ candidates for each product category i. Should any candidate fail to utilize its available capacity $L_i + h$ fully it is discarded and we seek a full utilization candidate instead. We do this by replacing the G_s^{app} as weights in the corresponding knapsack problem by V_s, which are designed to ensure that a full utilization set of even the least attractive items performs better than any other set of items which leave capacity unused. This approach ensures that, for each $L_i + h$, we are getting the best of all feasible full utilization solutions. Which of these candidates proves the most successful will depend upon the attractiveness of the product categories relative to each other and on the remaining capacity. We terminate the approach by taking each of the $P = \prod_{i=1}^{k} (o_i + 1)$ combinations of candidate sets and filling up any remaining capacity with items which are yet to be included to optimize an index-based objective.

Knapsack Method

I. Meeting the product category constraints.

For $i = 1, \ldots, k$ do the following:

1. Solve the optimization problem:

$$\min \ |A_i|,$$

subject to

$$A_i \subseteq K_i, \ \sum_{s \in A_i} c_s \geq L_i \text{ and } G_s^{app} \geq \max \left\{ G_j^{app}, j \in A_i \diagdown K_i \right\} \forall s \in A_i.$$

Define $o_i := \left(\sum_{s \in A_i} c_s \right) - L_i$ and observe that $0 \leq o_i \leq (\max_{s \in S} c_s) - 1$.

2. For each integer h, $0 \leq h \leq o_i$, solve the optimization problem:

$$\max \sum_{s \in K_i} G_s^{app} c_s x_s,$$

subject to

$$\sum_{s \in K_i} c_s x_s \leq L_i + h, x_s \in \{0, 1\}$$

and write $x\,(i,h) \in \{0,1\}^{|K_i|}$ for an optimal solution. If

$$\sum_{s \in K_i} c_s x_s\,(i,h) < L_i + h$$

then discard the solution $x\,(i,h)$ and proceed to Step 3.

3. For each h for which the solution $x\,(i,h)$ was discarded at Step 2, proceed as follows:

   (a) Let $\nu \in \{1,\ldots,S\}^{|K_i|}$ reference the items in category i. Let $\zeta_i \in \{1,\ldots,|K_i|\}^{|K_i|}$ be a permutation such that $G^{app}_{\nu(\zeta_i)}$ yields a list of indices in ascending order.

   (b) Choose δ less than $\big((L_i+h)|K_i|\big)^{-1}$ but greater than machine precision.

   (c) Define $V_s := 1 + \delta\ \zeta_i^{-1}(\nu^{-1}(s))$, where $\zeta_i^{-1}(\nu^{-1}(s))$ simply is the rank of item s within category i according to ascending G^{app}_s. Solve the optimization problem:

   $$\max \sum_{s \in K_i} V_s c_s x_s,$$

   subject to

   $$\sum_{s \in K_i} c_s x_s \le L_i + h, x_s \in \{0,1\}$$

   and write $x\,(i,h) \in \{0,1\}^{|K_i|}$ for an optimal solution.

4. Using the solutions $x(i,h)$ obtained from Steps 2 and 3, define $I_{i,h} := \big\{s; s \in K_i \text{ and } x_s(i,h) = 1\big\}$, $0 \le h \le o_i$. Set I_{i,L_i+o_i} is guaranteed non-empty.

II. Utilizing the remaining common area.

1. Write $P = \prod_{i=1}^{k}(o_i+1)$ and let $\chi \in \{1,\ldots,\max_i o_i + 1\}^{k\times P}$ reference the list of $h-$values of all (i,h) pairs considered in Steps 2 and 3 above when compiled into a single list. In particular χ_{ij} is the $h-$value of the (i,h) pair within the j^{th} listed combination.

2. Following **I** we have a range of candidate sets of residual products given by $R_j := \{1,\ldots,S\} \setminus \cup_{i=1}^{k} I_{i,\chi_{ij}}$ with associated remaining capacity $c_{R_j} = C_{\max} - \sum_{i=1}^{k} \sum_{s \in I_{i,\chi_{ij}}} c_s, 1 \le j \le P$.

3. For each $j, 1 \leq j \leq P$, solve the optimization problem:

$$\max \sum_{s \in R_j} G_s^{app} c_s x_s$$

subject to

$$\sum_{s \in R_j} c_s x_s \leq c_{R_j}, x_s \in \{0,1\}$$

and write $x(j) \in \{0,1\}^{|R_j|}$ for an optimal solution. The set $I_{0,j} := \{s : s \in R_j \text{ and } x_s(j) = 1\}$ is the corresponding set of items to be included to utilize the common area.

4. Write

$$\sigma_j := \sum_{s \in I_{0,j}} G_s^{app} c_s + \sum_{i=1}^{k} \sum_{s \in I_{i,\chi_{ij}}} G_s^{app} c_s$$

with $j^* \in \arg\max_{1 \leq j \leq P} \sigma_j$.

The set of items to be included by the heuristic is $I := I_{0,j^*} \cup \left(\cup_{i=1}^{k} I_{i,\chi_{i,j^*}} \right)$.

5 Numerical Results

Our numerical study first focusses on an evaluation of the performance of the heuristic index policies for dynamic assortment planning introduced in the last section. To this end we use generated data sets similar to those in the peer literature. This is the focus of subsection 5.1. We expand this study in subsection 5.2 by an exploration of the impact of demand uncertainty and length of selling season on heuristic policy performance. In the following subsection 5.3 we develop the marginal value of shelf-space, explore how it changes with store size and proceed in 5.4 to understand the impact of demand uncertainty on it. We conclude in subsection 5.5 with a study of how optimal store size is impacted by the length of the selling season and hence by the rapidity of change in product offerings.

Methodology

Given the Bayesian nature of the learning model, we can generate the sales data for our simulation directly using the predictive negative binomial distribution. Besides the randomization of sales data we average over the reward per product $\boldsymbol{r}$. These values are drawn from a single uniform distribution in the case of one category and from one or multiple different uniform distributions in the general case. The simulations of the policy performance are run as often as it takes to get the relative estimation error as low as 0.02% or less. If, for the sake of computational brevity, less runs were carried out, the precision is stated alongside the results.

A crucial part of the numerical study is the computation of H^*, which from (4) is an upper bound on the optimal return. We shall see that this bound is sufficiently tight that in many of our studies we are able to use it as a surrogate for the optimal return itself. Note from (4) that it is computed as a minimum of the (non-differentiable) Lagrangian function with respect to the non-negative Lagrange multipliers. We use the Nelder-Mead simplex algorithm for this task. See Nelder and Mead (1965). This computational effort involved in this task dominates that for all other aspects of the study and grows exponentially with T, the length of the sales season. Please note that, throughout the numerical study reported numerical values of the upper bound (H^*) are always factored by the length of the selling season (T). Hence reported values are in fact H^*T^{-1}.

