A complete algebraic solution to the optimal dynamic rationing policy in the stock-rationing queue with two demand classes

Quan-Lin Li[1] · Yi-Meng Li[2] · Jing-Yu Ma[3] · Heng-Li Liu[2]

Abstract

In this paper, we study a stock-rationing queue with two demand classes by means of the sensitivity-based optimization, and develop a complete algebraic solution to the optimal dynamic rationing policy. We show that the optimal dynamic rationing policy must be of transformational threshold type. Based on this finding, we can refine three sufficient conditions under each of which the optimal dynamic rationing policy is of threshold type (i.e., critical rationing level). To do this, we use the performance difference equation to characterize the monotonicity and optimality of the long-run average profit of this system, and thus establish some new structural properties of the optimal dynamic rationing policy by observing any given reference policy. Finally, we use numerical experiments to demonstrate our theoretical results of the optimal dynamic rationing policy. We believe that the methodology and results developed in this paper can shed light on the study of stock-rationing queue and open a series of potentially promising research.

Keywords Stock-rationing queue · Inventory rationing · Optimal dynamic rationing policy · Sensitivity-based optimization · Markov decision process

---

Quan-Lin Li and Jing-Yu Ma have contributed to the work equally and should be regarded as co-first authors

---

✉ Jing-Yu Ma
  mjy0501@126.com

1 School of Economics and Management, Beijing University of Technology, Beijing 100124, China

2 School of Economics and Management, Yanshan University, Qinhuangdao 066004, China

3 Business School, Xuzhou University of Technology, Xuzhou 221018, China

1 Introduction

In this paper, we consider a stock-rationing queueing problem of a warehouse with one type of products and two classes of demands, which may be viewed as coming from retailers with two different priority levels. Now, such a stock-rationing warehouse system becomes more and more important in many large cities under the current COVID-19 environment. For example, Beijing has seven super-large warehouses, which always supply various daily necessities, such as vegetables, meat, eggs, seafood to more than 40 million people every day. In the warehouses, each type of daily necessities are supplied by lots of different companies in China and the other countries, which lead to that the successive supply stream of each type of products can be well described as a Poisson process. In addition, the two retailers may be regarded as a large supermarket group and a community retail store group, respectively. Typically, the large supermarket group has a higher supply priority than the community retail store group. When the COVID-19 is at a serious warning in Beijing, the stock-rationing management of the warehouses plays a key role in strengthening the fine management of the warehouses such that every family at Beijing can have a very comprehensive life guarantee.

From the perspective of practical applications, such a stock-rationing queueing problem with multiple demand classes can always be encountered in many different real areas, for example, assemble-to-order systems, make-to-stock queues and multiechelon inventory systems by Ha (1997a); manufacturing by Zhao et al. (2005); airline by Wang et al. (2021); rental business by Altug and Ceryan (2021) and Jain et al. (2015); health care by Papastavrou et al. (2014) and Baron et al. (2019); and so forth. All the studies above show that the stock-rationing queues with multiple demand classes are not only necessary and important in many practical applications, but also have their own theoretical interest.

In the stock-rationing queueing systems, the stock rationing policies always assign different supply priorities to multiple classes of demands. In the early literature, the so-called critical rationing level was imagined intuitively, and its existence was further proved by Veinott (1965) and Topkis (1968). Once the critical rationing level is given and the on-hand inventory falls below the critical rationing level, a low priority demand may be either rejected, back-ordered or discarded such that the left on-hand inventory will be reserved to supply the future high priority demands. Thus designing and optimizing the critical rationing levels becomes a basic management way of inventory rationing across multiple demand classes. So far, analysis of the critical rationing levels has been interesting but difficult and challenging in the study of stock-rationing queues with multiple demand classes.

Some studies have applied the MDPs to discuss inventory rationing and stock-rationing queues across multiple demand classes by means of the submodular (or supermodular) technique, among which important examples include Ha (1997a), Ha (1997b), Ha (2000), Gayon et al. (2009), Benjaafar and ElHafsi (2006) and Nadar et al. (2014). To this end, it is a key that the structural properties of the optimal rationing policy need to be identified by using a set of structured value functions that is preserved under an optimal operator. Based on this, the optimal rationing policy of the inventory rationing across multiple demand classes can be further described and expressed by

the structural properties. In many more general cases, it is not easy and even very difficult to set up the structural properties of the optimal rationing policy. For this reason, some stronger model assumptions have to be further added to guarantee the existence of structural properties of the optimal policy. To improve the applicability of the MDPs, we propose a new algebraic method to find a complete algebraic solution to the optimal rationing policy by means of the sensitivity-based optimization.

The sensitivity-based optimization may be regarded as a new research branch of the MDPs, which grows out of infinitesimal perturbation analysis of discrete event dynamic systems, e.g., see Cao (2007). Note that one key of the sensitivity-based optimization is to set up and use the so-called performance difference equation, which is based on the perturbation realization factor as well as the performance potential related to the Poisson equation. To the best of our knowledge, this paper is the first to apply the sensitivity-based optimization to study the stock-rationing queues with multiple demand classes.

Based on the above analysis, we summarize the main contributions of this paper as follows:

(1) *A complete algebraic solution* This paper develops a complete algebraic solution to the optimal dynamic rationing policy of the stock-rationing queue by means of the sensitivity-based optimization, and shows that the optimal dynamic rationing policy must be of transformational threshold type, which can lead to refining three sufficient conditions under each of which the optimal dynamic rationing policy is of threshold type. In addition, it is worthwhile to note that our transformational threshold type results are sharper than the bang-bang control given in Ma et al. (2019), Ma et al. (2021) and Xia et al. (2021). Therefore, our algebraic method provides not only a necessary complement of policy spatial structural integrity but also a new way of optimality proof when comparing to the frequently-used submodular (or supermodular) technique of MDPs. Also, the complete algebraic solution to the optimal dynamic rationing policy can provide more effective support for numerical computation of the optimal policy and the optimal profit of this system.

(2) *A unified computational framework* To the best of our knowledge, this paper is the first to apply the sensitivity-based optimization to analyze the stock-rationing queues with multiple demand classes. It is necessary and useful to describe the three key steps: (a) Setting up a policy-based Markov process. (b) Constructing a policy-based Poisson equation, whose general solution can be used to characterize the monotonicity and optimality of the long-run average profit of this system. (c) Finding the optimal dynamic rationing policy in the three different areas of the penalty cost. In addition, the computational framework can sufficiently support numerical solution of stock-rationing queues with multiple demand classes while the submodular (or supermodular) technique of MDPs is very difficult to deal with more general stock-rationing queues.

(3) *Two different methods can sufficiently support each other* Our algebraic method sets up a complete algebraic solution to the optimal dynamic rationing policy, thus it can provide not only a necessary complement of policy spatial structural

integrity but also a new way of optimality proof when comparing to the frequently-used submodular (or supermodular) technique of MDPs. On the other hand, since our algebraic method and the submodular (or supermodular) technique are all important parts of the MDPs (the former is to use the poisson equations; while the latter is to apply the optimality equation), it is clear that the two different methods will sufficiently support each other in the study of inventory rationing and stock-rationing queues with multiple demand classes.

The remainder of this paper is organized as follows. Section 2 provides a literature review. Section 3 gives model description for the stock-rationing queue with two demand classes. Section 4 establishes an optimization problem to find the optimal dynamic rationing policy, in which we set up a policy-based birth-death process and define a more general reward function. Section 5 establishes a policy-based Poisson equation and provides its general solution with two free constants. Section 6 provides an explicit expression for the perturbation realization factor, and discusses the solution of the linear equation in the penalty cost. Section 7 discusses the monotonicity and optimality of the long-run average profit of this system, and finds the optimal dynamic rationing policy in three different areas of the penalty cost. Section 8 analyzes the stock-rationing queue under a threshold type (statical) rationing policy. Section 9 uses numerical experiments to demonstrate our theoretical results of the optimal dynamic rationing policy. Finally, some concluding remarks are given in Sect. 10.

2 Literature review

The inventory rationing across multiple demand classes was first analyzed by Veinott (1965) in the context of inventory control theory. From then on, some authors have discussed the inventory rationing problems. Readers may refer to a book by Möllering (2007); survey papers by Li et al. (2019a); and a research classification by Teunter and Haneveld (2008), Möllering and Thonemann (2008), Van Foreest and Wijngaard (2014) and Alfieri et al. (2017).

In the inventory rationing system, a critical rationing level was imagined from early research and practical experience. Veinott (1965) first proposed such a critical rationing level; while Topkis (1968) proved that the critical rationing level really exists and it is an optimal policy. It is a most basic problem how to mathematically prove whether a rationing inventory system has such a critical rationing level. Ha (1997a) made a breakthrough by applying the MDPs to analyze the inventory rationing policy for a stock-rationing queue with exponential production times, Poisson demand arrivals, lost sales and multiple demand classes.

Since the seminal work of Ha (1997a), it has been interesting to extend and generalize the way to apply the MDPs to deal with the stock-rationing queues and the rationing inventory systems. Important examples include the Erlang production times by Ha (2000) and Gayon et al. (2009); the backorders with two demand classes by Ioannidis et al. (2021); the capacity allocation by Shen and Yu (2019); omni-channel retailing by Goedhart et al. (2022); the batch ordering by Huang and Iravani (2008), the batch production by Pang et al. (2014); the utilization of information by Gayon

et al. (2009); an assemble-to-order production system by Elhafsi et al. (2015); ElHafsi et al. (2018) and Nadar et al. (2014); supply chain by van Wijk et al. (2019); dynamic price by Ding et al. (2016), Schulte and Pibernik (2017); and so forth.

In the inventory rationing literature, there exist two kinds of rationing policies: The static rationing policy and the dynamic rationing policy. Note that the dynamic rationing policy allows a threshold rationing level to change in time, depending on the number and ages of outstanding orders. In general, the static rationing policy is possible to miss some chances to further improve system performance, while the dynamic rationing policy should reflect better by means of various continuously updated information, the system performance can be improved dynamically. Deshpande et al. (2003) indicated that the optimal dynamic rationing policy may significantly reduce the inventory cost compared with the static rationing policy.

If there exist multiple replenishment opportunities, then the ordering policies are taken as two different types: Continuous review and periodic review. Therefore, our literature analysis for inventory rationing focuses on four different classes through combining the rationing policy (static vs. dynamic) with the inventory review (continuous vs. periodic) as follows: Static-continuous, static-periodic, dynamic-continuous and dynamic-periodic.

2.1 The static rationing policy (periodic vs. continuous)

The periodic review Veinott (1965) is the first to introduce an inventory rationing across different demand classes and propose a critical rationing level (i.e., the static rationing policy) in a periodic review inventory system with backorders. Subsequent research further investigated the periodic review inventory system with multiple demand classes, for example, the (s, S) policy by Cohen et al. (1988) and Tempelmeier (2006); the $(S - 1, S)$ policy by Ha (1997a, b); the lost sales by Dekker et al. (2002); the backorders by Möllering and Thonemann (2008); and the anticipated critical levels by Wang et al. (2013).

The continuous review Nahmias and Demmy (1981) is the first to propose and develop a constant critical level $(Q, r, \mathbf{C})$ policy in a continuous review inventory model with multiple demand classes, where Q is the fixed batch size, r is the reorder point and $\mathbf{C} = (C_1, C_2, \ldots, C_{n-1})$ is a set of critical rationing levels for n demand classes. From that time on, some authors have discussed the constant critical level $(Q, r, \mathbf{C})$ policy in continuous review inventory systems. Readers may refer to recent publications for details, among which are Melchiors et al. (2000), Dekker et al. (1998), Deshpande et al. (2003), Isotupa (2006), Arslan et al. (2007), Möllering and Thonemann (2008, 2010) and Escalona et al. (2015, 2017). In addition, the $(S - 1, S, \mathbf{C})$ inventory system was discussed by Dekker et al. (2002), Kranenburg and van Houtum (2007) and so on.

2.2 The dynamic rationing policy (continuous vs. periodic)

The continuous review Topkis (1968) is the first to analyze the dynamic rationing policy and indicate that the optimal rationing policy is a dynamic policy. Melchiors

(2003) considered a dynamic rationing policy in a (s, Q) inventory system with a key assumption that there was at most one outstanding order. Teunter and Haneveld (2008) developed a continuous time approach to determine the dynamic rationing policy for two Poisson demand classes, analyzed the marginal cost to determine the optimal remaining time for each rationing level, and expressed the optimal threshold policy through a schematic diagram or a lookup table. Fadıloğlu and Bulut (2010) proposed a dynamic rationing policy: Rationing with Exponential Replenishment Flow (RERF), for continuous review inventory systems with either backorders or lost sales. Wang et al. (2013) developed a dynamic threshold mechanism to allocate backorders when the multiple outstanding orders for different demand classes exist for the (Q, R) inventory system.

The periodic review: For the dynamic rationing policy in a periodic review inventory system, readers may refer to, such as two demand classes by Sobel and Zhang (2001), Frank et al. (2003) and Tan et al. (2009); dynamic critical levels and lost sales by Haynsworth and Haynsworth and Price (1989); multiple demand classes by Hung and Hsiao (2013); two backorder classes by Chew et al. (2013); general demand processes by Hung et al. (2012); mixed backorders and lost sales by Wang and Tang (2014); uncertain demand and production rates by Turgay et al. (2015); and incremental upgrading demands by You (2003).

3 Model description

In this section, we describe a stock-rationing queue with two demand classes, in which a single class of products are supplied to stock at a warehouse, and the two classes of demands come from two retailers with different priorities. In addition, we provide system structure, operational mode and mathematical notations.

A stock-rationing queue The warehouse has the maximal capacity N to stock a single class of products, and the warehouse needs to pay a holding cost C_1 per product per unit time. There are two classes of demands to order the products, in which the demands of Class 1 have a higher priority than that of Class 2, such that the demands of Class 1 can be satisfied in any non-zero inventory; while the demands of Class 2 may be either satisfied or refused based on the inventory level of the products. Figure 1 depicts a simple physical system to understand the stock-rationing queue.

The supply process The supply stream of the products to the warehouse is a Poisson process with arrival rate λ, where the price of per product is C_3 paid by the warehouse to the external product supplier. If the warehouse is full of the products, then any new arriving product has to be lost. In this case, the warehouse will have an opportunity cost C_4 per product rejected into the warehouse.

The service processes The service times provided by the warehouse to satisfy the demands of Classes 1 and 2 are i.i.d. and exponential with service rates μ_1 and μ_2, respectively. The service disciplines for the two classes of demands are all First Come First Serve (FCFS). The warehouse can obtain the service price R when one product is sold to Retailer 1 or 2. Note that each demand of Class 1 or 2 is satisfied by one product every time.
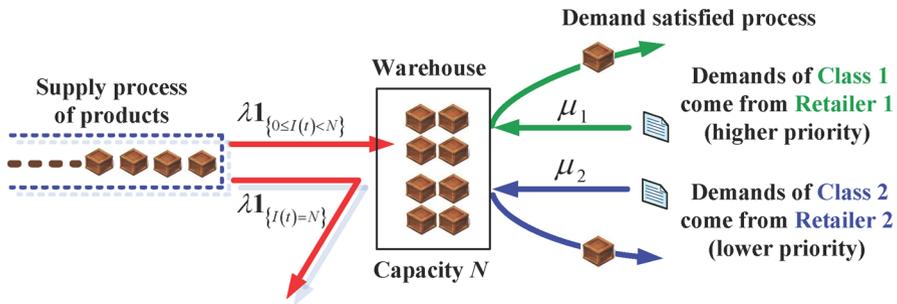
$I(t)$: **The total number of products in the stock-rationing queue at time t**

Fig. 1 A stock-rationing queue with two demand classes

The stock-rationing rule For the two classes of demands, each demand of Class 1 can always be satisfied in any non-zero inventory; while for satisfying the demands of Class 2, we need to consider three different cases as follows:

Case one: *The inventory level is zero.* In this case, there is no product in the warehouse. Thus any new arriving demand has to be rejected immediately. This leads to the the lost sales cost $C_{2,1}$ (resp. $C_{2,2}$) per unit time for any lost demand of Class 1 (resp. 2). We assume that $C_{2,1} > C_{2,2}$, which is used to guarantee the higher priority service for the demands of Class 1 when comparing to the lower priority for the demands of Class 2.

Case two: *The inventory level is low.* In this case, the number of products in the warehouse is not more than a key threshold K, where the threshold K is subjectively designed by means of some real experience. Note that the demands of Class 1 have a higher priority to receive the products than the demands of Class 2. Thus the warehouse will not provide any product to satisfy the demands of Class 2 under an equal service condition if the number of products in the warehouse is not more than K. Otherwise, such a service priority is violated (i.e., the demands of Class 2 are satisfied from a low stock), so that the warehouse must pay a penalty cost P per product supplied to the demands of Class 2 at a low stock. Note that the penalty cost P measures different priority levels to provide the products between the two classes of demands.

Case three: *The inventory level is high.* In this case, the number of products in the warehouse is more than the threshold K. Thus the demands of Classes 1 and 2 can be simultaneously satisfied due to enough products in the warehouse.

Independence We assume that all the random variables defined above are independent of each other.

In what follows, we use Table 1 to further summarize some above notations.

Remark 1 The penalty cost P is a necessary variable (setting up conditions of control classification) to dynamically control and optimize that the products are supplied to the demands of Class 2 at a low stock, and specifically, when the stock is not empty. While the lost sales costs, satisfying $C_{2,1} > C_{2,2}$, are introduced to provide some static and not enough penalties for that the products are supplied to the demands of Class 2 at a low stock.

