Ann Oper Res (2015) 231:207–228
DOI 10.1007/s10479-014-1573-y

A dynamic inventory rationing problem with uncertain demand and production rates

Zeynep Turgay · Fikri Karaesmen · E. Lerzan Örmeci

Published online: 18 March 2014

Abstract We investigate the structural properties of a finite horizon, discrete time single product inventory rationing problem, where we allow random replenishment (production) opportunities. In contrast to the standard models of dynamic capacity control in revenue management or production/inventory systems, we assume that the demand/production rates are not known with certainty but lie in some interval. To address this uncertainty, we formulate a robust stochastic dynamic program and show how the structural properties of the optimal policy propagate to the robust counterpart of the problem. Further, we explore how the optimal policy changes with respect to the uncertainty set. We also show that our results can be extended to certain alternative robust formulations.

Keywords Inventory control · Stochastic dynamic programming · Robust formulations

1 Introduction

In many practical optimization problems, the input parameters to the problem are not known with certainty and have to be approximated or estimated from limited data. This, if ignored, may cause significant suboptimality or infeasibility for the optimization problem considered. Robust optimization is a methodology that addresses this problem and has received considerable attention lately Ben-Tal et al. (2009).

Our focus on this paper is on a well-established dynamic capacity control problem in revenue management. Detailed information on revenue management can be found in Talluri and Ryzin (2005). In our problems, it is assumed that there are different customer classes that ran-

Z. Turgay · F. Karaesmen (✉) · E. L. Örmeci
Department of Industrial Eng., Koç University, Istanbul, Turkey
e-mail: fkaraesmen@ku.edu.tr

Z. Turgay
e-mail: zturgay@ku.edu.tr

E. L. Örmeci
e-mail: lormeci@ku.edu.tr

domly generate demands for a single-product inventory system. The classes are differentiated according to their rewards. In order to maximize the expected total profit, the system manager can ration the inventory by rejecting some of the arriving demands thereby reserving some inventory for future more profitable arrivals. The manager has to decide dynamically whether an incoming demand must be admitted (satisfied from stock) or not, taking into account the existing inventory position and the time left until the end of the horizon. If replenishment is allowed, the manager also has to determine when to produce.

The admission control literature for inventory problems is very rich and in this paper, we revisit two well-known specific problems. The first problem is a stock rationing problem for a production/inventory system that is first suggested by Ha (1997b). In this case, both the production system and the rationing system are parts of the model. In particular, the production system is modeled as a single-server queue that processes items one-by-one. The manager has to decide dynamically whether to produce an additional item in addition to the rationing decision for the incoming demand. Several papers have studied this problem and characterized the structure of the optimal policy under different assumptions [see Vericourt et al. (2002) or Gayon et al. (2009)]. The second related problem we consider is the dynamic admission control problem of Lautenbacher and Stidham (1999) for a single-product inventory which may be considered as a subcase of the former one. In the context of revenue management, this problem is typically analyzed under the assumption of a finite horizon and replenishments are not allowed [see Talluri and Ryzin (2005) or Aydin et al. (2009)].

Most papers on these topics make the assumption that demand and production rates are known with certainty except for a few exceptions (Birbil et al. 2009; Lan et al. 2008). We consider the case where the demand arrival and production rates are not known with certainty but lie in an interval. This is a relevant case since, in practice, demand and processing rates have to be estimated from limited data or are supported by subjective assessment which are inaccurate. For the case without replenishment, we contribute to the recent literature addressing parameter uncertainty by investigating the structure of optimal policies. For the case with replenishment, to our knowledge, the parameter uncertainty issue has not been addressed before.

There are several ways of incorporating input uncertainty and developing robust solutions in an optimization problem. Among them, the minimax approach—also known as absolute robust decision—leads to tractable stochastic dynamic programming formulations. This approach considers a dynamic game between the controller and Nature. Such a formulation of a Markov decision process with an uncertain transition probability distribution goes back to Satia and Lave (1973) who proposed a solution by a policy iteration approach. More recently, Nilim and Ghaoui (2005) and Iyengar (2005) simultaneously studied robust stochastic dynamic programs and established the existence of a robust Bellman recursion whose solution yields the robust value function and the corresponding optimal policy. In addition, both papers emphasized that under appropriate choice of uncertainty sets, the additional complexity brought by the robust formulation is reasonable if the standard formulation has a tractable solution. Following such a formulation in the context of robust demand admission control, the controller's aim is to maximize the expected profit by choosing the allowable actions (admission and production), whereas Nature tries to minimize the expected profit by choosing the worst-possible parameters (arrival and production rates) and acts upon observing the controller choice. This formulation is known as the robust counterpart of the classical problem. The robust optimal policy designates the policy which yields the highest expected profit after minimization by Nature.

In order to represent the connection between the real problem data and the prior estimations, a number of uncertainty models have been proposed. One relatively simple uncertainty model is to define uncertainty intervals where parameters are allowed to lie between a lower bound and an upper bound Satia and Lave (1973). Ben-Tal et al. (2009) and Nilim and Ghaoui (2005) propose a number of more sophisticated yet tractable uncertainty sets. Some recent applications of such uncertainty sets include Lim and Shanthikumar (2007), Jain et al. (2010) which employ entropy based models of uncertainty in robust dynamic pricing problems. We employ the simpler model of interval uncertainty in this paper because our focus is on exploring the structure of optimal policies rather than developing efficient uncertainty sets or fitting data to existing uncertainty sets. Bertsimas and Thiele (2006) consider robust formulations of some deterministic inventory control problems considering a known demand that is subject to interval uncertainty in each period. Some recent papers investigate the admission control problem under demand rate uncertainty. An absolute robust approach for both the static and dynamic versions of this problem is suggested by Birbil et al. (2009) where they employ an ellipsoidal model of uncertainty. It is shown that this uncertainty model leads to tractable solutions of the problem. By considering an interval uncertainty model, we complement the results of Birbil et al. (2009) by obtaining additional properties of optimal robust policies. Finally, Lan et al. (2008) and Ball and Queyranne (2009) propose and explore an alternative approach for addressing uncertainty based on a competitive analysis method.

In contrast to the above papers we obtain results on the structure of the optimal policy for both the cases with or without replenishment. We also explore how the optimal policies change with respect to uncertainty sets. The dynamic queueing and inventory control literature has a strong tradition in characterizing the structure of optimal policies. This is in part due to the computational efficiency of structured policies. However, the main reason for looking for structured policies is that they are usually expressed in a few parameters and tend to be easy to understand, communicate and implement. There are known general approaches to investigate the structure of the solution of a stochastic dynamic program [for instance Koole (1998, 2006), Cil et al. (2009)]. The current paper can be seen as an exploration such properties in the context of a robust stochastic dynamic program.

The organization of the paper is as follows. In Sect. 2, we analyze the single item inventory problem with random production referred to as the inventory rationing problem. In Sect. 3, we focus on an important special case of the problem in Sect. 2 where there are no replenishment opportunities. This model constitutes the building block for most dynamic revenue management problems. In Sect. 4, we present some numerical results and explore some of the quantitative trade-offs. Finally, our conclusions are provided in Sect. 5.

2 Inventory rationing for a production/inventory system

2.1 The nominal problem

In this section, we consider the inventory rationing problem for a production/inventory system. In this problem, a production system with limited capacity produces a single item that is demanded by different classes of customers. The demands that arrive when inventory is unavailable are assumed to be lost. The customer classes are differentiated by their arrival rates and their profit margins. Because the profit margins are class-dependent, it may be profitable for the firm to reject arriving demands from certain customer classes in order to reserve inventory for future more profitable arrivals. The decision problem of the firm is then to determine, based on the inventory on-hand, whether to produce an additional item to increase the inventory level and whether to admit an arriving demand from a given class.

We investigate a discrete-time model that is motivated by earlier works for the stock rationing problem for an M/M/1 make-to-stock queue (Ha 1997a, c; Vericourt et al. 2002; Gayon et al. 2009). This is a model that has received significant attention and was extended in several directions.

Ha's (1997b) inventory rationing problem with lost sales can be described as follows. There are n classes of customers whose demands arrive according to independent Poisson processes with rate λ_i ($i = 1, 2, ..., n$). A single server whose processing time is exponentially distributed with rate μ produces items one by one. If a demand of class i is admitted when there is at least one unit of inventory on hand, it is immediately satisfied and a class-dependent instant reward of R_i is obtained. If inventory is empty, all arriving demands are assumed to be lost. The classes are ordered such that if $i < j$ then $R_i > R_j$. The fictitious event is introduced to the problem which corresponds to the probability of no arrival and no product completion with λ_{n+1} and can also be considered as a special class with $R_{n+1} = 0$. At any time t, the inventory level is denoted by $X(t)$ (where $X(t) \in \mathcal{Z}^+$) and the inventory holding cost rate is $h(X(t))$ per unit of time. The holding cost function $h(x)$ is increasing and convex in x. Ha (1997b) considers an infinite-horizon discounted profit maximization objective with a discount rate of α.

