OPERATIONS RESEARCH
Vol. 62, No. 3, May–June 2014, pp. 696–711
ISSN 0030-364X (print) | ISSN 1526-5463 (online)

http://dx.doi.org/10.1287/opre.2014.1272
© 2014 INFORMS

Generalized Restless Bandits and the Knapsack Problem for Perishable Inventories

Darina Graczová
Department of Applied Mathematics and Statistics, Comenius University, 842 48 Bratislava, Slovakia, darina.graczova@gmail.com

Peter Jacko
Department of Management Science, Lancaster University Management School, Lancaster, LA1 4YX, United Kingdom;
and BCAM Basque Center for Applied Mathematics, 48009 Bilbao, Spain, p.jacko@lancaster.ac.uk

In this paper we introduce the *knapsack problem for perishable inventories* concerning the optimal dynamic allocation of a collection of products to a limited knapsack. The motivation for designing such a problem comes from retail revenue management, where different products often have an associated lifetime during which they can only be sold, and the managers can regularly select some products to be allocated to a limited promotion space that is expected to attract more customers than the standard shelves. Another motivation comes from scheduling of requests in modern multiserver data centers so that quality-of-service requirements given by completion deadlines are satisfied. Using the Lagrangian approach we derive an optimal index policy for the *Whittle relaxation* of the problem in which the knapsack capacity is used only on average. Assuming a certain structure of the optimal policy for the single-inventory control, we prove indexability and derive an efficient, linear-time algorithm for computing the index values. To the best of our knowledge, our paper is the first to provide indexability analysis of a restless bandit with bi-dimensional state (lifetime and inventory level). We illustrate that these index values are numerically close to the true index values when such a structure is not present. We test two index-based heuristics for the original, nonrelaxed problem: (1) a conventional *index rule*, which prescribes to order the products according to their current index values and promotes as many products as fit in the knapsack, and (2) a recently proposed *index-knapsack heuristic*, which employs the index values as a proxy for the price of promotion and proposes to solve a deterministic knapsack problem to select the products. By a systematic computational study we show that the performance of both heuristics is nearly optimal, and that the index-knapsack heuristic outperforms the conventional index rule.

Subject classifications: Markov decision processes; resource allocation; knapsack problem; index policies; Whittle index; Lagrangian relaxation.
Area of review: Stochastic Models.
History: Received September 2011; revisions received October 2012, August 2013; accepted December 2013. Published online in *Articles in Advance* April 28, 2014.

1. Introduction

In this paper we introduce the *knapsack problem for perishable inventories* concerning the optimal dynamic allocation of a collection of products to a limited knapsack. This is an extension of the classic knapsack problem to a multiperiod dynamic setting, in which units are allowed to randomly disappear, so that the knapsack capacity can be reallocated in every period. Such problems of stochastic dynamic resource allocation arise in different fields (see, e.g., Jacko 2013). For instance, modern data centers are composed of a large number of servers (or virtual machines), which must be allocated to a given set of requests to be scheduled so that the quality-of-service contracts be satisfied (Yang et al. 2011, Dance and Gaivoronski 2012, Glazebrook et al. 2011). Requests coming from different users are heterogeneous and may be given as a number of subrequests (an "inventory") that must be completed sequentially before a specified deadline. In the following we describe in more detail a similar problem in the implementation of product promotion for optimal revenue management in the retail industry. For concreteness, we shall focus on this framework throughout the paper.

1.1. Motivation

The main interest of retail companies is net revenue maximization. Managers are facing the problem of "proper" choice of products to sell or proper setting of prices of sales in order to obtain the maximal revenues. The assortment in many branches is changing very fast and the products can become "not topical" like seasonal products or can get "obsolete" or "perished" by a deadline after which they cannot be sold anymore and cause a cost (e.g., lost profit, product removal costs, penalty, etc.).

The retailers are therefore dealing with a trade-off between obtaining revenues from selling the products before their deadlines and incurring deadline costs from perished products. To maximize the expected revenue, the managers apply different promotion techniques. One way is to use price discounts to attract customers and incentivize them to buy. Of course, the price should not be permanently lower than the marginal costs since it would lead to negative net profit, so it must be increased once the demand accelerates. Such *dynamic pricing*, however, often cannot be implemented because of clauses in the contracts with suppliers or for brand image strategic reasons. Although research studies lead to mixed conclusions, retailers may also consider a possible negative effect of ubiquitous dynamic pricing on customer satisfaction and perception of price fairness. (See Talluri and van Ryzin 2004, for an overview of theory and practice of dynamic pricing.)

An alternative way is to advertise some of the products without price changes, e.g., announce them on large posters in front of the shop or to allocate them to a *promotion space* close to cashiers or at display shelves, where they can be more easily seen and bought. In this paper we introduce a dynamic and stochastic revenue management model, in which the prices do not change, but the dynamic allocation of products to a limited promotion space is used to increase the revenues. In the static (one-period) case this would correspond to the *knapsack problem* (Dantzig 1957), which is believed to be one of the easier NP-hard problems (Pisinger 2005). The principal possible implementation of our revenue management model is in grocery stores and supermarkets for product promotion, also known as display shelves assortment. As a special case we cover a variant of the dynamic assortment problem, whose practical importance in addition to dynamic pricing has been shown recently by Bernstein et al. (2011).

As another example of practical interest, consider an assortment problem in a car shop offering cars of the same brand or similar quality levels (so that brand effect is negligible), where only a few cars can be allocated in the showroom. Customers can see and even test the exhibits and can decide either to buy one of them, or to buy another type (which is not exhibited, but is offered), or, of course, not to buy any car. Moreover, there are more units of every type in the stock, which are for sale as well. Similar are shops with furniture, where one can allocate entire "living-rooms," "kitchens," "bathrooms," etc. Customers can choose from the offer of the product that is either exhibited or offered in a shop catalog. The assortment question is Which product types shall we exhibit in order to maximize the revenue?

Related problems of shelf-space allocation for perishing inventories have been investigated in the literature motivated by its practical relevance, and, typically, numerical solution techniques were proposed because of the problem complexity. Kar et al. (2001) and Bai and Kendall (2008) considered a *single-period* problem of deteriorating inventory and shelf-space allocation, which was optimally solved using the generalized reduced gradient algorithm. Approximate solutions such as greedy heuristics and metaheuristics were used for similar problems by, e.g., Urban (1998), Bai et al. (2008). Most of the literature considers *random lifetimes* (see Goyal and Giri 2001, for a survey), which allows for analytical approaches especially when the decay is exponential. We believe that because of the current regulation of food safety and because of the existing outlet mechanisms in other industries, it is more realistic to consider deterministic lifetimes. Further, we focus on addressing the problem over multiple periods of stochastic demand.

1.2. Modeling Approach

In this paper we present a model that we call the *knapsack problem for perishable inventories* (KPPI). We consider a collection of products, with a nonempty inventory, each of which may be perishable by a finite deadline, or nonperishable. In each discrete time period, a decision maker must select some of the products and decide how many units of the selected products should be allocated to a knapsack with limited volume.

We formulate the problem in the framework of Markov decision processes (MDP) with a sample-path knapsack capacity constraint. Because of this constraint, the dynamic programming approach does not render an optimal solution analytically, and numerically this approach is intractable (*curse of dimensionality*). Because of such intractability, we aim to obtain a well-grounded nearly optimal solution of the problem. We therefore employ the Whittle relaxation (Whittle 1988), which is to relax the family of sample-path constraints by a single one of allocating the knapsack capacity over the planning horizon only in expectation.

Such a relaxed problem can then be solved optimally by Lagrangian methods and the optimal solution to the Whittle relaxation is, under certain conditions, an *index policy*. For our problem it means that one can attach to each product an *index*, which is a function of its current inventory level and remaining lifetime, and then there is a particular value of the Lagrangian multiplier such that it is optimal for the Whittle relaxation to select all the products whose current index value is larger than the value of the Lagrangian multiplier. As usual for the Lagrangian multiplier, index values have an economic interpretation of a *fair charge*, measuring the efficiency, or the marginal rate of resource usage. Because of the assumption of mutually independent product demands, the index values can be obtained by solving optimally a parametric single-product subproblem independently of other products.

If the action space is binary, such problems are known as *restless bandits* (Whittle 1988, Niño-Mora 2007), and existence of such index values (so-called *indexability* property) must be established. For restless bandits, Whittle (1988) showed that this allows to define a heuristic for the original, nonrelaxed problem, called *index rule*, which prescribes to order the products according to their current index values and selects products until the resource capacity is filled. Such a heuristic has been reported to have an exceptional performance in a variety of problems (Niño-Mora 2007), it was proved optimal under a *frozen-if-not-allocated* assumption for an infinite horizon problem with a single-capacity resource in the celebrated work by Gittins (1979), and asymptotically optimal under certain technical assumptions in the time-average case as both the number of products and the resource capacity increase (Weber and Weiss 1990).

The challenge in our problem is then to study whether and how such results extend to the case of finite horizon (due to perishability), bi-dimensional state space (both inventory level and remaining lifetime are taken into account), and nonhomogeneous capacity utilization (due to different product volumes).

1.3. Paper Structure and Contributions

In §2 we propose an MDP model of perishable product with inventory. We further formulate the KPPI and then relax and decompose the problem into parameterized single-product subproblems. Section 3 is devoted to the study of indexability of the problem, where, in §3.4, we prove indexability and give a linear-time recursive characterization of the index values under certain optimal policy structures. The index values are obtained as a function of the product's deadline, inventory, profit margin, expected salvage value, product's volume, and selling probabilities with and without promotion. We further implement a general algorithm for testing indexability and for index computation, and we conjecture based on our experiments that the subproblem is indexable. Moreover, our experiments suggest that the index values are numerically close to the ones computed recursively.

Section 4 focuses on the solution to the original (nonrelaxed) KPPI. In §4.1 we introduce the dynamic programming formulation of KPPI to be used to compute an optimal policy in tractable cases, which are found to be those with no more than five products and planning horizon of no more than 32 periods. Employing the index values obtained in this paper, we obtain in §4.2 the index rule and design a new *index-knapsack heuristic*, which extends the one proposed in Jacko (2013). This heuristic solves at every period a knapsack problem, in which the index values are used as a proxy for item prices, exploiting thus their economic interpretation as the marginal rate of revenue from promotion.

The performance of the two heuristics is evaluated in a systematic computational study in §5, which suggests their near optimality, and further shows that the index-knapsack heuristic (with suboptimality below 0.7%) outperforms the classic index rule (with suboptimality below 3%) in all tested instances. Section 6 concludes and the proofs are deferred to the e-companion (available as supplemental material at http://dx.doi.org/10.1287/opre.2014.1272).

From the methodological point of view, the results of this paper contribute to several lines of research on the frontiers of operations research and applied probability:

1. We significantly expand the modeling framework of multiarmed bandits (Gittins 1979) and restless bandits (Whittle 1988) to more general *discretely divisible resource* allocation problems, and we design a new index-knapsack heuristic for such problems with an outstanding performance. We further generalize the model of Glazebrook and Minty (2009), who studied nonrestless bandits with general resource requirements and proposed a generalization of the Gittins (1979) solution.

2. We characterize a new family of restless bandits that are indexable. This is to the best of our knowledge the first restless bandit model with a general bi-dimensional state, for which indexability is analyzed; we are only aware of the previous work in Jacko and Niño-Mora (2008), in which model one of the state dimensions was only binary.

3. We develop a linear-time algorithm for computing the index values in the above-mentioned case, to be contrasted with the general state-of-the-art algorithm for restless bandits that take cubic time, i.e., two orders of magnitude (both in lifetime and inventory level) higher complexity (Niño-Mora 2007).