Table 1 Some costs and prices in the stock-rationing queue

C_1	The holding cost per unit time per product stored in the warehouse
$C_{2,1}$	The lost sales cost of each lost demand of Class 1
$C_{2,2}$	The lost sales cost of each lost demand of Class 2
C_3	The price of per product paid by the warehouse to the external product supplier
C_4	The opportunity cost per product rejected into the warehouse
P	The penalty cost per product supplied to the demands of Class 2 at a low stock
R	The service price of the warehouse paid by each satisfied demand

4 Optimization model formulation

In this section, we establish an optimization problem to find the optimal dynamic rationing policy in the stock-rationing queue. To do this, we set up a policy-based birth-death process, and define a more general reward function with respect to both states and policies of the policy-based birth-death process.

To study the stock-rationing queue with two demand classes, we first need to define both 'states' and 'policies' to express stochastic dynamics of the stock-rationing queue.

Let $I(t)$ be the number of products in the warehouse at time t. Then it is regarded as the state of this system at time t. Obviously, all the cases of State $I(t)$ form a state space as follows:

$$\mathbf{\Omega} = \{0, 1, 2, \ldots, N\}.$$

Also, State $i \in \mathbf{\Omega}$ is regarded as an inventory level of this system.

From the states, some policies are defined with a little bit more complicated. Let d_i be a policy related to State $i \in \mathbf{\Omega}$, and it expresses whether or not the warehouse prefers to supply some products to the demands of Class 2 when the inventory level is not more than the threshold K for $0 < K \leq N$. Thus, we have

$$d_i = \begin{cases} 0, & i = 0, \\ 0, 1, & i = 1, 2, \ldots, K, \\ 1, & i = K+1, K+2, \ldots, N, \end{cases} \tag{1}$$

where $d_i = 0$ and 1 represents that the warehouse rejects and satisfies the demands of Class 2, respectively. Obviously, not only does the policy d_i depend on State $i \in \mathbf{\Omega}$, but also it is controlled by the threshold K. Of course, for a special case, if $K = N$, then $d_i \in \{0, 1\}$ for $1 \leq i \leq N$.

Corresponding to each state in $\mathbf{\Omega}$, we define a time-homogeneous policy of the stock-rationing queue as

$$\mathbf{d} = (d_0; d_1, d_2, \ldots, d_K; d_{K+1}, d_{K+2}, \ldots, d_N).$$

Fig. 2 State transition relations of the policy-based Markov process

It follows from (1) that

$$\mathbf{d} = (0; d_1, d_2, \ldots, d_K; 1, 1, \ldots, 1). \tag{2}$$

Thus Policy $\mathbf{d}$ depends on $d_i \in \{0, 1\}$, which is related to State i for $1 \le i \le K$. Let all the possible policies of the stock-rationing queue, given in (2), form a policy space as follows:

$$\mathcal{D} = \{\mathbf{d} : \mathbf{d} = (0; d_1, d_2, \ldots, d_K; 1, 1, \ldots, 1), d_i \in \{0, 1\}, 1 \le i \le K\}.$$

Remark 2 In general, the threshold K is subjective and is designed by means of the real experience of the warehouse manager. If $K = N$, then the policy is expressed as

$$\mathbf{d} = (0; d_1, d_2, \ldots, d_N).$$

Thus our K-based policy $\mathbf{d} = (0; d_1, d_2, \ldots, d_K; 1, 1, \ldots, 1)$ is more general than Policy $\mathbf{d} = (0; d_1, d_2, \ldots, d_N)$.

Let $I^{(\mathbf{d})}(t)$ be the state of the stock-rationing queue at time t under any given policy $\mathbf{d} \in \mathcal{D}$. Then $\left\{ I^{(\mathbf{d})}(t) : t \ge 0 \right\}$ is a continuous-time policy-based Markov process on the state space $\mathbf{\Omega}$ whose state transition relations are depicted in Fig. 2.

It is easy to see from Fig. 2 that $\left\{ I^{(\mathbf{d})}(t) : t \ge 0 \right\}$ is a policy-based birth-death process. Based on this, the infinitesimal generator is given by

$$\mathbf{B}^{(\mathbf{d})} = \begin{pmatrix} -\lambda & \lambda & & & & & \\ v(d_1) & -[\lambda + v(d_1)] & \lambda & & & & \\ \ddots & \ddots & \ddots & & & & \\ & v(d_K) & -[\lambda + v(d_K)] & \lambda & & & \\ & & v(1) & -[\lambda + v(1)] & \lambda & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & v(1) & -[\lambda + v(1)] & \lambda \\ & & & & & v(1) & -v(1) \end{pmatrix},$$

$$\tag{3}$$

where $v(d_i) = \mu_1 + d_i \mu_2$ for $i = 1, 2, \ldots, K$, and $v(1) = \mu_1 + \mu_2$. It is clear that $v(d_i) > 0$ for $i = 1, 2, \ldots, K$. Thus the policy-based birth-death process $\mathbf{B}^{(\mathbf{d})}$ must be irreducible, aperiodic and positive recurrent for any given policy $\mathbf{d} \in \mathcal{D}$. In this

case, we write the stationary probability vector of the policy-based birth-death process $\left\{ I^{(\mathbf{d})}(t) : t \geq 0 \right\}$ as

$$\pi^{(\mathbf{d})} = \left(\pi^{(\mathbf{d})}(0); \pi^{(\mathbf{d})}(1), \ldots, \pi^{(\mathbf{d})}(K); \pi^{(\mathbf{d})}(K+1), \ldots, \pi^{(\mathbf{d})}(N) \right). \qquad (4)$$

Obviously, the stationary probability vector $\pi^{(\mathbf{d})}$ is the unique solution to the system of linear equations: $\pi^{(\mathbf{d})} \mathbf{B}^{(\mathbf{d})} = \mathbf{0}$ and $\pi^{(\mathbf{d})} \mathbf{e} = 1$, where $\mathbf{e}$ is a column vector of ones with a suitable dimension. We write

$$\xi_i^{(\mathbf{d})} = \begin{cases} \xi_0 = 1, \ i = 0, \\ \dfrac{\lambda^i}{\prod\limits_{j=1}^{i} v(d_j)}, & i = 1, 2, \ldots, K, \\ \dfrac{\lambda^i}{(\mu_1 + \mu_2)^{i-K} \prod\limits_{j=1}^{K} v(d_j)}, & i = K+1, K+2, \ldots, N, \end{cases} \qquad (5)$$

and

$$h^{(\mathbf{d})} = 1 + \sum_{i=1}^{N} \xi_i^{(\mathbf{d})}.$$

It follows from Subsection 1.1.4 of Chapter 1 in Li (2010) that

$$\pi^{(\mathbf{d})}(i) = \begin{cases} \dfrac{1}{h^{(\mathbf{d})}}, & i = 0 \\ \dfrac{1}{h^{(\mathbf{d})}} \xi_i^{(\mathbf{d})}, & i = 1, 2, \ldots, N. \end{cases} \qquad (6)$$

By using the policy-based birth-death process $\mathbf{B}^{(\mathbf{d})}$, now we define a more general reward function in the stock-rationing queue. It is seen from Table 1 that the reward function with respect to both states and policies is defined as a profit rate (i.e. the total system revenue minus the total system cost per unit time). By observing the impact of Policy $\mathbf{d}$ on the profit rate, the reward function at State i under Policy $\mathbf{d}$ is given by

$$\begin{aligned} f^{(\mathbf{d})}(i) = {}& R \left(\mu_1 1_{\{i>0\}} + \mu_2 d_i \right) - C_1 i - C_{2,1} \mu_1 1_{\{i=0\}} - C_{2,2} \mu_2 (1 - d_i) \\ & - C_3 \lambda 1_{\{i<N\}} - C_4 \lambda 1_{\{i=N\}} - P \mu_2 d_i 1_{\{1 \leq i \leq K\}}, \end{aligned} \qquad (7)$$

where, $1_{\{\cdot\}}$ represents the indicator function whose value is one when the event occurs; otherwise it is zero. By using the indicator function, satisfying and rejecting the demands of Class 1 are expressed as $1_{\{i>0\}}$ and $1_{\{i=0\}}$, respectively; the external products enter or are lost by the warehouse according to $1_{\{i<N\}}$ and $1_{\{i=N\}}$, respectively; and a penalty cost paid by the warehouse is denoted by means of $1_{\{1 \leq i \leq K\}}$ due to that the warehouse supplies the products to the demands of Class 2 at a low stock.

For the convenience of readers, it is necessary and useful to explain the reward function (7) from four different cases as follows:

Case (a): For $i = 0$,

$$f(0) = -C_{2,1}\mu_1 - C_{2,2}\mu_2 - C_3\lambda. \tag{8}$$

In Case (a), there is no product in the warehouse, thus it has to reject any demand of Classes 1 and 2.

Case (b): For $1 \leq i \leq K$,

$$f^{(\mathbf{d})}(i) = R(\mu_1 + \mu_2 d_i) - C_1 i - C_{2,2}\mu_2(1 - d_i) - C_3\lambda - P\mu_2 d_i. \tag{9}$$

In Case (b), since the inventory level is low for $1 \leq i \leq K$, the penalty cost is paid by the warehouse when it supplies the products to the demands of Class 2.

Differently from Cases (a) and (b), the inventory level is high for $K + 1 \leq i \leq N$ in Cases (c) and (d), thus it can synchronously satisfy the demands of Classes 1 and 2.

Case (c): For $K + 1 \leq i \leq N - 1$,

$$f(i) = R(\mu_1 + \mu_2) - C_1 i - C_3\lambda. \tag{10}$$

Case (d): For $i = N$,

$$f(N) = R(\mu_1 + \mu_2) - C_1 N - C_4\lambda. \tag{11}$$

Note that C_3 is the price per product paid by the warehouse to the external product supplier; while C_4 is the opportunity cost per product rejected into the warehouse.

Based on the above analysis, we define an $(N + 1)$-dimensional column vector composed of the elements $f(0)$, $f^{(\mathbf{d})}(i)$ for $1 \leq i \leq K$, and $f(j)$ for $K + 1 \leq j \leq N$ as follows:

$$\mathbf{f}^{(\mathbf{d})} = \left(f(0); f^{(\mathbf{d})}(1), f^{(\mathbf{d})}(2), \ldots, f^{(\mathbf{d})}(K); f(K+1), f(K+2), \ldots, f(N)\right)^T. \tag{12}$$

Now, we consider the long-run average profit of the stock-rationing queue (or the continuous-time policy-based birth-death process $\{I^{(\mathbf{d})}(t) : t \geq 0\}$) under any given policy $\mathbf{d}$. Let

$$\eta^{\mathbf{d}} = \lim_{T \to \infty} E\left\{\frac{1}{T}\int_0^T f^{(\mathbf{d})}\left(I^{(\mathbf{d})}(t)\right) dt\right\}.$$

Then

$$\eta^{\mathbf{d}} = \pi^{(\mathbf{d})}\mathbf{f}^{(\mathbf{d})}, \tag{13}$$

where $\pi^{(\mathbf{d})}$ and $\mathbf{f}^{(\mathbf{d})}$ are given by (4) and (12), respectively.

To further observe the long-run average profit $\eta^{\mathbf{d}}$, here we show how $\eta^{\mathbf{d}}$ depends on the penalty cost P, and particularly, $\eta^{\mathbf{d}}$ is linear in P. To do this, we write that for $i = 0$,

$$A_0 = 0, \quad B_0 = -C_{2,1}\mu_1 - C_{2,2}\mu_2 - C_3\lambda;$$

for $i = 1, 2, \ldots, K$,

$$A_i^{(\mathbf{d})} = \mu_2 d_i, \ B_i^{(\mathbf{d})} = R(\mu_1 + \mu_2 d_i) - C_1 i - C_{2,2}\mu_2(1 - d_i) - C_3\lambda;$$

for $i = K + 1, K + 2, \ldots, N - 1$,

$$A_i = 0, \ B_i = R(\mu_1 + \mu_2) - C_1 i - C_3\lambda;$$

for $i = N$,

$$A_i = 0, \ B_N = R(\mu_1 + \mu_2) - C_1 N - C_4\lambda.$$

Then it follows from (8) to (11) that for $i = 0$,

$$f(0) = B_0; \tag{14}$$

for $i = 1, 2, \ldots, K$,

$$f^{(\mathbf{d})}(i) = B_i^{(\mathbf{d})} - PA_i^{(\mathbf{d})}; \tag{15}$$

for $i = K + 1, K + 2, \ldots, N$,

$$f(i) = B_i. \tag{16}$$

It follows from (6) and (14) to (16) that

$$
\begin{aligned}
\eta^{\mathbf{d}} &= \pi^{(\mathbf{d})}\mathbf{f}^{(\mathbf{d})} \\
&= \pi^{(\mathbf{d})}(0) f(0) + \sum_{i=1}^{K} \pi^{(\mathbf{d})}(i) f^{(\mathbf{d})}(i) + \sum_{i=K+1}^{N} \pi^{(\mathbf{d})}(i) f(i) \\
&= D^{(\mathbf{d})} - PF^{(\mathbf{d})},
\end{aligned}
\tag{17}
$$

where

$$D^{(\mathbf{d})} = \pi^{(\mathbf{d})}(0) B_0 + \sum_{i=1}^{K} \pi^{(\mathbf{d})}(i) B_i^{(\mathbf{d})} + \sum_{i=K+1}^{N} \pi^{(\mathbf{d})}(i) B_i$$

and

$$F^{(\mathbf{d})} = \sum_{i=1}^{K} \pi^{(\mathbf{d})}(i) A_i^{(\mathbf{d})}.$$

Hence the long-run average profit $\eta^{\mathbf{d}}$ is linear in the penalty cost P.

We observe that when the inventory level is low, supplying the products to the demands of Class 2 leads to that both the total system revenue and the total system cost increase synchronously. Thus there is a tradeoff between the total system revenue and the total system cost. This motivates us to find an optimal dynamic rationing

policy such that the warehouse has the maximal profit. Therefore, our objective is to find an optimal dynamic rationing policy $\mathbf{d}^*$ such that the long-run average profit $\eta^{\mathbf{d}}$ is maximal for $\mathbf{d} = \mathbf{d}^*$, that is,

$$\mathbf{d}^* = \arg \max_{\mathbf{d} \in \mathcal{D}} \left\{ \eta^{\mathbf{d}} \right\}. \tag{18}$$

In fact, it is difficult and challenging not only to analyze some interesting structural properties of the optimal rationing policies $\mathbf{d}^*$, but also to provide some effective algorithms for computing the optimal dynamic rationing policy $\mathbf{d}^*$.

In the remainder of this paper, we apply the sensitivity-based optimization to study the optimal policy problem (18), where the Poisson equations will play a key role in the study of MDPs and sensitivity-based optimization.

5 A policy-based poisson equation

In this section, for the stock-rationing queue, we set up a policy-based Poisson equation which is derived by means of the law of total probability and analysis on some stopping times of the policy-based birth-death process $\left\{ I^{(\mathbf{d})}(t), t \geq 0 \right\}$. It is worth noting that the policy-based Poisson equation provides a useful relation between the sensitivity-based optimization and the MDPs, see, e.g. Puterman (2014) and Cao (2007).

For any given policy $\mathbf{d} \in \mathcal{D}$, it follows from Chapter 2 in Cao (2007) that for the continuous-time policy-based birth-death process $\left\{ I^{(\mathbf{d})}(t), t \geq 0 \right\}$, we define the performance potential as

$$g^{(\mathbf{d})}(i) = E \left\{ \int_0^{+\infty} \left[f^{(\mathbf{d})} \left(I^{(\mathbf{d})}(t) \right) - \eta^{\mathbf{d}} \right] dt \, \middle| \, I^{(\mathbf{d})}(0) = i \right\}, \tag{19}$$

where $\eta^{\mathbf{d}}$ is defined in (13). It is seen from Cao (2007) that for Policy $\mathbf{d} \in \mathcal{D}$, $g^{(\mathbf{d})}(i)$ quantifies the contribution of the initial State i to the long-run average profit of the stock-rationing queue. Here, $g^{(\mathbf{d})}(i)$ is also called the relative value function or the bias in the traditional MDP theory, see, e.g. Puterman (2014). We further define a column vector $\mathbf{g}^{(\mathbf{d})}$ as

$$\mathbf{g}^{(\mathbf{d})} = \left(g^{(\mathbf{d})}(0); g^{(\mathbf{d})}(1), \ldots, g^{(\mathbf{d})}(K); g^{(\mathbf{d})}(K+1), \ldots, g^{(\mathbf{d})}(N) \right)^T. \tag{20}$$

To compute the vector $\mathbf{g}^{(\mathbf{d})}$, we define the first departure time of the policy-based birth-death process $\left\{ I^{(\mathbf{d})}(t) : t \geq 0 \right\}$ beginning from State i as

$$\tau = \inf \left\{ t \geq 0 : I^{(\mathbf{d})}(t) \neq i \right\},$$

where $I^{(\mathbf{d})}(0) = i$. Clearly, τ is a stopping time of the policy-based birth-death process $\left\{ I^{(\mathbf{d})}(t) : t \geq 0 \right\}$. Based on this, if $i = 0$, then it is seen from (3) that State 0

is a boundary state of the policy-based birth-death process $\mathbf{B}^{(\mathbf{d})}$, hence $I^{(\mathbf{d})}(\tau) = 1$. Similarly, for each State $i \in \mathbf{\Omega}$, a basic relation is established as follows:

$$I^{(\mathbf{d})}(\tau) = \begin{cases} 1, & i = 0, \\ i - 1 \text{ or } i + 1, & i = 1, 2, \ldots, N - 1, \\ N - 1, & i = N. \end{cases} \tag{21}$$

To compute the column vector $\mathbf{g}^{(\mathbf{d})}$, we derive a policy-based Poisson equation in terms of both the stopping time τ and the basic relation (17). Based on this, we set up the Poisson equation according to four parts as follows:

