UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 09-31
Statistics and Econometrics Series 11
May 2009

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

AN INDEX POLICY FOR DYNAMIC PRODUCT PROMOTION AND THE KNAPSACK PROBLEM FOR PERISHABLE ITEMS

Peter Jacko and José Niño-Mora

Abstract

This paper introduces the knapsack problem for perishable items (KPPI), which concerns the optimal dynamic allocation of a limited promotion space to a collection of perishable items. Such a problem is motivated by applications in a variety of industries, where products have an associated lifetime after which they cannot be sold. The paper builds on recent developments on restless bandit indexation and gives an optimal marginal productivity index policy for the dynamic (single) product promotion problem with closed-form indices that yield structural insights. The performance of the proposed policy for KPPI is investigated in a computational study.

Keywords: dynamic promotion, perishable items, index policies, knapsack problem, restless bandits, finite horizon, marginal productivity index

Department of Statistics, Universidad Carlos III de Madrid, Av. de Universidad 30, 289 11 Leganés (Madrid), Spain, e-mail: peter.jacko@uc3m.es, jnimora@alum.mit.edu

An index policy for dynamic product promotion and the knapsack problem for perishable items ⋆

Peter Jacko and José Niño-Mora

Department of Statistics
Universidad Carlos III de Madrid
Avda. Universidad 30, 28911 Leganés (Madrid), Spain
peter.jacko@uc3m.es, jose.nino@uc3m.es

Abstract. This paper introduces the knapsack problem for perishable items (KPPI), which concerns the optimal dynamic allocation of a limited promotion space to a collection of perishable items. Such a problem is motivated by applications in a variety of industries, where products have an associated lifetime after which they cannot be sold. The paper builds on recent developments on restless bandit indexation and gives an optimal marginal productivity index policy for the dynamic (single) product promotion problem with closed-form indices that yield structural insights. The performance of the proposed policy for KPPI is investigated in a computational study.

Key words: dynamic promotion, perishable items, index policies, knapsack problem, restless bandits, finite horizon, marginal productivity index

⋆ This work has been supported by the Spanish Ministry of Education and Science/Science and Innovation under projects MTM2004-02334 and MTM2007-63140, being part of the PhD thesis work of the first author funded by predoctoral research training grant no. BES-2005-7026 attached to the first project and supervised by the second author. Preliminary versions have been presented at the 6th Czech-Slovak International Symposium on Combinatorics, Graph Theory, Algorithms and Applications (Prague, 2006), where an extended abstract appeared in the proceedings (Electronic Notes in Discrete Mathematics 28 pp. 145–152, 2007), and at the 8th Annual INFORMS Revenue Management and Pricing Section Conference (Montreal 2008).

1 Introduction

Managers in the retail industry face hard decisions on product assortment, pricing, and promotion. Product assortment (selection of products and their shelf space and location) is a strategic decision, taken over a long-term planning horizon (Kök et al., 2006). In contrast, product pricing and promotion have shorter planning horizons, as they are used in day-to-day marketing decisions to dynamically respond to demand variations. In the food retail industry, the complexity of pricing and promotion decisions is magnified by the *perishability* of products.

A standard approach to increase the revenue obtained from perishable products is *dynamic pricing*, i.e., adjusting product prices to demand variations. See, e.g., Elmaghraby and Keskinocak (2003), which gives an extensive overview of dynamic pricing and its adoption in practice.

However, discrete-time decision making, implementation costs, and retail brand image strategy make practitioners not to like changing prices too often or in an "unsystematic" fashion, as prescribed by theoretical dynamic pricing models. In addition, price reductions must usually be done over all product units, thus losing possible profits from customers who are willing to pay the original, higher price. Revenue managers naturally try to avoid situations in which the product price is lower than the inventory costs, as they lead to a negative net profit. Such a behavior may result in conservative product orders, which increase the probability and duration of stockouts. Based on several studies, Campo and Gijsbrechts (2005) documented that consumers' reaction to stockouts may have significant negative impacts on retail sales and revenues.

The above suggests that there is a strong need by retail managers for a "softer" marketing tool, which dynamically allows them to improve sales and revenues, yet without altering product prices.[1]

This paper introduces a revenue management model in which demand is altered not by price changes, but instead by dynamically allocating product units to a limited *promotion space*, where they are more likely to attract customers. Examples of such a promotion space include shelves close to the cash register, promotion kiosks, or a depot used for selling via the Internet.

We thus address the problem of designing a *dynamic promotion policy* for allocating units of perishable products to a promotion space, in order to maximize the expected revenue before all units expire, which we term the *knapsack problem for perishable items* (KPPI). Throughout the paper, we use the example of a food retailer, focusing on the case when there is a single unit of each product.

Our approach relies on a decomposition of the problem to single products. Each product is assigned a promotion index. Such an index, which we derive in closed form, allows us to consider a promotion index policy, where higher priority for promition is awarded to items whose current index values are larger. Although such an index policy may be suboptimal, it has a sound economic interpretation and is computationally tractable, since the indices are given in closed form.

[1] An example of the practical interest of such a tool is provided by the cooperation of Capgemini, Intel, Cisco, and Microsoft on a decision support system which includes Dynamic Promotion Management as one of three key solution areas.

1.1 Paper Structure, Goals, and Contributions

Section 2 outlines the model and briefly reviews the work on related models in the literature. We further describe the milestones in the research on the bandit problem and its extensions, especially the restless bandit problem, which we use to set our model. The celebrated classic result on the bandit problem is the optimality of a dynamic solution defined by certain priority indices. Derivation of such indices for our model is one of the two main goals of this paper.

The KPPI is formalized in section 3. Since the dynamic programming formulation is intractable, we formulate it approximately using a Lagrangian relaxation, which allows us to decompose the problem into single-item subproblems.

An optimal promotion policy for a single item defined via marginal productivity indices is derived in section 5. The optimal dynamic promotion policy is shown to be time-monotonous: the value of promoting increases as the deadline approaches.

This leads us to our second goal and contribution, the development of a new index-based heuristic for the KPPI in section 6. A computational study is described in section 7.

Concluding remarks are given in section 8. Proofs are deferred to the Appendix.

2 Model Outline and Related Work

A *perishable item* is a product unit with an associated lifetime ending at a *deadline*. At the deadline (e.g., the "best before" date) the product can no longer be sold, and only a *salvage value* is received. If an item is sold before the deadline, it yields a *revenue* (profit margin). The probability of selling only depends on whether the item is being promoted or not.

The KPPI of concern is to dynamically select a subset of items to be included in a promotion space (knapsack), in order to maximize the expected total discounted revenue.

We set the model in discrete time as a Markov decision process. We assume that the decisions are made in some regular time moments, say, twice a day, and the problem parameters are adjusted to such time periods. In general, the KPPI defines a stochastic variant of the knapsack problem with multiple units. As time evolves, some items get sold accordingly to a stochastic time-homogeneous demand and some items perish deterministically at their deadlines.

A related though distinct model is the *dynamic and stochastic knapsack problem* (DSKP) (cf., e.g., Papastavrou et al. (1996) and Lu et al. (1999)), which concerns the dynamic control of admission to a knapsack of arriving items with random rewards and weights, in order to maximize the knapsack contents' value by a given deadline.

The KPPI is related to dynamic pricing problems, and we show that a simple dynamic pricing problem (of a single product) can be formulated in our framework as a dynamic promotion problem when one must pay for promotion.

We further assume in our model that the demand is time-homogeneous (cf. Elmaghraby and Keskinocak (2003) for a justification of such an assumption).

2.1 Bandit Problem Literature

A convenient mathematical setting for the KPPI problem is the multi-armed bandit problem (cf. Gittins, 1979), in which one wants to dynamically choose between various *bandits* (reward-yielding processes) one in an optimal fashion. That model captures the fundamental trade-off between *exploitation* of current rewards and *exploration* of possible future rewards. Gittins (1979) showed that the multi-armed bandit problem can be optimally solved using his indices, which can be calculated in polynomial time.

The bandits are here the perishable items. In our model, however, there are four complications: the bandits are *restless*, because the items can get sold regardless of being in the knapsack or not, the time horizon is *finite* due to perishability, and we are to select *more than one* item for the knapsack, which is allowed to be filled partially, due to the *heterogeneity* of the items. We thus mix up two models: the bandit problem and the knapsack problem.