4. As a special case (when the product volume is unity and the probability of selling the product is zero if not selected for the knapsack), we obtain a linear-time characterization of both the finite-horizon and infinite-horizon variants of the Gittins index (Gittins 1979) for this model, which is two orders of magnitude (both in lifetime and inventory level) faster than the general state-of-the-art algorithm, which takes cubic time (Niño-Mora 2011).

2. Modeling of Knapsack Problem for Perishable Inventories

In this section we formally describe the problem and formulate it in MDP framework.

2.1. Problem

Consider a retailer that has available $I \geqslant 1$ perishable products, labeled by $i \in \mathscr{I}$, with inventories of $K_i \geqslant 1$ units of each product i. At every time period $s \in \mathscr{H} := \{0, 1, \ldots, H-1\}$, where $1 \leqslant H \leqslant \infty$ is the planning horizon, the retailer can decide what products to promote. Not all the products can be promoted at the same time because the promotion space (knapsack) has limited volume $C > 0$. Every unit of product i has volume $0 < W_i \leqslant C$ and we will typically have $\sum_i K_i W_i \gg C$. The retailer is dealing with a problem of dynamic allocation of a subset of $\sum_i K_i$ product units to a knapsack at every time period s.

Each product can only be sold during its lifetime, which consists of periods $0, 1, \ldots, T_i$, where $T_i \in [1, H]$ is the deadline for all the units of product i at which they perish. The product units can be sold until the end of period $T_i - 1$, when they are removed as perished and cannot be sold anymore. If a unit of product i is sold, it yields an expected revenue (profit margin) $R_i > 0$ at that period. Otherwise, a salvage value (revenue) is received in period T_i, whose expected value is denoted by $\alpha_i R_i$ for some (possibly negative) coefficient $\alpha_i \leqslant 1$. Revenues and the salvage value are discounted over time with factor $0 \leqslant \beta \leqslant 1$.

2.2. MDP Model of Perishable Inventory with Bernoulli Demand

For transparency, in the following we assume that only a single unit of each product can be demanded by customers, which is formalized by Bernoulli arrivals. We believe that this assumption is almost without loss of generality, since the length of the period can be taken arbitrarily small. As an important consequence we have that it is enough to promote at most one unit of each product in one time period, which keeps the problem analytically tractable.

If product i is not promoted in a given period, then a unit of this product is sold with probability $1 - q_i$ per period. If (a unit of) product i is promoted (selected for the knapsack) in a given period, then the probability of selling is increased to $1 - p_i$ (with $0 < p_i < q_i \leqslant 1$). In other words, p_i is the probability that no unit of product i is sold when promoted in a period, and q_i is the probability that no unit of product i is sold when not promoted in a period; and the practically more relevant products with $q_i > p_i$ are interpreted as having *positive promotion power*.

Inventory of K_i units of product i perishable at deadline T_i is defined independently of other products as the tuple $(\mathcal{N}_i, (\mathbf{W}_i^a)_{a\in\mathcal{A}}, (\mathbf{R}_i^a)_{a\in\mathcal{A}}, (\mathbf{P}_i^a)_{a\in\mathcal{A}})$, where

- the state space is $\mathcal{N}_i := (\mathcal{T}_i \times \mathcal{K}_i) \cup \{0\}$, where $\mathcal{T}_i := \{1, \ldots, T_i\}$, $\mathcal{K}_i := \{1, \ldots, K_i\}$ and $\mathcal{T}_i \times \mathcal{K}_i$ is Cartesian product of $\mathcal{T}_i$ and $\mathcal{K}_i$; $t \in \mathcal{T}_i$ represents the number of remaining periods before the deadline, and $k \in \mathcal{K}_i$ represents the remaining inventory (the number of remaining units) of product i;
- the action space for states in $\mathcal{T}_i \times \mathcal{K}_i$ is $\mathcal{A} := \{0, 1\}$. Action 1 means to promote a unit of the product and action 0 means not to promote;
- the expected one-period capacity occupation (volume) $W_{i,n}^a$ in state n under action a is as follows. For any state $n = (t, k) \in \mathcal{T}_i \times \mathcal{K}_i$,

$$W_{i,(t,k)}^1 := W_i, \quad W_{i,(t,k)}^0 := 0, \quad W_{i,0}^1 = W_{i,0}^0 := 0;$$

- the expected one-period revenue $R_{i,n}^a$ in state n under action a is as follows. For any state (t, k), where $k \in \mathcal{K}_i$ and $t \in \mathcal{T}_i \backslash \{1\}$,

$$\begin{aligned}
&R_{i,(t,k)}^1 := R_i(1 - p_i),\\
&R_{i,(1,k)}^1 := R_i(1 - p_i) + \beta\alpha_i R_i(p_i + k - 1), \quad R_{i,0}^1 := 0,\\
&R_{i,(t,k)}^0 := R_i(1 - q_i),\\
&R_{i,(1,k)}^0 := R_i(1 - q_i) + \beta\alpha_i R_i(q_i + k - 1), \quad R_{i,0}^0 := 0;
\end{aligned}$$

- the one-period transition probability matrix $\mathbf{P}_i^{1|\mathcal{N}_i}$ under promoting for $K_i = 2$ is

$$\begin{array}{c|c|cccc|cccc}
 & 0 & (1,1) & \ldots & (T_i-1,1) & (T_i,1) & (1,2) & \ldots & (T_i-1,2) & (T_i,2)\\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
\hline
(1,1) & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
(2,1) & 1-p_i & p_i & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
\vdots & \vdots & 0 & \ddots & 0 & 0 & 0 & 0 & 0 & 0\\
(T_i,1) & 1-p_i & 0 & 0 & p_i & 0 & 0 & 0 & 0 & 0\\
\hline
(1,2) & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
(2,2) & 0 & 1-p_i & 0 & 0 & 0 & p_i & 0 & 0 & 0\\
\vdots & 0 & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 & 0\\
(T_i,2) & 0 & 0 & 0 & 1-p_i & 0 & 0 & 0 & p_i & 0
\end{array}.$$

The matrix for larger K_i is obtained analogously using the same matrix blocks. The one-period transition matrix $\mathbf{P}_i^{0|\mathcal{N}_i}$ can be obtained from $\mathbf{P}_i^{1|\mathcal{N}_i}$ by replacing p_i by q_i.

To capture the dynamics of product i, we consider the state process $X_i(\cdot)$, where the state $X_i(s) \in \mathcal{N}_i$ at the beginning of period $s \in \mathcal{H}$ is $X_i(s) = (T_i - s, k)$, if $s < T_i$ (i.e., the product has not perished yet) and $0 < k \leqslant K_i$ units of the product still remain in the stock, and the state $X_i(s) = 0$ otherwise (i.e., if all the units of product i have been sold ($k = 0$) or perished ($s \geqslant T_i$)).

2.3. Formulation of KPPI

Consider now the collection of I products with inventories. Let $\Pi_{\mathbf{X},\mathbf{a}}$ be the space of randomized and nonanticipative policies depending on the joint state-process $\mathbf{X}(\cdot) := (X_i(\cdot))_{i\in\mathcal{I}}$ and deciding the joint action-process $\mathbf{a}(\cdot) := (a_i(\cdot))_{i\in\mathcal{I}}$. Let $\mathbb{E}_0^{\boldsymbol{\pi}}$ denote the expectation over the state-process $\mathbf{X}(\cdot)$ and over action process $\mathbf{a}(\cdot)$, conditioned on the initial joint state $\mathbf{X}(0) = (T_i, K_i)_{i\in\mathcal{I}}$ and on policy $\boldsymbol{\pi} \in \Pi_{\mathbf{X},\mathbf{a}}$.

For any discount factor β, the KPPI problem is to find a joint policy $\boldsymbol{\pi}$ maximizing the expected β-discounted aggregate revenue starting from the initial period 0 subject to the family of sample path knapsack capacity allocation constraints, i.e.,

$$\begin{aligned}
&\max_{\boldsymbol{\pi}\in\Pi_{\mathbf{X},\mathbf{a}}} \quad \mathbb{E}_0^{\boldsymbol{\pi}}\left[\sum_{i\in\mathcal{I}}\sum_{s=0}^{H} \beta^s R_{i,X_i(s)}^{a_i(s)}\right]\\
&\text{subject to} \quad \sum_{i\in\mathcal{I}} W_{i,X_i(s)}^{a_i(s)} \leqslant C \qquad \text{(KPPI)}\\
&\qquad\qquad\qquad \text{at each time period } s \in \mathcal{H}.
\end{aligned}$$

2.4. Relaxations and Decomposition

An exact dynamic programming formulation of the problem is possible (see §4.1), however, its analytical solution is inaccessible because of the sample path constraint. Even the numerical resolution of this problem via dynamic programming becomes quickly intractable because of the curse of dimensionality. Therefore, we analyze the Whittle relaxation of the problem via the Lagrangian approach (Whittle 1988). Because of the mutual independence of the product demand, this approach essentially allows to decompose the problem into tractable single-product parametric subproblems. This approach has been developed in the literature following Whittle (1988); see the general setting in Jacko (2009) for more details.

Thus, the product-i parametric optimization subproblem of *expected total discounted net revenue* is

$$\max_{\pi \in \Pi_{X_i, a_i}} \left\{ \mathbb{E}_0^{\pi}\left[\sum_{s=0}^{H} \beta^s R_{X_i(s)}^{a_i(s)}\right] - \nu\, \mathbb{E}_0^{\pi}\left[\sum_{s=0}^{H} \beta^s W_{X_i(s)}^{a_i(s)}\right]\right\}, \tag{1}$$

where, as before, $X_i(s)$ is the state of product i at time period s and $a_i(s) \in \{0, 1\}$ is action applied in time period s. However, the optimization is now over all policies $\pi \in \Pi_{X_i, a_i}$, where Π_{X_i, a_i} is the space of randomized and nonanticipative policies depending on the product-i state-process $X_i(\cdot)$ and deciding the product-i action-process $a_i(\cdot)$. The real-valued parameter ν that appears in the formulation is the Lagrangian multiplier.

3. Optimal Dynamic Promotion of Perishable Product with Inventory

In this section, we identify an optimal solution to the problem of promoting a single perishable product when one must pay for promoting. We interpret ν as a *cost of promoting* that must be paid for each space unit occupied in every period in which the product is promoted. Since we are now considering each product i in isolation, we drop the product's subscript i.

We focus on stationary deterministic policies, since it is known from the MDP theory that there exists an optimal policy that is stationary, deterministic, and independent of the initial state (Puterman 2005, Chap. 6). Let $\mathscr{S} \subseteq \mathscr{T} \times \mathscr{K}$ be an $\mathscr{S}$-active set representing a stationary policy, such that all states with active action (action 1) belong to $\mathscr{S}$ and states with passive action (action 0) to set $(\mathscr{T} \times \mathscr{K}) \setminus \mathscr{S}$. Without loss of generality, action 0 is taken at state 0 (notice that both actions have the same one-period consequence, therefore the two actions are equivalent).