Part (a): For $i = 0$, we have

$$\begin{aligned} g^{(\mathbf{d})}(0) &= E\left\{ \int_0^{+\infty} \left[f^{(\mathbf{d})}\left(I^{(\mathbf{d})}(t)\right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(0) = 0 \right\} \\ &= E\left\{ \tau \,\middle|\, I^{(\mathbf{d})}(0) = 0 \right\} \left[f(0) - \eta^{\mathbf{d}} \right] \\ &\quad + E\left\{ \int_\tau^{+\infty} \left[f^{(\mathbf{d})}\left(I^{(\mathbf{d})}(t)\right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(\tau) \right\} \\ &= \frac{1}{\lambda}\left[f(0) - \eta^{\mathbf{d}} \right] + E\left\{ \int_0^{+\infty} \left[f^{(\mathbf{d})}\left(I^{(\mathbf{d})}(t)\right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(0) = 1 \right\} \\ &= \frac{1}{\lambda}\left[f(0) - \eta^{\mathbf{d}} \right] + g^{(\mathbf{d})}(1), \end{aligned}$$

where for the policy-based birth-death process $\left\{ I^{(\mathbf{d})}(t) : t \geq 0 \right\}$, it is easy to see from Fig. 2 that by using $I^{(\mathbf{d})}(t) = 0$ for $0 \leq t < \tau$,

$$\int_0^\tau \left[f^{(\mathbf{d})}\left(I^{(\mathbf{d})}(t)\right) - \eta^{\mathbf{d}} \right] dt = \tau\left[f(0) - \eta^{\mathbf{d}} \right],$$

$$E\left\{ \tau \,\middle|\, I^{(\mathbf{d})}(0) = 0 \right\} = \frac{1}{\lambda}.$$

We obtain

$$-\lambda g^{(\mathbf{d})}(0) + \lambda g^{(\mathbf{d})}(1) = \eta^{\mathbf{d}} - f(0). \tag{22}$$

Part (b): For $i = 1, 2, \ldots, K$, it is easy to see from Fig. 2 that

$$\begin{aligned} g^{(\mathbf{d})}(i) &= E\left\{ \int_0^{+\infty} \left[f^{(\mathbf{d})}\left(I^{(\mathbf{d})}(t)\right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(0) = i \right\} \\ &= E\left\{ \tau \,\middle|\, I^{(\mathbf{d})}(0) = i \right\} \left[f^{(\mathbf{d})}(i) - \eta^{\mathbf{d}} \right] \\ &\quad + E\left\{ \int_\tau^{+\infty} \left[f^{(\mathbf{d})}\left(I^{(\mathbf{d})}(t)\right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(\tau) \right\} \\ &= \frac{1}{v(d_i) + \lambda} \left[f^{(\mathbf{d})}(i) - \eta^{\mathbf{d}} \right] \end{aligned}$$

$$
+ \frac{\lambda}{v(d_i) + \lambda} E \left\{ \int_0^{+\infty} \left[f^{(\mathbf{d})} \left(I^{(\mathbf{d})}(t) \right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(0) = i + 1 \right\}
$$

$$
+ \frac{v(d_i)}{v(d_i) + \lambda} E \left\{ \int_0^{+\infty} \left[f^{(\mathbf{d})} \left(I^{(\mathbf{d})}(t) \right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(0) = i - 1 \right\}
$$

$$
= \frac{1}{v(d_i) + \lambda} \left[f^{(\mathbf{d})}(i) - \eta^{\mathbf{d}} \right] + \frac{\lambda}{v(d_i) + \lambda} g^{(\mathbf{d})}(i + 1)
$$

$$
+ \frac{v(d_i)}{v(d_i) + \lambda} g^{(\mathbf{d})}(i - 1),
$$

where

$$
E \left\{ \tau \,\middle|\, I^{(\mathbf{d})}(0) = i \right\} = \frac{1}{v(d_i) + \lambda}.
$$

We obtain

$$
v(d_i) g^{(\mathbf{d})}(i - 1) - [v(d_i) + \lambda] g^{(\mathbf{d})}(i) + \lambda g^{(\mathbf{d})}(i + 1) = \eta^{\mathbf{d}} - f^{(\mathbf{d})}(i). \quad (23)
$$

Part (c): For $i = K + 1, K + 2, \ldots, N - 1$, by using Fig. 2 we have

$$
g^{(\mathbf{d})}(i) = E \left\{ \int_0^{+\infty} \left[f^{(\mathbf{d})} \left(I^{(\mathbf{d})}(t) \right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(0) = i \right\}
$$

$$
= E \left\{ \tau \,\middle|\, I^{(\mathbf{d})}(0) = i \right\} \left[f(i) - \eta^{\mathbf{d}} \right]
$$

$$
+ E \left\{ \int_\tau^{+\infty} \left[f^{(\mathbf{d})} \left(I^{(\mathbf{d})}(t) \right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(\tau) \right\}
$$

$$
= \frac{1}{\mu_1 + \mu_2 + \lambda} \left[f(i) - \eta^{\mathbf{d}} \right]
$$

$$
+ \frac{\lambda}{\mu_1 + \mu_2 + \lambda} E \left\{ \int_0^{+\infty} \left[f^{(\mathbf{d})} \left(I^{(\mathbf{d})}(t) \right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(0) = i + 1 \right\}
$$

$$
+ \frac{\mu_1 + \mu_2}{\mu_1 + \mu_2 + \lambda} E \left\{ \int_0^{+\infty} \left[f^{(\mathbf{d})} \left(I^{(\mathbf{d})}(t) \right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(0) = i - 1 \right\}
$$

$$
= \frac{1}{\mu_1 + \mu_2 + \lambda} \left[f(i) - \eta^{\mathbf{d}} \right] + \frac{\lambda}{\mu_1 + \mu_2 + \lambda} g^{(\mathbf{d})}(i + 1)
$$

$$
+ \frac{\mu_1 + \mu_2}{\mu_1 + \mu_2 + \lambda} g^{(\mathbf{d})}(i - 1),
$$

where

$$
E \left\{ \tau \,\middle|\, I^{(\mathbf{d})}(t) = i \right\} = \frac{1}{\mu_1 + \mu_2 + \lambda}.
$$

We obtain

$$
(\mu_1 + \mu_2) g^{(\mathbf{d})}(i - 1) - (\mu_1 + \mu_2 + \lambda) g^{(\mathbf{d})}(i) + \lambda g^{(\mathbf{d})}(i + 1) = \eta^{\mathbf{d}} - f(i). \quad (24)
$$

Part (d): For $i = N$, by using Fig. 2 we have

$$
\begin{aligned}
g^{(\mathbf{d})}(N) &= E\left\{ \int_0^{+\infty} \left[f^{(\mathbf{d})}\left(I^{(\mathbf{d})}(t)\right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(0) = N \right\} \\
&= E\left\{ \tau \,\middle|\, I^{(\mathbf{d})}(0) = N \right\} \left[f(N) - \eta^{\mathbf{d}} \right] \\
&\quad + E\left\{ \int_\tau^{+\infty} \left[f^{(\mathbf{d})}\left(I^{(\mathbf{d})}(t)\right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(\tau) \right\} \\
&= \frac{1}{\mu_1 + \mu_2} \left[f(N) - \eta^{\mathbf{d}} \right] \\
&\quad + E\left\{ \int_0^{+\infty} \left[f^{(\mathbf{d})}\left(I^{(\mathbf{d})}(t)\right) - \eta^{\mathbf{d}} \right] dt \,\middle|\, I^{(\mathbf{d})}(0) = N - 1 \right\} \\
&= \frac{1}{\mu_1 + \mu_2} \left[f(N) - \eta^{\mathbf{d}} \right] + g^{(\mathbf{d})}(N-1),
\end{aligned}
$$

where

$$
E\left\{ \tau \,\middle|\, I^{(\mathbf{d})}(t) = N \right\} = \frac{1}{\mu_1 + \mu_2}.
$$

We obtain

$$
(\mu_1 + \mu_2)\, g^{(\mathbf{d})}(N-1) - (\mu_1 + \mu_2)\, g^{(\mathbf{d})}(N) = \eta^{\mathbf{d}} - f(N). \tag{25}
$$

Thus it follows from (22), (23), (24) and (25) that

$$
-\mathbf{B}^{(\mathbf{d})} \mathbf{g}^{(\mathbf{d})} = \mathbf{f}^{(\mathbf{d})} - \eta^{\mathbf{d}} \mathbf{e}, \tag{26}
$$

where $\mathbf{B}^{(\mathbf{d})}$, $\mathbf{f}^{(\mathbf{d})}$ and $\eta^{\mathbf{d}}$ are given in (3), (12) and (13), respectively.

In what follows we provide an effective method to solve the policy-based Poisson equation, and show that there exist infinitely-many solutions with two free constants of additive terms. This leads to a general solution with the two free constants of the policy-based Poisson equation.

To solve the system of linear equations (26), it is easy to see that $\mathrm{rank}\left(\mathbf{B}^{(\mathbf{d})}\right) = N$ and $\det\left(\mathbf{B}^{(\mathbf{d})}\right) = 0$ due to the fact that the size of the matrix $\mathbf{B}^{(\mathbf{d})}$ is $N + 1$. Hence, this system of linear equations (26) exists infinitely-many solutions with a free constant of an additive term.

Let $\mathcal{B}$ be a matrix obtained through omitting the first row and the first column vectors of the matrix $\mathbf{B^{(d)}}$. Then,

$$
\mathcal{B} = \begin{pmatrix}
-[\lambda + v(d_1)] & \lambda \\
v(d_2) & -[\lambda + v(d_2)] & \lambda \\
& \ddots & \ddots & \ddots \\
& v(d_K) & -[\lambda + v(d_K)] & \lambda \\
& & v(1) & -[\lambda + v(1)] & \lambda \\
& & & \ddots & & \ddots & \ddots \\
& & & & v(1) & -[\lambda + v(1)] & \lambda \\
& & & & & v(1) & -v(1)
\end{pmatrix}.
$$

Obviously, $\mathrm{rank}(\mathcal{B}) = N$. Since the size of the matrix $\mathcal{B}$ is N, the matrix $\mathcal{B}$ is invertible, and $(-\mathcal{B})^{-1} > 0$.

Let $\mathbf{H^{(d)}}$ and $\varphi^{(d)}$ be two column vectors of size N obtained through omitting the first elements of the two column vectors $\mathbf{f^{(d)}} - \eta^{\mathbf{d}}\mathbf{e}$ and $\mathbf{g^{(d)}}$ of size $N+1$, respectively. Then,

$$
\mathbf{H^{(d)}} = \begin{pmatrix} H_1^{(d)} \\ H_2^{(d)} \\ \vdots \\ H_K^{(d)} \\ H_{K+1}^{(d)} \\ \vdots \\ H_N^{(d)} \end{pmatrix} = \begin{pmatrix} f^{(d)}(1) - \eta^{\mathbf{d}} \\ f^{(d)}(2) - \eta^{\mathbf{d}} \\ \vdots \\ f^{(d)}(K) - \eta^{\mathbf{d}} \\ f(K+1) - \eta^{\mathbf{d}} \\ \vdots \\ f(N) - \eta^{\mathbf{d}} \end{pmatrix} = \begin{pmatrix} \left[B_1^{(d)} - D^{(d)}\right] - P\left[A_1^{(d)} - F^{(d)}\right] \\ \left[B_2^{(d)} - D^{(d)}\right] - P\left[A_2^{(d)} - F^{(d)}\right] \\ \vdots \\ \left[B_K^{(d)} - D^{(d)}\right] - P\left[A_K^{(d)} - F^{(d)}\right] \\ \left[B_{K+1} - D^{(d)}\right] - P\left[A_{K+1} - F^{(d)}\right] \\ \vdots \\ \left[B_N - D^{(d)}\right] - P\left[A_N - F^{(d)}\right] \end{pmatrix}
$$

and

$$
\varphi^{(d)} = \left(g^{(d)}(1), g^{(d)}(2), \ldots, g^{(d)}(K); g^{(d)}(K+1), g^{(d)}(K+2), \ldots, g^{(d)}(N)\right)^T.
$$

Therefore, it follows from (26) that

$$
-\mathcal{B}\varphi^{(d)} = \mathbf{H^{(d)}} + v(d_1)\mathbf{e}_1 g^{(d)}(0), \tag{27}
$$

where $\mathbf{e}_1$ is a column vector with the first element be one and all the others be zero. Note that the matrix $-\mathcal{B}$ is invertible and $(-\mathcal{B})^{-1} > 0$, thus the system of linear equations (27) always has one unique solution

$$
\varphi^{(d)} = (-\mathcal{B})^{-1}\mathbf{H^{(d)}} + v(d_1)(-\mathcal{B})^{-1}\mathbf{e}_1 \cdot \Im, \tag{28}
$$

where $g^{(\mathbf{d})}(0) = \Im$ is any given constant. Let's take a convention

$$\binom{a}{\mathbf{b}} = (a, \mathbf{b})^T ,$$

where $\mathbf{b}$ may be a column vector. Then we have

$$
\begin{aligned}
\mathbf{g}^{(\mathbf{d})} &= \left(g^{(\mathbf{d})}(0) , \varphi^{(\mathbf{d})} \right)^T \\
&= \left(\Im, (-\mathcal{B})^{-1} \mathbf{H}^{(\mathbf{d})} + v(d_1)(-\mathcal{B})^{-1} \mathbf{e}_1 \cdot \Im \right)^T \\
&= \left(0, (-\mathcal{B})^{-1} \mathbf{H}^{(\mathbf{d})} \right)^T + \left(1, v(d_1)(-\mathcal{B})^{-1} \mathbf{e}_1 \right)^T \Im.
\end{aligned}
\tag{29}
$$

Note that $\mathbf{B}^{(\mathbf{d})}\mathbf{e} = 0$, thus a general solution to the policy-based Poisson equation is further given by

$$\mathbf{g}^{(\mathbf{d})} = \left(0, (-\mathcal{B})^{-1} \mathbf{H}^{(\mathbf{d})} \right)^T + \left(1, v(d_1)(-\mathcal{B})^{-1} \mathbf{e}_1 \right)^T \Im + \xi \mathbf{e}, \tag{30}$$

where $\Im$ and ξ are two free constants.

Based on the above analysis, the following theorem summarizes the general solution of the policy-based Poisson equation.

Theorem 1 *For the Poisson equation $-\mathbf{B}^{(\mathbf{d})}\mathbf{g}^{(\mathbf{d})} = \mathbf{f}^{(\mathbf{d})} - \eta^{\mathbf{d}}\mathbf{e}$, there exists a key special solution $\mathbf{g}_{Sp}^{\mathbf{d}} = \left(0, (-\mathcal{B})^{-1} \mathbf{H}^{(\mathbf{d})} \right)^T$, and its general solution is related to two free constants $\Im$ and ξ such that*

$$\mathbf{g}^{(\mathbf{d})} = \mathbf{g}_{Sp}^{\mathbf{d}} + \left(1, v(d_1)(-\mathcal{B})^{-1} \mathbf{e}_1 \right)^T \Im + \xi \mathbf{e},$$

where ξ is a potential displacement constant, and $\Im$ is a solution-free constant.

Remark 3 (1) To our best knowledge, this is the first to provide the general solution of the Poisson equations in the MDPs by means of two different free constants.

(2) Note that $\pi^{(\mathbf{d})}\mathbf{g}^{(\mathbf{d})} = \eta^{\mathbf{d}}$ and the matrix $-\mathbf{B}^{(\mathbf{d})} + \mathbf{e}\pi^{(\mathbf{d})}$ is invertible, thus the Poisson equation $-\mathbf{B}^{(\mathbf{d})}\mathbf{g}^{(\mathbf{d})} = \mathbf{f}^{(\mathbf{d})} - \eta^{\mathbf{d}}\mathbf{e}$ can become

$$\left(-\mathbf{B}^{(\mathbf{d})} + \mathbf{e}\pi^{(\mathbf{d})} \right) \mathbf{g}^{(\mathbf{d})} = \mathbf{f}^{(\mathbf{d})}.$$

This gives a solution of the Poisson equation as follows:

$$\mathbf{g}^{(\mathbf{d})} = \left(-\mathbf{B}^{(\mathbf{d})} + \mathbf{e}\pi^{(\mathbf{d})} \right)^{-1} \mathbf{f}^{(\mathbf{d})} + \xi \mathbf{e},$$

which is a special solution of the Poisson equation by comparing with that in Theorem 1.

(3) To further understand the solution of the Poisson equation, readers may refer to, for example, Cao (2007) and Ma et al. (2019) for more details.

6 Impact of the penalty cost

In this section, we provide an explicit expression for the perturbation realization factor of the policy-based birth-death process. Based on this, we can set up a linear equation with respect to the penalty cost, which is well related to the performance difference equation. Furthermore, we discuss some useful properties of policies in the set $\mathcal{D}$ by means of the solution of the linear equation in the penalty cost.

6.1 The perturbation realization factor

We define a perturbation realization factor as

$$G^{(\mathbf{d})}(i) \overset{\text{def}}{=} g^{(\mathbf{d})}(i-1) - g^{(\mathbf{d})}(i), \, i = 1, 2, \ldots, N. \tag{31}$$

It is easy to see from Cao (2007) that $G^{(\mathbf{d})}(i)$ quantifies the difference among two adjacent performance potentials $g^{(\mathbf{d})}(i)$ and $g^{(\mathbf{d})}(i-1)$, and measures the effect on the long-run average profit of the stock-rationing queue when the system state is changed from $i-1$ to i. By using the policy-based Poisson equation (26), we can derive a new system of linear equations, which can be used to directly express the perturbation realization factor $G^{(\mathbf{d})}(i)$ for $i = 1, 2, \ldots, N$.

By using (30), we can directly express the perturbation realization factor $G^{(\mathbf{d})}(i)$ for $i = 1, 2, \ldots, N$. On the other hand, by observing the special structure of the policy-based Poisson equation (26), we can propose a new method of sequence to set up an explicit expression for $G^{(\mathbf{d})}(i)$.