The multi-armed restless bandit problem (which is a generalization of the multi-armed bandit problem considered in Gittins (1979)) over the infinite horizon was proven to be P-space hard even in its deterministic version (Papadimitriou and Tsitsiklis, 1999). The research focus thus shifts to the design of well-grounded, tractable heuristic policies. For the analysis we use the framework and methodology proposed for restless bandits by Niño-Mora (2002, 2006). That work provided a sufficient condition for a restless bandit to be *indexable* together with an adaptive-greedy algorithm, which in $O(n^3)$ operations computes corresponding *marginal productivity indices* that extend earlier indices of Gittins (classical bandits, 1979) and Whittle (restless bandits, 1988). In our problem, the marginal productivity index can be interpreted as the *promotion priority index*.

The indexability property of a single item modeled as a restless bandit means that there exist promotion priority indices such that the optimal solution is to promote the item whenever its promotion priority index is higher than the cost of promotion space occupation (promotion cost). When coupling the bandits back into a multi-armed (non-restless) bandit problem, the promotion priority indices define an optimal policy: At every decision epoch choose the bandit of highest promotion priority index (*promotion-priority-index policy*). Such a priority policy is in general not optimal for the restless case, in which it becomes a well-grounded, efficient and practical heuristic.

Regarding the bandit problems with finite horizon, interesting results of index nature appear very sporadically, because of the intractability of the model, and therefore other methods (such as dynamic programming) are usually used. Even then, the problem is computationally intractable. Nevertheless, there is a tractable instance, the so-called *deteriorating case*, first presented for an infinite-horizon bandit problem by Gittins (1979), which was also successfully applied in a problem with finite-horizon objective (Manor and Kress, 1997). In that setting, the bandits were, however, not restless. The same is the case for the index policies for the finite-horizon multi-armed (non-restless) bandit problem: Niño-Mora (2005) showed that such a problem is indexable and provided a tractable algorithm.

The bandit problem framework was used to analyze adaptive marketing strategies by various authors. Caro and Gallien (2007) developed a model for dynamic assortment of seasonal goods and proposed an assortment-priority-index policy using approximate indices that they obtained in a closed form. An adaptive model for interactive marketing environment was introduced in Bertsimas and Mersereau (2007). Both works build upon the classical multi-armed bandit problem (in which the bandits are *not* restless) and propose heuristics based on an approximate problem decomposition.

3 The Knapsack Problem for Perishable Items

A retailer has a finite collection of perishable items to sell, labeled by $k \in \mathbb{K} \triangleq \{1, \ldots, K\}$. Item k can only be sold during its lifetime, which consists of time periods $t = 0, 1, \ldots, T_k - 1$, where $T_k < \infty$ is the item's deadline. If the item is sold, it yields a reward $R_k > 0$. Otherwise, a salvage value $\alpha_k R_k$ is obtained, for some (possibly negative) coefficient $\alpha_k < 1$. Rewards and costs are discounted over time with factor $0 < \beta \leq 1$.

The retailer can increase the probability that item k is sold during a period, from $1 - q_k$ to $1 - p_k$ (with $0 < q_k < p_k < 1$), by placing it in a *promotion knapsack*. We assume that such Bernoulli demand processes are independent across items. The advantages of promotion must be balanced by the following factors. First, the knapsack has a limited capacity, having room for holding an integer number W of physical space units. Item k occupies $W_k < W$ units, and we assume that $\sum_k W_k > W$. Second, usage of the knapsack incurs holding costs at rate ν per unit capacity used per period.

To formulate such a model as an MDP, we consider the *state* $X_k(t)$ of item k at the beginning of period t to be $X_k(t) = T_k - t$ (the number of remaining periods to the deadline) if $t < T_k$ and the item has not yet been sold, and to be $X_k(t) = 0$ if either $t \geq T_k$ or the item has been sold. Further, we use binary *action* (active/passive) processes $a_k(t) \in \{0, 1\}$, where the *active action* $a_k(t) = 1$ corresponds to promoting item k into the knapsack at the beginning of period t, and is only allowed when $t < T_k$ and the item remains unsold.

As for the state dynamics, the positive transition probabilities under the active action are given by $p_k^1(s, s-1) \triangleq p_k$ and $p_k^1(s, 0) \triangleq 1 - p_k$ for $2 \leq s \leq T_k$, and $p_k^1(1, 0) \triangleq 1$, whereas those under the *passive action* ($a_k(t) = 0$) are $p_k^0(s, s-1) \triangleq q_k$ and $p_k^0(s, 0) \triangleq 1 - q_k$ for $2 \leq s \leq T_k$, and $p_k^0(1, 0) \equiv p_k^0(0, 0) \triangleq 1$. All other state transition probabilities are zero.

The immediate expected rewards for item k in each state s are as follows. Under the active action, $R_k^1(s) \triangleq (1 - p_k) R_k$ for $2 \leq s \leq T_k$, and $R_k^1(1) \triangleq (1 - p_k) R_k + p_k \alpha_k R_k$. Under the passive action, $R_k^0(s) \triangleq (1 - q_k) R_k$ for $2 \leq s \leq T_k$, $R_k^0(1) \triangleq (1 - q_k) R_k + q_k \alpha_k R_k$, and $R_k^0(0) \triangleq 0$.

Action choice is prescribed by a dynamic *promotion policy* $\boldsymbol{\pi}$, to be chosen from the corresponding class $\boldsymbol{\Pi}(W)$ of history-dependent policies, which satisfy the sample-path knapsack capacity constraint

$$\sum_{k \in \mathbb{K}} W_k a_k(t) \leq W, \quad t = 0, 1, 2, \ldots \tag{1}$$

The KPPI is to find an admissible policy that maximizes the expected total discounted reward earned. We can formulate such a problem as

$$V^*(\mathbf{T}) = \max_{\boldsymbol{\pi} \in \boldsymbol{\Pi}(W)} \mathbb{E}_{\mathbf{T}}^{\boldsymbol{\pi}} \left[\sum_{k \in \mathbb{K}} \sum_{t=0}^{\max_k T_k} R_k^{a_k(t)}\big(X_k(t)\big) \beta^t \right], \tag{2}$$

where $\mathbb{E}_{\mathbf{T}}^{\boldsymbol{\pi}}[\cdot]$ denotes expectation under policy $\boldsymbol{\pi}$ conditioned on the initial joint state being equal to $\mathbf{T} = (T_k)$.

Being a finite-horizon MDP, standard results ensure existence of an optimal policy for (2) that is Markov deterministic. Yet, its computation via numerical solution of the resulting DP equations becomes intractable as the number K of items grows,due to the *curse of dimensionality*.

4 Work-Reward Restless Bandit Formulation of KPPI

We next formulate the KPPI as a variant of the multi-armed restless bandit problem, where the restless bandit is replaced by what we call the *work-reward restless bandit*. We set out to obtain a tractable index rule based on the marginal productivity indices.

In the multi-armed restless bandit problem (cf. Whittle, 1988), all bandits have the same requirement on the resource. In our model, however, we admit non-uniform resource (i.e., knapsack space) requirements, which is a special case of the model in Niño-Mora (2002). In the Whittle's restless bandit framework, the immediate *work* is assumed to be 1 for the active action and 0 for being passive. Nevertheless, in our case the active action (promoting) requires a non-uniform utilization of the knapsack and we need to reflect this feature in our model. Therefore, we define the immediate (promotion) work of an item i in state $t \in \mathscr{T}_i$ under the active action by its volume, $W^1_{i,t} := W_i$, and $W^0_{i,t} := 0$ under the passive action. We further define $W^0_{i,0} := 0$.

We arrive to the following formulation of the KPPI,

$$\begin{aligned} &\max_{\pi\in\Pi} \mathbb{E}^{\pi}_{\boldsymbol{T}} \left[\sum_{i\in\mathscr{I}} \sum_{s=0}^{\infty} \beta^s R^{a_i(s)}_{i,X_i(s)} \right] \\ &\text{subject to} \quad \sum_{i\in\mathscr{I}} W^{a_i(s)}_{i,X_i(s)} \le W \text{ at each time } s = 0, 1, \ldots, \infty \end{aligned} \tag{RB}$$

where, as before, $X_i(s)$ denotes the state of item i at time period s, starting at state $X_i(0) = T_i$.