The actions in this problem correspond to production control where it is assumed that the processor can be stopped and started at any time t and demand admission control where demands are admitted or rejected from the system upon arrival. Ha (1997b) shows that, after uniformization, the equivalent discrete time problem can be expressed as follows. Let $\gamma = \bar{\mu} + \sum_{i=1}^{n+1} \bar{\lambda_i} + \bar{\alpha}$ be the uniformization rate which can be set to 1 without loss of generality and by setting $\mu = \bar{\mu}/\gamma$, $\lambda_i = \bar{\lambda_i}/\gamma$ and $\alpha = \bar{\alpha}/\gamma$. Here, $\bar{\mu}$, $\bar{\lambda_i}$ and $\bar{\alpha}$ are the rates of the corresponding events in the original system, whereas μ, λ_i and α are the probabilities of these events in the uniformized model. We let $v_t(x)$ denote the expected optimal profit with t transitions to go until the end of the horizon. We obtain:

$$\begin{aligned} v_t(x) &= \mu \max\{v_{t-1}(x+1), v_{t-1}(x)\} \\ &\quad + \sum_{i=1}^{n+1} \lambda_i \max\{v_{t-1}(x-1) + R_i v_{t-1}(x)\} - h(x) \quad \text{if } x > 0, \end{aligned}$$

and

$$v_t(0) = \mu \max\{v_{t-1}(1), v_{t-1}(0)\} + \sum_{i=1}^{n+1} \lambda_i v_{t-1}(0) - h(0), \qquad \text{if } x = 0. \tag{1}$$

Let us denote the difference function by $\Delta v(x) = v(x) - v(x-1)$. Clearly, a class-$i$ customer is accepted at stage t if and only if $R_i - \Delta v_{t+1}(x) \geq 0$. Setting $\alpha' = (1-\alpha)$, Eq. (1) can alternatively be represented as follows:

$$\begin{aligned} v_t(x) &= \mu(\Delta v_{t-1}(x+1))^+ \\ &\quad + \sum_{i=1}^{n+1} \lambda_i \, (R_i - \Delta v_{t-1}(x))^+ + \alpha' v_{t-1}(x) - h(x), \quad \text{if } x > 0, \\ v_t(x) &= \mu(\Delta v_{t-1}(1))^+ + \alpha' v_{t-1}(0) - h(0), \quad \text{if } x > 0, \end{aligned}$$

where a^+ denotes $\max(0, a)$ for any real number a and $v_0(x) = 0$ for all x.

Ha (1997b) established that the value function $v_t(x)$ is concave for all finite t and for the infinite horizon value function as t tends to infinity. This implies that the optimal inventory rationing policy is of threshold type. In each period t, there is an admission threshold for each class. If the inventory available in the period is above the threshold, then the arriving demandis admitted. In addition, the optimal production policy is of target level type. There is a target production level below which the system should produce and at or above which the system should stop production. These results can be extended to a number of more complicated cases including Erlang processing times (Ha 2000; Gayon et al. 2009), batch arrivals (Huang and Iravani 2008; Cil et al. 2009). A different stream of literature considers the related multi-item production scheduling problem (Vericourt et al. 2000; Bertsimas and Paschalidis 2001).

One interesting extension to the main structural results is to compare the optimal production and inventory rationing systems that have different parameters for their demand and processing probabilities. Let the vector of arrival and production probabilities be given by $(\boldsymbol{\lambda}, \mu) = (\lambda_1, ..., \lambda_{n+1}, \mu)$.

To compare different input vectors, let us employ the following partial order. An arrival probability vector $\boldsymbol{\lambda}$ is said to be *preferred* (denoted by the $\succeq$ operator) over another $\boldsymbol{\lambda}'$ if it receives higher ordered classes with higher probability:

$$\boldsymbol{\lambda} \succeq \boldsymbol{\lambda}' \ \Leftrightarrow \ \sum_{i=1}^{k} \lambda_i \geq \sum_{i=1}^{k} \lambda_i' \ \text{ for } k = 1, 2, \ldots, n+1.$$

Similarly, a production probability μ is said to be *preferred* over another (μ'), if $\mu \geq \mu'$.

Theorem 1 *Consider two problems that are identical in all other parameters except their arrival and production probabilities in period t. Let $(\boldsymbol{\lambda}, \mu)$ and $(\boldsymbol{\lambda}', \mu')$ be two arrival and production probability vectors, and $v_t(x)$ and $v_t'(x)$ be the corresponding value functions respectively:*

If $\boldsymbol{\lambda} \succeq \boldsymbol{\lambda}'$ and $\mu \preceq \mu'$ then $\Delta v_t(x) \geq \Delta v_t'(x)$ for all x, t.

Proof The proof can be found in Appendix 1. □

Theorem 1 implies that optimal admission and production thresholds can be compared if the above preference orders hold. A more preferred arrival vector (together with a lower or equal production rate) leads to higher inventory targets and higher admission thresholds in each period, symmetrically a higher production rate (together with a less preferred or equal arrival vector) leads to lower inventory targets and lower admission thresholds.

2.2 The robust problem

In this section, we consider the discrete-time model in Sect. 2.1 but assume that the arrival and production rates (equivalently probabilities) are not known with certainty. We first consider a model of interval uncertainty for the finite horizon case where each rate parameter is estimated independently and is assumed to lie in an interval between upper and lower bounds rather than taking a single value.

In order to model decision making under such an uncertainty, we employ the max–min formulation and formulate a robust dynamic program. The robust dynamic programming framework with transition uncertainty was established by Nilim and Ghaoui (2005) and Iyengar (2005) and a revenue management example is studied in Birbil et al. (2009). Under the max–min robust formulation, the controller plays a game against nature. It is assumed that nature selects the worst possible probability distribution from the uncertainty set in each state and time after observing the controller's action. To achieve this, we let nature choose an independent arrival vector for each state, time and action as in Nilim and Ghaoui (2005).

Let us now define the action space of the problem. Let $\mathbf{a} = (a_0, a_1, a_2, \ldots, a_n)$, where a_0 describes the production decision (i.e. $a_0 = 0$ denotes controller's action 'not to produce' and $a_0 = 1$ denotes controller's action 'to produce') and similarly a_i for $i = 1, 2, ..., n$ describes the admission decision for class i (i.e. $a_i = 0$ denotes controller's action 'to reject' a class i demand and $a_i = 1$ denotes controller's action 'to admit'). $\mathcal{A}$ denotes the set of admissible actions (combinations of production and admission decisions for each class) of the controller. We redefine the arrival and production completion probabilities as $\lambda_{i,t}(x, \boldsymbol{a})$ and $\mu_t(x, \boldsymbol{a})$ which denote the probability of a class-i arrival and production completion respectively, at time t and when the system is in state x and takes action $\boldsymbol{a}$. Let an arrival probability vector be $\boldsymbol{\lambda}_t(x, \boldsymbol{a}) = (\lambda_{1,t}(x, \boldsymbol{a}), ..., \lambda_{n+1,t}(x, \boldsymbol{a}))$. We assume that the combined arrival and production completion probability vector $\boldsymbol{p}_t(x, \boldsymbol{a}) = (\boldsymbol{\lambda}_t(x, \boldsymbol{a}), \mu_t(x, \boldsymbol{a}))$ belongs to an uncertainty set which does not depend on state x and action $\boldsymbol{a}$. This appears to be a reasonable assumption in the inventory rationing context.

Let us define $\mathcal{P} \neq \emptyset$ an interval uncertainty set for the demand arrival—production probability vector:

$$\mathcal{P} = \left\{ \boldsymbol{z} = (z_0, \ldots, z_{n+1}) : \quad 0 \leq \underline{z}_i \leq z_i \leq \bar{z}_i, \ 0 \leq q \leq \sum_{i=1}^{n} z_i \leq 1 \right\},$$

where $\underline{z}_i$ and $\bar{z}_i$ upper and lower bounds on individual event probabilities and q is a lower bound on the total probability of arrival and production.

We define the robust value function $w_t(x)$ [see for example Nilim and Ghaoui (2005)] as:

$$\begin{aligned} w_t(x) &= \max_{\boldsymbol{a} \in \mathcal{A}} \min_{\boldsymbol{p}_t(x,\boldsymbol{a}) \in \mathcal{P}} \{\mu_t(x, \boldsymbol{a})\,(w_{t-1}(x + a_0)) \\ &\quad + \sum_{i=1}^{n+1} \lambda_{i,t}(x, \boldsymbol{a})\,(a_i R_i + w_{t-1}(x - a_i))\} - h(x), \quad \text{if } x > 0 \qquad (2) \\ w_t(x) &= \max_{\boldsymbol{a} \in \mathcal{A}} \min_{\boldsymbol{p}_t(x,\boldsymbol{a}) \in \mathcal{P}} \{\mu_t(0)\,(w_{t-1}(a_0))\} - h(0) \ \text{ if } x = 0. \end{aligned}$$

Note that the uncertainty set includes additional constraints representing sample information in addition to the default constraints $\lambda_{i,t}(x, \boldsymbol{a}), \mu_t(x, \boldsymbol{a}) \geq 0$ for all i and t and $\sum_{i=1}^{n} \lambda_{i,t}(x, \boldsymbol{a}) + \mu_t(x, \boldsymbol{a}) \leq 1$ for all t.