Experiment Description

In subsections 5.1 and 5.2 below we explore the quality of performance of the index heuristics proposed in section 4. There will be three distinct testing setups. The first will be a single category case ($k = 1$) where the lower shelf-space constraint is met trivially. Rewards r_s are drawn from the interval $[2, 8]$. The second setup uses a moderate number of categories ($k \in \{3, 4, 5\}$), with rewards drawn as for the first setup. In the third setup we develop the configuration for the second by generating rewards from distinct categories from distinct subintervals of $[2, 8]$. Hence when $k = 3$ we use the subintervals $([2, 4], [4, 6], [6, 8])$, for $k = 4$ we use $([2, 4], [3.5, 5.5], [4.5, 6.5], [6, 8])$ and for $k = 5$ we use $([2, 4], [3, 5], [4, 6], [5, 7], [6, 8])$. These examples create a tension between differentially attractive categories and the need to meet the lower shelf-space constraints for strategic reasons.

Throughout the experiments, the number of products is set to a moderate number of $S = 720$. In order to eliminate effects due to a change in the average number of products on shelf with respect to S, the total shelf space $C_{\max}$ is not set to be a fixed number, but to be proportional to the average shelf space need. Therefore $C_{\max} := 30 \frac{1}{S} \sum_{s=1}^{S} c_s$. The lower thresholds $L_i, i = 1, \ldots, k$ are throughout set such that the common shared area covers about 30% of the maximum capacity. Therefore $L_i := \lfloor 0.7\ C_{\max}/k \rfloor$.

The heuristic policies studied in the following two subsections all combine a choice of approximate index (greedy (GDY), Brezzi-Lai (BL) or Caro-Gallien (CG)) with an index-based heuristic approach to shelf-space allocation (Top-down, Mixed or Knapsack).

		GDY	CG	BL	Upper bound
{1, 2, 3}	Top-down	0.38	0.13	0.12	5154.7 ± 0.02%
	Knapsack	0.37	0.13	0.12	
{2, 4, 6}	Top-down	0.38	0.13	0.12	10309.5 ± 0.02%
	Knapsack	0.37	0.13	0.13	
{1, 2, 3, 4, 5}	Top-down	0.43	0.18	0.17	7734.8 ± 0.05%
	Knapsack	0.42	0.18	0.17	
{3, 4, 5}	Top-down	1.06	0.81	0.81	10315.8 ± 0.06%
	Knapsack	0.42	0.18	0.18	
{2, 4, 7}	Top-down	0.67	0.42	0.40	11163.9 ± 0.05%
	Knapsack	0.34	0.09	0.09	

Table 1: Suboptimalities (in % below the dual DP upper bound) of three indices combined with two heuristic methods for shelf-space allocation for various $\boldsymbol{c}$.

5.1 Impact of Product Categories (k) and Space Requirement ($\mathbf{c}$) on Policy Performance

This set of experiments examines the impact of different ranges of space demands, that is different vectors $\boldsymbol{c}$. Each of the sets $\{1,2,3\}$, $\{2,4,6\}$, $\{1,2,3,4,5\}$, $\{3,4,5\}$, $\{2,4,7\}$ serves as a $\boldsymbol{c}$-generating set, which means that the values of the set's elements are allocated equally often to the S items as their capacity need. The time horizon is $T = 10$ and the prior demand for all products is given by $\mathbb{E}_\gamma = 10$ and $\mathbb{V}_\gamma = 5$.

The results for the first setup (only one category) are shown in Table 1. First of all, the suboptimality levels for each of the heuristics are very low. That is, the adapted upper bound is exceptionally tight. For the given parameters the greedy index performs worst and the Brezzi-Lai index best.

In all three cases, where the greatest common divisor of the $\boldsymbol{c}$-generating sets lies within the set, i. e. $\{1,2,3\}$, $\{2,4,6\}$ and $\{1,2,3,4,5\}$, the top-down and the knapsack approach perform identically up to small error terms. In the other two cases, i. e. $\{3,4,5\}$ and $\{2,4,7\}$, the knapsack approach is clearly superior. Moreover the difference in suboptimality is approximately identical to the percentage reduction in shelf space usage in the top-down approach. Table 2 shows that the knapsack method makes use of all shelf space at all times, whereas the top-down approach does not. The difference in this average resource exploitation divided by the available amount gives the usage difference in the next to bottom line. The bottom line then shows that this lack of exploitation is responsible for the top-down approach's inferiority up to an amount of the order of the error terms. This result illuminates the advantage of the knapsack approach: It makes better use of the available shelf space.

Policy: CG		{2, 3, 4}	{2, 3, 5}	{2, 4, 5}	{2, 4, 7}	{3, 4, 5}	{3, 5, 7}
Opt. gap (%)	Top-down	0.50	0.49	0.52	0.42	0.81	0.64
	Knapsack	0.11	0.11	0.13	0.09	0.18	0.12
	Difference	0.39	0.38	0.40	0.34	0.63	0.52
Shelf usage (%)	Top-down	99.57	99.60	99.59	99.65	99.35	99.45
	Knapsack	100.00	100.00	100.00	100.00	100.00	100.00
	Difference	0.43	0.40	0.41	0.35	0.66	0.55
Usage diff. - opt. gap diff. (%)		0.04	0.02	0.01	0.02	0.02	0.03

Table 2: Knapsack approach outperforms top-down alternative due to shelf space usage. Tabulated values are rounded.

In the second setup (several categories, uniform reward draws) most of the findings from the first setup are verified. This includes the relation between suboptimality difference from top-down to knapsack and the shelf space usage. The same is true for the general relative performance of the CG, the BL and the greedy indices.

Although we consider now more than one category, the increasing number of categories k has little influence on the suboptimality gaps. Surprisingly, even the mixed and knapsack inclusion modes show little difference in performance (Table 3). Indeed, a large amount of additional sampling is needed until we find a statistically significant difference, which for the case of $\{2,4,7\}$, $k=3$ and the CG-policy is $0.007\% \pm 0.002\%$ in favor of the knapsack policy for a sample size of about 4 million.

Things change somewhat for the third setup (several categories, separated ranges for reward draws). Table 4 shows the numerical results for the same setting as before, but for unequally spread rewards, that is some categories are generally more profitable than others. The most significant observation is with respect to the relative performances of the different heuristic approaches to shelf-space allocation. The knapsack based combinatorial heuristic now clearly outperforms the mixed approach. There is roughly the same gap between the mixed and the top-down approach, but this now cannot be explained by capacity usage gaps alone. It is rather that more shelf space is allocated to items from clearly inferior categories than is demanded by the lower thresholds L_i. There is now a significant, albeit moderate difference between suboptimalities for different k. The main driver for that is very likely the particular choice of the intervals, as for an increasing number of intervals, the best items of the mid-level attractive categories get successively worse. The differences between policy performance due to the choice of index is consistent with earlier results. Both CG and BL continue to perform outstandingly well and modestly better than GDY.