We define the *expected total discounted revenue* under policy $\mathscr{S}$ if starting from state (t, k) as

$$\mathbb{R}_{(t,k)}^{\mathscr{S}} := \mathbb{E}_{(t,k)}^{\mathscr{S}}\left[\sum_{s=0}^{H} \beta^s R_{X_i(s)}^{a_i(s)}\right], \tag{2}$$

and the *expected total discounted promotion work* under policy $\mathscr{S}$ if starting from state (t, k) as

$$\mathbb{W}_{(t,k)}^{\mathscr{S}} := \mathbb{E}_{(t,k)}^{\mathscr{S}}\left[\sum_{s=0}^{H} \beta^s W_{X_i(s)}^{a_i(s)}\right]. \tag{3}$$

Let $\mathscr{F} := 2^{\mathscr{T} \times \mathscr{K}}$ be the family of all stationary policies. Then (1) can be rewritten as the following optimization problem

$$\mathbb{V}_{(T,K)}^{*} := \max_{\mathscr{S} \in \mathscr{F}} \left\{\mathbb{R}_{(T,K)}^{\mathscr{S}} - \nu \mathbb{W}_{(T,K)}^{\mathscr{S}}\right\}, \tag{4}$$

where $\mathbb{V}_{(t,k)}^{*}$ is called the *value function*.

3.1. General Properties of Optimal Solution

We first approach the problem by studying the Bellman equation, which allows to characterize properties of the value function and the optimal solution stated in the following two theorems. The proofs are provided in the electronic companion sections EC.1 and EC.2 (available as supplemental material at opre.2014.1272), respectively.

Theorem 1 (Value Function Properties). *The value function is independent of T and K and it satisfies the following*:

(i) $\mathbb{V}_0^* = 0$;

(ii) *if* $\alpha \geqslant 0$, *then* $\mathbb{V}_{(t,k)}^* \geqslant 0$ *and* $\mathbb{V}_{(t,k)}^*$ *is nondecreasing in k for all t*;

(iii) *if* $\mathbb{V}_{(s,1)}^* \leqslant 0$ *for some s (for which necessarily $\alpha \leqslant -(1-q)/\beta q$), then* $\mathbb{V}_{(t,k)}^*$ *is nonincreasing in* $k \geqslant \max\{1, t-s+1\}$ *for all t*;

(iv) *if* $\nu \geqslant 0$, *then* $\mathbb{V}_{(t,1)}^* \leqslant R$ *and* $\mathbb{V}_{(t,k)}^* - \mathbb{V}_{(t,k-1)}^*$ *is nonincreasing in k and all t*;

(v) *if* $\alpha \leqslant (1-q)/(1-\beta q)$, *then* $\mathbb{V}_{(t,1)}^*$ *is nondecreasing in t*;

(vi) *if* $\beta = 1$ *or* $\alpha \leqslant 0$, *then* $\mathbb{V}_{(t,k)}^*$ *is nondecreasing in t for all k*;

(vii) $\mathbb{V}_{(t,k+1)}^* - \mathbb{V}_{(t,k)}^* = \beta^t \alpha R$ *for all* $k \geqslant t$.

Theorem 2 (Optimal Solution Properties). *The optimal solution is independent of T and K and it satisfies*:

(i) *for* $\nu \geqslant 0$: *if it is optimal to promote in state* (t, k), *then it is optimal to promote in state* $(t, k+1)$;

(ii) *for* $\nu \geqslant R(q-p)(1-\beta)/(W(1-\beta q))$ *or for* $\alpha \leqslant (1-q)/(1-\beta q)$: *if it is optimal to promote in state* $(t, 1)$, *then it is optimal to promote in state* $(t-1, 1)$;

(iii) *for* $\beta = 1$ *and* $\nu \geqslant 0$ *or for* $\nu \geqslant (R/W)(q-p)$: *if it is optimal to promote in state* (t, k), *then it is optimal to promote in state* $(t-1, k)$;

(iv) *for* $\beta = 1$ *and* $\nu \geqslant 0$ *or for* $0 \leqslant \nu \leqslant (R/W)(q-p)$: *if it is optimal to promote in state* $(t-1, k-1)$, *then it is optimal to promote in state* (t, k);

(v) *if* $\nu \leqslant (R/W)(q-p)(1-\beta^t\alpha)$, *then it is optimal to promote in any state* (k, t) *with* $k \geqslant t$ *and if* $\nu \geqslant (R/W)(q-p)(1-\beta^t\alpha)$, *then it is optimal to not to promote in any state* (k, t) *with* $k \geqslant t$.

It can also be observed directly from the Bellman equations that the structure of the optimal policy (i.e., the family of active sets $\mathscr{S}$ as ν varies) does not depend on R, W, and the value function depends on R, W only via their ratio R/W. The claims of the two theorems provide an idea about the complexity of the problem and diversity of the structure of its optimal solution, strongly depending on whether $\beta = 1$ or $\beta < 1$, the sign of α and the value of ν.

3.2. Solvability by Index Policies

To better identify the optimal solution, we will be interested in characterizing an optimal policy in terms of index values, which indicate if the perishable product is worth promoting.

Figure 1. A scheme of a general algorithm for computing index values.

```
{Input $R, W, \alpha, K, T, \mathbf{P}^{0|\mathcal{N}}, \mathbf{P}^{1|\mathcal{N}}, \beta$}
$\mathcal{S} := \mathcal{T} \times \mathcal{K}$;
while $\mathcal{S} \neq \varnothing$
   pick $n \in \arg\max\{\nu_n^{\mathcal{S}}\colon\ n \in \mathcal{S}\}$;
   $\nu_n^* := \nu_n^{\mathcal{S}}$;
   $\mathcal{S} := \mathcal{S} \backslash \{n\}$;
end;
{Output $(\nu_n^*)_{n \in \mathcal{T} \times \mathcal{K}}$};
```

Definition 1 (Indexability). We say that ν-parameter problem (4) is *indexable* (or that the product is indexable), if there exist unique values $-\infty \leqslant \nu_n^* \leqslant \infty$ for all $n \in \mathcal{T} \times \mathcal{K}$ such that the following holds:

1. if $\nu_n^* \geqslant \nu$, then it is optimal to promote in state n, and
2. if $\nu_n^* \leqslant \nu$, then it is optimal not to promote in state n.

The function $n \mapsto \nu_n^*$ is called the (*Whittle*) *index*, and ν_n^*'s are called the (*Whittle*) *index values*.

Notice that the indexability property implies that for each value of ν, the problem is optimally solved by an $\mathcal{S}(\nu)$-active set (i.e., a stationary policy), and moreover these $\mathcal{S}(\nu)$-active sets monotonically diminish (by removing one or more states from the set) as ν grows. As a direct consequence of Theorem 2(v), we have the following proposition.

Proposition 1. *The index value of any state* (k, t) *with* $k \geqslant t$ *is* $\nu_{(k,t)}^* = (R/W)(q-p)(1-\beta^t\alpha)$.

We can see that for $t \leqslant k$ the index $\nu_{(t,k)}^*$ does not depend on the actual inventory k. This is because at most one unit of the product can be sold per period, so at most t units can be sold before the product perishes, so the remaining $k-t$ units get surely perished. Thus $t < k$ is a degenerate case, in which a preceding inventory replenishment control had failed or the realized demand in previous periods was extremely low.

In general, one can test indexability of a product numerically and compute index values using a general algorithm given in Niño-Mora (2007), whose fastest known implementation runs in $\mathcal{O}(T^3K^3)$. The algorithmic scheme is presented in Figure 1, where the complexity comes from computing quantities $\nu_n^{\mathcal{S}}$ at each step, which must also satisfy certain properties in order to imply indexability (see the details in section EC.3). According to our findings in numerical experiments, we conjecture that all perishable products are indexable. However, the structure of the optimal policy can considerably differ depending on the sign of α and values of parameters K, T, p, q, β, as established in §3.1, and further conjectured next.

Conjecture 1. *The perishable product is indexable for all values of the parameters* $K, T \geqslant 1$ *and* $R > 0$, $W > 0$, $0 < p < q \leqslant 1$, $\alpha \leqslant 1$, $0 \leqslant \beta \leqslant 1$.

However, the ordering in which the states are being removed from the active sets may in general depend on product parameters. This lack of universal structure complicates the indexability analysis of the problem; nevertheless, we study indexability and characterize the index values for general families of products in the following three subsections.

3.3. Myopic Policy

It is interesting to study problem (4) under the myopic criterion, i.e., optimizing only the current period revenue, which is obtained by setting $\beta = 0$. This leads to a policy, which may provide a practical though suboptimal solution for the original problem with $\beta > 0$.

Theorem 3 (Myopic Policy). *Under the myopic criterion* ($\beta = 0$), (4) *is optimally solved by the following policy*:

1. *if* $\nu \leqslant (R/W)(q-p)$, *then it is optimal to promote in all states* (t, k);
2. *if* $\nu \geqslant (R/W)(q-p)$, *then it is optimal not to promote in all states* (t, k).

The myopic policy is thus en extremely simple index policy, yielding the optimality of two active sets: the empty set $\varnothing$ and the full set of states $\mathcal{T} \times \mathcal{K}$, which is the trivial example of the structure of optimal policies. Consequently, the corresponding myopic index value is $\nu_{(t,k)}^{\text{MYOPIC}} = (R/W)(q-p)$ for all states (t, k).

3.4. Nonpositive Expected Salvage Value

This subsection focuses on the case $\alpha \leqslant 0$, for which we identify a family of products that are provably indexable and for which we provide an efficient algorithm for computation of the index values.

As the simplest nontrivial example of the structure of optimal policies, let us denote by $\mathcal{F}^{(1)}$ the *strongly time-monotonous* family of nested active sets

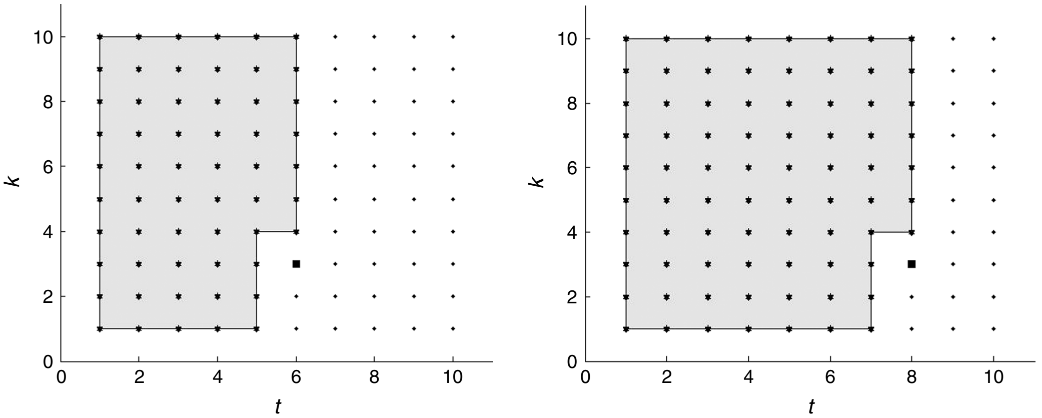
$$\mathcal{F}^{(1)} := \{\mathcal{S}_{(t,k)}^{(1)} \text{ for all } (t,k) \in \mathcal{T} \times \mathcal{K}\},$$
$$\text{where } \mathcal{S}_{(t,k)}^{(1)} := \{(s,l)\colon (s < t \text{ and } 1 \leqslant l \leqslant K) \text{ or } (s = t \text{ and } k < l \leqslant K)\}.$$

Figure 2 graphically illustrates two instances of stationary policies $\mathcal{S}_{(t,k)}^{(1)}$, and a product instance optimally solvable by policies belonging to $\mathcal{F}^{(1)}$ is presented in Table 1.

The active sets belonging to $\mathcal{F}^{(1)}$ are such that if it is optimal to promote at state (t, k), then it is optimal to promote at all the remaining periods regardless of the actual inventory, and also at period t if the inventory was larger than k. This family is very restrictive, but it gives a unique and total ordering of active sets in which they monotonically diminish.