2.3 Structural properties

In order to obtain some structural properties, we first begin by noting that the order of max and min in Eq. (2) can be switched without loss of generality. A proof of this property can be found in Turgay et al. (2013). The proof is based on the below intuition: Let $(\boldsymbol{\lambda}', \mu') \in \mathcal{P}$ be any parameter probability vector, then for all such vectors, the optimal action at stage t must be such that if $w_{t-1}(x + 1) \geq w_{t-1}(x)$ then it is optimal to produce at state x, and if $w_{t-1}(x - 1) + R_i \geq w_{t-1}(x)$ then it is optimal to admit a class i demand at state x. Hence, the optimal action is independent of Nature's posteriori choice of input probability vector. Using this property we can rewrite Eq. 2 as:

$$\begin{aligned} w_t(x) &= \min_{\boldsymbol{p}_t(x) \in \mathcal{P}} \left\{ \mu_t(x)\,(\Delta w_{t-1}(x + 1))^+ \right. \\ &\quad \left. + \sum_{i=1}^{n+1} \lambda_{i,t}(x)\,(R_i - \Delta w_{t-1}(x))^+ \right\} + \alpha' w_{t-1}(x) - h(x), \ \text{if } x > 0 \qquad (3) \\ w_t(0) &= \min_{\boldsymbol{p}_t(0) \in \mathcal{P}} \left\{ \mu_t(0)\,(\Delta w_{t-1}(1))^+ \right\} + \alpha' w_{t-1}(0) - h(0), \ \text{if } x = 0, \qquad (4) \end{aligned}$$

where $\boldsymbol{a}$ is dropped from the notation of parameters. Note that $\boldsymbol{p}_t(x) = (\boldsymbol{\lambda}_t(x), \mu_t(x))$

Theorem 2 *The robust value function $w_t(x)$ is concave in x for all t.*

Proof Suppose that $\boldsymbol{\lambda}_t(x)$ and $\mu_t(x)$ are the optimal solutions of the Nature for state x. We use an induction argument and assume that $w_{t-1}(x)$ is concave in x. Next, we have to show that under this assumption $w_t(x)$ preserves concavity. Using the concavity assumption, the following inequality holds if the arrival and rates are identical at states $x-1$, x and $x+1$ and are equal to $\boldsymbol{\lambda}_t(x)$ and $\mu_t(x)$ using the existing results (Ha 1997b; Cil et al. 2009):

$$\begin{aligned}
&\mu_t(x)\{\Delta w_{t-1}(x+1)\}^+ + \textstyle\sum_{i=1}^{n+1} \lambda_{i,t}(x)\{R_i - \Delta w_{t-1}(x)\}^+ + \alpha' w_{t-1}(x)\\
&\quad - h(x) \geq 1/2\{\mu_t(x)\{\Delta w_{t-1}(x)\}^+ + \textstyle\sum_{i=1}^{n+1} \lambda_{i,t}(x)\{R_i - \Delta w_{t-1}(x-1)\}^+\\
&\quad + \ \alpha' w_{t-1}(x-1) - h(x-1)\} + \ 1/2\{\mu_t(x)\{\Delta w_{t-1}(x+2)\}^+\\
&\quad + \textstyle\sum_{i=1}^{n+1} \lambda_{i,t}(x)\{R_i - \Delta w_{t-1}(x+1)\}^+\\
&\quad + \ \alpha' w_{t-1}(x+1) - h(x+1)\}.
\end{aligned}$$

Now let us relax the assumption that the arrival and production rates are equal for all states. Because Nature's objective is to minimize the robust value function $w_t(x)$, the following holds:

$$\begin{aligned}
&\mu_t(x)\{\Delta w_{t-1}(x+1)\}^+ + \textstyle\sum_{i=1}^{n+1} \lambda_{i,t}(x)\{R_i - \Delta w_{t-1}(x)\}^+ + \alpha' w_{t-1}(x)\\
&\quad - h(x) \geq 1/2\{\mu_t(x-1)\{\Delta w_{t-1}(x)\}^+ + \textstyle\sum_{i=1}^{n+1} \lambda_{i,t}(x-1)\{R_i - \Delta w_{t-1}(x-1)\}^+\\
&\quad + \ \alpha' w_{t-1}(x-1) - h(x-1)\} + \ 1/2\{\mu_t(x+1)\{\Delta w_{t-1}(x+2)\}^+\\
&\quad + \textstyle\sum_{i=1}^{n+1} \lambda_{i,t}(x+1)\{R_i - \Delta w_{t-1}(x+1)\}^+\\
&\quad + \ \alpha' w_{t-1}(x+1) - h(x+1)\}.
\end{aligned}$$

This establishes that $w_{t-1}(x)$ is concave in x. □

Theorem 2 implies that as in the standard inventory rationing problem of Ha (1997b), the optimal demand admission policy of the robust problem is of threshold type. Hence, in each period t and for each class i, there is an admission threshold $l_{i,t}$. Similarly, there is a target level S_t for each period t, such that the controller stops production if the inventory on hand reaches this level. This is a surprising result because the event probabilities in the robust problem are state (i.e. inventory level) dependent and time dependent by definition. While concavity can be extended to the time-dependent case in the standard problem, state dependence is very problematic and concavity does not hold in general when event probabilities are state dependent. Theorem 2 establishes that concavity survives when state dependent event probabilities are induced by the robust formulation. An illustrative example to such state dependence is provided in Sect. 4.

We should also note that while concavity of the robust value function can be established, there are no corresponding results for supermodularity/submodularity of the value function $w_t(x)$ in x, t or the monotonicity of thresholds over time.

Finally, if the uncertainty in event probabilities pertains to only one type of operator, i.e. either admission operators only or the production operator only, Nature's solution is independent of the state x for all stages t. We explore this further in the next section.

2.4 Behavior of the optimal policy for nested uncertainty sets

In this subsection, we explore the effects of increasing or decreasing uncertainty on the optimal robust value function and the optimal policy. The results we provided for the revenue management problem of Sect. 3 also hold for this problem under certain additional conditions.

Lemma 1 *Consider two problems that are identical in their parameters except their uncertainty sets that represent their arrival probabilities at stage t. Let $\mathcal{P}$ and $\mathcal{P}^\varepsilon$ be two uncertainty sets, $\boldsymbol{\lambda}_t$ and $\boldsymbol{\lambda}_t^\varepsilon$ be the corresponding Nature's optimal solutions respectively. Further suppose that $\mathcal{P} \subseteq \mathcal{P}^\varepsilon$ then:*

1. *If only arrival probabilities are uncertain (i.e. μ_t is exactly known), then $\boldsymbol{\lambda}_t \succeq \boldsymbol{\lambda}_t^\varepsilon$.*
2. *If only production probability is uncertain (i.e. λ_i is exactly known for all customer classesi), then $\mu_t \succeq \mu_t^\varepsilon$.*

Proof Please note that in either way (1 and 2) solution of Nature is independent of state x.

Proof of Part 1: According to definition of $\mathcal{P}$, $\mathcal{P}^\varepsilon \supseteq \mathcal{P}$ if and only if all of the following conditions hold:

1. $q^\varepsilon \leq q$
2. $\underline{z}_i^\varepsilon \leq \underline{z}_i$
3. $\bar{z}_i^\varepsilon \geq \bar{z}_i$

The corresponding solution of Nature is given in Theorem 3. For the first case the inequality is clear since $\sum_{i=1}^{i=k} \lambda_{i,t}^\varepsilon \leq \sum_{i=1}^{i=k} \lambda_{i,t}$ for all k. For the second and third cases, solution assigns probabilities in increasing order of rewards, thereby implying higher probabilities to lower revenue classes for $\mathcal{P}^\varepsilon$. Hence,

$$\sum_{i=k}^{n+1} \lambda_{i,t}^\varepsilon \geq \sum_{i=k}^{n+1} \lambda_{i,t}, \quad \text{for all } k$$

which implies that $\boldsymbol{\lambda}_t \succeq \boldsymbol{\lambda}_t^\varepsilon$.

Proof of Part 2: Proof of this part is more obvious since $\mu_{i,t}^\varepsilon \leq \mu_{i,t}$. □

Corollary 1 *Consider two problems that are identical in their parameters except their uncertainty sets that represent their arrival probabilities at stage t. Let $\mathcal{P}$ and $\mathcal{P}^\varepsilon$ be two uncertainty sets, $w_t(x)$ and $w_t^\varepsilon(x)$ be the corresponding value functions respectively. If $\mathcal{P} \subseteq \mathcal{P}^\varepsilon$ then:*

1. *$w_t(x) \geq w_t^\varepsilon(x)$ for all t, x,*
2. *(a) If only the arrival probabilities are uncertain, then $\Delta w_t(x) \geq \Delta w_t^\varepsilon(x)$ for all t, x,*
   (b) If only the production rate is uncertain, then $\Delta w_t(x) \leq \Delta w_t^\varepsilon(x)$ for all t, x.

Proof (1) This can be formally proven as in Paschalidis and Kang (2008).
(2) In order to prove the second part we use similar arguments to Theorem 1 and the complete proof is included in Appendix 1. □

The two properties of Corollary 1 have the following implications. As we enlarge the uncertainty set, the optimal robust value function decreases. Besides, if arrival rates are uncertain but the production rate is fixed, enlarging the uncertainty set part of the uncertainty set leads to lower optimal admission thresholds. Likewise, optimal production target levels decrease when the uncertainty set is enlarged. Finally, if the arrival rates are fixed, as the uncertainty set representing the production rates is enlarged, both the optimal admission thresholds and the production target levels increase.

Next, we investigate the structure of optimal policies under a more general robust dynamic programming formulation. The so-called S-Robust Policy framework was proposed by Xu and Mannor (2012) who propose a weighted optimization approach between multiple uncertainty sets that have a nested structure. In particular, in this approach, it is assumed that the transition probability vector belongs to a concentration set $\mathcal{P}^{\mho} \subseteq \mathcal{P}$ with probability δ

and it belongs to the larger set $\mathcal{P}$ with probability of 1, for all t. Here, the concentration set $\mathcal{P}^{\mho}$ can be viewed as a prior distribution and δ is a measure of reliance on that distribution. The concentration set weights could be represented as a vector $\boldsymbol{\delta} = (\delta_1, \delta_2, \delta_3, ..., \delta_T)$ if δ is allowed to vary between stages.