		GDY	CG	BL	Upper bound
	Top-down	0.74	0.50	0.47	
$\{2, 4, 7\}$; k=3	Mixed	0.41	0.14	0.12	11169.4 ± 0.03%
	Knapsack	0.41	0.16	0.12	
	Top-down	0.78	0.53	0.50	
$\{2, 4, 7\}$; k=4	Mixed	0.44	0.16	0.16	11169.4 ± 0.03%
	Knapsack	0.42	0.18	0.14	
	Top-down	0.82	0.58	0.52	
$\{2, 4, 7\}$; k=5	Mixed	0.48	0.21	0.18	11169.4 ± 0.03%
	Knapsack	0.46	0.21	0.16	
	Top-down	1.01	0.77	0.72	
$\{3, 4, 5\}$; k=3	Mixed	0.37	0.13	0.09	10310.1 ± 0.03%
	Knapsack	0.37	0.14	0.09	
	Top-down	1.03	0.79	0.75	
$\{3, 4, 5\}$; k=4	Mixed	0.41	0.15	0.12	10310.1 ± 0.03%
	Knapsack	0.41	0.15	0.12	
	Top-down	1.05	0.81	0.76	
$\{3, 4, 5\}$; k=5	Mixed	0.42	0.17	0.14	10310.1 ± 0.03%
	Knapsack	0.42	0.18	0.11	

Table 3: Suboptimalities (in % below the dual DP upper bound) of three indices combined with three heuristic methods for shelf-space allocation for various $\boldsymbol{c}, k$ and equally spread rewards.

		GDY	CG	BL	Upper bound
	Top-down	2.30	2.02	1.98	
$\{2, 4, 7\}$; k=3	Mixed	1.90	1.61	1.56	9439.8 ± 0.02%
	Knapsack	0.56	0.22	0.19	
	Top-down	2.59	2.35	2.36	
$\{2, 4, 7\}$; k=4	Mixed	2.17	1.94	1.96	9464.5 ± 0.02%
	Knapsack	0.58	0.26	0.27	
	Top-down	3.08	2.86	2.80	
$\{2, 4, 7\}$; k=5	Mixed	2.67	2.48	2.42	9410.1 ± 0.02%
	Knapsack	0.71	0.41	0.37	
	Top-down	2.38	2.09	2.05	
$\{3, 4, 5\}$; k=3	Mixed	1.61	1.31	1.29	8695.6 ± 0.02%
	Knapsack	0.53	0.21	0.16	
	Top-down	2.65	2.38	2.37	
$\{3, 4, 5\}$; k=4	Mixed	1.88	1.61	1.61	8681.0 ± 0.02%
	Knapsack	0.54	0.22	0.21	
	Top-down	2.96	2.69	2.64	
$\{3, 4, 5\}$; k=5	Mixed	2.21	1.95	1.90	8747.0 ± 0.02%
	Knapsack	0.68	0.35	0.31	

Table 4: Suboptimalities (in % below the dual DP upper bound) of three indices combined with three heuristic methods for shelf-space allocation for various $\boldsymbol{c}, k$ and unequally spread rewards.

5.2 Impact of Length of Selling Season (T) and Prior Demand Uncertainty ($\mathbb{V}_\gamma$) on Policy Performance

This series of experiments studies the effect of longer time horizons ($T = 20$) and the impact of higher uncertainty of the a priori demand estimates ($\mathbb{V}_\gamma = 50$). The $\boldsymbol{c}$-generating set is fixed to be $\{2, 4, 7\}$. In the first setup (one category) we take the following approach: From initial time and uncertainty values ($T = 10$, $\mathbb{V}_\gamma = 5$), we explore first the time sensitivity of policy performance (increase to $T = 20$) and then – starting again from the basic values – sensitivity to prior demand uncertainty (increase to $\mathbb{V}_\gamma = 50$). Table 5 shows the numerical results for these three settings. The observed relative performance continues to be related to the relative use made of the key shelf-space resource. Alterations in T or $\mathbb{V}_\gamma$ change the upper bound and the policy performance, but not the difference in suboptimality levels between the knapsack and the top-down approach. The upper bound H^* continues to be very tight. Enhanced prior uncertainty seriously undermines the performance of the greedy index.

		GDY	CG	BL	Upper bound
	Top-down	0.67	0.43	0.42	
T = 10; $\mathbb{V}_\gamma = 5$	Knapsack	0.34	0.09	0.09	11163.9 ± 0.05%
	Difference	0.33	0.34	0.33	
	Top-down	1.31	0.62	0.69	
T = 20; $\mathbb{V}_\gamma = 5$	Knapsack	0.99	0.28	0.35	11545.8 ± 0.06%
	Difference	0.32	0.34	0.34	
	Top-down	4.81	1.06	0.82	
T = 10; $\mathbb{V}_\gamma = 50$	Knapsack	4.47	0.74	0.50	15663.9 ± 0.06%
	Difference	0.34	0.32	0.32	

Table 5: Suboptimalities (in % below the dual DP upper bound) of three indices combined with two heuristic methods for shelf-space allocation for various T and $\mathbb{V}_\gamma$.

Tables 6 and 7 give results for respectively the second and third experimental setups. The principal features are broadly consistent with those of Table 5. Figure 1 shows two outtakes of Table 7. The left hand plot compares the heuristic fill-up methods *top-down*, *mixed* and *knapsack* for the configuration $k = 3$, $T = 10$, $\mathbb{V}_\gamma = 50$, $\boldsymbol{c}$-generating set $\{2, 4, 7\}$ and unequally spread rewards using the Caro-Gallien (CG) index. The plot shows that the knapsack method not only outperforms the other two, but gets very close to the upper bound. The right hand plot compares three index policies for the same configuration as before using the knapsack fill-up method along with each of the policies. The plot shows that the greedy policy does have difficulties with the high uncertainty, whereas the other two come very close to the upper bound. This is particularly noteworthy, as managers often

		GDY	CG	BL	Upper bound
	Top-down	0.74	0.47	0.50	
T = 10; $\mathbb{V}_\gamma = 5$	Mixed	0.41	0.12	0.14	11169.4 ± 0.03%
	Knapsack	0.41	0.12	0.16	
	Top-down	1.32	0.70	0.65	
T = 20; $\mathbb{V}_\gamma = 5$	Mixed	0.96	0.36	0.28	11543.5 ± 0.05%
	Knapsack	0.98	0.38	0.25	
	Top-down	4.79	0.93	0.98	
T = 10; $\mathbb{V}_\gamma = 50$	Mixed	4.51	0.55	0.76	15664.8 ± 0.05%
	Knapsack	4.49	0.43	0.80	

Table 6: Suboptimalities (in % below the dual DP upper bound) of three indices combined with three heuristic methods for shelf-space allocation for various T and $\mathbb{V}_\gamma$ and rewards equally spread.

		GDY	CG	BL	Upper bound
	Top-down	2.30	1.98	2.02	
T = 10; $\mathbb{V}_\gamma = 5$	Mixed	1.90	1.56	1.61	9439.8 ± 0.02%
	Knapsack	0.56	0.19	0.22	
	Top-down	3.44	2.68	2.61	
T = 20; $\mathbb{V}_\gamma = 5$	Mixed	3.02	2.29	2.21	9877.2 ± 0.02%
	Knapsack	1.88	1.05	0.93	
	Top-down	7.09	2.90	3.04	
T = 10; $\mathbb{V}_\gamma = 50$	Mixed	6.76	2.46	2.63	13404.1 ± 0.03%
	Knapsack	5.49	1.08	1.34	

Table 7: Suboptimalities (in % below the dual DP upper bound) of three indices combined with three heuristic methods for shelf-space allocation for various T and $\mathbb{V}_\gamma$ and rewards unequally spread.

tend to opt for secure options in the assortment rather than to take calculated risks at the beginning of the selling season. Overall, the new direction in fast procurement must go together with a new attitude towards risk to realize its full potential.