For $i = 1$, it follows from (22) that

$$-\lambda \left[g^{(\mathbf{d})}(0) - g^{(\mathbf{d})}(1) \right] = -\lambda G^{(\mathbf{d})}(1),$$

we have

$$\lambda G^{(\mathbf{d})}(1) = f(0) - \eta^{\mathbf{d}}. \tag{32}$$

For $i = 2, 3, \ldots, K$, it follows from (23) that

$$v(d_i) \left[g^{(\mathbf{d})}(i-1) - g^{(\mathbf{d})}(i) \right] - \lambda \left[g^{(\mathbf{d})}(i) - g^{(\mathbf{d})}(i+1) \right]$$
$$= v(d_i) G^{(\mathbf{d})}(i) - \lambda G^{(\mathbf{d})}(i+1),$$

this gives

$$\lambda G^{(\mathbf{d})}(i+1) = v(d_i) G^{(\mathbf{d})}(i) + f^{(\mathbf{d})}(i) - \eta^{\mathbf{d}}. \tag{33}$$

For $i = K+1, K+2, \ldots, N-1$, it follows from (24) that

$$(\mu_1 + \mu_2) \left[g^{(\mathbf{d})}(i-1) - g^{(\mathbf{d})}(i) \right] - \lambda \left[g^{(\mathbf{d})}(i) - g^{(\mathbf{d})}(i+1) \right]$$
$$= (\mu_1 + \mu_2) G^{(\mathbf{d})}(i) - \lambda G^{(\mathbf{d})}(i+1),$$

we obtain

$$\lambda G^{(\mathbf{d})}(i+1) = (\mu_1 + \mu_2) G^{(\mathbf{d})}(i) + f(i) - \eta^{\mathbf{d}}. \tag{34}$$

For $i = N$, it follows from (25) that

$$(\mu_1 + \mu_2) G^{(\mathbf{d})}(N) = \eta^{\mathbf{d}} - f(N). \tag{35}$$

By using (32), (33), (34) and (35), we obtain a new system of linear equations satisfied by $G^{(\mathbf{d})}(i)$ as follows:

$$\begin{cases} \lambda G^{(\mathbf{d})}(1) = f(0) - \eta^{\mathbf{d}}, & i = 1, \\ \lambda G^{(\mathbf{d})}(i+1) = v(d_i) G^{(\mathbf{d})}(i) + f^{(\mathbf{d})}(i) - \eta^{\mathbf{d}}, & i = 2, 3, \ldots, K, \\ \lambda G^{(\mathbf{d})}(i+1) = (\mu_1 + \mu_2) G^{(\mathbf{d})}(i) + f(i) - \eta^{\mathbf{d}}, & i = K+1, K+2, \ldots, N-1, \\ (\mu_1 + \mu_2) G^{(\mathbf{d})}(N) = \eta^{\mathbf{d}} - f(N), & i = N. \end{cases} \tag{36}$$

Fortunately, the following theorem can provide an explicit expression for the perturbation realization factor $G^{(\mathbf{d})}(i)$ for $1 \le i \le N$.

Theorem 2 *For any given policy* $\mathbf{d}$*, the perturbation realization factor* $G^{(\mathbf{d})}(i)$ *is given by*

(a) *for* $1 \le i \le K$,

$$G^{(\mathbf{d})}(i) = \lambda^{-i} \left[f(0) - \eta^{\mathbf{d}} \right] \prod_{k=1}^{i-1} v(d_k) + \sum_{r=1}^{i-1} \lambda^{r-i} \left[f^{(\mathbf{d})}(r) - \eta^{\mathbf{d}} \right] \prod_{k=r+1}^{i-1} v(d_k); \tag{37}$$

(b) *for* $K+1 \le i \le N$,

$$\begin{aligned} G^{(\mathbf{d})}(i) = {} & \lambda^{-i} \left[f(0) - \eta^{\mathbf{d}} \right] \prod_{k=1}^{K} v(d_k) \left[v(1) \right]^{i-K-1} \\ & + \sum_{r=1}^{K-1} \lambda^{r-K} \left[f^{(\mathbf{d})}(r) - \eta^{\mathbf{d}} \right] \prod_{k=r+1}^{K} v(d_k) \\ & + \sum_{r=K}^{i-1} \lambda^{r-i} \left[f(r) - \eta^{\mathbf{d}} \right] \left[v(1) \right]^{i-r-2}. \end{aligned}$$

Proof We only prove (a), since the proof of (b) is similar.

It follows from (36) that

$$G^{(\mathbf{d})}(1) = \frac{f(0) - \eta^{\mathbf{d}}}{\lambda}.$$

Similarly, we obtain

$$G^{(\mathbf{d})}(i+1) = \frac{v(d_i)}{\lambda} G^{(\mathbf{d})}(i) + \frac{f^{(\mathbf{d})}(i) - \eta^{\mathbf{d}}}{\lambda}, \quad i = 1, 2, \ldots, K.$$

By using (1.2.4) in Chapter 1 of Elaydi (1996), we can obtain the explicit expression of the perturbation realization factor as follows:

$$G^{(\mathbf{d})}(i) = \lambda^{-i}\left[f(0) - \eta^{\mathbf{d}}\right]\prod_{k=1}^{i-1}v(d_k) + \sum_{r=1}^{i-1}\lambda^{r-i}\left[f^{(\mathbf{d})}(r) - \eta^{\mathbf{d}}\right]\prod_{k=r+1}^{i-1}v(d_k)$$

for $i = 1, 2, \ldots, K$. This completes the proof. $\square$

6.2 The performance difference equation

For any given policy $\mathbf{d} \in \mathcal{D}$, the long-run average profit of the stock-rationing queue is given by

$$\eta^{\mathbf{d}} = \pi^{(\mathbf{d})}\mathbf{f}^{(\mathbf{d})},$$

and the policy-based Poisson equation is given by

$$\mathbf{B}^{(\mathbf{d})}\mathbf{g}^{(\mathbf{d})} = \eta^{\mathbf{d}}\mathbf{e} - \mathbf{f}^{(\mathbf{d})}.$$

It is seen from (3) and (12) that Policy $\mathbf{d}$ directly affects not only the elements of the infinitesimal generator $\mathbf{B}^{(\mathbf{d})}$ but also the reward function $\mathbf{f}^{(\mathbf{d})}$. Based on this, if Policy $\mathbf{d}$ changes to $\mathbf{d}'$, then the infinitesimal generator $\mathbf{B}^{(\mathbf{d})}$ and the reward function $\mathbf{f}^{(\mathbf{d})}$ can have their corresponding changes $\mathbf{B}^{(\mathbf{d}')}$ and $\mathbf{f}^{(\mathbf{d}')}$, respectively.

The following lemma provides a useful equation (called performance difference equation) for the difference $\eta^{\mathbf{d}'} - \eta^{\mathbf{d}}$ corresponding to any two different policies $\mathbf{d}, \mathbf{d}' \in \mathcal{D}$. Here, we only restate the performance difference equation without proof, readers may refer to Cao (2007) or Ma et al. (2019) for more details.

Lemma 1 *For any two policies* $\mathbf{d}, \mathbf{d}' \in \mathcal{D}$, *we have*

$$\eta^{\mathbf{d}'} - \eta^{\mathbf{d}} = \pi^{(\mathbf{d}')}\left[\left(\mathbf{B}^{(\mathbf{d}')} - \mathbf{B}^{(\mathbf{d})}\right)\mathbf{g}^{(\mathbf{d})} + \left(\mathbf{f}^{(\mathbf{d}')} - \mathbf{f}^{(\mathbf{d})}\right)\right]. \tag{38}$$

By using the performance difference Eq. (38), we can set up a partial order relation for the policies in the policy set $\mathcal{D}$ as follows. For any two policies $\mathbf{d}, \mathbf{d}' \in \mathcal{D}$, we write that $\mathbf{d}' \succ \mathbf{d}$ if $\eta^{\mathbf{d}'} > \eta^{\mathbf{d}}$; $\mathbf{d}' \approx \mathbf{d}$ if $\eta^{\mathbf{d}'} = \eta^{\mathbf{d}}$; and $\mathbf{d}' \prec \mathbf{d}$ if $\eta^{\mathbf{d}'} < \eta^{\mathbf{d}}$. Also, we write that $\mathbf{d}' \succeq \mathbf{d}$ if $\eta^{\mathbf{d}'} \geq \eta^{\mathbf{d}}$; and $\mathbf{d}' \preceq \mathbf{d}$ if $\eta^{\mathbf{d}'} \leq \eta^{\mathbf{d}}$.

Under this partial order relation, our research target is to find the optimal policy $\mathbf{d}^* \in \mathcal{D}$ such that $\mathbf{d}^* \succeq \mathbf{d}$ for any policy $\mathbf{d} \in \mathcal{D}$, i.e.,

$$\mathbf{d}^* = \arg\max_{\mathbf{d}\in\mathcal{D}}\left\{\eta^{\mathbf{d}}\right\}.$$

Note that the policy set $\mathcal{D}$ and the state set $\mathbf{\Omega}$ are all finite, thus an enumeration method using finite comparisons is feasible for finding the optimal policy $\mathbf{d}^*$ in the policy set $\mathcal{D}$.

To find the optimal policy $\mathbf{d}^*$, we define two policies $\mathbf{d}$ and $\mathbf{d}'$ with an interrelated structure at Position i as follows:

$$\mathbf{d} = \left(0; d_1, d_2, \ldots, d_{i-1}, \underline{d_i}, d_{i+1}, \ldots, d_K; 1, 1, \ldots, 1\right),$$
$$\mathbf{d}' = \left(0; d_1, d_2, \ldots, d_{i-1}, \underline{d_i'}, d_{i+1}, \ldots, d_K; 1, 1, \ldots, 1\right),$$

where $d_i', d_i \in \{0, 1\}$ with $d_i' \neq d_i$. Clearly, if the two policies $\mathbf{d}$ and $\mathbf{d}'$ have an interrelated structure at Position i, then only the difference between the two policies $\mathbf{d}$ and $\mathbf{d}'$ is at their ith elements: d_i and d_i'.

Lemma 2 *For the two policies $\mathbf{d}$ and $\mathbf{d}'$ with an interrelated structure at Position i: d_i and d_i', we have*

$$\eta^{\mathbf{d}'} - \eta^{\mathbf{d}} = \mu_2 \pi^{(\mathbf{d}')}(i)\left(d_i' - d_i\right)\left[G^{(\mathbf{d})}(i) + b\right], \tag{39}$$

where $b = R + C_{2,2} - P$.

Proof For the two policies $\mathbf{d}$ and $\mathbf{d}'$ with an interrelated structure at Position i: d_i and d_i', we have

$$\mathbf{d} = \left(0; d_1, d_2, \ldots, d_{i-1}, \underline{d_i}, d_{i+1}, \ldots, d_K; 1, 1, \ldots, 1\right),$$
$$\mathbf{d}' = \left(0; d_1, d_2, \ldots, d_{i-1}, \underline{d_i'}, d_{i+1}, \ldots, d_K; 1, 1, \ldots, 1\right).$$

It is easy to check from (3) that

$$\mathbf{B}^{(\mathbf{d}')} - \mathbf{B}^{(\mathbf{d})} = \begin{pmatrix} 0 & & & & & & \\ & 0 & \ddots & & & & \\ & & \ddots & \ddots & 0 & & \\ & & & \left(d_i' - d_i\right)\mu_2 & -\left(d_i' - d_i\right)\mu_2 & & \\ & & & & 0 & 0 & \\ & & & & & \ddots & \ddots \\ & & & & & & 0 \quad 0 \end{pmatrix}. \tag{40}$$

Also, from the reward function (9), we obtain

$$f^{(\mathbf{d})}(i) = \left(R + C_{2,2} - P\right)\mu_2 d_i + R\mu_1 - C_1 i - C_{2,2}\mu_2 - C_3\lambda$$

and

$$f^{(\mathbf{d}')}(i) = \left(R + C_{2,2} - P\right)\mu_2 d_i' + R\mu_1 - C_1 i - C_{2,2}\mu_2 - C_3\lambda.$$

This gives

$$\mathbf{f}^{(\mathbf{d}')} - \mathbf{f}^{(\mathbf{d})} = \left(0, 0, \ldots, 0, b\mu_2\left(d_i' - d_i\right), 0, \ldots, 0\right)^T. \tag{41}$$

Thus, it follows from Lemma 1, (40) and (41) that

$$
\begin{aligned}
\eta^{\mathbf{d}'} - \eta^{\mathbf{d}} &= \pi^{(\mathbf{d}')} \left[\left(\mathbf{B}^{(\mathbf{d}')} - \mathbf{B}^{(\mathbf{d})} \right) \mathbf{g}^{(\mathbf{d})} + \left(\mathbf{f}^{(\mathbf{d}')} - \mathbf{f}^{(\mathbf{d})} \right) \right] \\
&= \mu_2 \pi^{(\mathbf{d}')}(i) \left(d_i' - d_i \right) \left[g^{(\mathbf{d})}(i-1) - g^{(\mathbf{d})}(i) + b \right] \\
&= \mu_2 \pi^{(\mathbf{d}')}(i) \left(d_i' - d_i \right) \left[G^{(\mathbf{d})}(i) + b \right].
\end{aligned}
\tag{42}
$$

This completes the proof. □

For $d_i', d_i \in \{0, 1\}$ with $d_i' \neq d_i$, we have

$$
d_i' - d_i = \begin{cases} 1, & d_i' = 1, d_i = 0; \\ -1, & d_i' = 0, d_i = 1. \end{cases}
$$

Therefore, it is easy to see from (39) that to compare $\eta^{\mathbf{d}'}$ with $\eta^{\mathbf{d}}$, it is necessary to further analyze the sign of function $G^{(\mathbf{d})}(i) + b$. This will be developed in the next subsection.

6.3 The sign of $G^{(\mathbf{d})}(i) + b$

As seen from (42), the sign analysis of the performance difference $\eta^{\mathbf{d}'} - \eta^{\mathbf{d}}$ directly depends on that of $G^{(\mathbf{d})}(i) + b$. Thus, this subsection provides the sign analysis of $G^{(\mathbf{d})}(i) + b$ with respect to the penalty cost P.

Suppose that the inventory level is low. If the service priority is violated (i.e. the demands of Class 2 are served at a low stock), then the warehouse has to pay the penalty cost P for each product supplied to the demands of Class 2. Based on this, we study the influence of the penalty cost P on the sign of $G^{(\mathbf{d})}(i) + b$.

Substituting (14), (15), (16) and (17) into (37), we obtain that for $1 \leq i \leq K$,

$$
\begin{aligned}
G^{(\mathbf{d})}&(i) + b \\
&= R + C_{2,2} + \lambda^{-i} \left[B_0 - D^{(\mathbf{d})} \right] \prod_{k=1}^{i-1} v(d_k) + \sum_{r=1}^{i-1} \lambda^{r-i} \left[B_r^{(\mathbf{d})} - D^{(\mathbf{d})} \right] \prod_{k=r+1}^{i-1} v(d_k) \\
&\quad - P \left\{ 1 + \lambda^{-i} \left[A_0 - F^{(\mathbf{d})} \right] \prod_{k=1}^{i-1} v(d_k) + \sum_{r=1}^{i-1} \lambda^{r-i} \left[A_r^{(\mathbf{d})} - F^{(\mathbf{d})} \right] \prod_{k=r+1}^{i-1} v(d_k) \right\},
\end{aligned}
\tag{43}
$$

which is linear in the penalty cost P.

From $G^{(\mathbf{d})}(i) + b = 0$, we have

$$
P \left\{ 1 + \lambda^{-i} \left[A_0 - F^{(\mathbf{d})} \right] \prod_{k=1}^{i-1} v(d_k) + \sum_{r=1}^{i-1} \lambda^{r-i} \left[A_r^{(\mathbf{d})} - F^{(\mathbf{d})} \right] \prod_{k=r+1}^{i-1} v(d_k) \right\}
$$

Table 2 Numerical analysis of solutions for three different policies

$\mathfrak{P}_i^{(\mathbf{d})}$	$i = 0$	1	2	3	4	5	6	7	8	9	10
$\mathbf{d}_1$	11	11	38.3	20.7	21.4	21.3	20.9	20.7	20.4	20.1	19.8
$\mathbf{d}_2$	11	11	-6.5	14.3	88.4	384.9	$1.5e3$	$5.6e3$	$2.1e4$	$7.5e4$	$2.6e5$
$\mathbf{d}_3$	11	11	-7.1	23.0	-181.4	-80.4	-68.6	15.0	13.7	13.3	13.1

$$= R + C_{2,2} + \lambda^{-i} \left[B_0 - D^{(\mathbf{d})} \right] \prod_{k=1}^{i-1} v(d_k) + \sum_{r=1}^{i-1} \lambda^{r-i} \left[B_r^{(\mathbf{d})} - D^{(\mathbf{d})} \right] \prod_{k=r+1}^{i-1} v(d_k), \tag{44}$$

thus, the unique solution of the penalty cost P to Equation (44) is given by

$$\mathfrak{P}_i^{(\mathbf{d})} = \frac{R + C_{2,2} + \lambda^{-i} \left[B_0 - D^{(\mathbf{d})} \right] \prod_{k=1}^{i-1} v(d_k) + \sum_{r=1}^{i-1} \lambda^{r-i} \left[B_r^{(\mathbf{d})} - D^{(\mathbf{d})} \right] \prod_{k=r+1}^{i-1} v(d_k)}{1 + \lambda^{-i} \left[A_0 - F^{(\mathbf{d})} \right] \prod_{k=1}^{i-1} v(d_k) + \sum_{r=1}^{i-1} \lambda^{r-i} \left[A_r^{(\mathbf{d})} - F^{(\mathbf{d})} \right] \prod_{k=r+1}^{i-1} v(d_k)}. \tag{45}$$

It's easy to see from (43) that if $\mathfrak{P}_i^{(\mathbf{d})} > 0$ and $0 \le P \le \mathfrak{P}_i^{(\mathbf{d})}$, then $G^{(\mathbf{d})}(i) + b \ge 0$; while if $P \ge \mathfrak{P}_i^{(\mathbf{d})}$, then $G^{(\mathbf{d})}(i) + b \le 0$. Note that the equality can hold only if $P = \mathfrak{P}_i^{(\mathbf{d})}$

To understand the solution $\mathfrak{P}_i^{(\mathbf{d})}$ for $1 \le i \le K$, we use a numerical example to show the solutions in Table 2. To do this, we take the system parameters: $\lambda = 3$, $\mu_1 = 4$, $\mu_2 = 2$, $C_1 = 1$, $C_{2,1} = 4$, $C_{2,2} = 1$, $C_3 = 5$ and $C_4 = 1$. Further, we observe three different policies:

$$\mathbf{d}_1 = (0; 1, 1, 1, 1, 1, 1, 1, 1, 1, 1; 1, 1, 1, 1, 1),$$
$$\mathbf{d}_2 = (0; 0, 0, 0, 0, 0, 0, 0, 0, 0, 0; 1, 1, 1, 1, 1),$$
$$\mathbf{d}_3 = (0; 0, 0, 0, 0, 0, 1, 1, 1, 1, 1; 1, 1, 1, 1, 1).$$

In the stock-rationing queue, we define two critical values related to the penalty cost P as

$$P_H(\mathbf{d}) = \max_{\mathbf{d} \in \mathcal{D}} \left\{ 0, \mathfrak{P}_1^{(\mathbf{d})}, \mathfrak{P}_2^{(\mathbf{d})}, \ldots, \mathfrak{P}_K^{(\mathbf{d})} \right\}, \tag{46}$$

and

$$P_L(\mathbf{d}) = \min_{\mathbf{d} \in \mathcal{D}} \left\{ \mathfrak{P}_1^{(\mathbf{d})}, \mathfrak{P}_2^{(\mathbf{d})}, \ldots, \mathfrak{P}_K^{(\mathbf{d})} \right\}. \tag{47}$$

From Table 2, we see that it is possible to have $P_L(\mathbf{d}) < 0$ for Policy $\mathbf{d} = \mathbf{d}_2$ or $\mathbf{d} = \mathbf{d}_3$.