Xu and Mannor define an S-robust policy as the outcome of the following equation of optimality for $x > 0$ is:

$$\begin{aligned} w_t(x) = \max_{a \epsilon A} \Bigg\{ \delta_t \min_{\boldsymbol{p}_t^{\mho}(x,\boldsymbol{a}) \in \mathcal{P}^{\mho}} \Bigg[\mu_t^{\mho}(x)\,(w_{t-1}(x+a_0)) + \sum_{i=1}^{n+1} \lambda_{i,t}^{\mho}(x)\,(a_i R_i + w_{t+1}(x-a_i)) \Bigg] \\ + (1-\delta_t) \min_{\boldsymbol{p}_t(x,\boldsymbol{a}) \in \mathcal{P}} \Bigg[\mu_t(x)\,(w_{t-1}(x+a_0)) \\ + \sum_{i=1}^{n+1} \lambda_{i,t}(x)\,(a_i R_i + w_{t+1}(x-a_i)) \Bigg] \Bigg\} - h(x) \end{aligned}$$

The above equation is rewritten as:

$$\begin{aligned} w_t(x) = \delta_t \min_{\boldsymbol{p}_t^{\mho}(x) \in \mathcal{P}^{\mho}} \Bigg[\mu_t^{\mho}(x)\,(\Delta w_{t-1}(x+1))^+ + \sum_{i=1}^{n+1} \lambda_{i,t}^{\mho}(x)(R_i - \Delta w_{t+1}(x))^+ \Bigg] \\ + \ (1-\delta_t) \min_{\boldsymbol{p}_t(x) \in \mathcal{P}} \Bigg[\mu_t(x)\,(\Delta w_{t-1}(x+1))^+ + \sum_{i=1}^{n+1} \lambda_{i,t}(x)(R_i - \Delta w_{t+1}(x))^+ \Bigg] \\ + \ \alpha' w_{t-1}(x) - h(x). \end{aligned}$$

The next corollary establishes that robust value functions are monotone with respect to the reliance weight vector $\boldsymbol{\delta}$.

Corollary 2 *Consider two problems that are identical in their parameters except their reliance weight factors at stage t. Let δ^1 and δ^2 be two reliance weight factors, $w_t^1(x)$ and $w_t^2(x)$ be the corresponding value functions respectively. If ${\delta^1}_t \geq {\delta^2}_t$ then: t:*

1. *$w_t^1(x) \ \geq \ w_t^2(x)$ for all t, x,*
2. *(a) If only the arrival probabilities are uncertain, then $\Delta w_t^1(x) \ \geq \ \Delta w_t^2(x)$ for all t, x,*
   (b) If only the production rate is uncertain, $\Delta w_t^1(x) \ \leq \ \Delta w_t^2(x)$ for all t, x.

Proof The results follow from Corollary 1. □

At this point, it is useful to make some remarks. First, production costs can be added and/or multiple servers can be included to the model by using event based approach (Koole 1998, 2006; Cil et al. 2009) without much difficulty. In addition to this, we may let uncertainty set $\mathcal{P}$ to vary between stages without any violation to our established results. Finally, let us briefly discuss the infinite horizon extension. Iyengar (2005) and Nilim and Ghaoui (2005) establish that the respective controller and nature policies are stationary for the infinite horizon problem. Moreover, Nilim and Ghaoui (2005) show that the optimal value function of the infinite horizon problem with a discounted cost function can be obtained as the unique limit of the finite horizon problem. This suggests that the optimal policy structure can be extended to the infinite horizon case.

3 A robust dynamic revenue management problem

3.1 The nominal problem

Let us focus on a special case of the inventory rationing model in Sect. 2 where replenishments are not allowed. This special case is worthy of an additional discussion because it corresponds to a typical revenue management problem. Moreover, in this case we allow the probability of demand arrivals to depend on time t. A given inventory is to be sold over a finite time horizon to multiple customer classes with random demand and no replenishment opportunities. This is the typical situation for standard formulations in revenue management [see for example Lautenbacher and Stidham (1999), Talluri and Ryzin (2005) or Aydin et al. (2009)].

We use the identical notation as in Sect. 2. The nominal problem has the following optimality equation:

$$v_t(x) = \sum_{i=1}^{n} \lambda_{i,t} \max\{R_i + v_{t-1}(x-1), v_{t-1}(x)\} + \lambda_{n+1,t} v_{t-1}(x), \tag{5}$$

We can alternatively express (5) as follows:

$$v_t(x) = \sum_{i=1}^{n} \lambda_{i,t} \left(R_i - \Delta v_{t-1}(x)\right)^+ + v_{t-1}(x), \tag{6}$$

Because the above problem is a special case of the inventory rationing problem of Sect. 2, all structural results reported in that section hold for this version too. In addition to being concave Lautenbacher and Stidham (1999), $v_t(x)$ is also supermodular in t and x, i.e. $\Delta v_t(x) \geq \Delta v_{t-1}(x)$ for all t, x [see Talluri and Ryzin (2005) or Aydin et al. (2009)].

3.2 The robust problem

In this section, we focus on a robust formulation that takes into account arrival uncertainty for the discrete time revenue management formulation. Let us assume that the arrival probabilities—which may depend on x- $\lambda_{i,t}(x)$ are not known with certainty but are estimated to lie in some uncertainty set $\mathcal{P}_t$ in each period t, where:

$$\mathcal{P}_t = \left\{ \mathbf{y} = (y_0, \ldots, y_{n+1}) : \quad 0 \leq \underline{y}_{i,t} \leq y_i \leq \bar{y}_{i,t}, \; 0 \leq q_t \leq \sum_{i=1}^{n} y_i \leq 1 \right\},$$

where $\underline{y}_{i,t}$ and $\bar{y}_{i,t}$ are respectively lower and upper bounds on the arrival probability and q_t is a lower bound on the minimum total probability of demand arrival.

$\mathcal{P}_t$ is a fairly standard interval uncertainty set for a probability vector. However, some of the results we present in this section can be extended to a modified uncertainty set $\mathcal{C}_t \subseteq \mathcal{P}_t$ where $\mathcal{C}_t \neq \emptyset$ is defined as follows:

$$\mathcal{C}_t = \Bigg\{ \mathbf{y} = (y_1, \ldots, y_{n+1}) : \quad 0 \leq \underline{y}_{i,t} \leq y_i \leq \bar{y}_{i,t}, \\ \sum_{i=1}^{n} b_i y_{i,t} \geq Q_t, \; 0 \leq q_t \leq \sum_{i=1}^{n} y_i \leq 1 \Bigg\},$$

where Q_t is a lower bound on a linear combination of the decision variables y_i. In particular, using this constraint and taking $b_i = R_i$ one can bound the expected reward per stage which is useful for revenue management applications.

Given the uncertainty set $\mathcal{P}_t$, the robust value function, for $x > 0$, is given by:

$$w_t(x) = \max_{\boldsymbol{a}\in\mathcal{A}} \min_{\boldsymbol{\lambda}_t(x,\boldsymbol{a})\in\mathcal{P}_t} \left\{ \sum_{i=1}^{n} \lambda_{i,t}(x,\boldsymbol{a})\,(a_i R_i + w_{t-1}(x-a_i)) \right\}, \tag{7}$$

where we take the boundary conditions as $w_T(x) = 0$ for all x and $w_t(0) = 0$ for all t.

3.3 Structural properties

We first by presenting a property that does not necessarily hold when replenishments are allowed but facilitates the structure of the problem for the case without replenishments.

Theorem 3 *Consider the uncertainty set $\mathcal{P}_t$, then Nature's optimal choice of probability distribution can be obtained by a simple rule and is identical for all states at all times: $\boldsymbol{\lambda}_t(x) = \boldsymbol{\lambda}_t$ for all x.*

Proof Consider Nature's problem for a given x and t, which is a Linear Program with decision variables $\boldsymbol{\lambda}_t(x)$ and objective function coefficients $(R_1 - \Delta w_t(x))^+$, $(R_2 - \Delta w_t(x))^+$, ..., $(R_n - \Delta w_t(x))^+$. Since $R_1 \geq R_2 \geq ... \geq R_n$, $(R_1 - \Delta w_t(x))^+ \geq (R_2 - \Delta w_t(x))^+ \geq ... \geq (R_n - \Delta w_t(x))^+$ for all t, x. Please note that the problem can be also represented by the following equation through a transformation of the uncertainty set $\mathcal{P}_t$ to $\Delta\mathcal{P}_t$ (See Appendix 2:

$$\begin{aligned} w_t(x) = &\sum_{i=1}^{n} \underline{\lambda}_{i,t}(x)\,(R_i - \Delta w_{t-1}(x))^+ \\ &+ \min_{\Delta\boldsymbol{\lambda}_t(x)\in\Delta\mathcal{P}_t} \left\{ \sum_{i=1}^{n} \Delta\lambda_{i,t}(x)\,(R_i - \Delta w_{t-1}(x))^+ \right\} + w_{t-1}(x). \end{aligned}$$

The minimization term corresponds to a continuous Knapsack problem with upper bounds [with decision variables $\Delta\lambda_{i,t}(x)$]. In addition, the objective function coefficients $(R_i - \Delta w_{t-1}(x))^+$ are decreasing in i since $R_i > R_j$ if $i < j$ for any given state x. The optimal solution is then given by the following allocation where k denotes a class between $1, ..., n$:

$$\begin{aligned} \Delta\lambda_{i,t}(x) &= 0 \;\; if \;\; 1 \leq i < k \\ \Delta\lambda_{k,t}(x) &= q_t - \sum_{i=k+1}^{n} \Delta\lambda_{i,t}(x) \\ \Delta\lambda_{i,t}(x) &= \bar{\lambda}_{i,t} - \underline{\lambda}_{i,t} \;\; if \;\; k < i \leq n. \end{aligned}$$

The optimal solution clearly does not depend on x for all t. The result then follows. □

We have established that Nature's solution is identical for all states x for any stage t. Moreover, if the uncertainty set is not time dependent, i.e. $\mathcal{P}_t = \mathcal{P}$ for all t, then nature's optimal choice of probability distribution is identical for all states at all times: $\boldsymbol{\lambda}_t(x) = \boldsymbol{\lambda}$ for all x, t.