However, we shall consider a more general, *weakly time-monotonous* family $\mathcal{F}^{(2)}$ of active sets, for which the problem

Figure 2. Illustration of the ordering of states by a strongly time-monotonous family $\mathscr{F}^{(1)}$ of nested active sets for a particular value of parameter ν (in this case the index values can be computed analytically by (5)).



Notes. The state (t, k) at which both actions are optimal (to be added to the optimal active set) is marked by a square, the states at which it is optimal to promote (already in the active set) by stars, and the remaining states by dots. Set $\mathscr{S}^{(1)}_{(t,k)}$ is the filled area. The figures correspond to $(t, k) = (6, 3)$ (on the left) and $(t, k) = (8, 3)$ (on the right), respectively, of an instance with parameters $T = 10$; $K = 10$; $W = 1$; $R = 1$; $\alpha = -1/2$; $p = 0.9$; $q = 1$; $\beta = 0.4$.

Table 1. A product instance solvable by policies in $\mathscr{F}^{(1)}$ illustrating the structure of the optimal active sets: for any state (t, k), it is optimal to promote in states preceding it, and it is optimal not to promote in the remaining states.

(t, k)	$(1, 3)$	$(1, 2)$	$(1, 1)$	$(2, 3)$	$(2, 2)$	$(2, 1)$	$(3, 3)$	$(3, 2)$	$(3, 1)$	$(4, 3)$	$(4, 2)$	$(4, 1)$
$\nu^{(2)}_{(t,k)} = \nu^{*}_{(t,k)}$	650	650	650	590	590	588	554	554	551	532	532	529
(t, k)	$(5, 3)$	$(5, 2)$	$(5, 1)$	$(6, 3)$	$(6, 2)$	$(6, 1)$	$(7, 3)$	$(7, 2)$	$(7, 1)$	$(8, 3)$	$(8, 2)$	$(8, 1)$
$\nu^{(2)}_{(t,k)} = \nu^{*}_{(t,k)}$	519	519	517	512	511	510	507	507	506	504	504	503

Note. We further show that index values computed by the general algorithm and using the analytical characterization of $\nu^{(2)}_{(t,k)}$ are equal ($T = 8$, $K = 3$, $R = 1{,}000$, $W = 1$, $\alpha = -1/2$, $p = 0.95$, $q = 1$, $\beta = 0.6$).

can be solved analytically. The problem when considering $\mathscr{F}^{(1)}$ is a special case of that with $\mathscr{F}^{(2)}$. We define

$$\mathscr{F}^{(2)} := \{\mathscr{S} \subseteq \mathscr{T} \times \mathscr{K}\colon \text{for all } (t,k) \in \mathscr{S},\ \mathscr{S} \supseteq \mathscr{S}^{(2)}_{(t,k)}\},$$
$$\text{where } \mathscr{S}^{(2)}_{(t,k)} := \{(s,l)\colon 1 \leqslant s < t \text{ and } \max\{1, k-(t-s)\} \leqslant l \leqslant k\}.$$

Family $\mathscr{F}^{(2)}$ induces only a partial ordering of active sets in which they monotonically diminish (in contrast to the total ordering induced by $\mathscr{F}^{(1)}$). Figure 3 graphically illustrates two instances of active sets $\mathscr{S} \in \mathscr{F}^{(2)}$, together with instances of $\mathscr{S}^{(2)}_{(t,k)}$. Note that $\mathscr{S}^{(2)}_{(t,k)}$ is determined by the diagonal, since the states strictly below the diagonal are irrelevant because at most one product unit can be sold per period because of Bernoulli demand.

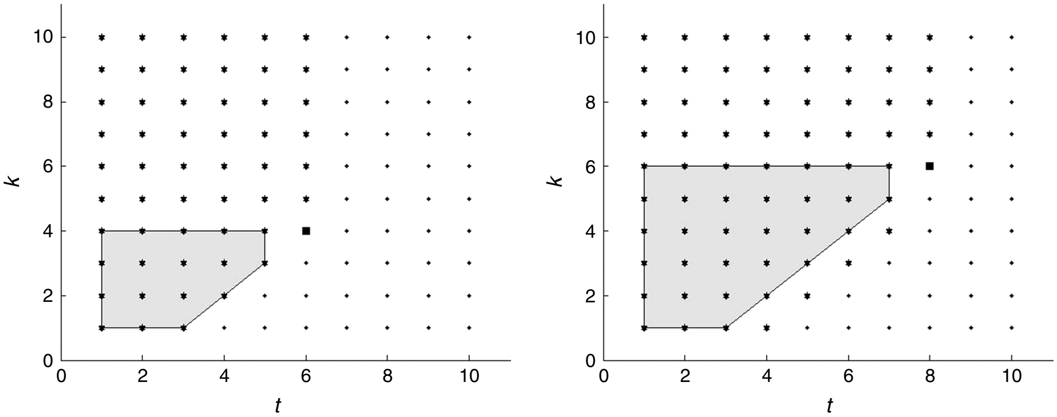
The following theorem is the main theoretical result of this paper. The proof is based on a work-reward analysis of the value function and on properties of binomial numbers, and, because of its extensive length, it is presented in section EC.4.

Theorem 4 (Indexability Under $\mathscr{F}^{(2)}$). *Suppose that $\alpha \leqslant 0$. If for every ν there is an optimal active set that belongs to $\mathscr{F}^{(2)}$, then the product is indexable, and the index value for its state $(t,k) \in \mathscr{T} \times \mathscr{K}$ is*

$$\nu^{(2)}_{(t,k)} = \begin{cases} \nu^{*}_{(t,k)} = \dfrac{R}{W}(1-p)\left[1 - \dfrac{1-q+\alpha(q-p)\beta^{t}}{1-p}\right], & t \leqslant k \\[2ex] \dfrac{R}{W}(1-p)\Bigg[1 - \Bigg(1-q+\alpha\beta^{t}(q-p)p^{t-k} \\ \quad \cdot \displaystyle\sum_{i=0}^{k-1} \binom{t-k-1+i}{i}(1-p)^{i}\Bigg) \\ \quad \cdot \Bigg(1-p-(q-p)\beta^{k}(1-p)^{k} \\ \quad \cdot \displaystyle\sum_{i=0}^{t-k-1} \binom{k-1+i}{i}(\beta p)^{i}\Bigg)^{-1}\Bigg], & t > k \end{cases} \tag{5}$$

Let us introduce in the following two propositions the recursive functions $f_{t,k}$ and $g_{t,k}$ that will be useful in the computation of the index values (5).

Figure 3. Illustration of the ordering of states by a weakly time-monotonous family $\mathcal{F}^{(2)}$ of nested active sets for a particular value of parameter ν (in this case the index values can be computed analytically by (5)).



Note. The state (t, k) at which both actions are optimal (to be added to the optimal active set) is marked by a square, the states at which it is optimal to promote (already in the active set) by stars, and the remaining states by dots. Set $\mathcal{S}^{(2)}_{(t,k)}$ is the filled area. The figures correspond to $(t, k) = (6, 4)$ (on the left) and $(t, k) = (8, 6)$ (on the right), respectively, for an instance with parameters $T = 10$; $K = 10$; $W = 1$; $R = 1$; $\alpha = -1/2$; $p = 0.7$; $q = 1$; $\beta = 0.9$.

Proposition 2 (Recursion for $f_{t,k}$). *Let $f_{t,k}$ denote the function*

$$f_{t,k} := \begin{cases} \beta^t p^{t-k} \sum_{i=0}^{k-1} \binom{t-k-1+i}{i} (1-p)^i, & \text{if } t > k, \\ \beta^t, & \text{if } t \leqslant k, \end{cases} \tag{6}$$

Then $f_{t,k}$ satisfies the following recursions:
(i) $f_{t,1} = \beta p f_{t-1,1}$ *for* $t > k$, $k = 1$,
(ii) $f_{t,k} = \beta p f_{t-1,k} + \beta(1-p) f_{t-1,k-1}$ *for* $t > k$, $k > 1$.

Proposition 3 (Recursion for $g_{t,k}$). *Let $g_{t,k}$ denote the function*

$$g_{t,k} := \begin{cases} \beta^k (1-p)^k \sum_{i=0}^{t-k-1} \binom{k-1+i}{i} (\beta p)^i, & \text{if } t > k, \\ 0, & \text{if } t \leqslant k. \end{cases} \tag{7}$$

Then $g_{t,k}$ satisfies the following recursions:
(i) $g_{t,1} = \beta p g_{t-1,1} + \beta(1-p)$ *for* $t > k$, $k = 1$,
(ii) $g_{t,k} = \beta p g_{t-1,k} + \beta(1-p) g_{t-1,k-1}$ *for* $t > k$, $k > 1$.

Because of these two propositions we have the following result giving a linear (in number of states TK) algorithm for computation of the index values. The improvement in the time complexity of the algorithm is of two orders of magnitude in both T and K, since the fastest known implementation of the general algorithm for index values computation requires $\mathcal{O}(T^3K^3)$ elementary operations (see the next subsection). Moreover, any algorithm must perform at least $\mathcal{O}(TK)$ elementary operations in order to compute TK index values, and this bound is achieved by our algorithm.

Proposition 4 (Fast Index Values Computation). *The index values in* (5) *can be characterized by*

$$\nu^{(2)}_{(t,k)} = \frac{R}{W}(1-p)\left[1 - \frac{1-q+\alpha(q-p)f_{t,k}}{1-p-(q-p)g_{t,k}}\right]. \tag{8}$$

Moreover, the index values, for all $t \in \mathcal{T}$ and $k \in \mathcal{K}$ can be computed using at most $\mathcal{O}(TK)$ elementary operations.

We can observe several remarkable properties of the index values.

Proposition 5 (Monotonicity). *The monotonicity properties of index values in* (5) *are*
(i) $\nu^{(2)}_{(t-1,k)} \geqslant \nu^{(2)}_{(t,k)}$ $\forall k \geqslant 1$, $\forall t > 1$;
(ii) $\nu^{(2)}_{(t,k-1)} \leqslant \nu^{(2)}_{(t,k)}$ $\forall k > 1$, $\forall t \geqslant 1$;
(iii) $\nu^{(2)}_{(s,l)} \geqslant \nu^{(2)}_{(t,k)}$ $\forall l \geqslant k$, $\forall s \leqslant t$.

Next we give an index characterization for a nonperishable product, obtained from index values (5) for perishable product in the limit $t \to \infty$.

Proposition 6 (Nonperishable Product). *Under assumptions of Theorem* 4, *the index for a nonperishable product with inventory $k \in \mathcal{K}$ is*

$$\nu^{(2)}_{(\infty,k)} = \begin{cases} \frac{R}{W}(1-p)[1-(1-q)\cdot(1-p-(q-p) \\ \quad \cdot(\beta(1-p)/(1-\beta p))^k)^{-1}], & \beta < 1 \\ 0, & \beta = 1. \end{cases}$$

Figure 4 illustrates the convergence of index values of perishable products as t grows, indicating that the formula given in Proposition 6 may be used as an approximation of the index value (5) for large t; see section EC.9 for more details. Note also that for large k, $\nu^{(2)}_{(\infty,k)}$ converges in case $\beta < 1$ to $(R/W)(q-p)$, which is the myopic index $\nu^{\text{MYOPIC}}_{(t,k)}$. That is, the myopic index value underestimates the efficiency of promotion for low t, and it overestimates it for high t. Nevertheless, it is a good approximation of the Whittle index value when both t and k are large in case $\beta < 1$ (although this is not true if $\beta = 1$).

Figure 4. Illustration of convergence of the index values $\nu^{(2)}_{(t,k)}$ as t grows in a product with parameters $T = 50$, $K = 7$, $W = 1$, $R = 1$, $\alpha = -1/2$, $p = 0.7$, $q = 0.85$.