5.3 The Marginal Value of Shelf Space

Shelf space is a key resource for the retailer. For a whole range of purposes access to the marginal value of shelf space, namely the increment in return achievable from an increment in the available shelf-space $C_{\max}$, is an important guide in decision-making. Such purposes could include the evaluation of plans to remodel an existing store or to move to another site of different capacity. Our numerical results to date suggest strongly that in attempting to study this marginal value we can for all practical purposes replace the optimal return

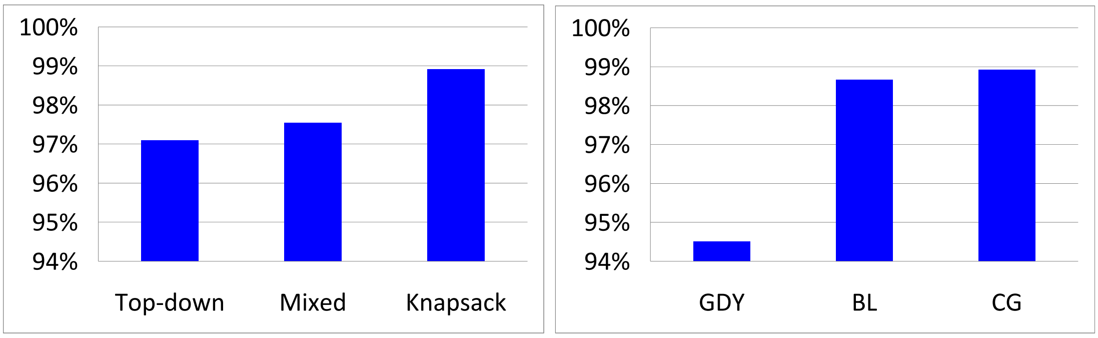

Figure 1: Graphical representation of parts of Table 7. Performance is expressed as a percentage of upper bound H^*.

J^* which we cannot compute by its tight upper bound H^* which we can. Henceforth we shall do so. Further we know from Proposition 1 that H^* is non-decreasing and concave in $C_{\max}$ and hence that the marginal value of shelf space as determined by H^* is positive and decreasing in $C_{\max}$.

We can shed considerably further light on the value of shelf-space in the single product category $(k = 1)$ case by developing the analysis of sections 3 and 4 and noting from $(3), (4)$ together with Proposition 2 and the comments following that

$$H_T^*(\mathbf{m}, \boldsymbol{\alpha}) = \min_{\lambda \geq 0} H_T^{\lambda}(\mathbf{m}, \boldsymbol{\alpha})$$

$$= \min_{\lambda \geq 0} \left\{ T\lambda C_{\max} + \mathbf{E}_{\pi_\lambda} \left[\sum_{\tau=1}^{T} \sum_{s=1}^{S} \left\{ \frac{r_s m_s(\tau)}{\alpha_s(\tau)} - \lambda c_s \right\} I\left[G_{s,T-\tau}\left\{ m_s(\tau), \alpha_s(\tau) \right\} > \lambda \right] \right] \right\}. \tag{15}$$

In (15), $\mathbf{E}_{\pi_\lambda}$ denotes an expectation taken over realisations of the system under policy π_λ for the Lagrangian relaxation which includes every product whose index exceeds λ in all assortments. This is the policy which achieves $H_T^{\lambda}(\mathbf{m}, \boldsymbol{\alpha})$. Note that we use $\{m_s(\tau), \alpha_s(\tau)\}$ for the (random) belief state (parameters of the posterior gamma distribution) for product s at time τ from the start of the selling season. It is trivial that $H_T^{\lambda}(\mathbf{m}, \boldsymbol{\alpha})$, as a function of λ, is convex and piecewise linear with hinge points at product index values. We write

$$C(\lambda) = \mathbf{E}_{\pi_\lambda} \left[\sum_{\tau=1}^{T} \sum_{s=1}^{S} c_s I\left[G_{s,T-\tau}\left\{ m_s(\tau), \alpha_s(\tau) \right\} > \lambda \right] \right] \tag{16}$$

for the total expected shelf-space used over the selling season by policy π_λ. $C(\lambda)$ is piecewise constant and decreasing in λ with points of discontinuity at product index values. We define

$$\lambda(C_{\max}) = \sup[\lambda; C(\lambda) > TC_{\max}]$$

for the value of the shelf-space charge λ which comes closest (from below) to achieving total expected shelf-space use over the season of $TC_{\max}$. The charge $\lambda(C_{\max})$ must be an index value, $G_{s,t}(m_s, \alpha_s)$, say. Use the notational shorthand $P(C_{\max})$ for the corresponding product / time / state combination (s, t, m_s, α_s) and $S(C_{\max})$ for the collection of product / time / state combinations whose associated indices exceed $\lambda(C_{\max})$. It is straightforward to show that

$$\lambda(C_{\max}) = \frac{R^*(S(C_{\max}) \cup P(C_{\max})) - R^*(S(C_{\max}))}{C^*(S(C_{\max}) \cup P(C_{\max})) - C^*(S(C_{\max}))}. \tag{17}$$

In (17), $R^*(Q)$ and $C^*(Q)$ denote respectively the expected return and the expected shelf-space use when only product / time / state combinations in Q are included in assortments. We can then infer that the upper bound on expected revenues $H_T^*(\mathbf{m}, \boldsymbol{\alpha})$, regarded as a function of $C_{\max}$ can be expressed as

$$H_T^*(\mathbf{m}, \alpha, C_{\max}) = \lambda(C_{\max})\{TC_{\max} - C^*(S(C_{\max}))\} + R^*(S(C_{\max})),$$

with $T\lambda(C_{\max})$ then recovered as the marginal value of shelf-space. Hence from (17) the marginal value of shelf-space is recovered as the product of T and the ratio of the increment in total expected return when adding the marginal product / time/ state combination $P(C_{\max})$ to the candidate such combinations for inclusion in assortments to the corresponding increment in total expected shelf-space usage. As $C_{\max}$ increases so the index associated with $P(C_{\max})$ decreases and, since this is the ratio $\lambda(C_{\max})$, also the marginal value of shelf-space.

Please find in Figure 2 a plot of the marginal value of shelf space as given by the discrete derivative of H^* with respect to $C_{\max}$. The model used for the computation has $k = 1, T = 10, \mathbf{E}[\gamma] = 10, \mathbb{V}_\gamma = 5, S = 720$ and $\mathbf{c} = \{1, 2, 3\}$. The plot is over the full extent of the range of $C_{\max}$, namely $0 \le C_{\max} \le \sum_{s=1}^{S} c_s = 1440$. To avoid the need for averaging over many samples, product rewards were not chosen randomly, but rather distributed linearly over the interval $[2, 8]$. Figure 2 confirms numerically the theoretical result in Proposition 1. The accompanying Figure 3 gives a corresponding plot of the second discrete derivative of the upper bound H^*. While all these observations about the behavior of the marginal value with respect to $C_{\max}$ fully coincide with intuition, please note that it is the fact that we are able to determine its exact value numerically that allows us to draw conclusions about optimal store sizes in the following sections.