Corollary 3 *Consider the uncertainty set $\mathcal{C}_t$, if $b_i \leq b_j$ for all $i \leq j$ then Nature's optimal choice of probability distribution can be obtained by a simple rule and is identical for all states at all times: $\boldsymbol{\lambda}_t(x) = \boldsymbol{\lambda}_t$ for all x.*

Proof Since $b_1 \leq b_2 \leq \ldots \leq b_n$, it is easy to conclude that:

$$\frac{(R_1 - \Delta w_t(x))^+}{b_1} \geq \frac{(R_2 - \Delta w_t(x))^+}{b_2} \geq \ldots \geq \frac{(R_n - \Delta w_t(x))^+}{b_n}.$$

for all t, x. Since $\mathcal{C}_t \subseteq \mathcal{P}_t$ the resultant optimal solution is given as:

$$\begin{aligned}
\Delta\lambda_{i,t}(x) &= 0 \;\; if \;\; 1 \leq i < k \\
\Delta\lambda_{k,t}(x) &= \frac{\Delta Q_t - \sum_{i=k+1}^{n} b_i \Delta\lambda_{i,t}(x)}{b_k} \\
\Delta\lambda_{i,t}(x) &= \bar{\lambda}_{i,t} - \underline{\lambda}_{i,t} \;\; if \;\; k < i \leq n.
\end{aligned}$$

As in Theorem 3, the optimal solution does not depend on x at all stages t. □

Corollary 4 *Consider the uncertainty set $\mathcal{C}_t$, if $b_i = R_i$ for all i then Nature's optimal choice of probability distribution is identical for all states at all stages: $\boldsymbol{\lambda}_t(x) = \boldsymbol{\lambda}_t$ for all x.*

Proof We know that:

$$\frac{(R_1 - \Delta w_{t-1}(x))^+}{R_1} \geq \frac{(R_2 - \Delta w_{t-1}(x))^+}{R_2} \geq \ldots \geq \frac{(R_n - \Delta w_{t-1}(x))^+}{R_n}$$

Using a similar argument to the one in the Proof of Corollary 3 we find that Nature's optimal solution does not depend on x. This establishes the result. □

Under the conditions of Theorem 3, Nature's probability distribution cannot be state dependent. The controller is then playing a game against a state-independent arrival distribution which makes the problem a standard Markov decision processes as in Talluri and Ryzin (2005) or Aydin et al. (2009). The next theorem establishes that all structural results of the nominal problem propagate to the robust counterpart.

Theorem 4 *Consider the uncertainty set $\mathcal{P}_t$, then the robust value has function has the following properties:*

1. *$w_t(x)$ is nondecreasing (ND) in x for all t,*
2. *$w_t(x)$ is concave in x for all t,*
3. *$w_t(x)$ is supermodular in x, t for all x, t.*

Proof Due to Theorem 3, nature's optimal policy does not depend on x. The controller's problem then becomes a Markov decision process with state independent demand arrival rates. The proof then follows from the results on the nominal problem given in Aydin et al. (2009). □

Theorem 4 establishes that, unlike in Sect. 2, all major principal structural properties of the optimal polices also hold for the robust counterpart of the problem defined in Eq. (7). This implies the optimality of threshold policies as in the nominal problem. Besides, supermodularity implies that the thresholds are also monotone over time.

Next we establish the structural results for the uncertainty set $\mathcal{C}$ under certain conditions.

Corollary 5 *Consider the uncertainty set $\mathcal{C}_t$, if $b_i \leq b_j$ for all $i \leq j$ or $b_i = R_i$ for all i, then the robust value function has the following properties:*

1. *$w_t(x)$ is nondecreasing (ND) in x for all t,*
2. *$w_t(x)$ is concave in x for all t,*
3. *$w_t(x)$ is supermodular in x, t for all x, t.*

Proof Due to Corollaries 3 and 4, nature's optimal policy does not depend on x. Once again, the proof then follows from the results on the nominal problem given in Aydin et al. (2009). □

4 Numerical results

4.1 An illustrative example

The robust inventory rationing problem of Sect. 2 has many interesting features and presents some challenges. To shed further light onto some of these features, let us explore a numerical example. In order to represent the controller's optimal actions in the problem, let us use a different (but equivalent) representation as in the following:

$$v_t(x) = \mu \max_{a_0}\{a_0 v_{t-1}(x+1) + (1-a_0)v_{t-1}(x)\} + \sum_{i=1}^{n+1} \lambda_i \max_{a_i}\{a_i(v_{t-1}(x-1)+R_i) + (1-a_i)v_{t-1}(x)\} - h(x) \quad \text{if } x > 0$$

and

$$v_t(0) = \mu \max_{a_0}\{a_0 v_{t-1}(1) + (1-a_0)v_{t-1}(0)\} + \sum_{i=1}^{n+1} \lambda_i v_{t-1}(0) - h(0), \qquad \text{if } x = 0. \tag{8}$$

To obtain numerical results, let us focus on a particular case of the above problem with two customer classes and $\mu_t = \mu$, $\lambda_{1,t} = \lambda_1$, $\lambda_{2,t} = \lambda_2$, $\lambda_{3,t} = 0$ (i.e. there is no fictitious class) for all t. Recall that for each demand class $a_i = 1$ denotes the action that admits the arriving demand and $a_i = 0$ corresponds to rejecting the customer. Similarly $a_0 = 1$ denotes an order to produce whereas $a_0 = 0$ denotes an order not to stop production. Remember, the nominal problem is then represented by for $x > 0$:

$$\begin{aligned} v_t(x) = {} & \mu \max_{a_0}\{a_0 v_{t-1}(x+1) + (1-a_0)v_{t-1}(x)\} \\ & + \lambda_1 \max_{a_1}\{a_1(v_{t-1}(x-1)+R_1) + (1-a_1)v_{t-1}(x)\} \\ & + \lambda_2 \max_{a_2}\{a_2(v_{t-1}(x-1)+R_2) + (1-a_2)v_{t-1}(x)\} - hx \quad \text{if } x > 0 \end{aligned} \tag{9}$$

Let us next consider the robust version where the uncertainty set $\mathcal{P}$ is such that $\mu_t(x) + \lambda_{1,t}(x) + \lambda_{2,t}(x) = 1$, $0.20 \le \mu_t(x) \le 0.50$, $0.20 \le \lambda_{1,t}(x) \le 0.50$ and $0.20 \le \lambda_{2,t}(x) \le 0.50$. Let's denote by $\boldsymbol{p}_t(x, \boldsymbol{a}) = (\mu_t(x, \boldsymbol{a}), \lambda_{1,t}(x, \boldsymbol{a}), \lambda_{2,t}(x, \boldsymbol{a}))$.

$$\begin{aligned} w_t(x) = \max_{\boldsymbol{a}}\{ & \min_{\boldsymbol{p}_t(x,\boldsymbol{a})\in\mathcal{P}} \mu_t(x,\boldsymbol{a})\,(a_0 w_{t-1}(x+1) + (1-a_0)w_{t-1}(x)) \\ & + \lambda_{1,t}(x,\boldsymbol{a})\,(a_1(w_{t-1}(x-1)+R_1) + (1-a_1)w_{t-1}(x)) \\ & + \lambda_{2,t}(x,\boldsymbol{a})\,(a_2(w_{t-1}(x-1)+R_2) + (1-a_2)w_{t-1}(x))\} - hx \text{ if } x > 0 \end{aligned} \tag{10}$$

with $w_0(x) = 0$, for all x, and $w_t(0) = 0$ for all t.

Let us further assume that $T = 10$, $R_1 = 10$, $R_2 = 1$ and the holding cost $h(x) = 0.01x$. We solve the resulting robust MDP numerically. Let $\boldsymbol{a}^*{}_t(x) = (a_0, a_1, a_2)$ denote the optimal action selected by controller at time t and state x and $\boldsymbol{p}^*{}_t(x) = (\mu, \lambda_1, \lambda_2)$ denote the optimal event probability distribution selected by Nature for that action. Table 1 reports $\boldsymbol{p}^*{}_t(x)$ and $\boldsymbol{a}^*{}_t(x)$ for $t = 5, 10, 15$ and $x = 1, 2..., 5$.

Table 1 Controller's and nature's optimal policies for the example

$x \downarrow \|t \rightarrow$	15	10	5
0	(1, 0, 0), (0.2, 0.3, 0.5)	(1, 0, 0), (0.2, 0.3, 0.5)	(1, 0, 0), (0.2, 0.3, 0.5)
1	(1, 1, 0), (0.3, 0.2, 0.5)	(1, 1, 0), (0.3, 0.2, 0.5)	(1, 1, 0), (0.2, 0.3, 0.5)
2	(1, 1, 0), (0.3, 0.2, 0.5)	(1, 1, 0), (0.3, 0.2, 0.5)	(1, 1, 0), (0.3, 0.2, 0.5)
3	(1, 1, 1), (0.5, 0.2, 0.3)	(1, 1, 0), (0.3, 0.2, 0.5)	(1, 1, 0), (0.3, 0.2, 0.5)
4	(0, 1, 1), (0.5, 0.2, 0.3)	(1, 1, 1), (0.3, 0.2, 0.5)	(1, 1, 0), (0.3, 0.2, 0.5)
5	(0, 1, 1), (0.5, 0.2, 0.3)	(1, 1, 1), (0.5, 0.2, 0.3)	(1, 1, 1), (0.3, 0.2, 0.5)

Table 2 Demand and reward parameters for the numerical example

Customer class	Reward	Nominal arrival probability	Interval
1	\$80 per item	0.075	(0.05, 0.10)
2	\$35 per item	0.075	(0.05, 0.10)
3	\$25 per item	0.15	(0.10, 0.20)

In summary, Nature is allowed to choose the probability distributions at each stage, state and action from $\mathcal{P}$. It can be observed from Table 1 that Nature's solution is dependent on the state of the system. The event probabilities are varying over time and over the available inventory levels. However, as established in Sect. 2, $w_t(x)$ is concave in x for all t and the optimal policy of the controller is of threshold type.