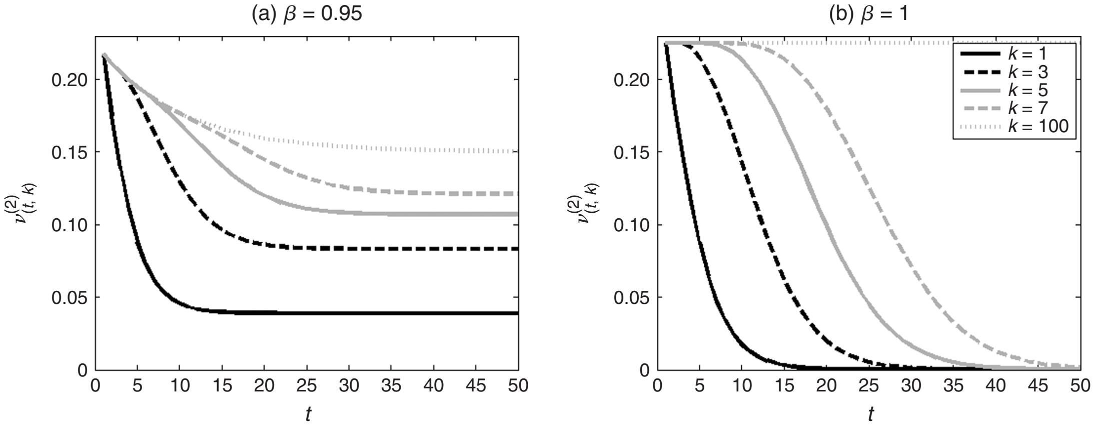
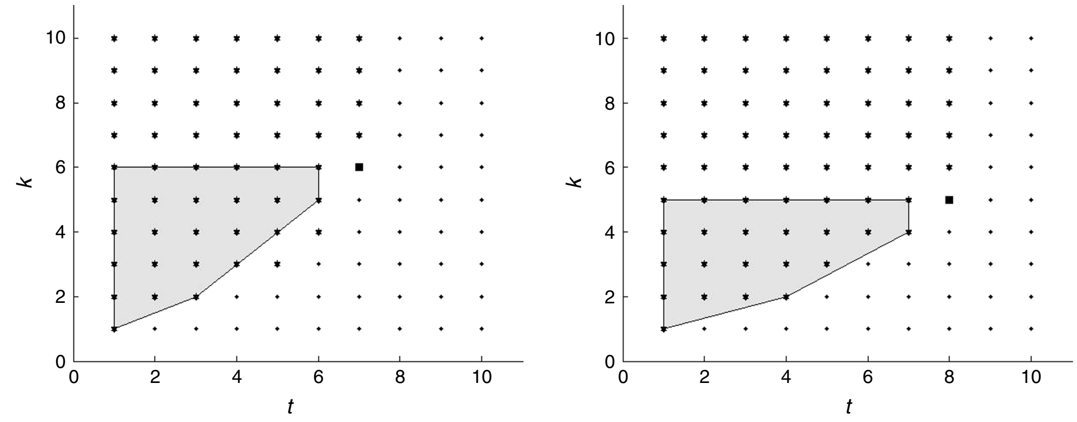


Figure 5. Illustration of the ordering of states by the conjectured family $\mathscr{F}^{(3)}$ of nested active sets for a particular value of parameter ν (in this case we propose to use the analytical index (5) as an approximation of the exact one).



Notes. The state (t, k) at which both actions are optimal (to be added to the optimal active set) is marked by a square, the states at which it is optimal to promote (already in the active set) by stars, and the remaining states by dots. Set $\mathscr{S}^{(3)}_{(t,k)}$ is the filled area. The figures correspond to $(t, k) = (7, 6)$ (on the left) and $(t, k) = (8, 5)$ (on the right), respectively, of an instance with parameters $T = 10$; $K = 10$; $W = 1$; $R = 1$; $\alpha = -1$; $p = 0.7$; $q = 0.85$; $\beta = 0.95$.

However, $\mathscr{F}^{(2)}$ does not always characterize the optimal policies entirely. We thus define a more relaxed, partially ordered family of active sets

$$\mathscr{F}^{(3)} := \{\mathscr{S} \subseteq \mathscr{T} \times \mathscr{K}\colon \text{for all } (t,k) \in \mathscr{S},\ \mathscr{S} \supseteq \mathscr{S}^{(3)}_{(t,k)}\}, \quad \text{where}$$
$$\mathscr{S}^{(3)}_{(t,k)} := \{(s,l)\colon 1 \leqslant s < t \text{ and } l(s) \leqslant l \leqslant k, \text{ where}$$
$$\max\{1, k-(t-s)\} \leqslant l(s) \text{ is some real-valued}$$
$$\text{nondecreasing function over } s \in [1, t-1]\}.$$

Notice that we have $\mathscr{S}^{(3)}_{(t,k)} \subseteq \mathscr{S}^{(2)}_{(t,k)} \subseteq \mathscr{S}^{(1)}_{(t,k)}$ for all (t, k), and therefore $\mathscr{F}^{(1)} \subseteq \mathscr{F}^{(2)} \subseteq \mathscr{F}^{(3)}$. The differences between these sets can be easily seen from a comparison of Figures 2, 3, and 5. The latter further illustrates the structure of active sets of family $\mathscr{F}^{(3)}$. We conjecture that $\mathscr{F}^{(3)}$ always completely characterizes the optimal polices.

Conjecture 2. *If $\alpha \leqslant 0$, then we believe the following is true*:

1. *the index is nonincreasing in t for every k* (*i.e., Proposition* 5(i) *holds*);
2. *family $\mathscr{F}^{(3)}$ contains an optimal policy, and function $l(s)$ is convex for sufficiently low β, and it is concave for $\beta = 1$*;
3. *if $\beta = 1$, then the structure of the optimal policy* (*i.e., the order of the added states to the active set $\mathscr{S}$ as ν decreases*) *does not depend on α.*

Unfortunately, the index values under $\mathscr{F}^{(3)}$ cannot be derived analytically, since the exact structure of optimal policies is not easily identifiable. For practical purposes we therefore propose to use the index values characterized in §3.4 as approximate index values in general. Although it is not guaranteed that optimal policies for any value of ν belong to family $\mathscr{F}^{(2)}$, the index values computed are numerically close to those obtained by Theorem 4 according to our numerical experiments. In particular, the relative deviation of index values given by Theorem 4 from true index values is below 1 % on average and below 5 % in the worst case. See section EC.10 for more details.

Two perishable product instances are presented in Tables 1 and 2, where the order of states is according to the index values computed using the general algorithm, and index values given by Theorem 4 are also shown for comparison. In Table 1 a product optimally solvable by policies in family $\mathcal{F}^{(1)}$ is shown, and strong time monotonicity holds (all the states with shorter time to deadline are in the active set). This instance therefore obeys the policy structures in both $\mathcal{F}^{(2)}$ and $\mathcal{F}^{(3)}$, and the index values given by Theorem 4 equal to those computed numerically. In the more common case, illustrated in Table 2, the structure of optimal policies obeys family $\mathcal{F}^{(2)}$ only until state $(5,4)$ (including). Note that until that row, the numerical and analytical index values are equal. After that row, the optimal policy structure belongs to $\mathcal{F}^{(3)}$, but not to $\mathcal{F}^{(2)}$. For instance, states $(5,3)$, $(4,2)$, $(6,4)$, $(3,1)$ (that lie on the same diagonal) would have to be added in the order $(3,1)$, $(4,2)$, $(5,3)$, $(6,4)$ instead in order to have policies belonging to $\mathcal{F}^{(2)}$; similarly, the remaining states must also be reordered.

However, the numerical differences between index values are very small and the only place in which the order of states obtained by our index values differs from the true order is having states $(4,2)$ and $(6,4)$ swapped. To sum up, our index values may well approximate the true index values despite assuming different structures of optimal policies. This encourages us to implement our index characterization by Theorem 4 in heuristics for KPPI in the following section.

3.5. Positive Expected Salvage Value

For the case $1 \geqslant \alpha > 0$, we consider a complementary, *weakly counter-time-monotonous* family $\mathcal{F}^{(-2)}$ of active sets, for which the problem can in part be solved analytically. We define

$$\mathcal{F}^{(-2)} := \{\mathcal{S} \subseteq \mathcal{T} \times \mathcal{K}\text{: for all } (t,k) \notin \mathcal{S},\ \mathcal{S} \cap \mathcal{S}^{(2)}_{(t,k)} = \varnothing\}.$$

Family $\mathcal{F}^{(-2)}$ induces only a partial ordering of active sets in which they monotonically diminish, but in the direction contrary to the family $\mathcal{F}^{(2)}$. Figure 6 on the left graphically illustrates one instance of active sets $\mathcal{S} \in \mathcal{F}^{(-2)}$.

Analogously to Theorem 4 we can obtain a slightly weaker result. The proof is provided in section EC.11.

Theorem 5 (Index Values Under $\mathcal{F}^{(-2)}$). *Suppose that $1 \geqslant \alpha > 0$. If there is ν_0 such that for every $\nu \geqslant \nu_0$ there is an optimal active-set $\mathcal{S}$ that belongs to $\mathcal{F}^{(-2)}$, and if the product is indexable, then the index value for its state $(t,k) \in \mathcal{S}$ with $\nu^*_{(t,k)} \geqslant \nu_0$ is*

$$\nu^{(-2)}_{(t,k)} = \begin{cases} \nu^*_{(t,k)} = \dfrac{R}{W}(q-p)[1-\alpha\beta^t], & t \leqslant k \\ \dfrac{R}{W}(q-p)\Bigg[1-\beta^k(1-q)^k \\ \quad \cdot \displaystyle\sum_{i=0}^{t-k-1}\binom{k-1+i}{i}(\beta q)^i - \alpha\beta^t q^{t-k} \\ \quad \cdot \displaystyle\sum_{i=0}^{k-1}\binom{t-k-1+i}{i}(1-q)^i\Bigg], & t > k. \end{cases} \tag{9}$$

Analogously to the case $\alpha \leqslant 0$, it is possible to obtain a fast algorithm for computation of the index values and establish monotonicity properties. We provide these results and further discussion in section EC.12.

The main limitation of Theorem 5 is that it relies on the assumption of existence of value ν_0. Indeed, in view of Theorem EC.2 and the results in Jacko (2013), $\mathcal{F}^{(-2)}$ may not contain optimal policy for all values of ν except for the less practically relevant case $\alpha \geqslant (1-q)/(1-\beta q)$. However, it is possible to observe numerically that such ν_0 often exists, and so formula (9) characterizes those index values with $\nu^*_{(t,k)} \geqslant \nu_0$.

Since the deadline T is fixed, it turns out that formula (9) gives exact characterization of index values for all the states with k large enough for a given t, or, equivalently, for all the states with t small enough for given k. On the other hand, Theorem EC.2 and the results in Jacko (2013) give exact characterization of index values for all the states with $k=1$. Thus, relatively few states (in our numerical experiments with $T=K=10$ typically those with $k=2,3$) are left without exact characterization, and these could be either approximated or computed exactly numerically.

If $\alpha > 0$, the structure of the optimal policies is relatively irregular, and we conjecture it has the following properties.