4.1 Problem Relaxation and Lagrangian Relaxation Decomposition

Whittle (1988) proposed for restless bandits what has become known as the *Whittle's relaxation*: replace the infinite set of resource constraints by one constraint requiring to use the full resource *in expectation*. In our case, the Whittle's relaxation of the (RB) is the following:

$$\begin{aligned} &\max_{\pi\in\Pi} \mathbb{E}^{\pi}_{\boldsymbol{T}} \left[\sum_{i\in\mathscr{I}} \sum_{s=0}^{\infty} \beta^s R^{a_i(s)}_{i,X_i(s)} \right] \\ &\text{subject to} \quad \mathbb{E}^{\pi}_{\boldsymbol{T}} \left[\sum_{i\in\mathscr{I}} \sum_{s=0}^{\infty} \beta^s W^{a_i(s)}_{i,X_i(s)} \right] = \frac{W}{1-\beta}. \end{aligned} \tag{WR}$$

The Whittle relaxation (WR) can be solved by the Lagrangian method. Let κ be a Lagrangian multiplier for the constraint, then the Lagrangian of (WR) is

$$L(\pi,\kappa) = \mathbb{E}^{\pi}_{\boldsymbol{T}} \left[\sum_{i\in\mathscr{I}} \sum_{s=0}^{\infty} \beta^s R^{a_i(s)}_{i,X_i(s)} \right] - \kappa \left(\mathbb{E}^{\pi}_{\boldsymbol{T}} \left[\sum_{i\in\mathscr{I}} \sum_{s=0}^{\infty} \beta^s W^{a_i(s)}_{i,X_i(s)} \right] - \frac{W}{1-\beta} \right)$$

and can be rewritten as

$$L(\pi,\kappa) = \sum_{i\in\mathscr{I}} \left(\mathbb{E}^{\pi}_{\boldsymbol{T}} \left[\sum_{s=0}^{\infty} \beta^s R^{a_i(s)}_{i,X_i(s)} \right] - \kappa \mathbb{E}^{\pi}_{\boldsymbol{T}} \left[\sum_{s=0}^{\infty} \beta^s W^{a_i(s)}_{i,X_i(s)} \right] \right) + \kappa \frac{W}{1-\beta} \tag{L}$$

For a given penalizing parameter κ, (L) can be decomposed and analyzed separately for each item, which we will do in section 5. We interpret the parameter κ as the competitive market price of space, the resource provided by the knapsack. Indeed, the term $\kappa W/(1-\beta)$ can be viewed as the *money budget* allocated for the knapsack space we expect to be using (κW per period). Since we only consider the space utilization in expectation, we in fact assume existence of a space market, where we permit to "buy" some amount of space if necessary or to "sell" some amount of space if it is not used. Then, there is an optimal market price κ^* which balances expected supply (selling free space) and expected demand (buying necessary space). If this price is known, then $\max_{\pi\in\Pi} L(\pi,\kappa^*)$ solves (WR).

This *perfect market assumption* reflected at the Whittle's relaxation is sufficient for the KPPI to be solved efficiently. Its solution, however, is not feasible for the original problem (RB), because no such space market is in practice available. The optimal solution is a dynamic adaptive-knapsack policy; in the original problem a dynamic fixed-knapsack policy is sought. Nevertheless, the optimal solution to the Whittle's relaxation yields a tractable bound for the original problem.

The optimal dynamic adaptive-knapsack policy, however, may be relevant in some applications. One could think of adjusting the knapsack's space dynamically as the optimal perfect market solution requires, e.g. by reserving a necessary number of promotion shelves, where the price κ^* must be paid as a promoter's wage for each reserved space unit.

5 Optimal Promotion of Perishable Item

The aim of this section is to obtain the marginal productivity indices (MPIs) for a perishable item in isolation, which will be the building block for the KPPI solution developed in section 6. The MPIs capture the marginal rate of promotion and define an index policy, which furnishes an optimal control of a perishable item by indicating when it is worth promoting. For that end, we introduce a per-period *promotion cost* $\nu_i \geq 0$, which must be paid in every period when the item is being promoted.

Since we are considering item i in isolation, in the following we drop the item's subscript i. Starting from state t, we define the *expected total discounted net revenue* under policy π as

$$\mathbb{E}_t^\pi\left[\sum_{s=0}^{\infty}\beta^s R_{X(s)}^{a(s)}\right] - \nu\mathbb{E}_t^\pi\left[\sum_{s=0}^{\infty}\beta^s W_{X(s)}^{a(s)}\right], \tag{3}$$

where, as before, $X(s)$ is the state at time period s and $a(s)$ is the action applied in time period s, counting promoting as 1 and not promoting as 0. The symbol $\mathbb{E}_t^\pi$ denotes the expectation under policy π if starting from state t.

Denote by Π the set of all non-anticipative policies for such a problem. The goal is to find a policy $\pi^* \in \Pi$ that maximizes (3) for $t = T$ among all such policies, and thus optimally resolves the trade-off between the expected total discounted revenue and the expected total discounted promotion cost.

The perishable item as defined above falls to the concept of restless bandit (a binary-action MDP). Under some circumstances, one can identify its optimal control in terms of marginal productivity indices (MPIs). Next we show that such an optimal MPI policy for perishable item exists under a demand regularity condition, and we identify it.

5.1 Assumption

We will continue under the following assumption, which requires the promotion power $q - p$ to be positive.

Assumption 1. $q - p > 0$.

Assumption 1 is a consistency requirement on promotion power which rules out uninteresting items that should never be promoted. Indeed, the optimal action in all the states for an item with promotion power $q - p \leq 0$ is not promoting (as long as $\nu \geq 0$). We will show that promoting is optimal if promoting was optimal in the previous period, i.e., the efficiency of promoting nondecreases over time.

5.2 Marginal Productivity Index

Now we examine the economics of promoting the perishable item. In particular, we examine the efficiency of promoting the item in its current state if one must pay for promotion. We will identify circumstances in which it is worth to promote the item, by assigning the *marginal productivity index* (MPI), which captures the marginal rate of promoting, to each controllable state. The optimal MPI policy is: "Promote the item if and only if the MPI of the actual state is greater than the promotion cost ν." We employ the MPI calculation method as described for restless bandits in Niño-Mora (2006).

We show in Proposition 1 that a perishable item is *indexable*, that is, the optimal decisions are prescribed by the MPI policy, using marginal productivity indices assigned to controllable states.

Proposition 1. *The perishable item is indexable, and the marginal productivity index for its controllable state t is*

$$\nu_t^* = \frac{R}{W} \frac{(q-p)\left[(1-\beta)\frac{1-(\beta p)^{t-1}}{1-\beta p} + (1-\beta\alpha)(\beta p)^{t-1}\right]}{1-(\beta q - \beta p)\frac{1-(\beta p)^{t-1}}{1-\beta p}}. \tag{4}$$

The proof of Proposition 1 is presented in the Appendix together with a more detailed description of the work-reward analysis. Next we list the most appealing properties of the MPI, which have insightful interpretation and define heuristical priorities for promotion if various items compete for a limited promotion space. In section 6 we implement MPI as promotion priority measure: the higher the MPI, the higher the promotion priority.

Proposition 2 (MPI Properties). *For any controllable state t,*

(i) the MPI is nonnegative and proportional to R/W;
(ii) an item with lower probability of being sold when not promoted ($(1-q)$'s), ceteris paribus, has higher MPI;
(iii) (Time Monotonicity) the MPI of an item is nondecreasing as t diminishes (i.e., as the deadline approaches).

MPI resolves the trade-off between immediate and postponed promotion. Proposition 2(iii) is a crucial property of MPIs, which demonstrates that the marginal productivity index is nondecreasing as the deadline approaches. Armed with this result, we can look for an *optimal promotion starting time* τ^*,

$$\tau^* := \max\{\tau \in \mathscr{T} : \nu_t^* > \nu \text{ for all } t \in \mathscr{T} \text{ such that } t \geq \tau\}. \tag{5}$$

In other words, (if τ^* is finite,) τ^* is the threshold time period, from which the MPI is larger than the promotion cost ν, i.e. from which it is optimal to start to promote the item, ceasing to promote it at the deadline. If τ^* is not finite, then it is never optimal to promote the item.

Proposition 3. *The optimal starting time τ^* is finite if and only if*

$$\frac{R}{W}\frac{(1-\beta\alpha)}{\nu}(q-p) > 1.$$

Further,

(i) if τ^* is finite, then promoting is optimal in all time periods from τ^* to 1 and not promoting is optimal in the remaining time periods;
(ii) if τ^* is not finite, then not promoting is optimal in all time periods.