The following proposition uses the two critical values $P_H(\mathbf{d})$ and $P_L(\mathbf{d})$, together with the penalty cost P, to provide some sufficient conditions under which the function $G^{(\mathbf{d})}(i) + b$ is either positive, zero or negative.

Proposition 1 (1) *If $P \geq P_H(\mathbf{d})$ for any given policy $\mathbf{d} \in \mathcal{D}$, then for each $i = 1, 2, \ldots, K$,*

$$G^{(\mathbf{d})}(i) + b \leq 0.$$

(2) *If $P_L(\mathbf{d}) > 0$ and $0 \leq P \leq P_L(\mathbf{d})$ for any given policy $\mathbf{d} \in \mathcal{D}$, then for each $i = 1, 2, \ldots, K$,*

$$G^{(\mathbf{d})}(i) + b \geq 0.$$

Proof (1) For any given policy $\mathbf{d} \in \mathcal{D}$, if $P \geq P_H(\mathbf{d})$, then it follows from (46) that for each $i = 1, 2, \ldots, K$,

$$P \geq \mathfrak{P}_i^{(\mathbf{d})},$$

this leads to that $G^{(\mathbf{d})}(i) + b \leq 0$.

(2) For any given policy $\mathbf{d} \in \mathcal{D}$, if $P_L(\mathbf{d}) > 0$ and $0 \leq P \leq P_L(\mathbf{d})$, then it follows from (47) that for each $i = 1, 2, \ldots, K$,

$$0 \leq P \leq \mathfrak{P}_i^{(\mathbf{d})},$$

this gives that $G^{(\mathbf{d})}(i) + b \geq 0$. This completes the proof. □

However, for the case with $P_L(\mathbf{d}) < P < P_H(\mathbf{d})$ for any given policy $\mathbf{d} \in \mathcal{D}$, it is a little bit complicated to determine the sign of $G^{(\mathbf{d})}(i) + b$ for each $i = 1, 2, \ldots, K$. For this reason, our discussion will be left in the next section.

For any two policies $\mathbf{d}, \mathbf{c} \in \mathcal{D}$,

$$\mathbf{d} = (0; d_1, d_2, \ldots, d_{i-1}, d_i, d_{i+1}, \ldots, d_K; 1, 1, \ldots, 1),$$
$$\mathbf{c} = (0; c_1, c_2, \ldots, c_{i-1}, c_i, c_{i+1}, \ldots, c_K; 1, 1, \ldots, 1).$$

we write

$$S(\mathbf{d}, \mathbf{c}) = \{i : d_i \neq c_i, i = 1, 2, \ldots, K - 1, K\}$$

and its complementary set

$$\overline{S(\mathbf{d}, \mathbf{c})} = \{i : d_i = c_i, i = 1, 2, \ldots, K - 1, K\}.$$

Then

$$S(\mathbf{d}, \mathbf{c}) \cup \overline{S(\mathbf{d}, \mathbf{c})} = \{1, 2, \ldots, K - 1, K\}.$$

The following lemma sets up a policy sequence such that any two adjacent policies of them have the difference at the corresponding position of only one element. The proof is easy and is omitted here.

Lemma 3 *For any two policies* $\mathbf{d}, \mathbf{c} \in \mathcal{D}$, $S(\mathbf{d}, \mathbf{c}) = \{i_1, i_2, i_3, \ldots, i_{n-1}, i_n\}$, *then there exists a policy sequence:* $\mathbf{d}^{(k)}$ *for* $k = 1, 2, 3, \ldots, n-1, n$, *such that*

$$S\left(\mathbf{d}, \mathbf{d}^{(1)}\right) = \{j_1\},$$

$$S\left(\mathbf{d}^{(1)}, \mathbf{d}^{(2)}\right) = \{j_2\},$$

$$\vdots$$

$$S\left(\mathbf{d}^{(n-1)}, \mathbf{d}^{(n)}\right) = \{j_n\},$$

where $\mathbf{d}^{(n)} = \mathbf{c}$, *and* $\{i_1, i_2, i_3, \ldots, i_{n-1}, i_n\} = \{j_1, j_2, j_3, \ldots, j_{n-1}, j_n\}$. *Also, for* $k = 1, 2, 3, \ldots, n-1, n$, *we have*

$$S\left(\mathbf{d}, \mathbf{d}^{(k)}\right) = \{j_1, j_2, j_3, \ldots, j_k\}.$$

The following theorem provides a class property of the policies in the set $\mathcal{D}$ by means of the function $G^{(\mathbf{c})}(i) + b$ for any policy $\mathbf{c} \in \mathcal{D}$ and for each $i \in S(\mathbf{d}, \mathbf{c})$, where Policy $\mathbf{d}$ is any given reference policy in the set $\mathcal{D}$. Note that the class property will play a key role in developing some new structural properties of the optimal dynamic rationing policy.

Theorem 3 (1) *If* $P \geq P_H(\mathbf{d})$ *for any given policy* $\mathbf{d}$, *then for any policy* $\mathbf{c} \in \mathcal{D}$ *and for each* $i \in S(\mathbf{d}, \mathbf{c})$,

$$G^{(\mathbf{c})}(i) + b \leq 0.$$

(2) *If* $P_L(\mathbf{d}) > 0$ *and* $0 \leq P \leq P_L(\mathbf{d})$ *for any given policy* $\mathbf{d}$, *then for any policy* $\mathbf{c} \in \mathcal{D}$ *and for each* $i \in S(\mathbf{d}, \mathbf{c})$,

$$G^{(\mathbf{c})}(i) + b \geq 0.$$

Proof We only prove (1), while (2) can be proved similarly.

If $P \geq P_H(\mathbf{d})$ for any given policy $\mathbf{d}$, then it follows from (1) of Proposition 1 that for $i = 1, 2, \ldots, K$,

$$G^{(\mathbf{d})}(i) + b \leq 0.$$

From Policy $\mathbf{d}$, we observe any different policy $\mathbf{c} \in \mathcal{D}$. If the two policies $\mathbf{d}$ and $\mathbf{c}$ have n different elements: $d_{i_l} \neq c_{i_l}$ for $l = 1, 2, \ldots, n$, then $S(\mathbf{d}, \mathbf{c}) = \{i_l : l = 1, 2, \ldots, n\}$.

Note that the performance difference Eq. (39) can only be applied to two policies $\mathbf{d}'$ and $\mathbf{d}$ with an interrelated structure at Position $i : d_i', d_i \in \{0, 1\}$ with $d_i' \neq d_i$, thus for a policy $\mathbf{c} \in \mathcal{D}$ with $S(\mathbf{d}, \mathbf{c}) = \{i_1, i_2, i_3, \ldots, i_{n-1}, i_n\}$, our following discussion needs to use the policy sequence: $\mathbf{d}^{(k)}$ for $k = 1, 2, 3, \ldots, n-1, n$, given in Lemma

3. To this end, our further proof is to use the mathematical induction in the following three steps:

Step one: Analyzing the two policies $\mathbf{d}$ and $\mathbf{d}^{(1)}$.

For each $j_1 \in \{i_1, i_2, i_3, \ldots, i_{n-1}, i_n\}$, we take $S\left(\mathbf{d}, \mathbf{d}^{(1)}\right) = \{j_1\}$. It follows from the performance difference Eq. (39) that

$$\eta^{\mathbf{d}^{(1)}} - \eta^{\mathbf{d}} = \mu_2 \pi^{\left(\mathbf{d}^{(1)}\right)} (j_1) \left(d_{j_1}^{\left(\mathbf{d}^{(1)}\right)} - d_{j_1} \right) \left[G^{(\mathbf{d})} (j_1) + b \right]. \tag{48}$$

Similarly, we have

$$\eta^{\mathbf{d}} - \eta^{\mathbf{d}^{(1)}} = \mu_2 \pi^{(\mathbf{d})} (j_1) \left(d_{j_1} - d_{j_1}^{\left(\mathbf{d}^{(1)}\right)} \right) \left[G^{\left(\mathbf{d}^{(1)}\right)} (j_1) + b \right]. \tag{49}$$

It is easy to see from (48) and (49) that

$$G^{\left(\mathbf{d}^{(1)}\right)} (j_1) + b = \frac{\pi^{\left(\mathbf{d}^{(1)}\right)} (j_1)}{\pi^{(\mathbf{d})} (j_1)} \left[G^{(\mathbf{d})} (j_1) + b \right] \leq 0. \tag{50}$$

Therefore, for Policy $\mathbf{d}^{(1)} \in \mathcal{D}$, $G^{\left(\mathbf{d}^{(1)}\right)} (j_1) + b \leq 0$ for each $j_1 \in \{i_1, i_2, i_3, \ldots, i_{n-1}, i_n\}$.

Step two: Analyzing the two policies $\mathbf{d}^{(1)}$ and $\mathbf{d}^{(2)}$.

For each $j_2 \in \{i_1, i_2, i_3, \ldots, i_{n-1}, i_n\}$, we take $S\left(\mathbf{d}^{(1)}, \mathbf{d}^{(2)}\right) = \{j_2\}$. It is easy to see from (50) that

$$G^{\left(\mathbf{d}^{(2)}\right)} (j_2) + b = \frac{\pi^{\left(\mathbf{d}^{(2)}\right)} (j_2)}{\pi^{\left(\mathbf{d}^{(1)}\right)} (j_2)} \left[G^{\left(\mathbf{d}^{(1)}\right)} (j_2) + b \right] \leq 0.$$

Therefore, for Policy $\mathbf{d}^{(2)} \in \mathcal{D}$, $G^{\left(\mathbf{d}^{(2)}\right)} (j_2) + b \leq 0$ for each $j_2 \in \{i_1, i_2, i_3, \ldots, i_{n-1}, i_n\}$.

Step three: Assume that for $l = 3, 4, \ldots, k-2, k-1$, we have obtained that for Policy $\mathbf{d}^{(l)} \in \mathcal{D}$ with $S\left(\mathbf{d}^{(l-1)}, \mathbf{d}^{(l)}\right) = \{j_l\}$, $G^{\left(\mathbf{d}^{(l)}\right)} (j_l) + b \leq 0$ for each $j_l \in \{i_1, i_2, i_3, \ldots, i_{n-1}, i_n\}$. Now, we prove the next case with $l = k$.

For each $j_k \in \{i_1, i_2, i_3, \ldots, i_{n-1}, i_n\}$, we take $S\left(\mathbf{d}^{(k-1)}, \mathbf{d}^{(k)}\right) = \{j_k\}$. It is easy to see from (50) that

$$G^{\left(\mathbf{d}^{(k)}\right)} (j_k) + b = \frac{\pi^{\left(\mathbf{d}^{(k)}\right)} (j_k)}{\pi^{\left(\mathbf{d}^{(k-1)}\right)} (j_k)} \left[G^{\left(\mathbf{d}^{(k-1)}\right)} (j_k) + b \right] \leq 0.$$

This gives that for Policy $\mathbf{d}^{(k)} \in \mathcal{D}$, $G^{\left(\mathbf{d}^{(k)}\right)} (j_k) + b \leq 0$ for each $j_k \in \{i_1, i_2, i_3, \ldots, i_{n-1}, i_n\}$. Thus, this result holds for the case with $l = k$.

Following the above analysis, we can prove by induction that for Policy $\mathbf{d}^{(n)} \in \mathcal{D}$, $G^{\left(\mathbf{d}^{(n)}\right)} (j_n) + b \leq 0$ for each $j_n \in \{i_1, i_2, i_3, \ldots, i_{n-1}, i_n\}$. Since $\mathbf{c} = \mathbf{d}^{(n)}$, we obtain

that for Policy $\mathbf{c} \in \mathcal{D}$, $G^{(\mathbf{c})}(i) + b \leq 0$ for each $i \in \{i_1, i_2, i_3, \ldots, i_{n-1}, i_n\}$. This completes the proof. $\square$

7 Monotonicity and optimality

In this section, we analyze the optimal dynamic rationing policy in the three different areas of the penalty cost: $P \geq P_H(\mathbf{d})$; $P_L(\mathbf{d}) > 0$ and $0 < P \leq P_L(\mathbf{d})$; and $P_L(\mathbf{d}) < P < P_H(\mathbf{d})$, which are studied as three different subsections, respectively. For the three areas, some new structural properties of the optimal dynamic rationing policy are given by using our algebraic method. Also, it is easy to see that for the first two areas: $P \geq P_H(\mathbf{d})$; and $P_L(\mathbf{d}) > 0$ and $0 < P \leq P_L(\mathbf{d})$, the optimal dynamic rationing policy is of threshold type; while for the third area: $P_L(\mathbf{d}) < P < P_H(\mathbf{d})$, it may not be of threshold type but must be of transformational threshold type.

As seen from Lemma 2, to compare $\eta^{\mathbf{d}'}$ with $\eta^{\mathbf{d}}$, our aim is to focus on only Position i with $d_i' \neq d_i$ for $d_i', d_i \in \{0, 1\}$. Also, Lemma 3 provides a useful class property of the policies in the set $\mathcal{D}$ under the function $G^{(\mathbf{c})}(i) + b$ for any policy $\mathbf{c}, \mathbf{d} \in \mathcal{D}$ and for each $i \in S(\mathbf{d}, \mathbf{c})$. The two lemmas are very useful for our research in the next subsections.

7.1 The penalty cost $P \geq P_H(\mathbf{d})$

In this subsection, for the area of the penalty cost: $P \geq P_H(\mathbf{d})$ for any given policy $\mathbf{d}$, we find the optimal dynamic rationing policy of the stock-rationing queue, and further compute the maximal long-run average profit of this system.

The following theorem uses the class property of the policies in the set $\mathcal{D}$, given in (1) of Theorem 3, to set up some basic relations between any two policies. Thus, we find the optimal dynamic rationing policy of the stock-rationing queue.

Theorem 4 *If $P \geq P_H(\mathbf{d})$ for any given policy $\mathbf{d}$, then the optimal dynamic rationing policy of the stock-rationing queue is given by*

$$\mathbf{d}^* = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1).$$

This shows that if the penalty cost is higher with $P \geq P_H(\mathbf{d})$ for any given policy $\mathbf{d}$, then the warehouse can not supply any product to the demands of Class 2.

Proof If $P \geq P_H$ for any given policy $\mathbf{d}$, then our proof will focus on that for any policy $\mathbf{c} \in \mathcal{D}$, we can have

$$\mathbf{d}^* \succeq \mathbf{c}.$$

Based on this, we need to study some useful relations among the three policies: $\mathbf{d}$, $\mathbf{c}$ and $\mathbf{d}^*$, where $\mathbf{d}^*$ is deterministic with $d_i^* = 0$ for each $i = 1, 2, \ldots, K-1, K$.

To compare $\eta^{\mathbf{c}}$ with $\eta^{\mathbf{d}^*}$, let $S(\mathbf{d}^*, \mathbf{c}) = \{n_l : l = 1, 2, \ldots, n\}$ for $1 \leq n \leq K$. Then $c_{n_l} = 1$ for $l = 1, 2, \ldots, n$, since $d_i^* = 0$ for each $i = 1, 2, \ldots, K-1, K$.

For the two policies $\mathbf{d}$ and $\mathbf{c}$, we have $d_i, c_i \in \{0, 1\}$. Further, for the three elements: d_i, c_i and $d_i^* = 0$ for $i \in S(\mathbf{d}^*, \mathbf{c})$, we need to consider four different cases as follows:

Case one: $d_i = c_i = d_i^* = 0$. Since $c_i = d_i^*$, this case does not require any analysis by using Lemma 2.

Case two: $d_i = 1$ and $c_i = d_i^* = 0$. Since $c_i = d_i^*$, this case does not require any analysis by using Lemma 2.