Conjecture 3. *If $\alpha > 0$, then we believe the following is true*:

1. *For every k there is $1 \leqslant t_k \leqslant +\infty$ such that the index is nondecreasing in $t \leqslant t_k$ and nonincreasing in $t \geqslant t_k$, and t_k is increasing in k; moreover, the higher α, the higher t_k's, and the higher β, the lower t_k's.*
2. *Once it is optimal to promote in state $(1,1)$ (i.e., $\nu \leqslant R(q-p)(1-\alpha\beta)/W$), then family $\mathcal{F}^{(3)}$ contains an optimal policy, and function $l(s)$ is concave for all β.*
3. *If $\beta = 1$, then $t_k = 1$ for all k and the structure of the optimal policy (i.e., the order of the added states to the active set $\mathcal{S}$ as ν decreases) does not depend on α, and is the same as when $\alpha \leqslant 0$.*

We remark that Theorem EC.2 implies that $t_1 = +\infty$, which is in line with 3(i), and would further imply $t_k = +\infty$ for all k. Also, 3(ii) is in agreement with 4(iii)–(iv). Note also that $\mathcal{F}^{(-2)} \subseteq \mathcal{F}^{(3)}$. For illustration, see Figure 6 on the right. Figure 6 on the left however does not satisfy that property. This may be the case, for instance, of the problem of car showroom exhibition, in which it might be more profitable (as measured by the Whittle index value) to select a new car model with low inventory than to select an older car model with high inventory for the showroom.

4. Solutions to Knapsack Problem for Perishable Inventories

In this section we return to the unrelaxed KPPI problem and show how the index derived in the previous section can be implemented in order to obtain a nearly optimal solution.

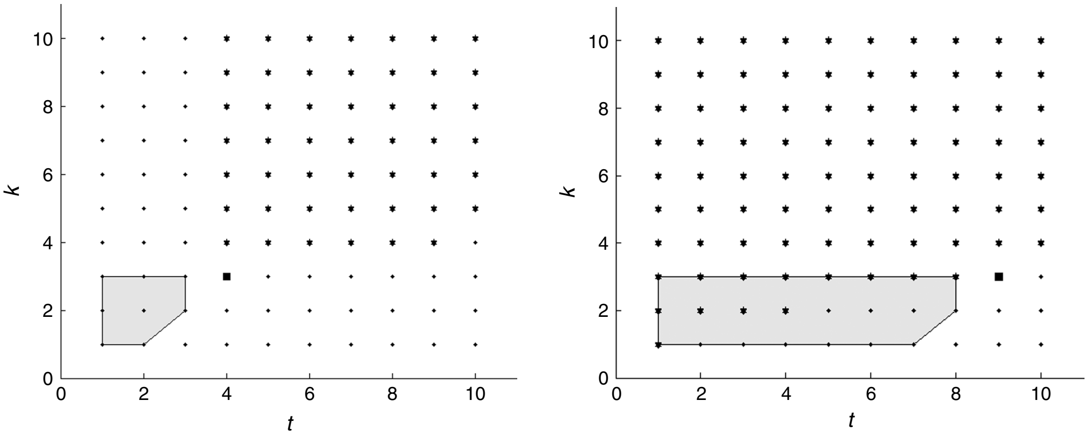
Table 2. A product instance solvable by policies in $\mathscr{F}^{(3)}$ (but not by those in $\mathscr{F}^{(1)}$ or $\mathscr{F}^{(2)}$) illustrating the structure of the optimal active sets: for any state (t,k), it is optimal to promote in states preceding it, and it is optimal not to promote in the remaining states.

(t,k)	$(1,4)$	$(1,3)$	$(1,2)$	$(1,1)$	$(2,4)$	$(2,3)$	$(2,2)$	$(3,4)$	$(3,3)$	$(2,1)$	$(4,4)$	$(3,2)$
$\nu^*_{(t,k)}$	9,833	9,833	9,833	9,833	9,675	9,675	9,675	9,525	9,525	9,402	9,382	9,318
$\nu^{(2)}_{(t,k)}$	9,833	9,833	9,833	9,833	9,675	9,675	9,675	9,525	9,525	9,402	9,382	9,318

(t,k)	$(4,3)$	$(5,4)$	$(5,3)$	$(4,2)$	$(6,4)$	$(3,1)$	$(6,3)$	$(5,2)$	$(4,1)$	$(6,2)$	$(5,1)$	$(6,1)$
$\nu^*_{(t,k)}$	9,228	9,133	8,780	8,765	8,758	8,578	8,253	8,067	7,647	7,417	7,053	6,799
$\nu^{(2)}_{(t,k)}$	9,228	9,133	8,749	8,708	8,746	8,578	8,129	7,945	7,647	7,307	7,053	6,799

Note. We further show the index values computed by the general algorithm ($\nu^*_{(t,k)}$) and using the analytical characterization of $\nu^{(2)}_{(t,k)}$, highlighting when they are not equal ($T=6$, $K=4$, $R=1{,}000$, $W=1$, $\alpha=-1/2$, $p=1/3$, $q=1$, $\beta=0.95$).

Figure 6. Illustration of the ordering of states by a weakly counter-time-monotonous family $\mathscr{F}^{(-2)}$ of nested active sets for a particular value of parameter ν such that $\nu \geqslant \nu_0$ (on the left; in this case the index value can be computed analytically by (9)) and such that $\nu < \nu_0$ (on the right).



Notes. The state (t,k) at which both actions are optimal (to be added to the optimal active set) is marked by a square, the states at which it is optimal to promote (already in the active set) by stars, and the remaining states by dots. Set $\mathscr{S}^{(2)}_{(t,k)}$ is the filled area. The figures correspond to $(t,k)=(4,3)$ (on the left) and $(t,k)=(9,3)$ (on the right), respectively, of an instance with parameters $T=10$; $K=10$; $W=1$; $R=1$; $\alpha=0.2$; $p=0.7$; $q=0.8$; $\beta=0.95$.

4.1. Dynamic Programming Formulation and Exact Solution of KPPI

We first outline an exact dynamic programming formulation of the problem. Recall from §2.3 that $\mathbf{X}(s)$ is the actual joint state of all the products at period $s\in[0,H]$. Let $\mathscr{M}^{(s)}$ be the set of a possible joint states at time period s. In particular, for $s=H$, we have a single possible joint state of having all the products perished, so $\mathscr{M}^{(s)}=\{\mathbf{0}\}$.

Then, the Bellman equation for the optimal value $D^{(s)}_{\mathbf{m}}$ for each joint state $\mathbf{m}\in\mathscr{M}^{(s)}$ and period $s\in[0,H-1]$ is as follows:

$$D^{(s)}_{\mathbf{m}} = \max_{\sum_{i\in\mathscr{I}} z_i(\mathbf{m},s)W_i\leqslant C}\left(\sum_{i\in\mathscr{I}} R_{i,m_i}^{z_i(\mathbf{m},s)} + \beta \sum_{\mathbf{m}'\in\mathscr{M}^{(s+1)}} \mathbb{P}_{\mathbf{m}}^{\mathbf{z}(\mathbf{m},s)} \cdot [\mathbf{X}(s+1)=\mathbf{m}']D^{(s+1)}_{\mathbf{m}'}\right),$$

$$D^{(H)}_{\mathbf{m}} = 0, \quad \text{for } \mathbf{m}\in\mathscr{M}^{(H)},$$

where $\mathbf{z}(\mathbf{m},s) := (z_i(\mathbf{m},s))_{i\in\mathscr{I}}$ is a vector of binary decision variables denoting whether a unit of product i is selected for promotion or not (i.e., $z_i(\mathbf{m},s)\in\{0,1\}$), and moreover such that $z_i(\mathbf{m},s)=0$ if $m_i=0$ (i.e., product i cannot be promoted if it is perished or sold out).

By $\mathbb{P}_{\mathbf{m}}^{\mathbf{z}(\mathbf{m},s)}[\mathbf{X}(s+1)=\mathbf{m}']$ we understand the probability of reaching the next-period joint state $\mathbf{X}(s+1)=\mathbf{m}'$ conditional on the actual-period joint state $\mathbf{X}(s)=\mathbf{m}$ and on employing actions according to $\mathbf{z}(\mathbf{m},s)$. This joint state dynamics is, obviously, a result of selling (or not selling) particular products and of their perishing.

In other words, $D^{(s)}_{\mathbf{m}}$ is the maximum expected total discounted revenue over all possible combinations of actions for the actual inventory $\mathbf{m}$ at time period s, according to the knapsack capacity C. We are interested in value $D^{(0)}_{\mathbf{m}^*} =: D^{\max}$ from the initial joint state $\mathbf{m}^* := ((T_i,K_i))_{i\in\mathscr{I}}$. Moreover, the optimal policy is obtained by vectors $\mathbf{z}(\mathbf{m},s) := (z_i(\mathbf{m},s))_{i\in\mathscr{I}}$ for each s and every $\mathbf{m}$ that maximize the right-hand side of the Bellman equations.

Thus, the above approach yields the exact solution to KPPI. However, it is intractable because of the combinatorial explosion of both the joint state space and the joint action space given the knapsack constraint, which is known as the *curse of dimensionality*. Indeed, in order to obtain an optimal solution and an optimal policy, it is necessary to obtain optimal policies for all possible future joint states in all future periods. Our implementation of the dynamic programming approach in Matlab leads to acceptable runtime (of below 10 minutes per problem instance) only for up to $I = 5$ products and up to horizon $H = 32$.

4.2. Heuristics for KPPI

Since the index captures the marginal rate of revenue from promotion, we next propose to implement index values as "a promotion price per volume unit." Recall expression (5) for the index calculation, which is to be used in the heuristics as explained next.

Consider a period $s \in [0, H-1]$, at which the joint state is $\mathbf{X}(s)$. Let $\tilde{\mathcal{I}} \subseteq \mathcal{I}$ be the set of products that are not perished and at least one unit is available (according to $\mathbf{X}(s)$), i.e., $i \in \tilde{\mathcal{I}}$ if and only if $X_i(s) \neq 0$. Then for all $i \in \tilde{\mathcal{I}}$, an index value by (5) exists, so we can define the following *knapsack problem prices*:

$$v_i^{(s)} := W_i \nu^*_{i,\,X_i(s)}. \tag{10}$$

And we consider the following 0-1 knapsack problem to solve

$$\begin{aligned} \max_{\mathbf{y}^{(s)}} \quad & \sum_{i \in \tilde{\mathcal{I}}} y_i^{(s)} v_i^{(s)} \\ \text{subject to} \quad & \sum_{i \in \tilde{\mathcal{I}}} y_i^{(s)} W_i \leqslant C \\ & y_i^{(s)} \in \{0, 1\} \quad \text{for all } i \in \tilde{\mathcal{I}}, \end{aligned} \tag{KP}$$

where $\mathbf{y}^{(s)} := (y_i^{(s)})_{i \in \tilde{\mathcal{I}}}$ is a vector of binary decision variables denoting whether a unit of product i is selected for promotion or not.

We propose the following heuristics for solving KPPI:

- *Index-knapsack* (*IK*) *heuristic*. Calculate the prices $v_i^{(s)}$ for $s = 0$ and then solve the knapsack problem (KP) optimally.
- *Index rule* (*IR*). Compute the index values $\nu^*_{i,\,X_i(s)}$ for $s = 0$ and then select the products for promotion in greedy manner (highest first) until either the capacity is filled or there are no more products.

Notice that the computational complexity of these two heuristics is dramatically lower than that of the optimal solution via dynamic programming; in particular, they do not suffer from the curse of dimensionality. Index rule is the simpler one, since it only requires for each product to run a linear-time algorithm computing the index values and then sorting them. The index-knapsack (IK) heuristic requires to compute the prices (at exactly the same time as computing the index values), and then solving a single knapsack problem, which can be solved in pseudo-polynomial time $\mathcal{O}(I \cdot C)$ in the worst case and for which practically efficient exact algorithms exist (Pisinger 2005).