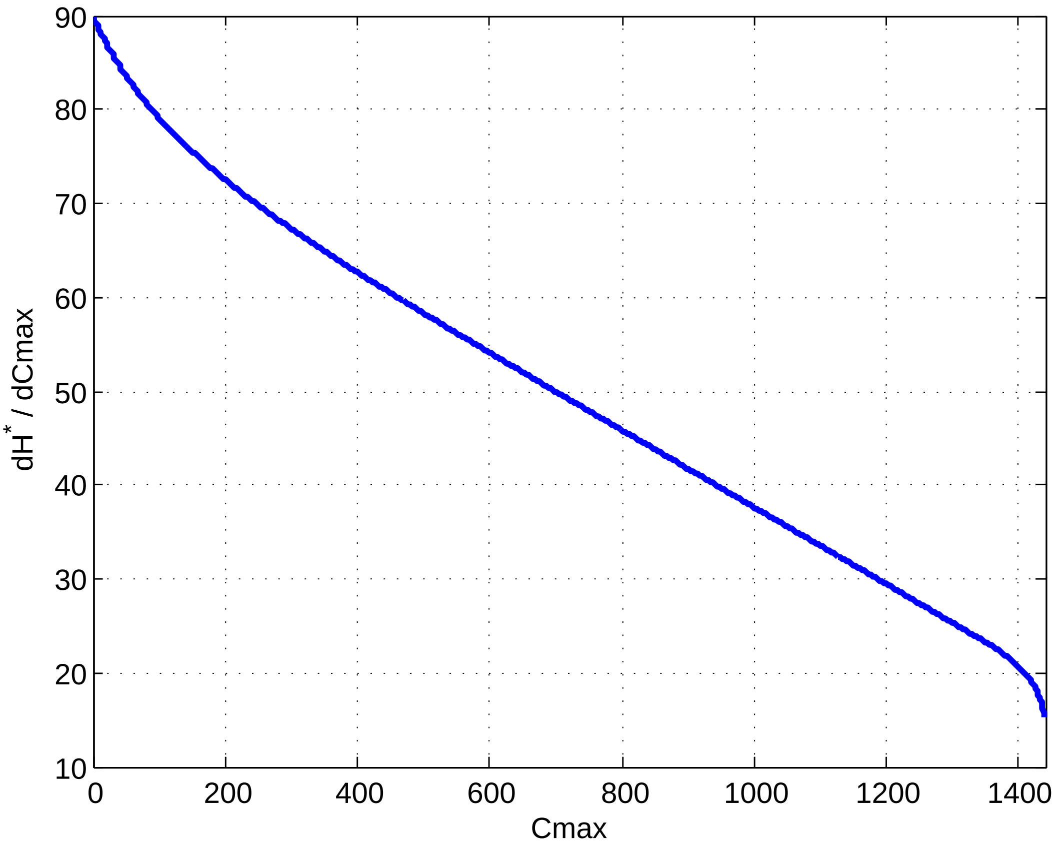

Figure 2: Marginal value of shelf space

5.4 The impact of learning on the marginal value of shelf space

We now explore how enhanced prior uncertainty about the demand for products impacts the marginal value of shelf-space and hence decisions to which that measure is relevant. We shall gain insights on questions such as whether store managers who take account of demand learning in their assortment planning are likely to need larger or smaller shops. Also on whether and how this depends on the degree of uncertainty about the demand concerned.

As a benchmark we consider natural analogues of the optimal return J^* and the Lagrangian upper bound H^* under no learning. We define

$$J^*_{T,\text{no-learn}}(\mathbf{m}, \alpha) = T \max_{\substack{\mathbf{u}\in\{0,1\}^S, \\ \sum_{s=1}^S c_s u_s \le C_{\max}}} \sum_{s=1}^{S} r_s \frac{m_s}{\alpha_s} u_s \tag{18}$$

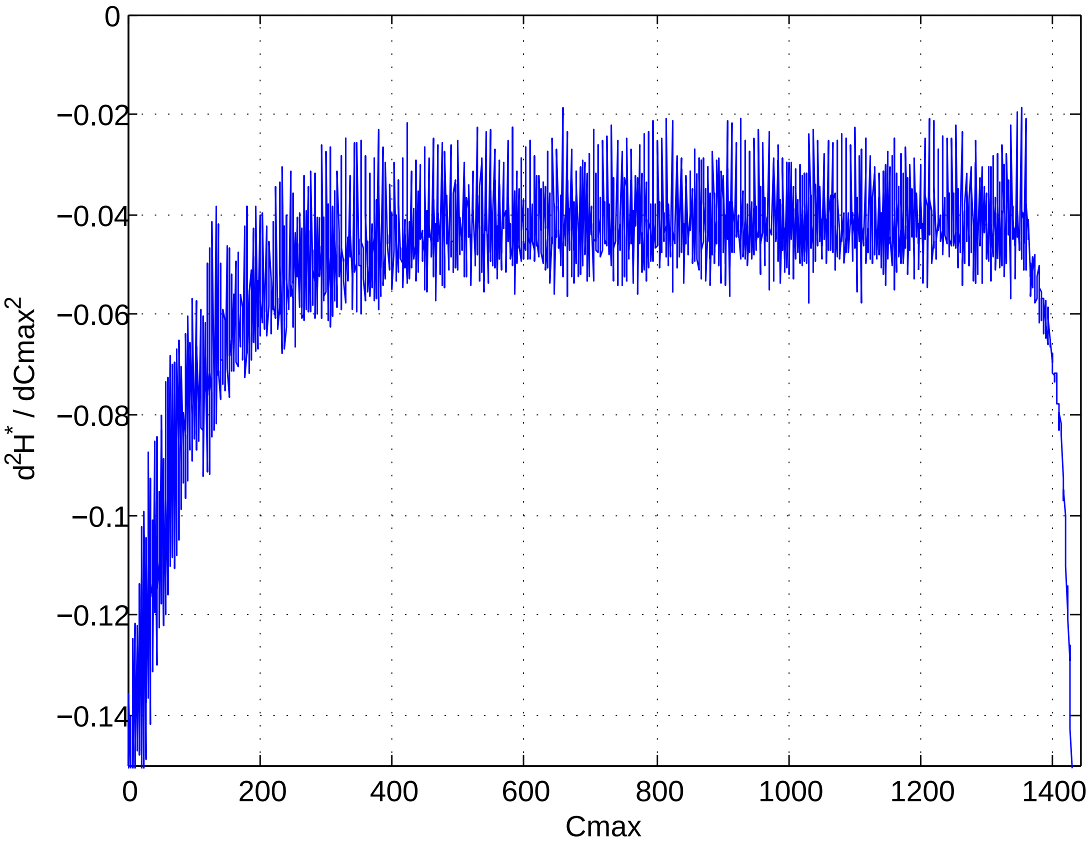

Figure 3: Second derivative of total value per maximum shelf space

and

$$H^*_{T,\text{no-learn}}(\mathbf{m}, \boldsymbol{\alpha}) = \min_{\lambda \geq 0} \left\{ T\lambda C_{\max} + T \sum_{s=1}^{S} \max \left\{ r_s \frac{m_s}{\alpha_s} - \lambda c_s, 0 \right\} \right\} \tag{19}$$

The differences $H^*_{\text{no-learn}} - J^*_{\text{no-learn}}$ are easy to compute, are small and relate to the space unit needs for the sales. Plainly, neither $J^*_{\text{no-learn}}$ nor $H^*_{\text{no-learn}}$ depend upon the prior variance $\mathbb{V}_\gamma$. They can be thought of as the values of J^* and H^* appropriate for the case $\mathbb{V}_\gamma = 0$, respectively.