The above result assures that promotion is to be made in a natural way: the item is selected for promotion only once and remains promoted while it is profitable to do it.

5.3 Special Cases and Further Remarks

We further give the marginal productivity index for certain classes of perishable items.

Proposition 4.

(i) [Undiscounted Case] Under quasi-nondecreasing promotion power, in the case $\beta = 1$, the marginal productivity index for controllable state t is

$$\nu_t^* = \frac{R}{W}\frac{(1-\alpha)(q-p)(1-p)p^{t-1}}{1-q+(q-p)p^{t-1}}.$$

(ii) [Reduction to $c\mu$-rule] If $q = 1$, $W = 1$, and the discount factor $\beta = 1$, then the perishable item is indexable, and the marginal productivity index for its controllable state $t \in \mathscr{T}$ is

$$\nu_t^* = R(1-\alpha)(1-p).$$

Proposition 4(ii) tackles the situation in which there is no possibility of selling the item if not promoted. Thus, only promoted item can be sold, and selling happens with probability $1-p$ in every period. Interpreting this probability as a service rate, the MPI reduces to the $c\mu$-rule, well-known in the queueing theory (see, e.g., Buyukkoc et al., 1985), where $c := R(1-\alpha)$ is the reduction in revenue if item is not sold during its lifetime. Such an MPI is constant over time and in particular it does not depend on the number of periods before the deadline. This rule is fittingly applied in assortment practice where products are chosen accordingly to their profitability and attractiveness.

5.4 Formulation of Dynamic Pricing Problem in our Framework

Suppose that we are given an additional parameter called *discount* (price markdown) $D \geq 0$, so that the revenue is $R - D$ instead of R if the item is promoted. Thus, $1 - q$ can be interpreted as the probability of selling the item priced at R, and $1 - p$ can be interpreted as the probability of selling the item priced at $R - D$. Let $\widetilde{\nu}$ be the per-period cost of maintaining (or informing about) the lower price. We are thus addressing a simple case of the classic *dynamic pricing problem*.

In particular, we would like to have the following revenues:

$$
\begin{aligned}
\widetilde{R}_t^0 &:= \beta R(1-q), && \text{for } t \in \mathscr{T} \setminus \{1\}; \\
\widetilde{R}_t^0 &:= \beta R(1-q) + \beta\alpha R q, && \text{for } t = 1; \\
\widetilde{R}_t^1 &:= \beta (R-D)(1-p), && \text{for } t \in \mathscr{T} \setminus \{1\}; \\
\widetilde{R}_t^1 &:= \beta (R-D)(1-p) + \beta\alpha R p, && \text{for } t = 1; \\
\widetilde{R}_0^0 &:= 0,
\end{aligned}
$$

Denote by $\widetilde{D} := \beta D(1-p)$. In order to cast the above problem into our framework, we define the one-period revenue for all $t \in \mathscr{T}$ as follows:

$$
R_t^0 := \widetilde{R}_t^0, \qquad R_t^1 := \widetilde{R}_t^1 + \widetilde{D}, \qquad R_0^0 := \widetilde{R}_0^0,
$$

and the promotion cost $\nu := \widetilde{\nu} + \widetilde{D}$.

Then, this modified model is effectively the same as the original one, and hence the optimal policy for the dynamic problem is the following: "Price the product at the lower price if and only if the MPI of the actual state is greater than the cost $\widetilde{\nu} + \widetilde{D}$."

6 Index-Based Heuristics for KPPI

Together with the relaxation, Whittle (1988) proposed a simple priority policy for the multi-armed restless bandit problem employing the indices once an optimal index policy for each bandit is available: "Promote the bandit of highest index". Heuristics based on this simple idea showed an exceptionally accurate performance in various problems formulated in the framework of the multi-armed restless bandit problem. In the KPPI problem, however, this heuristic may not be the best proposal, since it assumes the same space requirement of all perishable items ($W_i = 1$).

In a more general model of Niño-Mora (2002), the resource requirements were assumed to be non-uniform and stochastic. Thus, his indices differ from the Whittle's and it was shown that the latter may be suboptimal in the Niño-Mora (2002)'s model. We applied this approach to the perishable item in section 5, where the item's MPI is volume-adjusted (since it includes a division by the volume W_i), so that the heuristic is: "Promote the item of highest volume-adjusted index".

Since the marginal productivity index can be interpreted as measuring the marginal rate of substitution (i.e., the price per unit of space requirement) of promoting the item, we propose the following heuristic construction for the KPPI: "Promote the items that are given by an optimal solution to the knapsack subproblem defined below with item's MPIs multiplied by volumes as the objective function price coefficients and item volumes as the knapsack constraint weights".

Notice that the greedy solution to the knapsack problem arising in our heuristic reduces it to "Promote the items of highest volume-adjusted index". It is well known that the greedy algorithm yields an optimal solution of a knapsack problem when all the weights are uniform; however, in the general case it is suboptimal. Our simulation study presented in section 7 suggests that the latter heuristic reveals an analogous performance: it is inferior and converging to ours.

In the following we assume that $q_i > p_i$ holds for all items i (otherwise not promoting is always optimal for such an item). Recall expression (4) for the MPI calculation, which is to be used in the heuristics via

$$v_i := W_i \nu^*_{i,T_i}. \tag{6}$$

Heuristic MPI–OPT: Calculate the prices v_i and then solve the knapsack subproblem optimally.

Heuristic MPI–GRE: Calculate the price/volume ratios v_i/W_i and then select the items for promotion in a greedy manner (highest first).

6.1 Knapsack Subproblem

Suppose that the knapsack-problem prices v_i of all items are calculated using expression (6). Then we have the following 0-1 knapsack problem to solve:

$$\begin{aligned} \max_{\boldsymbol{z}} & \sum_{i \in \mathscr{I}} z_i v_i \\ \text{subject to} \quad & \sum_{i \in \mathscr{I}} z_i W_i \leq W \\ & z_i \in \{0, 1\} \text{ for all } i \in \mathscr{I} \end{aligned} \tag{KP}$$

where $\boldsymbol{z} = (z_i : i \in \mathscr{I})$ is the vector of binary decision variables denoting whether the item i is selected for the promotion knapsack or not.

The quality of the solution $\boldsymbol{z}$ is not guaranteed to be optimal. The experimental study in the next section, however, reveals its nearly-optimal behavior, systematically outperforming other considered heuristics.

Finally, the next proposition asserts that the KPPI is a generalization of the knapsack problem.

Proposition 5 (KPPI Reduction to KP). *If $T_i = 1, q_i = 1, p_i = 0$ for all $i \in \mathscr{I}$, then any optimal solution $\boldsymbol{z}^*$ of the knapsack problem (KP) is an optimal solution of the KPPI.*

7 Experimental Study

In this section we present results of computational experiments, in which we evaluate the performance of heuristics MPI–OPT and MPI–GRE. We further compare their performance to the greedy *Earlier-Deadline-First* policy (EDF–GRE), a benchmark policy often observed in practice.

Heuristic EDF–GRE: Select products in a greedy manner after sorting the items so that product i_1 is preferred to product i_2, if:

(i) $T_{i_1} < T_{i_2}$,
(ii) $T_{i_1} = T_{i_2}$ and $R_{i_1}(1-\alpha_{i_1}) > R_{i_1}(1-\alpha_{i_1})$,
(iii) $T_{i_1} = T_{i_2}$ and $R_{i_1}(1-\alpha_{i_1}) = R_{i_1}(1-\alpha_{i_1})$ and $W_{i_1} < W_{i_2}$.

The following is the worst-case (i.e., minimizing) solution of the knapsack subproblem whenever all the price coefficients are positive, which is our case.

Heuristic MIN: Leave the knapsack empty.

In each experiment we randomly generate 10^4 instances for each fixed pair (I, T), denoting the number of products and the time horizon, respectively, such that $I \in \{2, 3, 4, \ldots, 8\}$ and $T \in \{2, 4, 6, \ldots, 20\}$. For each product i we set $\alpha_i = 0.5$ and we assure that $T_1 := T$. We assume that the standard and the promotion demands are Poisson with the respective means λ_i^0, λ_i^1 and such that $\frac{1}{2}\lambda_i^a T_i \leq 1 < \frac{3}{2}\lambda_i^a T_i$ for both $a \in \{0, 1\}$. The last condition assures that each item has a non-extreme probability of being sold before the deadline. Thus, we define $q_i := \exp\{-\lambda_i^0\}$ and $p_i := \exp\{-\lambda_i^1\}$, and assure that $q_i > p_i$. We further generate the following uniformly distributed parameters:

$$W_i \in [10, 50]; \qquad R_i \in [10, 50]; \qquad T_i \in [2, T]; \qquad \lambda_i^0, \lambda_i^1 \in \left(\frac{2J_i}{3T_i}, \frac{2J_i}{T_i}\right].$$

Finally, a uniformly distributed knapsack volume is generated: $W \in [\max\{W_i\}, 30\% \cdot \sum_i W_i)$.