We remark that index rule can be itself seen as a greedy algorithm for the knapsack problem of IK heuristic, analogously to the Dantzig greedy algorithm for the classic knapsack problem (Dantzig 1957). One can therefore expect that IK heuristics should outperform index rule. Moreover, the absolute difference in their performance can be arbitrarily bad, but they should converge to each other as either (i) product volumes of all the products become equal or (ii) as problem granularity decreases (i.e., as both the number of products and the knapsack capacity increase).

5. Experimental Study

We have studied performance of the two proposed heuristics, comparing them to the optimal policy in numerical experiments for a variety of model parameters. In this section we present results of systematical computational experiments in which we evaluate the suboptimality gap of the two heuristics.

To summarize, the experimental study suggests near optimality of both the index-knapsack heuristic and the index rule. The former is, on average, always outperforming the latter, often significantly. Moreover, the difference in the performance of the index-knapsack heuristic with exact index value and that with approximate index values is negligible.

5.1. Performance Evaluation Measures

By solving the KPPI optimally we obtain the maximizing policy, which yields the optimal objective value $D^{\max}$. Via backward recursion, the objective value of other policies is obtained by employing the policy at each step, denoted by D^{π} for policy π. We next introduce performance evaluation measures we use to report the experimental results.

The *relative suboptimality gap* of policy π, conventionally used in literature, is defined as

$$\mathrm{rsg}(\pi) = \frac{D^{\max} - D^{\pi}}{D^{\max}}. \tag{11}$$

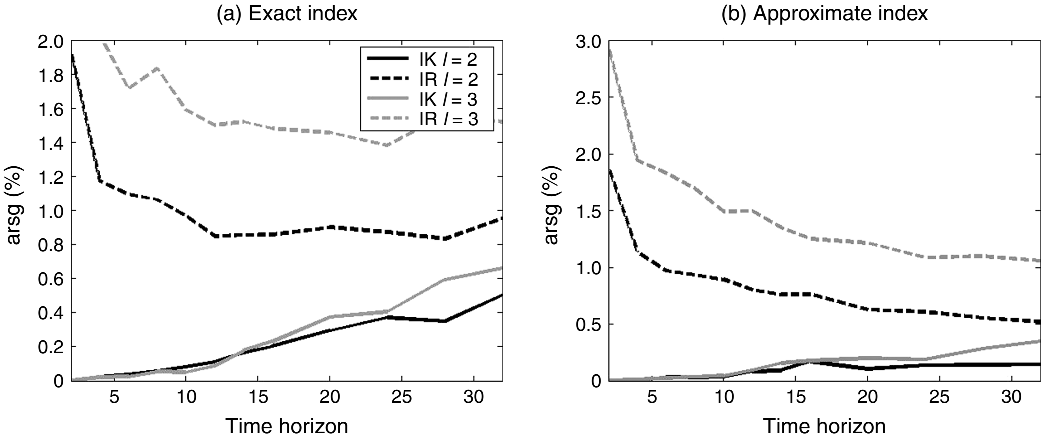
Clearly, as long as $D^{\max} > 0$, we have $0 \leqslant \mathrm{rsg}(\pi) \leqslant +\infty$, where $\mathrm{rsg}(\pi) = 0$ is obtained by the maximizing policy. However, if $\alpha_i \leqslant 0$ for some of the products, we may have $D^{\max} \leqslant 0$ and therefore may also be $\mathrm{rsg}(\pi) < 0$ or a problem of division by zero may appear. On the other hand, the worst-case policy of leaving the knapsack empty may have an objective value close to $D^{\max}$ in some instances. We can therefore conclude that this measure may overestimate the quality of π by reporting small or negative values even for the worst policies.

This motivates us to use another measure, the *adjusted relative suboptimality gap* of policy π, defined as

$$\mathrm{arsg}(\pi) = \frac{D^{\max} - D^{\pi}}{D^{\max} - D^{\min}}, \tag{12}$$

where $D^{\min}$ is the worst-case policy of leaving the knapsack empty. With this measure we always have $0 \leqslant \mathrm{arsg}(\pi) \leqslant 1$

Figure 7. Mean adjusted relative suboptimality gap of IK heuristic (solid line) and IR heuristic (dashed line) in problems with $I = 2$ (black line) and $I = 3$ (grey line) products.



(as long as $D^{\max} - D^{\min} \neq 0$), and both limiting values can be achieved.

5.2. Experimental Study Setting

For each pair (I, H), denoting the number of products and the problem time horizon, respectively, such that $I = \{2, 3, 4, 5\}$ and $H = \{2, 4, \ldots, 16\}$ or $I = \{2, 3\}$ and $H = \{20, 24, \ldots, 32\}$, we have randomly generated 5,000 (for $H \leqslant 16$) or 10,000 (for $H > 16$) instances. We set $\alpha_i = -1/2$ for each product i and we assure that $T_1 := H$. For each product we have randomly generated the number of units K_i such that $K_i \leqslant T_i$. To keep the instances tractable (solvable in less than 10 minutes) we have also set $K_i < 10$ and $\sum_i K_i \leqslant 20$.

We assume Poisson arrivals of customers for each product i, denoting by λ_i^0 and λ_i^1 the mean arrival rate for non-promoted and for promoted product, respectively. We restrict these values to $2/3K_i < \lambda_i^a T_i \leqslant 2K_i$ for both $a \in \{0, 1\}$, which assures that the probability of selling all the units of product i before the deadline for each product is within a reasonable range, since $\lambda_i^a T_i$ is the expected number of customer arrivals during the product's lifetime. The respective probabilities of not selling any unit are $q_i = e^{-\lambda_i^0}$ and $p_i = e^{-\lambda_i^1}$. Thus we generate the following uniformly distributed parameters:

$$W_i \in [10, 25]; \quad R_i \in [10, 50]; \quad T_i \in [2, H];$$

$$K_i \in [1, \max(T_i, 9)]; \quad \lambda_i^0, \lambda_i^1 \in \left(\frac{2}{3}\frac{K_i}{T_i}, \frac{2K_i}{T_i}\right].$$

Finally, a uniformly distributed knapsack volume C is generated in the interval

$$C \in \left[\max\{W_i\}; \max\left\{\max\{W_i\}, 40\% \cdot \sum_i W_i\right\}\right].$$

We focus on the case $\beta = 1$, which is most likely to be implemented in practice, but for $\beta \approx 1$ the results are similar, and the performance of the heuristics improves as $\beta \to 0$.

5.3. Scenarios with Two or Three Products

We first describe the results of experiments with $I = 2$ and $I = 3$ products, which are computationally the most accessible ones, and therefore can provide richer information. Figure 7 presents the mean adjusted relative suboptimality gap in identical sets of problem instances, where the numerically computed (exact) index was implemented in (a), and the analytically computed (approximate) index was taken for the two heuristics in (b). In all cases the mean gap is below 3%, and it is interesting to observe that the IK heuristic significantly outperforms the index rule (IR). The difference in performance is more prevalent for smaller time horizons, and seems to diminish as the horizon grows. This behavior suggests that the two heuristics become equivalent when the time horizon is large enough. Moreover, Figure 7 indicates that there is virtually no difference in performance between implementing the approximate and the exact index values.

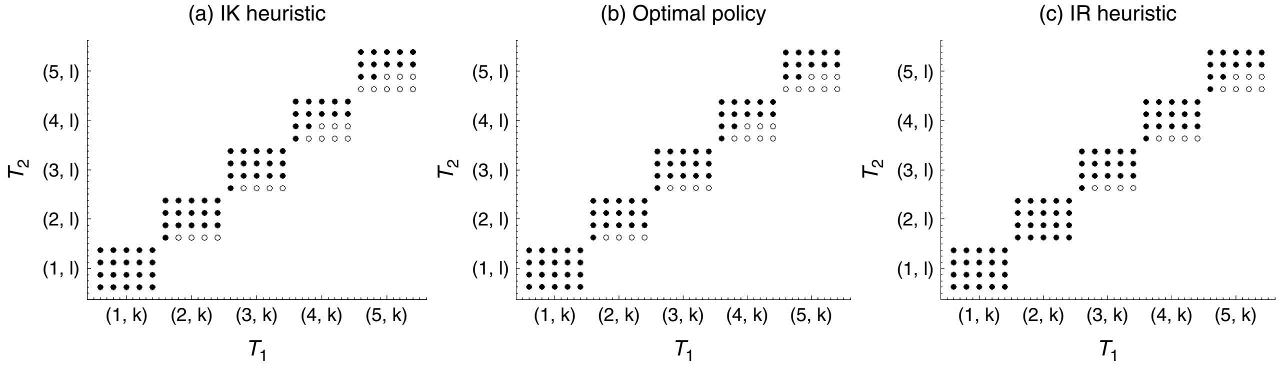
In Figure 8 we illustrate the structure of the optimal policy and of the IK and IR heuristics in a particular problem instance with $I = 2$ products and horizon $H = 5$, in which only one product fits in the knapsack. The figures show the switching curves of the respective policies. Whereas IK is the same as the optimal policy, IR heuristics tend to promote product 2 in more joint states. Note that the joint states in which IR takes suboptimal action are sometimes several at the same time to go. This is a typical situation that we have observed when both I and H are small: IK is optimal in more than 90% problem instances.

These results clearly indicate that the IK heuristic should be the preferred choice in this setting. Note that solving (KP) in the case of 2 or 3 products is a simple task and so it brings virtually no computational overhead with respect to IR.

5.4. Scenarios with Four or Five Products

With respect to the results of experiments with $I = 4$ and $I = 5$ products, we note that they become more computationally

Figure 8. The structure of policies in a problem with parameters $I = 2$, $T_1 = T_2 = 5$, $K_1 = 5$, $K_2 = 4$, $p_1 = 0.5$, $p_2 = 0.4$, $q_1 = 0.65$, $q_2 = 0.7$, $\alpha_1 = \alpha_2 = -1/2$, $\beta = 1$, $W_1 = 16$, $W_2 = 10$, $R_1 = 32$, $R_2 = 24$, $C = 20$.



Notes. Each block of $K_1 \times K_2$ points shows prescribed actions for a fixed number of time period to go, $t = T_1 = T_2$. A circle denotes promoting product 1, and a bullet denotes promoting product 2. The horizontal axis of each block refers to states $(t, 1), (t, 2), \ldots, (t, K_1)$, and the vertical axis of each block refers to states $(t, 1), (t, 2), \ldots, (t, K_2)$.

demanding as the horizon H grows, and so we were only able to run experiments for $H \leqslant 16$. The performance of the IK and IR heuristics is shown in Figures 9 and 10, respectively, where they are contrasted also with the results for $I = 2, 3$. It is interesting to see that IK's mean performance seems to be insensitive to the number of products considered, and remains extremely close to optimal (although there is an approximately linear dependence on the problem horizon).

Figure 9. Mean adjusted relative suboptimality gap of IK heuristic.