In Figure 4 find a number of plots of H^* against $C_{\max}$ for the range $0 \leq C_{\max} \leq \sum_{s=1}^{S} c_s = 1440$ for a scenario in which $k = 1, T = 10, \mathbf{E}[\gamma] = 10, S = 720$, $\mathbf{c} = \{1, 2, 3\}$ and where product rewards are distributed linearly over the interval $[2, 8]$. Each continuous plot corresponds to a different value of $\mathbb{V}_\gamma$ as indicated in Table 8. The dash-dotted magenta plot is of $H^*_{\text{no-learn}}$. Figures 5, 6 and 7 contain corresponding plots of respectively $H^* - H^*_{\text{no-learn}}$, the discrete derivative of H^* and the discrete derivative of $H^* - H^*_{\text{no-learn}}$. The plain message

$\mathbb{V}_\gamma$	color	m	α
100	green	1	0.1
50	red	2	0.2
25	cyan	4	0.4
12.5	black	8	0.8
5	blue	20	2

Table 8: The range of prior demand uncertainties used in the study

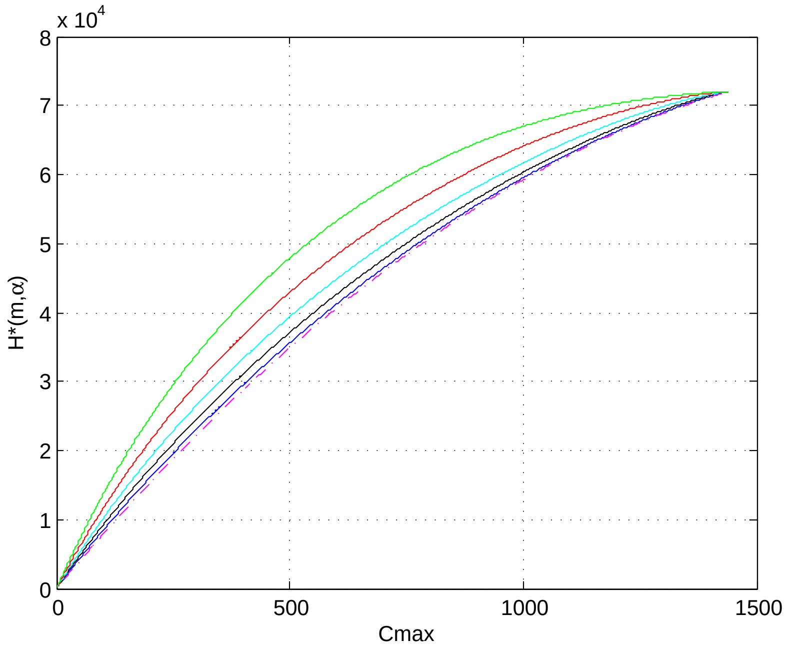

Figure 4: Revenue plotted against $C_{\max}$ for five prior demand uncertainties and a no-learning option

of Figure 4 is that expected revenues will be greatest when the demand uncertainty is largest, for given prior means. When demand uncertainty is high, our heuristic policies will enable us to learn about high return products whose true demand rates are higher than the prior mean and allow such products to feature strongly in assortments especially later in the sales season. Figure 6 makes it clear that for the instances studied the value of policies which feature active learning about demand rates is at its greatest for large prior variance and available shelf-space which can accommodate between 30% and 70% of the available products.

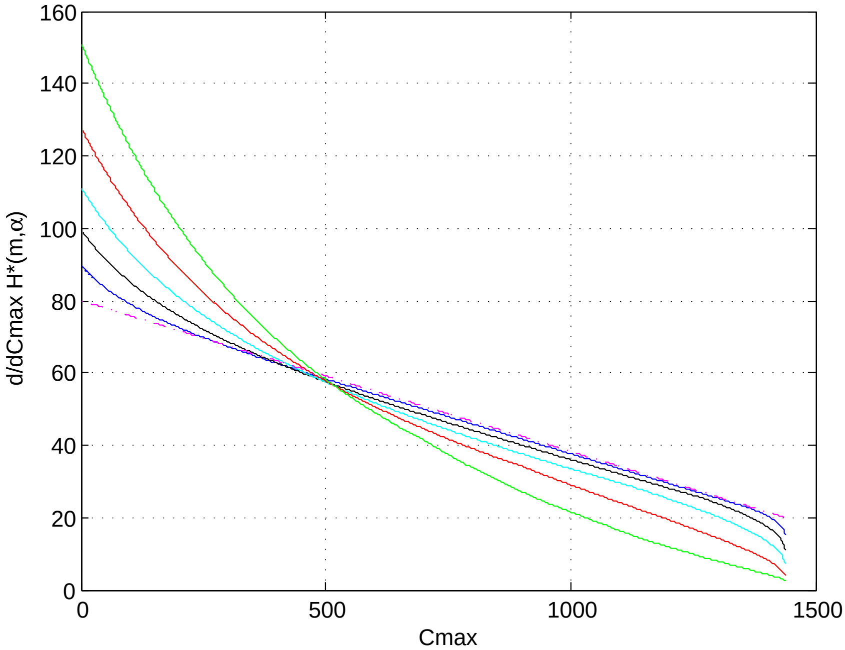

Figure 5: Marginal revenue plotted against $C_{\max}$ for five prior demand uncertainties and a no-learning option

Figure 4 which features plots of the marginal value of shelf-space would potentially yield solutions to problems of the form

$$\max_{C_{\max}} \{H^* (C_{\max}) - RC_{\max}\}$$

to determine the optimal store size ($C_{\max}$) when the rental charge for a single unit of shelf-space for one unit of time is R. Solutions to such problems may be read off from the plots in Figure 5 as an inverse mapping evaluated at R. We see that for retail markets in which there is demand uncertainty, if rents are high (and so optimal store sizes are small) then increased demand uncertainty means an increase in store size to support active learning among high earning products. If rents are low (and optimal store sizes are large) then increased demand uncertainty means a decrease in store size to mitigate the risk of poorly performing products.

Please note that the findings in the section have implications that concern the full setup of the supply chain. The choice towards implementing a system with demand learning necessarily requires investment into a logistics configuration that allows rapid replenishment. While outside the scope of this paper, out model might also be consulted in support of such decision and whether or not the move towards a more costly, but also more reactive, supply chain solution is beneficial for the firm.