4.2 Numerical analysis

In this section, we present some numerical results for the make-to-stock queue with multiple demand classes introduced in Sect. 2. Let us consider a system consisting of three customer classes. The holding cost is assumed to be \$5 per item per year (approximately \$0.0142 per item per day). The (daily) production probability is 0.2 and is certain. The (daily) demand arrival probabilities are assumed to be uncertain. In particular, we assume that there is best guess for the demand probability which we label as the nominal probability and an interval around this nominal probability. This data as well as the rewards of each class are presented in Table 2.

We experiment with the S-Robust Policy which includes the nominal policy and the pure robust policy as special cases. The nominal probabilities are taken as the concentration set and the optimal policies are obtained for different δ values where δ reflects the weight of the concentration set. Hence, $\delta = 1.0$ designates the nominal solution whereas $\delta = 0$ designates the pure robust solution. We solve the problem for different δ values between [0, 1] and compute the optimal $S - robust$ policy for different values of δ. Then we simulate the performance of these policies for demand data that is sampled from the uncertainty set. In particular, we generate the arrival probabilities to lie in their associated intervals uniformly, consequently with a mean equal to the nominal arrival probability.

In Figs. 1– 4 we present the long run results (using 1 million periods) of the problem. In Figs. 1 and 2 the admission thresholds of classes 2 and 3 and the target levels are depicted as a function of δ. Obviously, customer class 1 is the preferred customer in this problem and is always accepted to the system. Clearly, the admission thresholds increase as reliance on the nominal distribution increases as established in Corollary 2. Similarly, optimal target levels also increase as reliance on the nominal distribution increases. Therefore, at any given

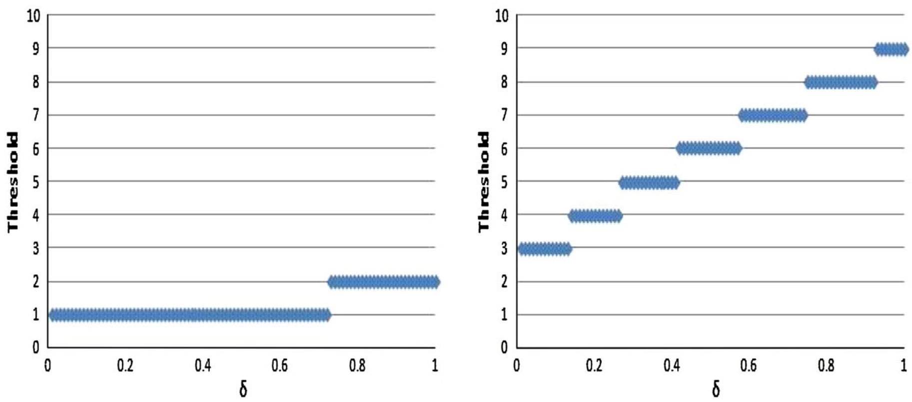


Fig. 1 Optimal admission thresholds of customer classes 2 and 3 as a function of δ

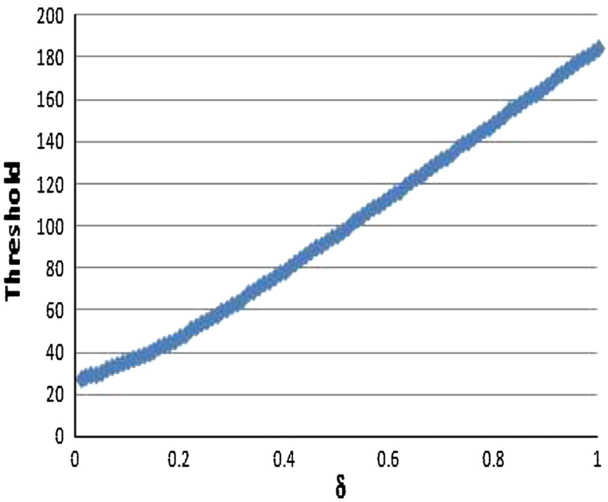


Fig. 2 Optimal base stock levels as a function of δ

inventory level, the controller becomes less willing to sell and more willing to produce when δ increases. Robustness in this problem requires setting lower thresholds and lower target inventory levels.

Figure 3 depicts the average profit per stage as a function of δ. To better understand how the average profit increases in δ, we next investigate the fill rates (demand satisfaction probabilities) for each class as a function of δ. Apparently, increasing robustness (measured by δ) requires treating customers similarly in terms of demand admission in addition to keeping lower target inventory levels. Figure 4 reports the fill rates of class 1 and class 3, this shows that the fill rate of class 1 is increasing and the fill rate of class 3 is decreasing in δ. Please note that decreasing δ, results in a decrease in the service quality of the class 1 customer as the controller uses a lower base stock level and admits more customers from other classes which increases the stock-out probability. On the other hand, class 3 has better access to the inventory and its fill rate improves when δ decreases.

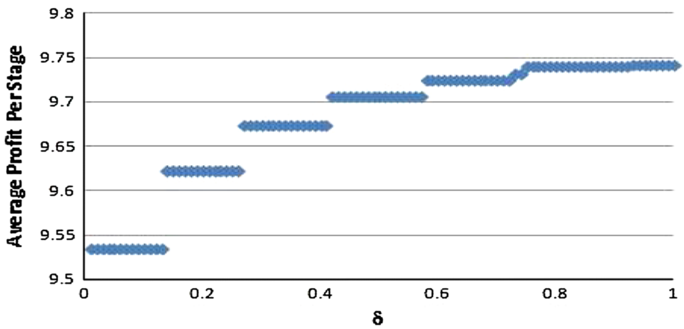


Fig. 3 Average profit as a function of δ

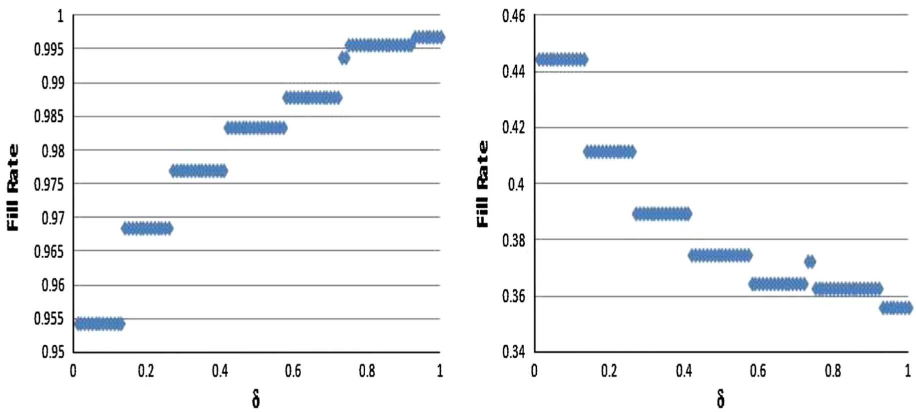


Fig. 4 Fill rates of class 1 and class 3 customers wrt δ

An interesting question is how robustness affects overall performance. For this investigation, we consider two measures of performance: the expected total profit and the variance of the total profit obtained by simulation. Next, we report results for these performance measures as a function of δ. In order to explore the effects of variability on the expected profit, we explored the total expected profit over a short horizon. We consider the case where the total horizon is 55 stages and the initial inventory is 0. With these parameters the expected sales over the horizon is approximately 9 units. As a benchmark, we also consider the case where replenishment is not allowed. In this case, we assume that the starting inventory is 9 (corresponding to the average sales with the above case). In Fig. 5, we present the expected profit versus the variance for both cases. It can be observed that there is a significant trade-off between expected profit and variance of the profit. In addition to this, in Fig. 6 we depict the "simulated expected total profit—standard deviation of the profit". Figure 6 depicts an interesting result, the performance of the nominal solution is sufficiently high in the case with production. However, when there is no production semi-robust solutions performs better than the nominal solution when variability is introduced to the objective.

Obviously the variance in the case without replenishment is less than the former case since the number of available items in the stock not affected by random production. Besides, in this case the opportunity to improve this variability is stronger. The total changes in the expected profit between the absolute robust and the nominal cases are nearly the same

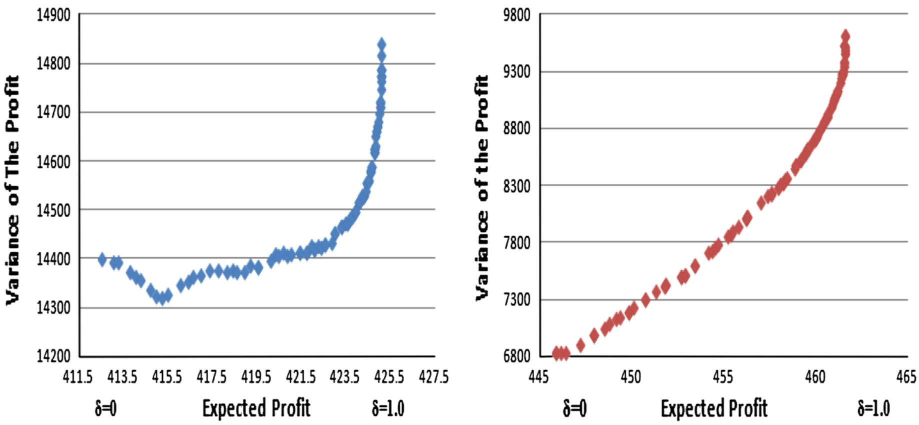


Fig. 5 Simulated expected total profit versus variance of the profit with production (*left*) and without production (*right*)