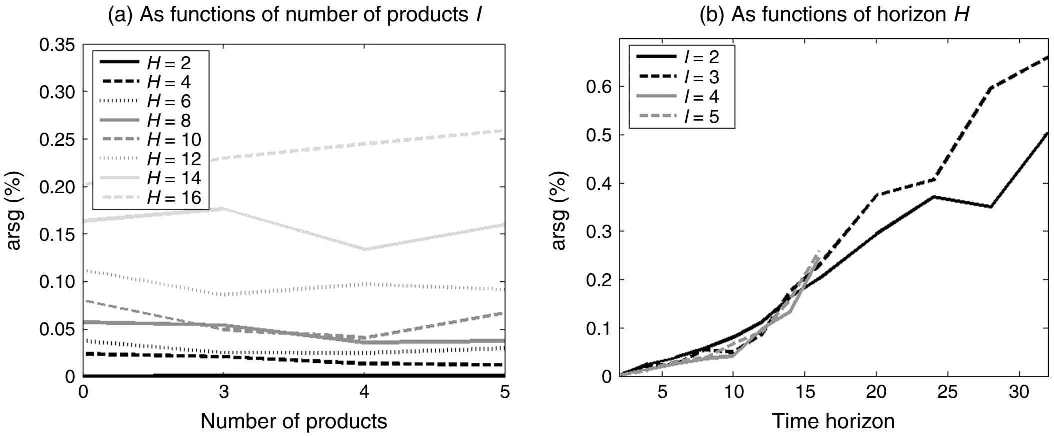
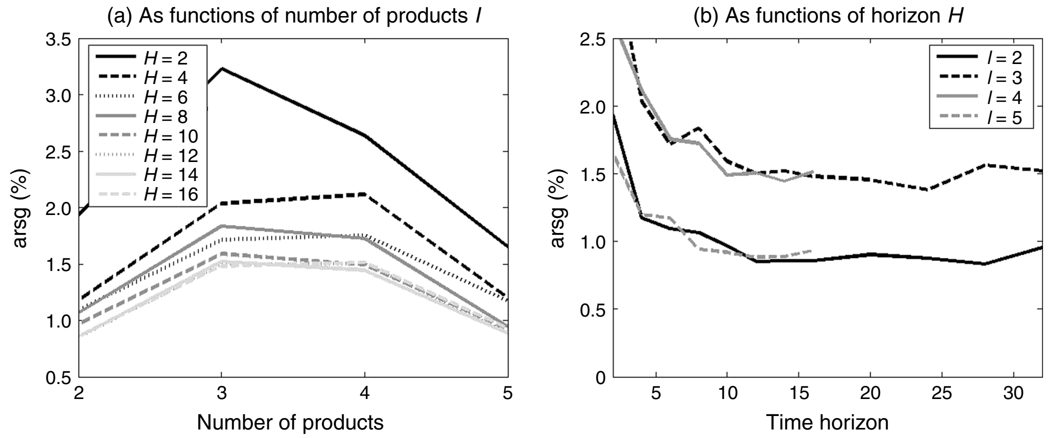


Figure 10. Mean adjusted relative suboptimality gap of IR heuristic.



On the other hand, IR's dependence on the number of products is inconclusive, but its performance improves with a higher horizon and stabilizes relatively quickly (for $H \geqslant 10$).

Finally we note that in these figures we exhibit heuristics with exact index values. Their performance with approximate index values is similar, and slightly inferior for IR. Nevertheless, IK significantly outperforms IR in all these instances.

6. Conclusion

We have introduced a model for the problem of resource allocation to inventories of perishable products and formulated it as an extension of the restless bandit problem. We have designed two index-based heuristics and shown in numerical experiments their nearly optimal performance. We have further derived an efficient algorithm for approximately computing the index values. We believe that we have provided an efficient, implementable solution to this complex problem that is intractable for optimal solution. This may both foster a further research in this direction and be applicable in operations management in different industries.

From the methodological point of view, our approach is a natural generalization of the celebrated (optimal) Gittins index policy for the multiarmed bandit problem (Gittins 1979) (which requires frozen-if-not-allocated assumption, so it cannot model finite-horizon problems), and the Whittle index rule for the restless bandit problem (Whittle 1988) (which assumes binary resource consumption $W_i = 1$, so it does not model heterogeneous product volumes). Indeed, we provide a model, solution approach and a new heuristic for general discretely divisible resource allocation problems, which can be viewed as an analogy of the extension by Glazebrook and Minty (2009) of the Gittins index policy. Note that the indexability property is still not completely understood and considered somewhat *mysterious* (Jacko 2010, Gittins et al. 2011). This paper contributes with the first indexability analysis of an MDP with a general bi-dimensional state space.

A challenge that deserves future research attention is to develop indexability analysis and obtain index values for MDPs where multiple actions are available. This paper gives a step toward such multiaction problems. An interesting extension would therefore be the problem, in which we can promote more than only one unit of each product, as a consequence of removing the Bernoulli arrivals assumption.

We have assumed in our model that the demand is independent across products. This may be a limitation from a product promotion point of view, where correlation between the products' demand often exists. For such a case, the method of the Whittle and Lagrangian relaxation would not lead to a decomposition of the problem into single products, but into independent groups of products. This brings the question of whether there is a way of defining an index with optimality properties; we are not aware of any research in this direction at all. Alternatively, one could propose to heuristically modify the optimal single-product index so that it captures interdependence: to increase it if there are complementary products and to decrease it if there are substitutable products with nonzero inventories. Nevertheless, we believe that in many practical situations, the numbers of complementary and substitutable products can roughly be the same across products, so the index values would be additively modified roughly by the same constant, which means that the index rule would remain roughly the same as in the case with independent products, whereas the index-knapsack heuristic would lead to a slightly different solution (tending to select for promotion a higher number of products if complementary products prevail, and a lower number of products if substitutable products prevail). It would therefore be interesting to investigate the performance of our two general-purpose heuristics as an alternative solution in particular problems with dependent demand when the optimal policy is intractable, e.g., Mahajan and van Ryzin (2001), Bernstein et al. (2011). Our approach should also be contrasted with ad-hoc proposals of indices in choice-model literature, e.g., Golrezaei et al. (2014) and with approximation algorithms, e.g., Levi and Shi (2013).

With respect to scheduling of requests in data centers, the results of our paper can be applied as follows. The user requests $i \in \mathcal{I}$ given as a number of subrequests K_i requiring per-slot processing capacity W_i that must be completed sequentially before a deadline of T_i time slots are available for scheduling. If scheduled on one of the standard servers (or virtual machines), say, in a processor-sharing way, the subrequest is completed with probability $1 - q_i$ within a slot. However, the limited processing capacity of C is available for faster processing, which increases the probability of completion of a subrequest to $1 - p_i$. For each subrequest completed before the deadline, the data center earns an expected revenue R_i, and for those completed after the deadline, $\alpha_i R_i$ is obtained. The solution provided in this paper indicates which requests should be processed by the faster servers in order to maximize the expected revenue of the data center. Finally, we believe that for this application it may be rather appropriate to assume independence in the completion of subrequests.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/opre.2014.1272.

Acknowledgments

The authors are grateful to U. Ayesta and K. Glazebrook for many fruitful discussions and the two anonymous referees for valuable suggestions that helped to improve the presentation and extend the results of the original paper. This research was mostly done during the first author's stay at the Basque Center for Applied Mathematics (BCAM) during January–May 2011 funded by the BCAM internship program. An earlier version of this work was presented at the Prague Economic Meeting in June 2011. Support of VEGA [Grant 1/0381/09] is acknowledged by the first author. Research of the second author was partially supported by [Grant MTM2010-17405] (Ministerio de Ciencia e Innovación, Spain) and [Grant PI2010-2] (Department of Education and Research, Basque Government).

References

Bai R, Kendall G (2008) A model for fresh produce shelf-space allocation and inventory management with freshness-condition-dependent demand. *INFORMS J. Comput.* 20(1):78–85.

Bai R, Burke EK, Kendall G (2008) Heuristic, meta-heuristic and hyper-heuristic approaches for fresh produce inventory control and shelf space allocation. *J. Oper. Res. Soc.* 59:1387–1397.

Bernstein F, Kök AG, Xie L (2011) Dynamic assortment customization with limited inventories. Technical report, Duke University, Durham, NC.

Dance C, Gaivoronski AA (2012) Stochastic optimization for real time service capacity allocation under random service demand. *Ann. Oper. Res.* 193(1):221–253.

Dantzig GB (1957) Discrete-variable extremum problems. *Oper. Res.* 5(2):266–277.

Gittins JC (1979) Bandit processes and dynamic allocation indices. *J. Roy. Statist. Soc., Ser. B* 41(2):148–177.

Gittins JC, Glazebrook K, Weber R (2011) *Multi-Armed Bandit Allocation Indices*, 2nd ed. (John Wiley & Sons, Chichester, UK).

Glazebrook K, Minty R (2009) A generalized Gittins index for a class of multiarmed bandits with general resource requirements. *Math. Oper. Res.* 34(1):26–44.

Glazebrook KD, Hodge DJ, Kirkbride C (2011) General notions of indexability for queueing control and asset management. *Ann. Appl. Probab.* 21:876–907.

Golrezaei N, Nazerzadeh H, Rusmevichientong P (2014) Real-time optimization of personalized assortments. *Management Sci.* Forthcoming.

Goyal SK, Giri BC (2001) Recent trends in modeling of deteriorating inventory. *Eur. J. Oper. Res.* 134:1–16.

Jacko P (2009) Adaptive greedy rules for dynamic and stochastic resource capacity allocation problems. *Medium for Econometric Applications* 17(4):10–16.

Jacko P (2010) *Dynamic Priority Allocation in Restless Bandit Models* (Lambert Academic Publishing, Saarbrücken, Germany).

Jacko P (2013) Resource capacity allocation to stochastic dynamic competitors: Knapsack problem for perishable items and index-knapsack heuristic. *Ann. Oper. Res.* To appear.

Jacko P, Niño-Mora J (2008) Admission control and routing to parallel queues with delayed information via marginal productivity indices. *Proc. Third Internat. Conf. Performance Evaluation Methodologies and Tools (ValueTools).* ACM International Conference Proceeding Series (ACM, New York).

Jacko P, Sansò B (2012) Optimal anticipative congestion control of flows with time-varying input stream. *Performance Evaluation* 69(2):86–101.

Kar S, Bhunia AK, Maiti M (2001) Inventory of multi-deteriorating items sold from two shops under single management with constraints on space and investment. *Comput. Oper. Res.* 28:1203–1221.

Levi R, Shi C (2013) Approximation algorithms for the stochastic lot-sizing problem with order lead times. *Oper. Res.* 61(3):593–602.

Mahajan S, van Ryzin G (2001) Stocking retail assortments under dynamic consumer substitution. *Oper. Res.* 49(3):334–351.

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

Niño-Mora J (2011) Computing a classic index for finite-horizon bandits. *INFORMS J. Comput.* 23(2):254–267.

Pisinger D (2005) Where are the hard knapsack problems? *Comput. Oper. Res.* 32:2271–2284.

Puterman ML (2005) *Markov Decision Processes: Discrete Stochastic Dynamic Programming* (John Wiley & Sons, Hoboken, NJ).

Talluri KT, van Ryzin G (2004) *The Theory and Practice of Revenue Management* (Springer, New York).

Urban TL (1998) An inventory-theoretic approach to product assortment and shelf-space allocation. *J. Retailing* 74(1):15–35.

Weber R, Weiss G (1990) On an index policy for restless bandits. *J. Appl. Probab.* 27(3):637–648.

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

Yang R, Bhulai S, van der Mei R (2011) Optimal resource allocation for multiqueue systems with a shared server pool. *Queueing Systems* 68:133–163.

Darina Graczová is a Ph.D. student in applied mathematics at Comenius University in Bratislava, Slovakia. She was an internship student at the Basque Center for Applied Mathematics, Spain in Jan 2011–May 2011. Her research interests are in stochastic optimization and modelling.

Peter Jacko is a lecturer in the Department of Management Science at Lancaster University Management School, United Kingdom. He was formerly coleader of the Networks research group at the Basque Center for Applied Mathematics, Spain. His main research interests are in modelling and designing tractable and well-performing scheduling and resource allocation solutions to dynamic and stochastic problems, especially those arising in the context of networks and with applications in business administration and in communications engineering.