Case three: $c_i = 1$ and $d_i = d_i^* = 0$. Note that $c_i \neq d_i$, by using (1) of Theorem 3, we obtain that $G^{(\mathbf{c})}(i) + b \leq 0$. On the other hand, since $c_i \neq d_i^*$, it follows from the performance difference Eq. (39) that for each $i \in S(\mathbf{d}^*, \mathbf{c})$,

$$\eta^{\mathbf{d}^*} - \eta^{\mathbf{c}} = \mu_2 \pi^{(\mathbf{d}^*)}(i)\left(d_i^* - c_i\right)\left[G^{(\mathbf{c})}(i) + b\right]$$
$$= -\mu_2 \pi^{(\mathbf{d}^*)}(i)\left[G^{(\mathbf{c})}(i) + b\right] \geq 0.$$

Thus $\eta^{\mathbf{d}^*} \geq \eta^{\mathbf{c}}$, this gives $\mathbf{d}^* \succeq \mathbf{c}$.

Case four: $d_i = c_i = 1$ and $d_i^* = 0$. Note that $d_i^* \neq d_i$, by using (1) of Theorem 3, we obtain that $G^{(\mathbf{d}^*)}(i) + b \leq 0$. On the other hand, since $c_i \neq d_i^*$, it follows from the performance difference Eq. (39) that for each $i \in S(\mathbf{d}^*, \mathbf{c})$,

$$\eta^{\mathbf{c}} - \eta^{\mathbf{d}^*} = \mu_2 \pi^{(\mathbf{c})}(i)\left(c_i - d_i^*\right)\left[G^{(\mathbf{d}^*)}(i) + b\right]$$
$$= \mu_2 \pi^{(\mathbf{c})}(i)\left[G^{(\mathbf{d}^*)}(i) + b\right] \leq 0.$$

Thus $\eta^{\mathbf{d}^*} \geq \eta^{\mathbf{c}}$, this gives $\mathbf{d}^* \succeq \mathbf{c}$.

Based on the above four discussions, we obtain that $\mathbf{d}^* \succeq \mathbf{c}$ for any policy $\mathbf{c} \in \mathcal{D}$. This completes the proof. □

For $\mathbf{d}^* = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1)$, let $\mathbf{d}^{(n)}$ be a policy in the policy set $\mathcal{D}$ with

$$S\left(\mathbf{d}^*, \mathbf{d}^{(n)}\right) = \{i_l : l = 1, 2, \ldots, n\}$$

for $1 \leq n \leq K$. To understand Policy $\mathbf{d}^{(n)}$, we take three examples: $S\left(\mathbf{d}^*, \mathbf{d}^{(1)}\right) = \{i_1\}$, $S\left(\mathbf{d}^*, \mathbf{d}^{(2)}\right) = \{i_1, i_2\}$, $S\left(\mathbf{d}^*, \mathbf{d}^{(3)}\right) = \{i_1, i_2, i_3\}$. Also, $S\left(\mathbf{d}^{(n-1)}, \mathbf{d}^{(n)}\right) = \{i_n\}$ for $1 \leq n \leq K$. Note that

$$\mathbf{d}^{(K)} = (0; 1, 1, \ldots, 1; 1, 1, \ldots, 1).$$

The following corollary provides a set-structured decreasing monotonicity of the policies $\mathbf{d}^{(n)} \in \mathcal{D}$ for $n = 1, 2, \ldots, K-1, K$. In fact, this monotonicity is guaranteed by the class property of policies in the set $\mathcal{D}$, given in (1) of Theorem 3. The proof is easy by using a similar analysis to that in Theorem 4, thus it is omitted here.

Corollary 5 *If* $P \geq P_H(\mathbf{d})$ *for any given policy* $\mathbf{d}$*, then*

$$\mathbf{d}^* \succeq \mathbf{d}^{(1)} \succeq \mathbf{d}^{(2)} \succeq \mathbf{d}^{(3)} \succeq \cdots \succeq \mathbf{d}^{(K-1)} \succeq \mathbf{d}^{(K)}.$$

In what follows we compute the maximal long-run average profit of the stock-rationing queue.

When $P \geq P_H(\mathbf{d})$ for any given policy $\mathbf{d}$, the optimal dynamic rationing policy is given by

$$\mathbf{d}^* = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1),$$

thus it follows from (5) that

$$\xi_0 = 1, \ i = 0,$$
$$\xi_i^{(\mathbf{d}^*)} = \begin{cases} \alpha^i, & i = 1, 2, \ldots, K, \\ \left(\dfrac{\alpha}{\beta}\right)^K \beta^i, & i = K+1, K+2, \ldots, N, \end{cases}$$

and

$$h^{(\mathbf{d}^*)} = 1 + \sum_{i=1}^{N} \xi_i^{(\mathbf{d}^*)} = 1 + \frac{\alpha\left(1 - \alpha^K\right)}{1 - \alpha} + \left(\frac{\alpha}{\beta}\right)^K \frac{\beta^{K+1}\left(1 - \beta^{N-K}\right)}{1 - \beta},$$

where $\alpha = \lambda/\mu_1$ and $\beta = \lambda/(\mu_1 + \mu_2)$. It follows from (6) that

$$\pi^{(\mathbf{d}^*)}(i) = \begin{cases} \dfrac{1}{h^{(\mathbf{d}^*)}}, & i = 0, \\ \dfrac{1}{h^{(\mathbf{d}^*)}} \xi_i^{(\mathbf{d}^*)}, & i = 1, 2, \ldots, N. \end{cases}$$

At the same time, it follows from (8) to (11) that

$$\begin{aligned} f(0) &= -C_{2,1}\mu_1 - C_{2,2}\mu_2 - C_3\lambda, & i = 0; \\ f^{(\mathbf{d}^*)}(i) &= R\mu_1 - C_1 i - C_{2,2}\mu_2 - C_3\lambda, & 1 \leq i \leq K; \\ f(i) &= R(\mu_1 + \mu_2) - C_1 i - C_3\lambda 1_{\{i<N\}} - C_4\lambda 1_{\{i=N\}}, & K+1 \leq i \leq N. \end{aligned}$$

Since

$$\eta^{\mathbf{d}^*} = \sum_{i=0}^{N} \pi^{(\mathbf{d}^*)}(i) f^{(\mathbf{d}^*)}(i), \tag{51}$$

we obtain

$$\begin{aligned} \eta^{\mathbf{d}^*} &= \frac{1}{h^{(\mathbf{d}^*)}} \left\{ -\left(C_{2,1}\mu_1 + C_{2,2}\mu_2 + C_3\lambda\right) + \sum_{i=0}^{K} \left(R\mu_1 - C_1 i - C_{2,2}\mu_2 - C_3\lambda\right)\alpha^i \right. \\ &\quad \left. + \sum_{i=K+1}^{N} \left[R(\mu_1 + \mu_2) - C_1 i - C_3\lambda 1_{\{i<N\}} - C_4\lambda 1_{\{i=N\}}\right]\left(\frac{\alpha}{\beta}\right)^K \beta^i \right\} \\ &= \frac{1}{h^{(\mathbf{d}^*)}} \left\{ -\gamma_1 + \gamma_2 \frac{\alpha\left(1 - \alpha^K\right)}{1 - \alpha} - C_1\left[\frac{\alpha\left(1 - \alpha^K\right)}{(1 - \alpha)^2} - \frac{K\alpha^{K+1}}{1 - \alpha}\right] \right. \end{aligned}$$

$$+ \left(\frac{\alpha}{\beta}\right)^K \gamma_3 \frac{\beta^{K+1}\left(1 - \beta^{N-K}\right)}{1 - \beta}$$
$$- \left(\frac{\alpha}{\beta}\right)^K C_1 \left[\frac{K\beta^{K+1} - N\beta^{N+1}}{1 - \beta} + \frac{\beta^{K+1}\left(1 - \beta^{N-K}\right)}{(1 - \beta)^2}\right]\Bigg\},$$

where

$$\gamma_1 = C_{2,1}\mu_1 + C_{2,2}\mu_2 + C_3\lambda,$$
$$\gamma_2 = R\mu_1 - C_{2,2}\mu_2 - C_3\lambda,$$
$$\gamma_3 = R\left(\mu_1 + \mu_2\right) - C_3\lambda 1_{\{i<N\}} - C_4\lambda 1_{\{i=N\}}.$$

7.2 The penalty cost $P_L\left(\mathbf{d}\right) > 0$ and $0 \leq P \leq P_L\left(\mathbf{d}\right)$

In this subsection, we consider the area of the penalty cost: $P_L\left(\mathbf{d}\right) > 0$ and $0 \leq P \leq P_L\left(\mathbf{d}\right)$ for any given policy $\mathbf{d}$. We first find the optimal dynamic rationing policy of the stock-rationing queue. Then we compute the maximal long-run average profit of this system.

The following theorem finds the optimal dynamic rationing policy of the stock-rationing queue in the area of the penalty cost: $P_L\left(\mathbf{d}\right) > 0$ and $0 \leq P \leq P_L\left(\mathbf{d}\right)$ for any given policy $\mathbf{d}$. The proof is similar to that of Theorem 4.

Theorem 6 *If $P_L\left(\mathbf{d}\right) > 0$ and $0 \leq P \leq P_L\left(\mathbf{d}\right)$ for any given policy $\mathbf{d}$, then the optimal dynamic rationing policy of the stock-rationing queue is given by*

$$\mathbf{d}^* = (0; 1, 1, \ldots, 1; 1, 1, \ldots, 1).$$

This shows that if the penalty cost is lower with $P_L\left(\mathbf{d}\right) > 0$ and $0 \leq P \leq P_L\left(\mathbf{d}\right)$, then the warehouse would like to supply the products to the demands of Class 2.

Proof If $P_L\left(\mathbf{d}\right) > 0$ and $0 \leq P \leq P_L\left(\mathbf{d}\right)$ for any given policy $\mathbf{d}$, then we prove that for any policy $\mathbf{c} \in \mathcal{D}$,

$$\mathbf{d}^* \succeq \mathbf{c}.$$

For this, we need to consider the three policies: $\mathbf{d}, \mathbf{c}$ and $\mathbf{d}^*$, where $\mathbf{d}^*$ is deterministic with $d_i^* = 1$ for each $i = 1, 2, \ldots, K - 1, K$.

To compare $\eta^{\mathbf{c}}$ with $\eta^{\mathbf{d}^*}$, let $S\left(\mathbf{d}^*, \mathbf{c}\right) = \{n_l : l = 1, 2, \ldots, m\}$ for $1 \leq m \leq K$. Then $c_{n_l} = 0$ for $l = 1, 2, \ldots, m$, since $d_i^* = 1$ for each $i = 1, 2, \ldots, K - 1, K$.

For the two policies $\mathbf{d}$ and $\mathbf{c}$, we have $d_i, c_i \in \{0, 1\}$. Based on this, for the three elements: d_i, c_i and $d_i^* = 1$ for $i \in S\left(\mathbf{d}^*, \mathbf{c}\right)$, we need to consider four different cases as follows:

Case one: $d_i = c_i = d_i^* = 1$. Since $c_i = d_i^*$, this case does not require any analysis according to Lemma 2.

Case two: $d_i = 0$ and $c_i = d_i^* = 1$. Since $c_i = d_i^*$, this case does not require any analysis by using Lemma 2.

Case three: $c_i = 0$ and $d_i = d_i^* = 1$. Note that $c_i \neq d_i$, by using (2) of Theorem 3, we obtain that $G^{(\mathbf{c})}(i) + b \geq 0$. On the other hand, since $c_i \neq d_i^*$, it follows from the performance difference Eq. (39) that for each $i \in S(\mathbf{d}^*, \mathbf{c})$,

$$
\begin{aligned}
\eta^{\mathbf{d}^*} - \eta^{\mathbf{c}} &= \mu_2 \pi^{(\mathbf{d}^*)}(i)\left(d_i^* - c_i\right)\left[G^{(\mathbf{c})}(i) + b\right] \\
&= \mu_2 \pi^{(\mathbf{d}^*)}(i)\left[G^{(\mathbf{c})}(i) + b\right] \geq 0.
\end{aligned}
$$

Thus $\eta^{\mathbf{d}^*} \geq \eta^{\mathbf{c}}$, this gives $\mathbf{d}^* \succeq \mathbf{c}$.

Case four: $d_i = c_i = 0$ and $d_i^* = 1$. Note that $d_i^* \neq d_i$, by using (2) of Theorem 3, we obtain that $G^{(\mathbf{d}^*)}(i) + b \geq 0$. On the other hand, since $c_i \neq d_i^*$, it follows from the performance difference Eq. (39) that for each $i \in S(\mathbf{d}^*, \mathbf{c})$,

$$
\begin{aligned}
\eta^{\mathbf{c}} - \eta^{\mathbf{d}^*} &= \mu_2 \pi^{(\mathbf{c})}(i)\left(c_i - d_i^*\right)\left[G^{(\mathbf{d}^*)}(i) + b\right] \\
&= -\mu_2 \pi^{(\mathbf{c})}(i)\left[G^{(\mathbf{d}^*)}(i) + b\right] \leq 0.
\end{aligned}
$$

Thus $\eta^{\mathbf{d}^*} \geq \eta^{\mathbf{c}}$, this gives $\mathbf{d}^* \succeq \mathbf{c}$. This completes the proof. $\square$

For $\mathbf{d}^* = (0; 1, 1, \ldots, 1; 1, 1, \ldots, 1)$, let $\mathbf{d}^{(n)}$ be a policy in the policy set $\mathcal{D}$ with $S\left(\mathbf{d}^*, \mathbf{d}^{(n)}\right) = \{k_l : l = 1, 2, \ldots, n\}$ for $1 \leq n \leq K$, where

$$
\widetilde{\mathbf{d}}^{(K)} = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1).
$$

The following corollary provides a set-structured decreasing monotonicity of the policies $\mathbf{d}^{(n)} \in \mathcal{D}$ for $n = 1, 2, \ldots, K - 1, K$. This monotonicity comes from the class property of the policies in the set $\mathcal{D}$, given in (2) of Theorem 3. The proof is easy and omitted here.

Corollary 7 *If $P_L(\mathbf{d}) > 0$ and $0 \leq P \leq P_L(\mathbf{d})$ for any given policy $\mathbf{d}$, then*

$$
\mathbf{d}^* \succeq \mathbf{d}^{(1)} \succeq \mathbf{d}^{(2)} \succeq \mathbf{d}^{(3)} \succeq \cdots \succeq \mathbf{d}^{(K-1)} \succeq \mathbf{d}^{(K)}.
$$

If $P_L(\mathbf{d}) > 0$ and $0 \leq P \leq P_L(\mathbf{d})$ for any given policy $\mathbf{d}$, then the optimal dynamic rationing policy is given by

$$
\mathbf{d}^* = (0; 1, 1, \ldots, 1; 1, 1, \ldots, 1).
$$

In this case, we obtain

$$
\begin{aligned}
\xi_0 &= 1, \quad i = 0, \\
\xi_i^{(\mathbf{d}^*)} &= \beta^i, \ i = 1, 2, \ldots, N,
\end{aligned}
$$

and

$$h^{(\mathbf{d}^*)} = 1 + \sum_{i=1}^{N} \xi_i^{(\mathbf{d}^*)} = 1 + \frac{\beta \left(1 - \beta^N\right)}{1 - \beta}.$$

It follows from Subsection 1.1.4 of Chapter 1 in Li (2010) that

$$\pi^{(\mathbf{d}^*)}(i) = \begin{cases} \frac{1}{h^{(\mathbf{d}^*)}} & i = 0, \\ \frac{\beta^i}{h^{(\mathbf{d}^*)}}, & i = 1, 2, \ldots, N, \end{cases}$$

At the same time, it follows from (8) to (11) that

$$\begin{aligned}
f(0) &= -C_{2,1}\mu_1 - C_{2,2}\mu_2 - C_3\lambda, & i = 0; \\
f^{(\mathbf{d}^*)}(i) &= R(\mu_1 + \mu_2) - C_1 i - C_3\lambda - P\mu_2, & 1 \le i \le K; \\
f^{(\mathbf{d}^*)}(i) &= R(\mu_1 + \mu_2) - C_1 i - C_3\lambda 1_{\{i<N\}} - C_4\lambda 1_{\{i=N\}}, & K+1 \le i \le N.
\end{aligned}$$

Thus we obtain

$$\begin{aligned}
\eta^{\mathbf{d}^*} &= \frac{1}{h^{(\mathbf{d}^*)}} \left\{ -\left(C_{2,1}\mu_1 + C_{2,2}\mu_2 + C_3\lambda\right) + \sum_{i=1}^{K} \left[R(\mu_1 + \mu_2) - C_1 i - C_3\lambda - P\mu_2\right]\beta^i \right. \\
&\quad \left. + \sum_{i=K+1}^{N} \left[R(\mu_1 + \mu_2) - C_1 i - C_3\lambda 1_{\{i<N\}} - C_4\lambda 1_{\{i=N\}}\right]\beta^i \right\} \\
&= \frac{1}{h^{(\mathbf{d}^*)}} \left\{ -\gamma_1 + \gamma_4 \frac{\beta\left(1 - \beta^K\right)}{1 - \beta} - C_1 \left[\frac{\beta\left(1 - \beta^K\right)}{(1-\beta)^2} - \frac{K\beta^{K+1}}{1-\beta}\right] \right. \\
&\quad \left. + \gamma_3 \frac{\beta^{K+1}\left(1 - \beta^{N-K}\right)}{1 - \beta} - C_1 \left[\frac{K\beta^{K+1} - N\beta^{N+1}}{1-\beta} + \frac{\beta^{K+1}\left(1 - \beta^{N-K}\right)}{(1-\beta)^2}\right] \right\},
\end{aligned}$$

where

$$\begin{aligned}
\gamma_1 &= C_{2,1}\mu_1 + C_{2,2}\mu_2 + C_3\lambda, \\
\gamma_2 &= R\mu_1 - C_{2,2}\mu_2 - C_3\lambda, \\
\gamma_3 &= R(\mu_1 + \mu_2) - C_3\lambda 1_{\{i<N\}} - C_4\lambda 1_{\{i=N\}}, \\
\gamma_4 &= R(\mu_1 + \mu_2) - C_3\lambda - P\mu_2.
\end{aligned}$$

7.3 The penalty cost $P_L(\mathbf{d}) < P < P_H(\mathbf{d})$

In this subsection, we discuss the third area of the penalty cost: $P_L < P < P_H$ for any given policy $\mathbf{d}$. Note that this analysis is a little more complicated than those in the previous two areas. To this end, we propose a new algebraic method to find the optimal dynamic rationing policy of the stock-rationing queue. Based on this, we show that

the optimal dynamic rationing policy may not be of threshold type, but it must be of transformational threshold type.