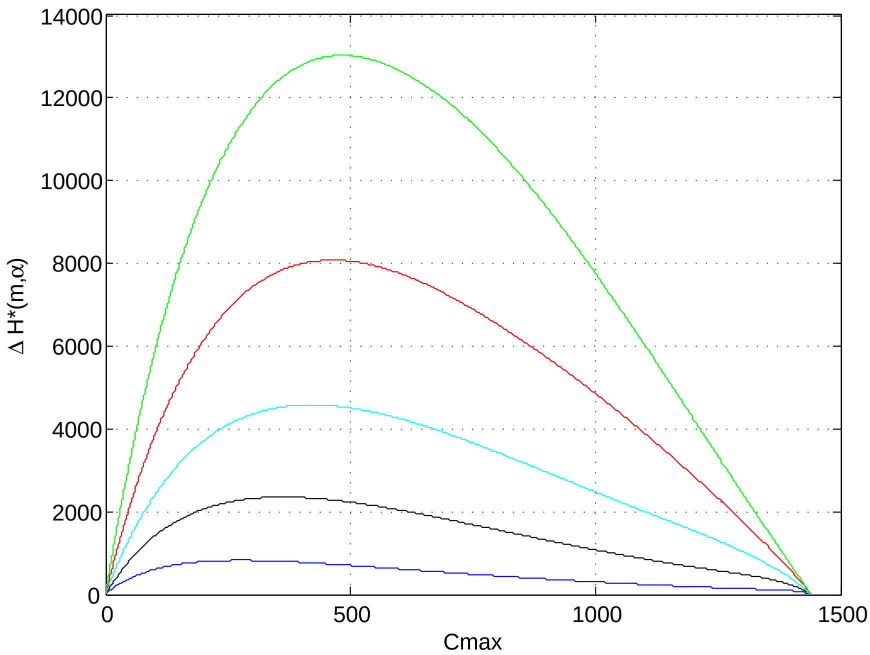

Figure 6: Surplus revenue of learning against no-learning plotted against $C_{\max}$ for five prior demand uncertainties

5.5 The Impact of Length of Selling Season (T) on Optimal Store Size

Figure 8 contains plots of the marginal values of shelf space for a range of T, the length of the selling season. The problem setting has $k = 1, \mathbf{E}[\gamma] = 10, \mathbb{V}_\gamma = 5$ and $S = 51$. Figure 9 has a corresponding typical plot of profit, namely revenue net of rent charged per unit of shelf space. We assume rental cost is linearly increasing with a increment of 50 monetary units per unit of shelf space. Recall that all reported revenues are per unit of time and hence we see from Figure 9 that profit per unit of time increases with the length of the selling season. This reflects the fact that mean revenues increase through the season as we gain more exploitable information about true demand rates and so our assortments can increasingly feature the best performing products. In this sense long selling seasons are a good thing. Correspondingly, it is unsurprising that from Figure 9 the store size ($C_{\max}$) maximising profit decreases with the length of the selling season. The rationale for this is that a short selling season requires more active learning in its early stages and hence larger stores. This is overwhelmingly the predominant insight. The cited example of small grocery shops in major metropolitan areas is supported by this finding. Such shops certainly face seasonality but once a store is established there is little further learning to be done about demand rates and selling seasons are close to infinite. Our analysis then suggests that optimal store sizes are small in such circumstances.

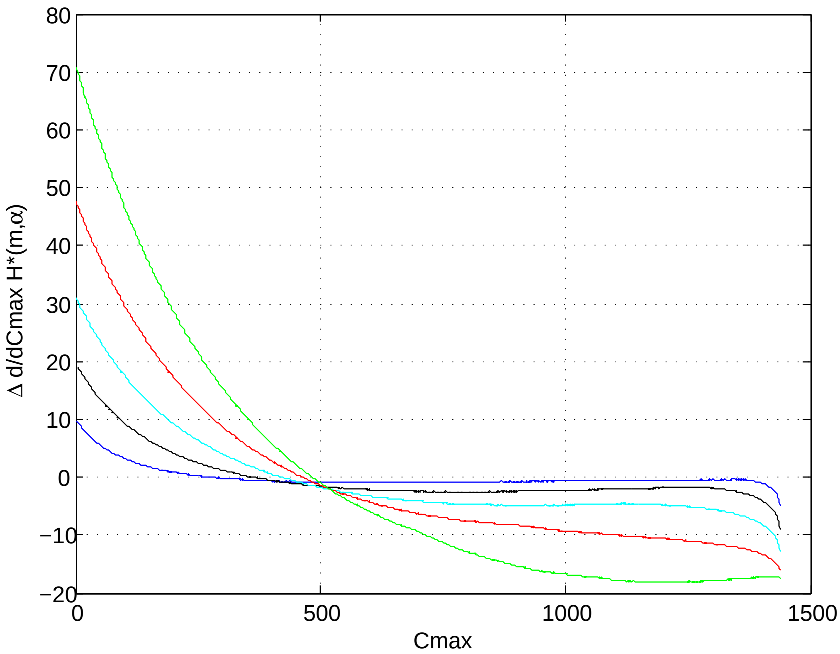

Figure 7: Marginal surplus revenue of learning against no-learning plotted against $C_{\max}$ for five prior demand uncertainties

Figure 8 points to a different analysis when rents are extremely high. In this case optimal store sizes are guaranteed small. If the selling season is also short then there will be very little opportunity for learning and very few products will ever be candidates for inclusion in assortments. A longer selling season opens up greater opportunity for profiting from demand learning, brings many more products into play and opens up the possibility that larger stores may be preferred.

6 Conclusions

The aim of this paper was to provide some guidance to managers on the interplay between shelf-space, demand learning and dynamic assortment problem decisions faced by fast fashion retailers. We have considered a stylized version of this problem, resulting in a finite horizon multi-armed bandit model with Bayesian learning, that takes the space requirement of the products into account. This has enabled us to calculate the value of capacity in such a challenging environment, leading us to be able to describe in detail the profit and loss resulting from capacity choice.

Further, we have numerically investigated the optimal store size for various uncertainty values and lengths of selling season. Our model then provides one possible explanation for the real-life observation that convenience stores, which allow long term learning for an

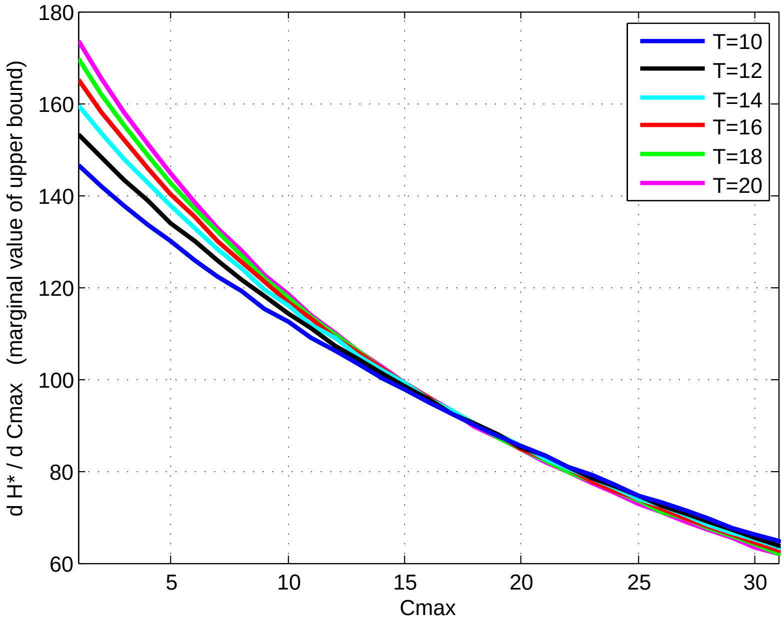

Figure 8: Marginal value plotted against $C_{\max}$ for six values of T, the length of the selling season