We focus on the discount factor $\beta = 1$, as this is the case most likely to be implemented in practice. Moreover, our experiments (not reported here) suggest that this is also the hardest case and the performance of index-based heuristics improves as the discount factor diminishes.

The experiments were performed on PC with 2.66 GHz CPU and 1.5 GB RAM working on Windows XP. A Delphi code was developed by the author, implementing a standard enumerative routine for the knapsack subproblem. Finally, the performance evaluation measures (see below) were calculated using Matlab, which also created the figures presented here.

7.1 Performance Evaluation Measures

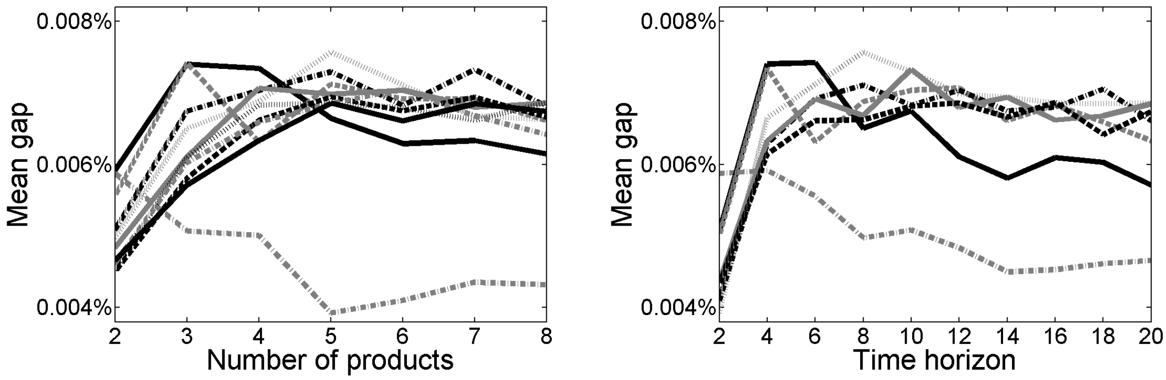
We obtain the maximizing policy solving the KPPI optimally, which also yields the optimal objective value D^{MAX}. The objective values of the other policies are also obtained via the Bellman equations, employing the respective heuristic at each step, denoted D^π for a policy π. We next introduce performance evaluation measures we use to report the experiment results.

The *relative suboptimality gap* of policy π is calculated via

$$\text{rsg}(\pi) = \frac{D^{\text{MAX}} - D^\pi}{D^{\text{MAX}}}. \tag{7}$$

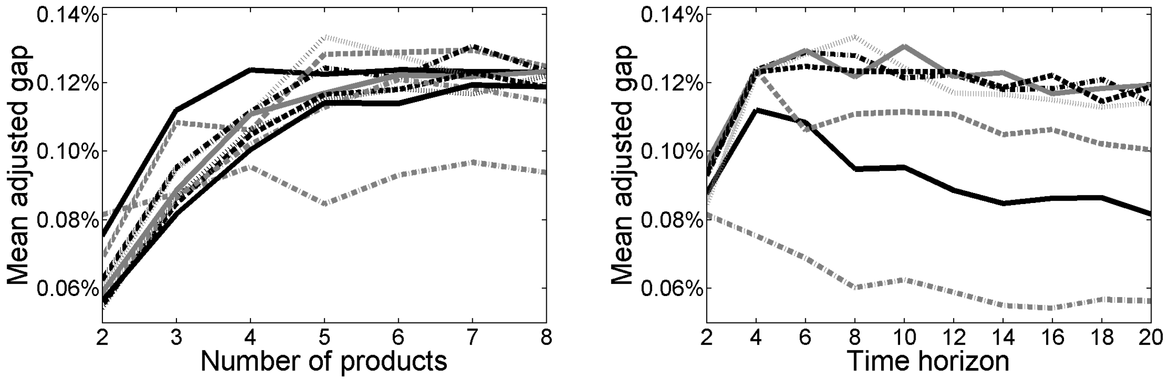
Clearly, we have $0 \leq \text{rsg}(\pi) \leq 1$, where $\text{rsg}(\pi) = 0$ is obtained by the maximizing policy. However, $\text{rsg}(\pi) = 1$ cannot be achieved unless $\alpha_i \leq 0$ at least for some item i. This motivates us to introduce the *adjusted relative suboptimality gap* of policy π, calculated via

$$\text{arsg}(\pi) = \frac{D^{\text{MAX}} - D^\pi}{D^{\text{MAX}} - D^{\text{MIN}}}. \tag{8}$$



(a) for values $T = 2, 4, 6, \ldots, 20$ (b) for values $I = 2, 3, \ldots, 8$

Figure 1. Mean relative suboptimality gap of heuristic MPI–OPT.



(a) for values $T = 2, 4, 6, \ldots, 20$ (b) for values $I = 2, 3, \ldots, 8$

Figure 2. Mean adjusted relative suboptimality gap of heuristic MPI–OPT.

In our case $0 \leq \mathrm{arsg}(\pi) \leq 1$, and both limiting values can be achieved.

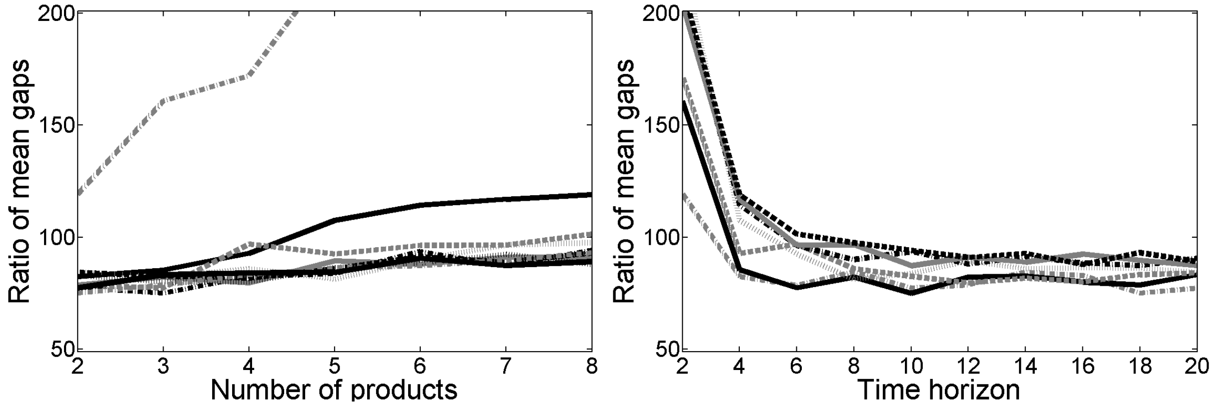
We further introduce a measure to be used to compare the mean performance of an alternative heuristic with respect to Heuristic MPI–OPT, as follows:

$$\mathrm{ratio}(\pi) = \frac{\mathrm{mean}(\mathrm{rsg}(\pi))}{\mathrm{mean}(\mathrm{rsg}(\text{MPI–OPT}))}. \tag{9}$$

This ratio captures the extent to which the mean absolute gap (i.e., the revenue loss) created by Heuristic MPI–OPT may be expected to be magnified if policy π is implemented instead. Thus, we have $\mathrm{ratio}(\pi) > 1$ if and only if policy π is on average worse than Heuristic MPI–OPT. An analogous ratio is used with the arsg measure.