For convenience of the readers, it is necessary and useful to simply recall several previous results as follows.

For any given policy $\mathbf{d} = (0; d_1, d_2, \ldots, \ldots, d_{K-1}, d_K; 1, 1, \ldots, 1)$, the unique solution of the linear equation $G^{(\mathbf{d})}(i) + b = 0$ in the penalty cost P is given by

$$
\mathfrak{P}_i^{(\mathbf{d})} = \frac{R + C_{2,2} + \lambda^{-i}\left[B_0 - D^{(\mathbf{d})}\right]\prod_{k=1}^{i-1} v\,(d_k) + \sum_{r=1}^{i-1} \lambda^{r-i}\left[B_r^{(\mathbf{d})} - D^{(\mathbf{d})}\right]\prod_{k=r+1}^{i-1} v\,(d_k)}{1 + \lambda^{-i}\left[A_0 - F^{(\mathbf{d})}\right]\prod_{k=1}^{i-1} v\,(d_k) + \sum_{r=1}^{i-1} \lambda^{r-i}\left[A_r^{(\mathbf{d})} - F^{(\mathbf{d})}\right]\prod_{k=r+1}^{i-1} v\,(d_k)},
$$

which is a fixed real number for $1 \le i \le K$.

We introduce a convention: If $\mathfrak{P}_{n-1}^{(\mathbf{d})} < \mathfrak{P}_n^{(\mathbf{d})} = \mathfrak{P}_{n+1}^{(\mathbf{d})} = \cdots = \mathfrak{P}_{n+i}^{(\mathbf{d})} = c$ and $\mathfrak{P}_{n-1}^{(\mathbf{d})} < P \le c$, then we write

$$
\mathfrak{P}_{n-1}^{(\mathbf{d})} < P \le \mathfrak{P}_n^{(\mathbf{d})} = \mathfrak{P}_{n+1}^{(\mathbf{d})} = \cdots = \mathfrak{P}_{n+i}^{(\mathbf{d})},
$$

that is, the penalty cost P is written in front of all the equal elements in the sequence $\left\{\mathfrak{P}_k^{(\mathbf{d})} : n \le k \le n+i\right\}$.

For the sequence $\left\{\mathfrak{P}_k^{(\mathbf{d})} : 1 \le k \le K\right\}$, we set up a new permutation from the smallest to the largest as follows:

$$
\mathfrak{P}_{i_1}^{(\mathbf{d})} \le \mathfrak{P}_{i_2}^{(\mathbf{d})} \le \cdots \le \mathfrak{P}_{i_{K-1}}^{(\mathbf{d})} \le \mathfrak{P}_{i_K}^{(\mathbf{d})},
$$

it is clear that $\mathfrak{P}_{i_1}^{(\mathbf{d})} = P_L(\mathbf{d})$ and $\mathfrak{P}_{i_K}^{(\mathbf{d})} = P_H(\mathbf{d})$. For convenience of the description, for the incremental sequence $\left\{\mathfrak{P}_{i_j}^{(\mathbf{d})} : 1 \le j \le K\right\}$, we write its subscript vector as $(i_1, i_2, \ldots, i_{K-1}, i_K)$. Note that the subscript vector $(i_1, i_2, \ldots, i_{K-1}, i_K)$ depends on Policy $\mathbf{d}$.

The following lemma shows how the penalty cost P is distributed in the sequence $\left\{\mathfrak{P}_k^{(\mathbf{d})} : 1 \le k \le K\right\}$.

Lemma 4 *If $P_L(\mathbf{d}) < P < P_H(\mathbf{d})$ for any given policy $\mathbf{d}$, then there exists the minimal positive integer $n_0 \in \{1, 2, \ldots, K\}$ such that either*

$$
\mathfrak{P}_{i_{n_0}}^{(\mathbf{d})} < P = \mathfrak{P}_{i_{n_0+1}}^{(\mathbf{d})}
$$

or

$$
\mathfrak{P}_{i_{n_0}}^{(\mathbf{d})} < P < \mathfrak{P}_{i_{n_0+1}}^{(\mathbf{d})}.
$$

Proof Note that

$$P_H\,(\mathbf{d}) = \max_{\mathbf{d}\in\mathcal{D}}\left\{0, \mathfrak{P}_1^{(\mathbf{d})}, \mathfrak{P}_2^{(\mathbf{d})}, \ldots, \mathfrak{P}_K^{(\mathbf{d})}\right\}$$

and

$$P_L\,(\mathbf{d}) = \max_{\mathbf{d}\in\mathcal{D}}\left\{\mathfrak{P}_1^{(\mathbf{d})}, \mathfrak{P}_2^{(\mathbf{d})}, \ldots, \mathfrak{P}_K^{(\mathbf{d})}\right\},$$

it is easy to see that $P_H\,(\mathbf{d})$ and $P_L\,(\mathbf{d})$ are two fixed real numbers. If $P_L < P < P_H\,(\mathbf{d})$ for Policy $\mathbf{d}$, then there exists the minimal positive integer $n_0 \in \{1, 2, \ldots, K-1, K\}$ such that

$$P_L\,(\mathbf{d}) \le \mathfrak{P}_{i_{n_0}}^{(\mathbf{d})} < P \le \mathfrak{P}_{i_{n_0+1}}^{(\mathbf{d})} < P_H\,(\mathbf{d}).$$

This shows that either for $P = \mathfrak{P}_{i_{n_0+1}}^{(\mathbf{d})}$,

$$\mathfrak{P}_{i_{n_0}}^{(\mathbf{d})} < P = \mathfrak{P}_{i_{n_0+1}}^{(\mathbf{d})};$$

or for $P < \mathfrak{P}_{i_{n_0+1}}^{(\mathbf{d})}$,

$$\mathfrak{P}_{i_{n_0}}^{(\mathbf{d})} < P < \mathfrak{P}_{i_{n_0+1}}^{(\mathbf{d})}.$$

This completes the proof.　　　　　　　　　　　　　　　　　　　　□

Now, our task is to develop a new method for finding the optimal dynamic rationing policy by means of the following two useful information: (a) The incremental sequence

$$P_L\,(\mathbf{d}) = \mathfrak{P}_{i_1}^{(\mathbf{d})} \le \mathfrak{P}_{i_2}^{(\mathbf{d})} \le \cdots \le \mathfrak{P}_{i_{K-1}}^{(\mathbf{d})} \le \mathfrak{P}_{i_K}^{(\mathbf{d})} = P_H\,(\mathbf{d});$$

and (b) the penalty cost P has a fixed position: $\mathfrak{P}_{i_{n_0}}^{(\mathbf{d})} < P \le \mathfrak{P}_{i_{n_0+1}}^{(\mathbf{d})}$, where n_0 is the minimal positive integer in the set $\{1, 2, \ldots, K-1, K\}$.

In what follows we discuss two different cases: A simple case and a general case. *Case one: A simple case with*

$$P_L\,(\mathbf{d}) = \mathfrak{P}_1^{(\mathbf{d})} \le \mathfrak{P}_2^{(\mathbf{d})} \le \cdots \le \mathfrak{P}_{K-1}^{(\mathbf{d})} \le \mathfrak{P}_K^{(\mathbf{d})} = P_H\,(\mathbf{d}). \tag{52}$$

In this case, the subscript vector is expressed as $\{1, 2, 3, \ldots, K-1, K\}$ depending on Policy $\mathbf{d}$.

If $P_L\,(\mathbf{d}) < P < P_H\,(\mathbf{d})$ for any given policy $\mathbf{d}$, then there exists the minimal positive integer $n_0 \in \{1, 2, \ldots, K-1, K\}$ such that

$$P_L\,(\mathbf{d}) = \mathfrak{P}_1^{(\mathbf{d})} \le \cdots \le \mathfrak{P}_{n_0-1}^{(\mathbf{d})} < P \le \mathfrak{P}_{n_0}^{(\mathbf{d})} \le \cdots \le \mathfrak{P}_K^{(\mathbf{d})} = P_H\,(\mathbf{d}).$$

Based on this, we take two different sets

$$\Lambda_1 = \left\{ \mathfrak{P}_1^{(\mathbf{d})}, \mathfrak{P}_2^{(\mathbf{d})}, \ldots, \mathfrak{P}_{n_0-1}^{(\mathbf{d})} \right\}$$

and

$$\Lambda_2 = \left\{ \mathfrak{P}_{n_0}^{(\mathbf{d})}, \mathfrak{P}_{n_0+1}^{(\mathbf{d})}, \ldots, \mathfrak{P}_K^{(\mathbf{d})} \right\}.$$

By using the two sets Λ_1 and Λ_2, we write

$$\overline{P}_H (\mathbf{d}; 1 \to n_0 - 1) = \max_{1 \le i \le n_0-1} \left\{ \mathfrak{P}_i^{(\mathbf{d})} \right\}$$

and

$$\overline{P}_L (\mathbf{d}; n_0 \to K) = \min_{n_0 \le j \le K} \left\{ \mathfrak{P}_j^{(\mathbf{d})} \right\}.$$

It is clear that $\overline{P}_H (\mathbf{d}; 1 \to n_0 - 1) = \mathfrak{P}_{n_0-1}^{(\mathbf{d})}$ and $\overline{P}_L (\mathbf{d}; n_0 \to K) = \mathfrak{P}_{n_0}^{(\mathbf{d})}$.

For this simple case, the following theorem finds the optimal dynamic rationing policy, which is of threshold type.

Theorem 8 *For the simple case with $P_L (\mathbf{d}) < P < P_H (\mathbf{d})$ for any given policy $\mathbf{d}$, if there exists the minimal positive integer $n_0 \in \{1, 2, \ldots, K - 1, K\}$ such that*

$$P_L (\mathbf{d}) = \mathfrak{P}_1^{(\mathbf{d})} \le \cdots \le \mathfrak{P}_{n_0-1}^{(\mathbf{d})} < P \le \mathfrak{P}_{n_0}^{(\mathbf{d})} \le \cdots \le \mathfrak{P}_K^{(\mathbf{d})} = P_H (\mathbf{d}), \qquad (53)$$

then the optimal dynamic rationing policy is given by

$$\mathbf{d}^* = \left(0; \underbrace{0, 0, \ldots, 0,}_{n_0-1 \; zeros} \underbrace{1, 1, \ldots, 1}_{K-n_0+1 \; ones}; 1, 1, \ldots, 1 \right).$$

Proof The proof follows that in Theorems 4 and 6.

On the one hand, in the set Λ_1, it is easy to see from (53) that $P > \overline{P}_H (\mathbf{d}; 1 \to n_0 - 1)$ for Policy $\mathbf{d}$. Now, our aim is to focus on a sub-policy

$$\widetilde{\mathbf{d}}_a = \left(0; d_1, d_2, \ldots, d_{n_0-1}, *, *, \ldots, *; 1, 1, \ldots, 1 \right).$$

For the sub-policy $\left(d_1, d_2, \ldots, d_{n_0-1} \right)$, it is easy to see from the set Λ_1 that $P > \overline{P}_H (\mathbf{d}; 1 \to n_0 - 1)$. Thus it follows from Theorem 4 that the optimal dynamic rationing sub-policy is given by

$$\widetilde{\mathbf{d}}_a^* = (0; 0, 0, \ldots, 0, *, *, \ldots, *; 1, 1, \ldots, 1).$$

On the other hand, it is seen from the set Λ_2 that $0 \leq P \leq \overline{P}_L\left(\mathbf{d}; n_0 \to K\right)$ for Policy $\mathbf{d}$. We consider another sub-policy

$$\widetilde{\mathbf{d}}_b = \left(0; *, *, \ldots, *, d_{n_0}, d_{n_0+1}, \ldots, d_K; 1, 1, \ldots, 1\right).$$

For the sub-policy $\left(d_{n_0}, d_{n_0+1}, \ldots, d_K\right)$, it is easy to see from the set Λ_2 that $0 \leq P \leq \overline{P}_L\left(\mathbf{d}; n_0 \to K\right)$. Thus it is easy to see from Theorem 6 that the optimal dynamic rationing sub-policy is given by

$$\widetilde{\mathbf{d}}_b^* = (0; *, *, \ldots, *, 1, 1, \ldots, 1; 1, 1, \ldots, 1).$$

Based on the above two discussions, from the total set $\Lambda_1 \cup \Lambda_2$, by observing the total policy $\left(d_1, d_2, \ldots, d_{n_0-1}; d_{n_0}, d_{n_0+1}, \ldots, d_K\right)$ or Policy $\mathbf{d}$, the optimal dynamic rationing policy is given by

$$\mathbf{d}^* = \widetilde{\left(\mathbf{d}_a^*\right)}_b = \widetilde{\left(\mathbf{d}_b^*\right)}_a = \left(0; \underbrace{0, 0, \ldots, 0}_{n_0-1 \text{ zeros}}, \underbrace{1, 1, \ldots, 1}_{K-n_0+1 \text{ ones}}; 1, 1, \ldots, 1\right).$$

This completes the proof.   □

Remark 4 It is easy to see that in Theorems 4, 6 and 8, the optimal dynamic rationing policy is of threshold type (i.e., critical rationing level).

Case two: A general case with

$$P_L\left(\mathbf{d}\right) = \mathfrak{P}_{i_1}^{(\mathbf{d})} \leq \mathfrak{P}_{i_2}^{(\mathbf{d})} \leq \cdots \leq \mathfrak{P}_{i_{K-1}}^{(\mathbf{d})} \leq \mathfrak{P}_{i_K}^{(\mathbf{d})} = P_H\left(\mathbf{d}\right).$$

For the incremental sequence $\left\{\mathfrak{P}_{i_j}^{(\mathbf{d})} : j = 1, 2, \ldots, K\right\}$, we write its subscript vector as $(i_1, i_2, \ldots, i_{K-1}, i_K)$, which depends on Policy $\mathbf{d}$. In the general case, we assume that $(i_1, i_2, \ldots, i_{K-1}, i_K) \neq (1, 2, \ldots, K-1, K)$.

If $P_L\left(\mathbf{d}\right) < P < P_H\left(\mathbf{d}\right)$ for any given policy $\mathbf{d}$, then there exists the minimal positive integer $n_0 \in \{1, 2, \ldots, K-1, K\}$ such that

$$P_L\left(\mathbf{d}\right) = \mathfrak{P}_{i_1}^{(\mathbf{d})} \leq \cdots \leq \mathfrak{P}_{i_{n_0-1}}^{(\mathbf{d})} < P \leq \mathfrak{P}_{i_{n_0}}^{(\mathbf{d})} \leq \cdots \leq \mathfrak{P}_{i_K}^{(\mathbf{d})} = P_H\left(\mathbf{d}\right).$$

Based on this, we take two sets

$$\Lambda_1^G = \left\{\mathfrak{P}_{i_1}^{(\mathbf{d})}, \mathfrak{P}_{i_2}^{(\mathbf{d})}, \ldots, \mathfrak{P}_{i_{n_0-1}}^{(\mathbf{d})}\right\}$$

and

$$\Lambda_2^G = \left\{\mathfrak{P}_{i_{n_0}}^{(\mathbf{d})}, \mathfrak{P}_{i_{n_0+1}}^{(\mathbf{d})}, \ldots, \mathfrak{P}_{i_K}^{(\mathbf{d})}\right\}.$$

For the two sets Λ_1^G and Λ_2^G, we write

$$\overline{P}_H^G (\mathbf{d}; 1 \to n_0 - 1) = \max_{1 \le k \le n_0 - 1} \left\{ \mathfrak{P}_{i_k}^{(\mathbf{d})} \right\}$$

and

$$\overline{P}_L^G (\mathbf{d}; n_0 \to K) = \min_{n_0 \le k \le K} \left\{ \mathfrak{P}_{i_k}^{(\mathbf{d})} \right\},$$

It is clear that $\overline{P}_H^G (\mathbf{d}; 1 \to n_0 - 1) = \mathfrak{P}_{i_{n_0-1}}^{(\mathbf{d})}$ and $\overline{P}_L^G (\mathbf{d}; n_0 \to K) = \mathfrak{P}_{i_{n_0}}^{(\mathbf{d})}$.

Corresponding to the subscript vector of the incremental sequence $\left\{ \mathfrak{P}_{i_k}^{(\mathbf{d})} : 1 \le k \le K \right\}$, we transfer Policy

$$\mathbf{d} = \left(0; d_1, d_2, \ldots, d_{n_0-1}, d_{n_0}, d_{n_0+1}, \ldots, d_K; 1, 1, \ldots, 1 \right)$$

into a new transformational policy

$$\mathbf{d} \, (\text{Transfer}) = \left(0; d_{i_1}, d_{i_2}, \ldots, d_{i_{n_0-1}}, d_{i_{n_0}}, d_{i_{n_0+1}}, \ldots, d_{i_K}; 1, 1, \ldots, 1 \right).$$

Therefore, a transformation of the optimal dynamic policy $\mathbf{d}^*$ is

$$(1, 2, \ldots, K - 1, K) \Rightarrow (i_1, i_2, \ldots, i_{K-1}, i_K) ;$$

and an inverse transformation of the optimal transformational dynamic policy $\mathbf{d}^*$ (Transfer) is

$$(i_1, i_2, \ldots, i_{K-1}, i_K) \Rightarrow (1, 2, \ldots, K - 1, K) .$$

For the general case, the following theorem finds the optimal dynamic rationing policy, which may not be of threshold type, but must be of transformational threshold type.