assortment that may not change over long periods, tend to be smaller in size than those shops which sell rapidly changing collections of seasonal items. Our model therefore provides valuable guidance for managers by identifying these factors as key drivers for shelf-space decisions. Managers in an environment where short term replenishment and assortment changes are possible should look at the level of demand uncertainty and the length of selling seasons when making initial decisions on store sizes. Overall, our paper shows that managers must consciously adopt a new attitude towards demand risk. In an environment where fast replenishment is possible and hence corrections to a store's assortment can be accomplished easily, demand uncertainty carries opportunities rather than risks. Managers who proceed myopically and choose an assortment that they think will sell well on expectation, a strategy that might have worked well in a supply chain setting with long lead times, will now leave substantial profit on the table. Of course, the idea of the trade-off of exploration versus exploitation and in particular the necessary change over time during the sales horizon if crucial for all store managers to understand and should be part of any corporate training effort.

While our stylized model already delivers valuable insight for managers who consider upsizing or downsizing stores, there might still be features that could be included to improve the quality of advice the model gives in practical situations. Future research, for example,

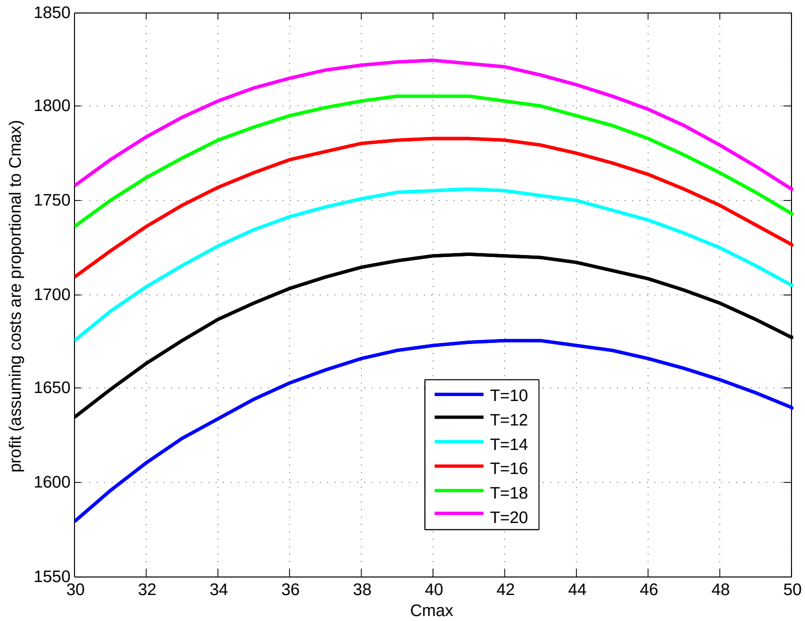

Figure 9: Profit plotted against $C_{\max}$ for six values of T, the length of the selling season

might aim to include substitution effects into the model. This seems to be a promising direction, both from a practical and theoretical standpoint.

References

Aviv, Y., A. Pazgal. 2005. Dynamic pricing of short life-cycle products through active learning. *Olin School Business, Washington Univ., St. Louis, MO* .

Brezzi, M., T.L. Lai. 2002. Optimal learning and experimentation in bandit problems. *Journal of Economic Dynamics and Control* **27**(1) 87–108.

Bultez, A., P. Naert. 1988. SH. ARP: shelf allocation for retailers' profit. *Marketing Science* **7**(3) 211–231.

Caro, F., J. Gallien. 2007. Dynamic assortment with demand learning for seasonal consumer goods. *Management Science* **53**(2) 276.

Chakrabarti, D., R. Kumar, F. Radlinski, E. Upfal. 2008. Mortal multi-armed bandits. Daphne Koller, Dale Schuurmans, Yoshua Bengio, Leon Bottou, eds., *NIPS – Neural Information Processing Systems*. MIT Press, 273–280.

Corstjens, M., P. Doyle. 1981. A model for optimizing retail space allocations. *Management Science* **27**(7) 822–833.

Fancher, L. 1991. Computerized space management: A strategic weapon. *Discount Merchandiser* **31**(3) 64.

Farias, V., R. Madan. 2008. The irrevocable multi-armed bandit problem. *Working paper* .

Fisher, M., A. Raman. 1996. Reducing the cost of demand uncertainty through accurate response to early sales. *Operations Research* 87–99.

Ginebra, J., M.K. Clayton. 1995. Response surface bandits. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**(4) 771–784.

Gittins, JC. 1979. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society. Series B (Methodological)* 148–177.

Gittins, John, Kevin Glazebrook, Richard Weber. 2011. *Multi-Armed Bandit Allocation Indices*. 2nd ed. Wiley.

Glazebrook, KD, C. Kirkbride, D. Ruiz-Hernandez. 2006. Spinning plates and squad systems: Policies for bi-directional restless bandits. *Advances in Applied Probability* 95–115.

Hariga, M.A., A. Al-Ahmari, A.R.A. Mohamed. 2007. A joint optimisation model for inventory replenishment, product assortment, shelf space and display area allocation decisions. *European Journal of Operational Research* **181**(1) 239–251.

Kök, A.G., M.L. Fisher, R. Vaidyanathan. 2008. Assortment planning: Review of literature and industry practice. *Retail Supply Chain Management* 1–55.

Mahajan, A., D. Teneketzis. 2007. Multi-armed bandit problems. *Foundations and Applications of Sensor Management* 121–151.

Nafari, M., J. Shahrabi. 2010. A temporal data mining approach for shelf-space allocation with consideration of product price. *Expert Systems with Applications: An International Journal* **37**(6) 4066–4072.

Nelder, JA, R. Mead. 1965. A simplex method for function minimization. *The Computer Journal* **7**(4) 308.

Saure, D., A. Zeevi. 2009. Optimal dynamic assortment planning. Tech. rep., Columbia Business School.

Wartenberg, F., W. Gaul, R. Decker. 1997. Computergestützte Regaloptimierung im Einzelhandel. *Der Markt* **36**(3) 185–196.

Weber, R.R., G. Weiss. 1990. On an index policy for restless bandits. *Journal of Applied Probability* **27**(3) 637–648.

Whittle, P. 1980. Multi-armed bandits and the Gittins index. *Journal of the Royal Statistical Society. Series B (Methodological)* **42**(2) 143–149.

Whittle, P. 1988. Restless bandits: Activity allocation in a changing world. *Journal of Applied Probability* **25** 287–298.

Yang, M.H. 2001. An efficient algorithm to allocate shelf space. *European Journal of Operational Research* **131**(1) 107–118.

Yang, M.H., W.C. Chen. 1999. A study on shelf space allocation and management. *International Journal of Production Economics* **60** 309–317.

Yücel, E., F. Karaesmen, F.S. Salman, M. Türkay. 2009. Optimizing product assortment under customer-driven demand substitution. *European Journal of Operational Research* **199**(3) 759–768.