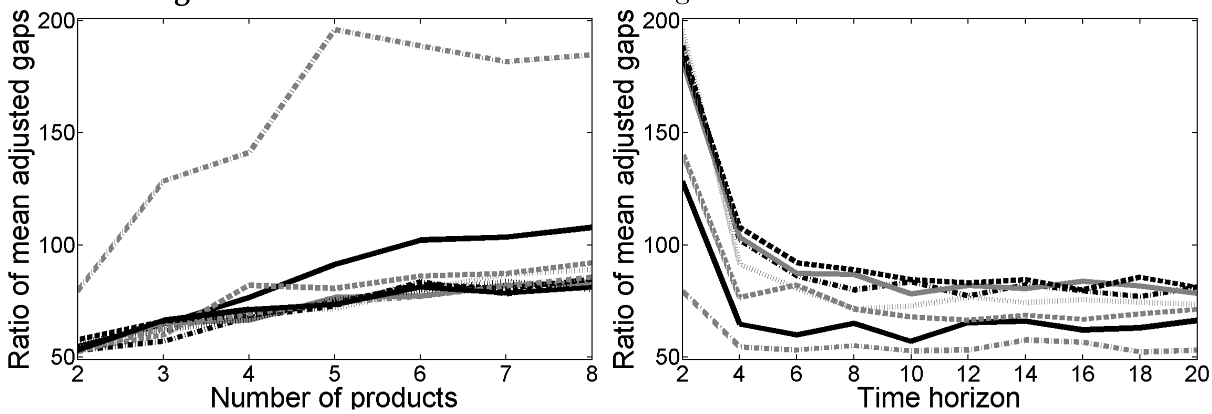
7.2 Results

Figure 1 exhibits two projections of the mean $\mathrm{rsg}(\text{MPI–OPT})$ as function of the number of products I and the time horizon T. The figure shows an excellent mean performance of heuristic MPI-OPT well below 0.01%, and further suggests that such a performance can be expected even



(a) for values $T = 2, 4, 6, \dots, 20$ (b) for values $I = 2, 3, \dots, 8$

Figure 3. Performance ratio in terms of rsg of EDF–GRE over MPI–OPT.



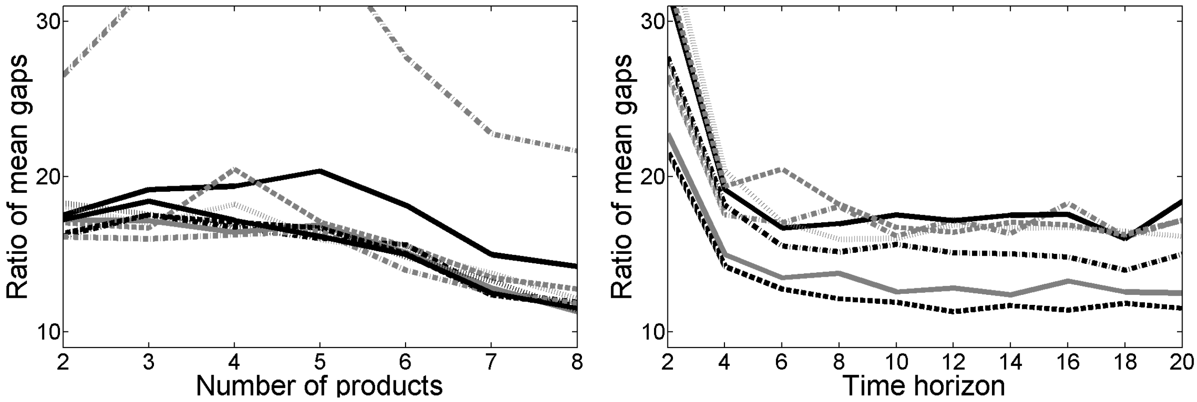
(a) for values $T = 2, 4, 6, \dots, 20$ (b) for values $I = 2, 3, \dots, 8$

Figure 4. Performance ratio in terms of arsg of EDF–GRE over MPI–OPT.

for higher values of I and T. These strong results are further confirmed in Figure 2 considering the arsg measure.

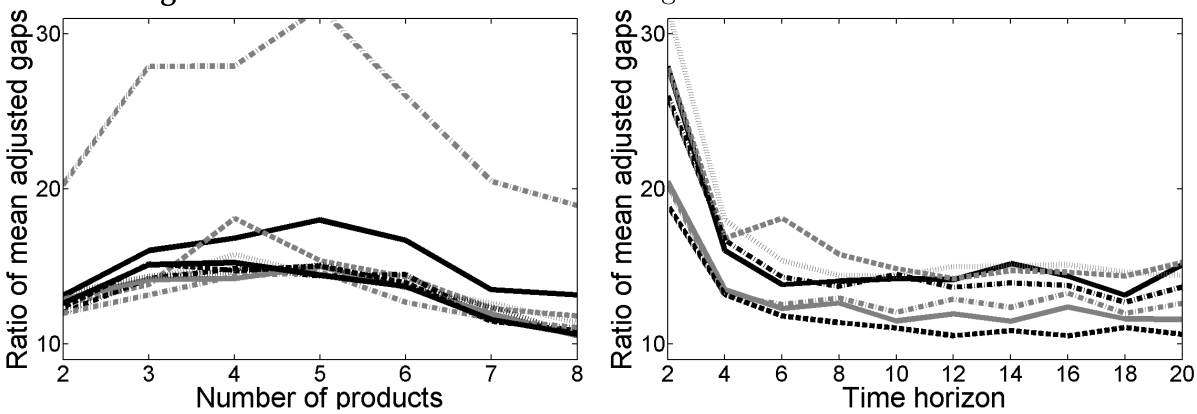
The ratio of the benchmark Heuristic EDF–GRE is presented in Figure 3 and Figure 4. The benchmark policy's mean gap is in all cases more than 50-times larger than that of Heuristic MPI–OPT, and the ratio grows with the number of items I. Further, in Figure 5 and Figure 6 we evaluate Heuristic MPI–GRE, whose mean performance is in all cases more than 10-times worse, though improving with higher I once this passes the value 5.

Finally, we remark that the worst-case performance achieved by the maximum rsg (arsg) values of Heuristic MPI–OPT are relatively small, ranging between 0.3% and 15% (4% and 14%). The maximum rsg (arsg) values of Heuristic MPI–GRE range between 1% and 8% (22% and 72%), that is, its worst-case performance is good in absolute terms, but is especially bad in the problems where promotion has small effect on total revenues. The worst-case performance of Heuristic EDF–GRE ranges between 3% and 8% (51% and 100%).



(a) for values $T = 2, 4, 6, \ldots, 20$ (b) for values $I = 2, 3, \ldots, 8$

Figure 5. Performance ratio in terms of rsg of MPI–GRE over MPI–OPT.



(a) for values $T = 2, 4, 6, \ldots, 20$ (b) for values $I = 2, 3, \ldots, 8$

Figure 6. Performance ratio in terms of arsg of MPI–GRE over MPI–OPT.

8 Conclusions

We have developed a dynamic and stochastic model of dynamic promotion and proposed a policy that has a natural economic interpretation and suggests itself to be easily implementable in practice. These advantages come at the cost of possible suboptimality of such a dynamic solution, which was, however, shown to be negligible and smaller than the cost of implementing a naïve marketing solution. The model has an appealing property of being extensible to a variety of ad-hoc requirements that managers or certain circumstances may impose.

A challenge showing itself is to extend the model presented in this paper to account for price changes, inventories with dependent demands and product assortment, and obtain an appealing index-based solution. The analysis of that problem is, however, more complex and the theoretical background must be extended in order to tackle such problems.

Our model offers a comprehensive modeling framework that may be used in other applications, since the items considered in knapsack problems are often perishable, either naturally or due to special restrictions. An application, for example, arises in surgery, when only a limited number of patients may be chosen to undertake an alternative treatment (e.g., a transplantation). Further, the task management problem, in which tasks have associated deadlines and one can work only on a subset of them at a time, also falls to the general KPPI setting.

References

Bertsimas, D. and Mersereau, A. J. (2007). A learning approach for interactive marketing to a customer segment. *Operations Research*, 55(6):1120–1135.

Buyukkoc, C., Varaiya, P., and Walrand, J. (1985). The $c\mu$ rule revisited. *Advances in Applied Probability*, 17(1):237–238.

Campo, K. and Gijsbrechts, E. (2005). Retail assortment, shelf and stockout management: Issues, interplay and future challenges. *Applied Stochastic Models in Business and Industry*, 21:383–392.

Caro, F. and Gallien, J. (2007). Dynamic assortment with demand learning for seasonal consumer goods. *Management Science*, 53(2):276–292.

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

Gittins, J. C. (1979). Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society, Series B*, 41(2):148–177.

Kök, A. G., Fisher, M., and Vaidyanathan, R. (2006). *Retail Supply Chain Management*, chapter Assortment Planning: Review of Literature and Industry Practice. Kluwer Publishers. (to appear).