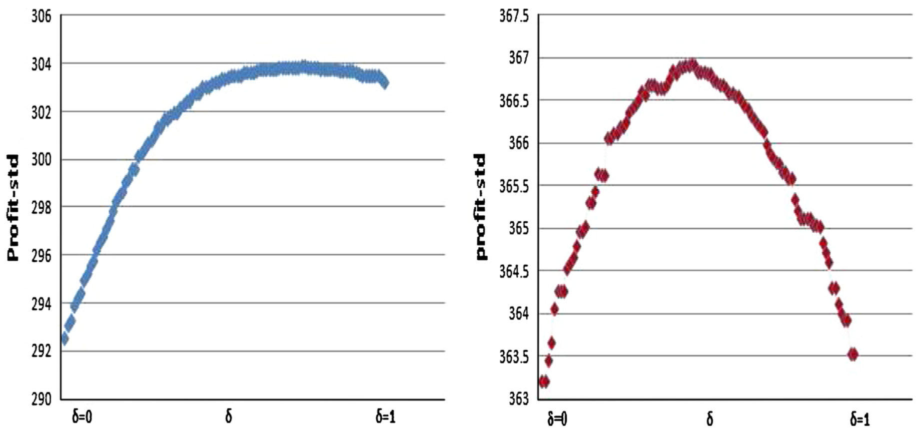
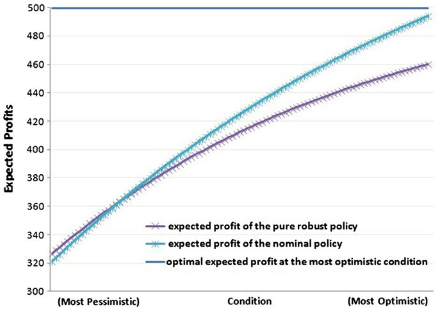


Fig. 6 Simulated expected total profit—standard deviation of the profit versus δ with production (*left*) and without production (*right*)

but the improvement in the variance is approximately 4 % in the case with production and 30 % without production. This preliminary exploration suggests that there may be useful links between robust policies and their applications in a risk-sensitive decision making environment where decision makers may also be concerned about controlling the variance of the return.

Our last comparison is on expected profit of the nominal and pure robust solutions under different conditions. In order to make such comparison we calculate the expected profits of pure robust and nominal policies between the most pessimistic [where arrival vector is $\boldsymbol{p}_{pes} = (0.05, 0.05, 0.10)$] and most optimistic [where arrival vector is $\boldsymbol{p}_{opt} = (0.10, 0.10, 0.20)$] conditions. Please note that the pure robust solution optimizes the most pessimistic condition, whereas nominal solution optimizes a probability vector which is simply the average of the $\boldsymbol{p}_{pes}$ and $\boldsymbol{p}_{opt}$. We compare the expected profit of the policies with respect to different weighted averages of these probability vectors $\boldsymbol{p}_{pes}$ and $\boldsymbol{p}_{opt}$ and the results are presented in Fig. 7. Although robust policy is advantageous over nominal policy in terms of variability and



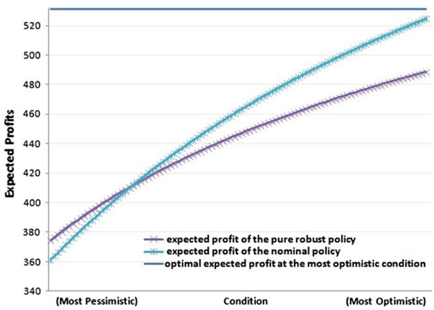


Fig. 7 Comparison of expected profits of robust and nominal policy for different probability vectors with production (*left*) and without production (*right*)

worst case situations, as it is clearly seen in the figure it significantly deviates from optimality as the condition improves.

5 Conclusion

We investigated the robust versions of two single-product dynamic demand admission problems: an inventory rationing and production control problem for a production/inventory system and a revenue management problem where a fixed inventory is allocated over time to different classes of customers. We showed that, under certain interval uncertainty models, the optimal policy structure of the corresponding nominal problems without parameter uncertainty extends to the robust case.

One drawback of a robust dynamic model is that the resulting policy may be too conservative. To alleviate this problem, we extended the analysis to a weighted optimization approach recently suggested by Xu and Mannor (2012). This approach can calibrate the level of robustness by choosing the appropriate weights between alternative objectives. We showed that the optimal policy structure is not affected by this formulation.

Finally, we presented numerical results that explore how robustness affects optimal admission and production policies. While expected profits may be affected negatively by taking a robust approach, there are situations where the gain in the variance of the profit may be significant. This may suggest a computational link between risk-sensitive decision making and robust optimal policies.

In future research, we aim to explore the optimal policy structure for robust formulations of more general production/inventory control problems. Risk-sensitive optimization and computational approaches also appear to be fruitful avenues for further exploration.

Appendix 1: A proof of theorem

Before proving Theorem 1 we introduce a simple algorithm that starts with $\boldsymbol{\lambda}$ and ends up with $\boldsymbol{\lambda}'$ (where $\boldsymbol{\lambda}_t \succeq \boldsymbol{\lambda}'_t$) by a sequence of reallocation of probabilities. By definition, $\sum_{i=1}^{n+1} \lambda_i \leq 1$ and $\sum_{i=1}^{n+1} \lambda'_i \leq 1$. Now consider the following sequence of vectors, $\boldsymbol{\lambda}^{(1)}, \boldsymbol{\lambda}^{(2)}, ..., \boldsymbol{\lambda}^{(n+1)}$. Let $\boldsymbol{\lambda}^{(1)} = \boldsymbol{\lambda}$.

Now let $\epsilon_1 = \lambda_1 - \lambda'_1$, we construct $\boldsymbol{\lambda}^{(2)} = \boldsymbol{\lambda}^{(1)} + (-\epsilon_1, \epsilon_1, 0, 0, ..., 0)$. Obviously $\epsilon_1 \geq 0$. In the next iteration, we let $\epsilon_2 = \lambda_2 + \lambda_1 - \lambda'_1 - \lambda'_2$, again $\epsilon_2 \geq 0$ and $\boldsymbol{\lambda}^{(3)} =$

$\boldsymbol{\lambda}^{(2)} + (0, -\epsilon_2, \epsilon_2, 0, ..., 0)$. We continue to iterate similarly for n steps. At step n we have: $\boldsymbol{\lambda}^{(n)} = \boldsymbol{\lambda}^{(n-1)} + (0, 0, ..., -\epsilon_n, \epsilon_n)$. At step n we have: $\boldsymbol{\lambda}^{(n+1)} = \boldsymbol{\lambda}^{(n-1)} + (0, 0, ..., -\epsilon_{n+1})$. By construction, we have $\boldsymbol{\lambda}^{(n+1)} = \boldsymbol{\lambda}'$. In addition, the sequence of vectors have the property $\boldsymbol{\lambda}^{(1)} \preceq \boldsymbol{\lambda}^{(2)} \preceq ... \preceq \boldsymbol{\lambda}^{(n+1)}$.

Proof We prove the desired result in two phases corresponding to stages t and $t+1$. Consider two systems that are identical and substitute $\boldsymbol{\lambda}_t$ with $\boldsymbol{\lambda}_t'$ and μ_t with μ_t'. In the first phase, we prove that $\Delta v_t(x) \geq \Delta v_t'(x)$ holds at t, then in the second phase we prove that $\Delta v_{t+1}(x) \geq \Delta v_{t+1}'(x)$. Please note that at stage t, we have $v_{t-1}(x) = v_{t-1}'(x)$, for all x and $\boldsymbol{\lambda}_t \neq \boldsymbol{\lambda}_t'$ and $\mu_{t+1} \neq \mu_{t+1}'$ whereas in stage $t+1$, $\boldsymbol{\lambda}_{t+1} = \boldsymbol{\lambda}_{t+1}'$ and $\mu_{t+1} = \mu_{t+1}'$.

Suppose $\Delta v_t(x) \geq \Delta v_t'(x)$ holds, then $v_t(x) - v_t'(x) \geq v_t(x-1) - v_t'(x-1)$. This implies that as we replace the $\boldsymbol{\lambda}_t$ with $\boldsymbol{\lambda}_t'$ at stage t the loss in $v_t(x)$ is greater than loss in $v_t(x-1)$. Whereas as we replace μ_t with μ_t' at stage t, associated gain in $v_t(x)$ is less than gain in $v_t(x-1)$. We again use the above algorithm in order to perform such a replacement. Hence, we consecutively decrease the arrival probability of a class i by ϵ and increase a class j by ϵ where $i < j$. Then we simply increase the production rate by $\mu_t' - \mu_t$.

$$\epsilon\,(R_i - \Delta v_{t-1}(x))^+ - \epsilon\left(R_j - \Delta v_{t-1}(x)\right)^+ \geq \epsilon\,(R_i - \Delta v_{t-1}(x-1))^+ \\ -\epsilon\left(R_j - \Delta v_{t-1}(x-1)\right)^+. \quad (11)$$

We prove the inequality case by case, note that A stands for admission and R stands for rejection. The non-trivial cases are listed below (we do not present the cases where all classes are accepted or all classes are rejected since these are obvious). Please note that accepting a lower class (j) means that a higher class (i) is always accepted. Also please note that, if a customer class is accepted at an inventory $x-1$ then it is also accepted at x, and if it is rejected at x then it is also rejected at $x-1$.

R_i $-\Delta v_{t-1}(x)$	$-R_j$ $-\Delta v_{t-1}(x)$	$\geq$ $\geq$	R_i $-\Delta v_{t-1}(x-1)$	$-R_j$ $-\Delta v_{t-1}(x-1)$	Result
(A)	(A)		(A)	(R)	$R_i - R_j \geq R_i - \Delta v_{t-1}(x-1)$
(A)	(R)		(A)	(R)	$R_i - \Delta v_{t-1}(x) \geq R_i - \Delta v_{t-1}(x-1)$
(A)	(A)		(R)	(R)	$R_i - R_j \geq 0$
(A)	(R)		(R)	(R)	$R_i - \Delta v_{t-1}(x) \geq 0$

Except for the case in the first row, all inequalities follow easily by concavity of $v(x)$ (a summary of the result is provided in the last column). Consider the case in the first row: because class j is rejected at $x-1$ for this case, we must have $R_j - \Delta v_{t-1}(x-1) \leq 0$. This implies that $R_i - R_j \geq R_i - \Delta v_{t-1}(x-1)$. Lastly, it is obvious that $(\Delta v_{t-1}(x+1))^+ \leq (\Delta v_{t-1}(x))^+$.