Lu, L. L., Chiu, S. Y., and Cox, L. A. J. (1999). Optimal project selection: Stochastic knapsack with finite time horizon. *Journal of the Operational Research Society*, 50:645–650.

Manor, G. and Kress, M. (1997). Optimality of the greedy shooting strategy in the presence of incomplete damage information. *Naval Research Logistics*, 44:613–622.

Niño-Mora, J. (2002). Dynamic allocation indices for restless projects and queueing admission control: A polyhedral approach. *Mathematical Programming, Series A*, 93(3):361–413.

Niño-Mora, J. (2005). Marginal productivity index policies for the finite-horizon multiarmed bandit problem. In *Proceedings of the 44th IEEE Conference on Decision and Control and European Control Conference ECC 2005 (CDC-ECC '05)*, pages 1718–1722.

Niño-Mora, J. (2006). Restless bandit marginal productivity indices, diminishing returns, and optimal control of make-to-order/make-to-stock $M/G/1$ queues. *Mathematics of Operations Research*, 31(1):50–84.

Niño-Mora, J. (2007). Dynamic priority allocation via restless bandit marginal productivity indices. *TOP*, 15(2):161–198.

Papadimitriou, C. H. and Tsitsiklis, J. N. (1999). The complexity of optimal queueing network. *Mathematics of Operations Research*, 24(2):293–305.

Papastavrou, J. D., Rajagopalan, S., and Kleywegt, A. J. (1996). The dynamic and stochastic knapsack problem with deadlines. *Management Science*, 42(12):1706–1718.

Whittle, P. (1988). Restless bandits: Activity allocation in a changing world. *A Celebration of Applied Probability, J. Gani (Ed.), Journal of Applied Probability*, 25A:287–298.

A Work-Reward Analysis

In order to prove Proposition 1, we describe some crucial points from the restless bandit framework in more detail. For a survey on this methodology refer to Niño-Mora (2007). Note that any set $\mathscr{S} \subseteq \mathscr{T}$ can represent a stationary policy, by being active in all the states belonging to $\mathscr{S}$ and being passive in all the states belonging to $\mathscr{T} \setminus \mathscr{S}$. We will call such a policy an *$\mathscr{S}$-active policy*, and $\mathscr{S}$ an *active set*. Note also that we can restrict our attention to stationary deterministic policies, since it is well-known from the MDP theory that there exists an optimal policy which is stationary, deterministic, and independent of the initial state.

Let $\mathscr{S} \subseteq \mathscr{T}$ be an active set. We can reformulate (3) as

$$f_t^{\mathscr{S}} - \nu g_t^{\mathscr{S}} := \mathbb{E}_t^{\mathscr{S}} \left[\sum_{s=0}^{t-1} \beta^s \boldsymbol{P}_{t,t-s}^{s|\mathscr{S}} R_{t-s}^{I_{\mathscr{S}}(t-s)} \right] - \nu \mathbb{E}_t^{\mathscr{S}} \left[\sum_{s=0}^{t-1} \beta^s \boldsymbol{P}_{t,t-s}^{s|\mathscr{S}} W_{t-s}^{I_{\mathscr{S}}(t-s)} \right], \tag{10}$$

where $\boldsymbol{P}_{i,j}^{j-i|\mathscr{S}}$ is the probability of moving from state $i \in \mathscr{X}$ to state $j \in \mathscr{X}$ in exactly $j-i$ periods under policy $\mathscr{S}$ and $I_{\mathscr{S}}(s)$ is the indicator function $I_{\mathscr{S}}(s) = \begin{cases} 1, & \text{if } s \in \mathscr{S}, \\ 0, & \text{if } s \notin \mathscr{S}. \end{cases}$ We will call $f_t^{\mathscr{S}}$ the *expected total discounted revenue* under policy $\mathscr{S}$ if starting from state t, and we will write it in a more convenient way as

$$f_t^{\mathscr{S}} = \mathbb{E}_t^{\mathscr{S}} \left[\sum_{s=1}^{t} \beta^{t-s} \boldsymbol{P}_{t,s}^{t-s|\mathscr{S}} R_s^{I_{\mathscr{S}}(s)} \right]. \tag{11}$$

Similarly, we will call $g_t^{\mathscr{S}}$ the *expected total discounted promotion work* under policy $\mathscr{S}$ if starting from state t, and we will write it in a more convenient way as

$$g_t^{\mathscr{S}} = \mathbb{E}_t^{\mathscr{S}} \left[\sum_{s=1}^{t} \beta^{t-s} \boldsymbol{P}_{t,s}^{t-s|\mathscr{S}} W_s^{I_{\mathscr{S}}(s)} \right]. \tag{12}$$

Let, further, $\langle a, \mathscr{S} \rangle$ be the policy which takes action $a \in \mathscr{A}$ in the current time period and adopts an $\mathscr{S}$-active policy thereafter. For any state $t \in \mathscr{T}$ and an $\mathscr{S}$-active policy, the $(t, \mathscr{S})$-*marginal revenue* is defined as

$$r_t^{\mathscr{S}} := f_t^{\langle 1, \mathscr{S} \rangle} - f_t^{\langle 0, \mathscr{S} \rangle}, \tag{13}$$

and the $(t, \mathscr{S})$*-marginal promotion work* as

$$w_t^{\mathscr{S}} := g_t^{\langle 1, \mathscr{S} \rangle} - g_t^{\langle 0, \mathscr{S} \rangle}. \tag{14}$$

These marginal revenue and marginal promotion work capture the change in the expected total discounted revenue and promotion work, respectively, which results from being active instead of passive in the first time period and following the $\mathscr{S}$-active policy afterwards. Finally, if $w_t^{\mathscr{S}} \neq 0$, we define the $(t, \mathscr{S})$*-marginal productivity rate* as

$$\nu_t^{\mathscr{S}} := \frac{r_t^{\mathscr{S}}}{w_t^{\mathscr{S}}}. \tag{15}$$

```
set  $\widehat{\mathscr{S}_0} := \emptyset$;
for  $k := 0$  to  $T+1$  do
   choose  $\tau_{k+1} \in \{\tau \in \mathscr{T} \setminus \widehat{\mathscr{S}_k};\ \nu_\tau^{\widehat{\mathscr{S}_k}} \geq \nu_t^{\widehat{\mathscr{S}_k}}$  for all $t \in \mathscr{T} \setminus \widehat{\mathscr{S}_k}\}$;
   set  $\widehat{\mathscr{S}_{k+1}} := \widehat{\mathscr{S}_k} \cup \{\tau_{k+1}\}$;
   set  $\widehat{\nu}_{\tau_{k+1}} := \nu_{\tau_{k+1}}^{\widehat{\mathscr{S}_k}}$;
end {for};
```

Figure 7. Adaptive-greedy algorithm for calculation of MPIs.

In order to verify that a perishable item satisfies PCL-indexability (which implies existence of MPIs), we need to postulate a family $\mathscr{F}$ of optimal active sets. We define it as a family of nested sets $\mathscr{F} := \{\mathscr{S}_0, \mathscr{S}_1, \ldots, \mathscr{S}_T\}$, where $\mathscr{S}_k := \{1, 2, \ldots, k\}$. The following assumption must be verified.

Assumption 2 (Positive Marginal Works). *The marginal promotion work $w_t^{\mathscr{S}} > 0$ for all states $t \in \mathscr{T}$ under any $\mathscr{S}$-active policy from a feasible family of active sets $\mathscr{F} \subseteq 2^{\mathscr{T}}$.*

If Assumption 2 holds and the quantities $\widehat{\nu}_{\tau_{k+1}}$ computed in the *adaptive-greedy algorithm* presented in Figure 7 are nonincreasing in k, then the marginal productivity indices exist and equal $\nu_t^* := \widehat{\nu}_t$ and $\mathscr{F}$ contains an optimal active set for any ν.