Next, we consider the second phase where we need to establish that $\Delta v_{t-1}(x) \geq \Delta v_{t-1}'(x)$. We use a similar approach here, but we consider only one operator T^i (admission decision for a single class) and T (production decision) at a time. The cases related to accept all and reject all for both systems at states $x-1$ and x are obvious. Please note that since $\Delta v_t(x) \geq \Delta v_t'(x)$, therefore $R_i - \Delta v_t(x) \leq R_i - \Delta v_t'(x)$, which means that any class accepted to the initial system will always be accepted to the second system. Except from the obvious cases (accept all, reject all) there are only four alternatives:

$T^i v_{t-1}(x)$	$-T^i v'_{t-1}(x)$	$\geq$	$T^i v_{t-1}(x-1)$	$-T^i v'_{t-1}(x-1)$	Result
(A)	(A)		(R)	(R)	$R_i + v_t(x-1) - R_i - v'_t(x-1) \geq v_t(x-1) - v'_t(x-1)$
(A)	(A)		(R)	(A)	$R_i + v_t(x-1) - R_i - v'_t(x-1) \geq v_t(x-1) - R_i - v'_t(x-2)$
(R)	(A)		(R)	(A)	$v_t(x) - R_i - v'_t(x-1) \geq v_t(x-1) - R_i - v'_t(x-2)$
(R)	(A)		(R)	(R)	$v_t(x) - v'_t(x-1) - R_i \geq v_t(x-1) - v'_t(x-1)$

The first case is clear. Consider the second case, since it is optimal to accept at $x-1$ for the second system, $R_i + v'_t(x-2) \geq v'_t(x-1)$. The third case is trivial, the LHS can be easily decreased by replacing the optimal action of the first system and the second case is attained. The last case is clear too, since the optimal action of the first system at state x is rejection, $v_t(x) \geq v_t(x-1) + R_i$.

For the production operator which we denote by T we need to show that it also preserves the inequality. Since $\Delta v_{t+1}(x) \geq \Delta v'_{t+1}(x)$ if it is not optimal to produce in the original system then it is also not optimal to produce in the perturbed system. By concavity we know that the base stock policy is optimal therefore the cases except from the produce at all of the conditions and not produce at all of the conditions the cases are as follows:

$Tv_{t+1}(x)$	$-Tv'_{t+1}(x)$	$\geq$	$Tv_{t+1}(x-1)$	$-Tv_{t+1}(x-1)$	Result
(NP)	(NP)		(P)	(P)	$v_t(x) - v'_t(x) \geq v_t(x) - v'_t(x)$
(P)	(NP)		(P)	(P)	$v_t(x+1) - v'_t(x) \geq v_t(x) - v'_t(x)$
(NP)	(NP)		(P)	(NP)	$v_t(x) - v'_t(x) \geq v_t(x) - v'_t(x-1)$
(P)	(NP)		(P)	(NP)	$v_t(x+1) - v'_t(x) \geq v_t(x) - v'_t(x-1)$

The first case is clear. In the second case it is optimal to make a production in the original system therefore $v_t(x+1) \geq v_t(x)$. The third case is similar too, since it is not optimal to produce at the second system we have $v'_t(x) \leq v'_t(x-1)$. For the last case $v_t(x+1) - v'_t(x) \geq v_t(x) - v'_t(x)$ and $v_t(x) - v'_t(x) \geq v_t(x) - v'_t(x-1)$. These results hold for each operator T_i and T, hence any convex combination of them satisfies the inequality. This completes the proof. □

Appendix 2: Definition of $\Delta\mathcal{P}_t$ and $\Delta\mathcal{C}_t$

Here we present the definitions of the transformed uncertainty sets employed in the proofs.

$$\Delta\mathcal{P}_t = \Bigg\{ \boldsymbol{\Delta y} = (\Delta y_0, \ldots, \Delta y_{n+1}) : \quad 0 \le \Delta y_i \le \bar{y}_{i,t} - \underline{y}_{i,t}, \\ 0 \le q - \sum_{i=1}^{n} \underline{y}_{i,t} \le \sum_{i=1}^{n} \Delta y_i \le 1 \Bigg\},$$

$$\Delta\mathcal{C}_t = \Bigg\{ \boldsymbol{\Delta y} = (\Delta y_1, \ldots, \Delta y_{n+1}) : \quad 0 \le \Delta y_{i,t} \le \bar{y}_i - \underline{y}_{i,t}, \\ \sum_{i=1}^{n} b_i \Delta y_i \ge Q - \sum_{i=1}^{n} b_i \Delta y_i,\ 0 \le q - \sum_{i=1}^{n} \underline{y}_{i,t} \le \sum_{i=1}^{n} \Delta y_i \le 1 \Bigg\},$$

References

Aydin, S., Akcay, Y., & Karaesmen, F. (2009). On the structural properties of a discrete-time single product revenue management problem. *Operations Research Letters*, *37*(4), 273–279.

Ball, M. O., & Queyranne, M. (2009). Toward robust revenue management: Competitive analysis of online booking. *Operations Research*, *57*(4), 950–963.

Ben-Tal, A., El Ghaoui, L., & Nemirovskiĭ, (2009). *Robust optimization*. Princeton, NJ: Princeton Univ Press.

Bertsimas, D., & Paschalidis, I. (2001). Probabilistic service level guarantees in make-to-stock manufacturing systems. *Operations Research*, *49*(1), 119–133.

Bertsimas, D., & Thiele, A. (2006). A robust optimization approach to inventory theory. *Operations Research*, *54*(1), 150–168.

Birbil, S. I., Frenk, J. B. G., Gromicho, J. A. S., & Zhang, S. (2009). The role of robust optimization in single-leg airline revenue management. *Management Science*, *55*(1), 148.

Cil, E. B., Ormeci, E. L., & Karaesmen, F. (2009). Effects of system parameters on the optimal policy structure in a class of queueing control problems. *Queueing Systems*, *61*(4), 273–304.

De Vericourt, F., Karaesmen, F., & Dallery, Y. (2000). Dynamic scheduling in a make-to-stock system: A partial characterization of optimal policies. *Operations Research*, *48*(5), 811–819.

De Vericourt, F., Karaesmen, F., & Dallery, Y. (2002). Optimal stock allocation for a capacitated supply system. *Management Science*, *48*, 1486–1501.

Gayon, J. P., de Véricourt, F., & Karaesmen, F. (2009). Stock rationing in an $M/E_r/1$ multi-class make-to-stock queue with backorders. *IIE Transactions*, *41*(12), 1096–1109.

Ha, A. Y. (1997). Inventory rationing in a make-to-stock production system with several demand classes and lost sales. *Management Science*, *43*(8), 1093–1103.

Ha, A. Y. (2000). Stock rationing in an $M/E_k/1$ make-to-stock queue. *Management Science*, *46*(1), 77–87.

Ha, A. Y. (1997). Inventory rationing in a make-to-stock production system with several demand classes and lost sales. *Management Science*, *43*, 1093–1103.

Ha, A. Y. (1997). Stock-rationing policy for a make-to-stock production system with two priority classes and backordering. *Naval Research Logistics (NRL)*, *44*(5), 457–472.

Huang, B., & Iravani, S. M. R. (2008). A make-to-stock system with multiple customer classes and batch ordering. *Operations Research*, *56*(5), 1312–1320.

Iyengar, G. N. (2005). Robust dynamic programming. *Mathematics of Operations Research*, *30*(2), 257.

Jain, A., Lim, A. E. B., & Shanthikumar, J. G. (2010). On the optimality of threshold control in queues with model uncertainty. *Queueing Systems*, *65*(2), 157–174.

Koole, G. (1998). Structural results for the control of queueing systems using event-based dynamic programming. *Queueing Systems*, *30*(3), 323–339.

Koole, G. (2006). *Monotonicity in Markov reward and decision chains: Theory and applications*. Boston: Now Publishers Inc.

Lan, Y., Gao, H., Ball, M., & Karaesmen, I. (2008). Revenue management with limited demand information. *Management Science*, *54*(9), 1594–1609.

Lautenbacher, C. J., & Stidham, S, Jr. (1999). The underlying Markov decision process in the single-leg airline yield-management problem. *Transportation Science*, *33*(2), 136.

Lim, A. E. B., & Shanthikumar, J. G. (2007). Relative entropy, exponential utility, and robust dynamic pricing. *Operations Research*, *55*(2), 198–214.

Nilim, A., & Ghaoui, L. E. (2005). Robust control of Markov decision processes with uncertain transition matrices. *Operations Research*, *53*(5), 780–798.

Paschalidis, I. C. & Kang, S. C. (2008). A robust approach to Markov decision problems with uncertain transition probabilities. *In Proceedings of the 17th IFAC World Congress*, (pp. 408–413).

Satia, J. K., & Lave, R. E, Jr. (1973). Markovian decision processes with uncertain transition probabilities. *Operations Research*, *21*(3), 728–740.

Talluri, K. T., & Van Ryzin, G. (2005). *The theory and practice of revenue management* (Vol. 68). New York: Springer Verlag.

Turgay, Z., Karaesmen, F., & Ormeci, L. (2013). *Structural properties of a class of robust inventory and queueing control problems. Technical report*. Istanbul: Koc University.

Xu, H., & Mannor, S. (2012). Distributionally robust markov decision processes. *Mathematics of Operations Research*, *37*, 288–300.