B Proof of Proposition 1

We next show that Assumption 2 holds and derive a closed-form expression for the MPI given in Proposition 1. Plugging (11) and (12) into (13) and (14), respectively, we obtain two expressions that will be used in the following analysis:

$$r_t^{\mathscr{S}} = \left(R_t^1 - R_t^0\right) - (\beta q - \beta p) \sum_{s=1}^{t-1} \beta^{t-s-1} \boldsymbol{P}_{t-1,s}^{t-s-1|\mathscr{S}} R_s^{I_{\mathscr{S}}(s)}, \tag{16}$$

$$w_t^{\mathscr{S}} = \left(W_t^1 - W_t^0\right) - (\beta q - \beta p) \sum_{s=1}^{t-1} \beta^{t-s-1} \boldsymbol{P}_{t-1,s}^{t-s-1|\mathscr{S}} W_s^{I_{\mathscr{S}}(s)}. \tag{17}$$

It is well known from the MDP theory that the transition probability matrix for multiple periods is obtained by multiplication of transition probability matrices for subperiods. Hence, given an active set $\mathscr{S} \subseteq \mathscr{T}$, we have

$$\boldsymbol{P}^{t-s|\mathscr{S}} = \left(\boldsymbol{P}^{1|\mathscr{S}}\right)^{t-s}, \tag{18}$$

where the matrix $\boldsymbol{P}^{1|\mathscr{S}}$ is an $(T+1) \times (T+1)$-matrix constructed so that its row $x \in \mathscr{X}$ is the row x of the matrix $\boldsymbol{P}^{1|\mathscr{T}}$ if $x \in \mathscr{S}$, and is the row x of the matrix $\boldsymbol{P}^{1|\emptyset}$ otherwise. For definiteness, we remark that $\boldsymbol{P}^{0|\mathscr{S}}$ is an identity matrix.

Lemma 1. *Let $t \in \mathscr{T}$ and consider any integer $0 \leq k \leq T$. Then,*

$$r_t^{\mathscr{S}_k} = \begin{cases} R(q-p)\left[(1-\beta)\dfrac{1-(\beta p)^{t-1}}{1-\beta p}\right. \\ \left. +(1-\beta\alpha)(\beta p)^{t-1}\right], & \text{if } k \geq t-1 \geq 0, \\ R(q-p)\left[(1-\beta)\dfrac{1-(\beta q)^{t-k-1}}{1-\beta q}\right. \\ +(1-\beta)(\beta q)^{t-k-1}\dfrac{1-(\beta p)^k}{1-\beta p} \\ \left. +(1-\beta\alpha)(\beta q)^{t-k-1}(\beta p)^k\right], & \text{if } T-1 \geq t-1 \geq k. \end{cases} \tag{19}$$

$$w_t^{\mathscr{S}_k} = \begin{cases} W\left[1-(\beta q-\beta p)\dfrac{1-(\beta p)^{t-1}}{1-\beta p}\right], & \text{if } k \geq t-1 \geq 0, \\ W\left[1-(\beta q-\beta p)(\beta q)^{t-k-1}\dfrac{1-(\beta p)^k}{1-\beta p}\right], & \text{if } T-1 \geq t-1 \geq k. \end{cases} \tag{20}$$

Proof. Under an active set $\mathscr{S}_k$, from (18) we get for $T \geq t-1 \geq s \geq 1$,

$$\boldsymbol{P}_{t-1,s}^{t-s-1|\mathscr{S}_k} = \begin{cases} p^{t-s-1}, & \text{if } k \geq t-1 \geq s \geq 0, \\ q^{t-s-1}, & \text{if } T \geq t-1 \geq s \geq k, \\ q^{t-k-1}p^{k-s}, & \text{if } T \geq t-1 \geq k \geq s \geq 0, \end{cases}$$

These expressions together with the definitions of R_t^a and W_t^a plugged into (16)–(17) after simplifying conclude the proof. □

Lemma 2. *For any integer $k \geq 0$ we have*

(i) $w_k^{\mathscr{S}_k} > 0$;
(ii) $w_t^{\mathscr{S}_k} > 0$ *for all* $t \in \mathscr{T}$.

Proof. Denote by

$$h(k) := (\beta q - \beta p)\frac{1-(\beta p)^k}{1-\beta p}, \tag{21}$$

so that, using (20), $w_k^{\mathscr{S}_k} = W\,[1-h(k)]$.

(i) For $k = 0$, we have $h(0) = 0$ by definition. For $k \geq 1$, $h(k) = (1-(\beta p)^k)\frac{\beta q - \beta p}{1-\beta p} < 1$. □
(ii) Implied by (i) and (20). □

Lemma 3. *In each step $k = 0, 1, \ldots, T-1$ of the adaptive-greedy algorithm presented in Figure 7, Assumption 2 is satisfied and the algorithm sets*

$$\tau_{k+1} = k+1; \qquad \widehat{\mathscr{S}}_{k+1} = \{1, \ldots, k+1\}; \qquad \widehat{\nu}_{k+1} = \frac{R(q-p)\left[(1-\beta)\dfrac{1-(\beta p)^k}{1-\beta p} + (1-\beta\alpha)(\beta p)^k\right]}{W\left[1-(\beta q-\beta p)\dfrac{1-(\beta p)^k}{1-\beta p}\right]}$$

Proof. Consider the step k. Having $\widehat{\mathscr{S}_k} = \{1, 2, \ldots, k\}$, for $t \in \mathscr{T} \setminus \widehat{\mathscr{S}_k}$, by Lemma 1 we have

$$r_t^{\widehat{\mathscr{S}_k}} = R(q-p)\left[(1-\beta)\frac{1-(\beta q)^{t-k-1}}{1-\beta q} + (1-\beta)(\beta q)^{t-k-1}\frac{1-(\beta p)^k}{1-\beta p} + (1-\beta\alpha)(\beta q)^{t-k-1}(\beta p)^k\right],$$

$$w_t^{\widehat{\mathscr{S}_k}} = W\left[1-(\beta q - \beta p)(\beta q)^{t-k-1}\frac{1-(\beta p)^k}{1-\beta p}\right] > 0,$$

where the positivity is due to Lemma 2 (which also holds for $t \in \widehat{\mathscr{S}_k}$), implying that Assumption 2 is satisfied. Furthermore, then $r_t^{\widehat{\mathscr{S}_k}}$ is nondecreasing and $w_t^{\widehat{\mathscr{S}_k}}$ is nonincreasing as t diminishes (i.e. as t gets closer to the deadline). Hence the maximum $\nu_t^{\widehat{\mathscr{S}_k}}$ is attained at $t = k+1$ and the algorithm sets what is stated. □

Lemma 3 is a crucial result in the proof of Proposition 1. It verifies that the family $\mathscr{F}$ required in Assumption 2 is the family of nested active sets $\mathscr{F} = \{\mathscr{S}_0, \mathscr{S}_1, \ldots, \mathscr{S}_T\}$. Finally, Proposition 2(iii) assures that the algorithm's output $\widehat{\nu}_t$ is nondecreasing as t diminishes. This concludes the proof of Proposition 1. □

C Proof of Proposition 2

(i) Immediate from (4). □

(ii) Formally, we are to prove the following statement: If the probability q is replaced by $q' \leq q$, then $\nu_t^{*\prime} \leq \nu_t^*$ for any $t \in \mathscr{T}$. It is straightforward to see that (4) is nondecreasing in q. □

(iii) In order to see that the marginal productivity index is nondecreasing as t diminishes, it is enough to compare the MPIs for $t \in \mathscr{T} \setminus \{T\}$ and $t+1$. Niño-Mora (2007, p. 172) showed that under positive marginal works, $\nu_t^* \geq \nu_{t+1}^*$ is equivalent to $\nu_t^{\mathscr{S}_{t+1}} \geq \nu_{t+1}^{\mathscr{S}_{t+1}}$, which is satisfied as shown in the proof of Lemma 3. □

D Proof of Proposition 3

In order the set $\{\tau \in \mathscr{T} : \nu_t^* > \nu \text{ for all } t \in \mathscr{T} \text{ such that } t \geq \tau\}$ to be nonempty, due to Proposition 2(iii) we need to have $\nu_1^* > \nu$, that is, $\frac{R}{W}(1-\beta\alpha)(q-p) > \nu$. □

E Proof of Proposition 4

(i) The MPI for $\beta = 1$ is given by the limit of the discounted MPI (4), if it exists. The limit exists and is equal to the stated expression. □

(ii) Straightforward from (4) after setting $q := 1$ and $\beta := 1$. □

F Proof of Proposition 5

Under the above assumptions, all the products perish within one period, and promoting is equivalent to avoiding the deadline cost. The problem thus reduces to a combinatory problem of choosing a subset of items not to be promoted that minimizes the aggregate cost of not promoted items while the remaining items do not occupy more than W. Since the aggregate cost of all items is constant, this problem is equivalent to choosing a subset of items to be promoted that maximizes the aggregate cost of promoted items while their aggregate volume is not greater than W, which is the knapsack problem. □