Theorem 9 *For the general case with $P_L (\mathbf{d}) < P < P_H (\mathbf{d})$ for any given policy $\mathbf{d}$, if there exists the minimal positive integer $n_0 \in \{1, 2, \ldots, K - 1, K\}$ such that*

$$P_L (\mathbf{d}) = \mathfrak{P}_{i_1}^{(\mathbf{d})} \le \cdots \le \mathfrak{P}_{i_{n_0-1}}^{(\mathbf{d})} < P \le \mathfrak{P}_{i_{n_0}}^{(\mathbf{d})} \le \cdots \le \mathfrak{P}_{i_K}^{(\mathbf{d})} = P_H (\mathbf{d}) ,$$

then the optimal transformational dynamic rationing policy is given by

$$\mathbf{d}^* \, (\text{Transfer}) = \left(0; \underbrace{0, 0, \ldots, 0}_{n_0 - 1 \, zeros}, \underbrace{1, 1, \ldots, 1}_{K - n_0 + 1 \, ones}; 1, 1, \ldots, 1 \right).$$

Proof From the set Λ_1^G, it is easy to see that $P > \overline{P}_H^G (\mathbf{d}; 1 \to n_0 - 1)$. Hence we consider the transformational sub-policy

$$\widetilde{\mathbf{d}}_a \text{ (Transfer)} = \left(0; d_{i_1}, d_{i_2}, \ldots, d_{i_{n_0-1}}, *, *, \ldots, *; 1, 1, \ldots, 1 \right).$$

By observing the transformational sub-policy $\left(d_{i_1}, d_{i_2}, \ldots, d_{i_{n_0-1}} \right)$ related to $P > \overline{P}_H^G(\mathbf{d}; 1 \to n_0 - 1)$, it is easy to see from the proof of Theorem 4 that the optimal transformational dynamic rationing sub-policy is given by

$$\widetilde{\mathbf{d}}_a^* \text{ (Transfer)} = (0; 0, 0, \ldots, 0, *, *, \ldots, *; 1, 1, \ldots, 1).$$

Similarly, from $0 \leq P \leq \overline{P}_L^G (\mathbf{d}; n_0 \to K)$ in the set Λ_2^G, we discuss the transformational sub-policy

$$\widetilde{\mathbf{d}}_b \text{ (Transfer)} = \left(0; *, *, \ldots, *, d_{i_{n_0}}, d_{i_{n_0+1}}, \ldots, d_{i_K}; 1, 1, \ldots, 1 \right).$$

By observing the transformational sub-policy $\left(d_{n_0}, d_{n_0+1}, \ldots, d_K \right)$ related to $0 \leq P \leq \overline{P}_L^G (\mathbf{d}; n_0 \to K)$, it is easy to see from the proof of Theorem 6 that the optimal transformational dynamic rationing sub-policy is given by

$$\widetilde{\mathbf{d}}_b^* \text{ (Transfer)} = (0; *, *, \ldots, *, 1, 1, \ldots, 1; 1, 1, \ldots, 1).$$

Therefore, by observing the total transformational sub-policy $(d_{i_1}, d_{i_2}, \ldots, d_{i_{n_0-1}}, d_{i_{n_0}}, d_{i_{n_0+1}}, \ldots, d_{i_K})$ in the total set $\Lambda_1^G \cup \Lambda_2^G$, the optimal transformational dynamic rationing policy is given by

$$\mathbf{d}^* \text{ (Transfer)} = \left(\widetilde{\mathbf{d}}_a^* \text{ (Transfer)} \right)_b^* \text{ (Transfer)} = \left(\widetilde{\mathbf{d}}_b^* \text{ (Transfer)} \right)_a^* \text{ (Transfer)}$$

$$= \left(0; \underbrace{0, 0, \ldots, 0}_{n_0-1 \text{ zeros}}, \underbrace{1, 1, \ldots, 1}_{K-n_0+1 \text{ ones}}; 1, 1, \ldots, 1 \right).$$

This completes the proof.   □

Remark 5 (1) For the general case, although the optimal dynamic rationing policy is not of threshold type, we show that it must be of transformational threshold type. Thus the optimal transformational dynamic policy of the stock-rationing queue has a beautiful form as follows:

$$\mathbf{d}^* \text{ (Transfer)} = \left(0; \underbrace{0, 0, \ldots, 0}_{n_0-1 \text{ zeros}}, \underbrace{1, 1, \ldots, 1}_{K-n_0 \text{ ones}}; 1, 1, \ldots, 1 \right).$$

(2) We use an inverse transformation of $\mathbf{d}^*$ (Transfer) to be able to restore the original optimal dynamic policy $\mathbf{d}^*$, since $\mathbf{d}^*$ (Transfer) is always obtained easily. To indicate such an inverse process, we take a simple example:

$$\mathfrak{P}_1^{(\mathbf{d})} \leq \mathfrak{P}_3^{(\mathbf{d})} \leq \mathfrak{P}_4^{(\mathbf{d})} \leq \mathfrak{P}_7^{(\mathbf{d})} < P \leq \mathfrak{P}_2^{(\mathbf{d})} \leq \mathfrak{P}_5^{(\mathbf{d})} \leq \mathfrak{P}_6^{(\mathbf{d})} \leq \mathfrak{P}_8^{(\mathbf{d})},$$

it is easy to check that

$$\mathbf{d}^* = (0; 0, 1, 0, 0, 1, 1, 0, 1; 1, 1, 1, 1).$$

Remark 6 The transformational version $\mathbf{d}^*$ (Transfer) of the optimal dynamic rationing policy $\mathbf{d}^*$ plays a key role in the applications of the sensitivity-based optimization to the study of stock-rationing queues. On the other hand, it is worthwhile to note that the RG-factorization of block-structured Markov processes can be extended and generalized to a more general optimal transformational version $\mathbf{d}^*$ (Transfer) in the study of stock-rationing block-structured queues. See Li (2010) and Ma et al. (2019) for more details.

Remark 7 The bang-bang control is an effective method to roughly describe the optimal dynamic policy, e.g., see Xia et al. (2021) and Ma et al. (2019, 2021). However, our optimal transformational dynamic policy $\mathbf{d}^*$ (Transfer) provides a more detailed result, and also can restore the original optimal dynamic policy $\mathbf{d}^*$ by means of an inverse transformation: $(i_1, i_2, \ldots, i_{K-1}, i_K) \Rightarrow (1, 2, \ldots, K-1, K)$. Therefore, our optimal transformational dynamic rationing policy is superior to the bang-bang control.

The following theorem provides a useful summarization for Theorems 4 to 9. Based on this, for the optimal dynamic policy of the stock-rationing queue, we provide a complete algebraic solution: (a) The optimal dynamic rationing policy can hold under each of the three different conditions; and (b) the optimal transformational dynamic rationing policy (no optimal dynamic rationing policy) can hold under one condition

Theorem 10 *For the stock-rationing queue with two demand classes, there must exist an optimal transformational dynamic rationing policy*

$$\mathbf{d}^* \text{ (Transfer)} = \left(0; \underbrace{0, 0, \ldots, 0}_{n_0 - 1 \text{ zeros}}, \underbrace{1, 1, \ldots, 1}_{K - n_0 \text{ ones}}; 1, 1, \ldots, 1 \right).$$

Based on this finding, we can achieve the following two useful results:

(a) *The optimal dynamic rationing policy $\mathbf{d}^*$ is of critical rationing level (i.e., threshold type) under each of the three conditions: (i) $P \geq P_H(\mathbf{d})$ for any given policy $\mathbf{d}$; (i i) $P_L(\mathbf{d}) > 0$ and $0 \leq P \leq P_L(\mathbf{d})$ for any given policy $\mathbf{d}$; and (i i i) $P_L(\mathbf{d}) < P < P_H(\mathbf{d})$ with the subscript vector $(1, 2, \ldots, K-1, K)$ depending on Policy $\mathbf{d}$.*

(b) *The optimal dynamic rationing policy is not of critical rationing level (i.e., threshold type) if $P_L(\mathbf{d}) < P < P_H(\mathbf{d})$ with the subscript vector $(i_1, i_2, \ldots, i_{K-1}, i_K) \neq (1, 2, \ldots, K-1, K)$ depending on Policy $\mathbf{d}$.*

7.4 A global optimal analysis

In this subsection, for a fixed penalty cost P, we discuss how to find a global optimal policy of the stock-rationing queue with two demand classes by means of Theorem 10. Note that if $\mathbf{d}^*$ is a global optimal policy of this system, then $\mathbf{d}^* \succeq \mathbf{c}$ for any $\mathbf{c} \in \mathcal{D}$. Also, we provide a simple effective method to be able to find the global optimal policy from the policy set $\mathcal{D}$.

In the policy set $\mathcal{D}$, we define two key policies:

$$\mathbf{d}_1 = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1)$$

and

$$\mathbf{d}_2 = (0; 1, 1, \ldots, 1; 1, 1, \ldots, 1).$$

Note that there are 2^k different policies in the set $\mathcal{D}$, we write

$$\mathcal{D} = \left\{ \mathbf{d}_1, \mathbf{d}_2; \mathbf{c}_3, \mathbf{c}_4, \ldots, \mathbf{c}_{2^k-1}, \mathbf{c}_{2^k} \right\}.$$

The following theorem describes a useful characteristics of the two key policies $\mathbf{d}_1$ and $\mathbf{d}_2$ by means of the class property of the policies in the set $\mathcal{D}$, given in Theorem 3. This characteristics makes us to be able to find the global optimal policy of the stock-rationing queue.

Theorem 11 (1) *If a fixed penalty cost $P \geq P_H(\mathbf{d})$ for any given policy $\mathbf{d}$, then $P \geq P_H(\mathbf{d}_1)$.*
(2) *If a fixed penalty cost $P_L(\mathbf{d}) > 0$ and $0 \leq P \leq P_L(\mathbf{d})$ for any given policy $\mathbf{d}$, then $P_L(\mathbf{d}_2) > 0$ and $0 \leq P \leq P_L(\mathbf{d}_2)$.*

Proof We only prove (1), while (2) can be proved similarly.

We assume the penalty cost: $P < P_H(\mathbf{d}_1)$ for Policy $\mathbf{d}_1 = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1)$. Then there exists the minimal positive integer $n_0 \in \{1, 2, \ldots, K-1, K\}$ such that

$$0 < P \leq \mathfrak{P}_{i_{n_0}}^{(\mathbf{d}_1)} \leq \cdots \leq \mathfrak{P}_{i_K}^{(\mathbf{d}_1)} = P_H(\mathbf{d}_1),$$

and also there exists at least a positive integer $m_0 \in \{n_0+1, n_0+2, \ldots, K-1, K\}$ such that

$$\mathfrak{P}_{i_{m_0-1}}^{(\mathbf{d}_1)} < \mathfrak{P}_{i_{m_0}}^{(\mathbf{d}_1)}. \tag{54}$$

Let

$$\overline{P}_L^G(\mathbf{d}_1, n_0 \to K) = \min \left\{ \mathfrak{P}_{i_{n_0}}^{(\mathbf{d}_1)}, \mathfrak{P}_{i_{n_0+1}}^{(\mathbf{d}_1)}, \ldots, \mathfrak{P}_{i_{K-1}}^{(\mathbf{d}_1)}, \mathfrak{P}_{i_K}^{(\mathbf{d}_1)} \right\} = \mathfrak{P}_{i_{n_0}}^{(\mathbf{d}_1)} > 0.$$

Then from $0 \leq P \leq \overline{P}_L^G (\mathbf{d}_1, n_0 \to K)$, we discuss the transformational sub-policy

$$\widetilde{(\mathbf{d}_1)}_b \,(\text{Transfer}) = \left(0; *, *, \ldots, *, d_{i_{n_0}}, d_{i_{n_0+1}}, \ldots, d_{i_K}; 1, 1, \ldots, 1\right).$$

By observing the transformational sub-policy $\left(d_{n_0}, d_{n_0+1}, \ldots, d_K\right)$ related to $0 \leq P \leq \overline{P}_L^G (\mathbf{d}_1, n_0 \to K)$, it is easy to see from the proof of Theorem 6 that the optimal transformational dynamic rationing sub-policy is given by

$$\widetilde{(\mathbf{d}_1)}_b^* \,(\text{Transfer}) = (0; *, *, \ldots, *, 1, 1, \ldots, 1; 1, 1, \ldots, 1).$$

This gives

$$\widetilde{(\mathbf{d}_1)}_b^* \,(\text{Transfer}) \succ \mathbf{d}_1 = \mathbf{d}^* \tag{55}$$

by using (54), where $\mathbf{d}^*$ is given in Theorem 4.

   Since for a fixed penalty cost $P \geq P_H (\mathbf{d})$ for Policy $\mathbf{d}$, it follows from Theorem 4 that the optimal dynamic rationing policy of the stock-rationing queue is given by

$$\mathbf{d}^* = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1).$$

By using (54), we obtain

$$\widetilde{(\mathbf{d}_1)}_b^* \,(\text{Transfer}) \prec \mathbf{d}^*. \tag{56}$$

   This makes a contradiction between (55) and (56), thus our assumption on the penalty cost: $P < P_H (\mathbf{d}_1)$ should not be correct. This completes the proof.   □

   Theorem 11 shows that to find the optimal dynamic rationing policy of the stock-rationing queue, our first step is to check whether there exists (a) the penalty cost $P \geq P_H (\mathbf{d}_1)$, or (b) the fixed penalty cost $P_L (\mathbf{d}_2) > 0$ and $0 \leq P \leq P_L (\mathbf{d}_2)$. Thus, the two special policies $\mathbf{d}_1$ and $\mathbf{d}_2$ are chosen as the first observation of our algebraic method on the the optimal dynamic rationing policy.

   The following theorem provides the global optimal solution to the optimal dynamic rationing policy of the stock-rationing queue.

Theorem 12 *In the stock-rationing queue with two demand classes, we have*

(1) *If a fixed penalty cost $P \geq P_H (\mathbf{d}_1)$, then $\mathbf{d}^* = \mathbf{d}_1 \succeq \mathbf{c}$ for any $\mathbf{c} \in \mathcal{D}$.*
(2) *If a fixed penalty cost $P_L (\mathbf{d}_2) > 0$ or $0 \leq P \leq P_L (\mathbf{d}_2)$, then $\mathbf{d}^* = \mathbf{d}_2 \succeq \mathbf{c}$ for any $\mathbf{c} \in \mathcal{D}$.*
(3) *If a fixed penalty cost P satisfies $P < P_H (\mathbf{d}_1)$ and $P > P_L (\mathbf{d}_2)$, then*

$$\mathbf{d}^* = \max \left\{ \widetilde{(\mathbf{d}_1)}_b^* \,(\text{Transfer}), \widetilde{(\mathbf{d}_2)}_a^* \,(\text{Transfer}), (\mathbf{c}_k)^* \,(\text{Transfer}) \text{ for } k = 3, 4, \ldots, K \right\}$$

and $\mathbf{d}^* \succeq \mathbf{c}$ for any $\mathbf{c} \in \mathcal{D}$.

Proof We only prove (3), while (1) and (2) are provided in those of Theorem 11.

If $P < P_H(\mathbf{d}_1)$ or $P > P_L(\mathbf{d}_2)$, then both $\mathbf{d}_1$ and $\mathbf{d}_2$ are not the optimal dynamic rationing policy of the system. In this case, by using Theorem 10, we indicate that the optimal dynamic rationing policy must be of transformational threshold type. Thus we have

$$\mathbf{d}^* = \max\left\{ \widetilde{(\mathbf{d}_1)}_b^* (\text{Transfer}), \widetilde{(\mathbf{d}_2)}_a^* (\text{Transfer}), (\mathbf{c}_k)^* (\text{Transfer}) \text{ for } k = 3, 4, \ldots, K \right\},$$

which is of transformational threshold type, since K is a finite positive integer. It is clear that $\mathbf{d}^* \succeq \mathbf{c}$ for any $\mathbf{c} \in \mathcal{D}$. This completes the proof. □

8 The static rationing policies

In this section, we analyze the static (i.e., threshold type) rationing policies of the stock-rationing queue with two demand classes, and discuss the optimality of the static rationing policies. Furthermore, we provide a necessary condition under which a static rationing policy is optimal. Based on this, we can intuitively understand some differences between the optimal static and dynamic rationing policies.

To study static rationing policy, we define a static policy subset of the policy set $\mathcal{D}$ as follows. For $\theta = 1, 2, \ldots, K, K+1$, we write $\mathbf{d}_{\triangle,\theta}$ as a static rationing policy $\mathbf{d}$ with $d_i = 0$ if $1 \leq i \leq \theta - 1$ and $d_i = 1$ if $\theta \leq i \leq K$. Clearly, if $\theta = 1$, then

$$\mathbf{d}_{\triangle,1} = (0; 1, 1, \ldots, 1; 1, 1, \ldots, 1);$$

if $\theta = K$, then

$$\mathbf{d}_{\triangle,K} = (0; 0, 0, \ldots, 0, 1; 1, 1, \ldots, 1);$$

and if $\theta = K + 1$, then

$$\mathbf{d}_{\triangle,K+1} = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1).$$

Let

$$\mathcal{D}^\triangle = \left\{ \mathbf{d}_{\triangle,\theta} : \theta = 1, 2, \ldots, K, K+1 \right\}.$$

Then

$$\mathcal{D}^\triangle = \left\{ \left(0; \underbrace{0, 0, \ldots, 0}_{\theta-1 \text{ zeros}}, 1, 1, \ldots, 1; 1, 1, \ldots, 1\right) : \theta = 1, 2, \ldots, K, K+1 \right\}.$$

It is easy to see that the static rationing policy set $\mathcal{D}^\triangle \subset \mathcal{D}$.

For a static rationing policy $\mathbf{d}_{\triangle,\theta} = \left(0; \underbrace{0, 0, \ldots, 0}_{\theta-1 \text{ zeros}}, 1, 1, \ldots, 1; 1, 1, \ldots, 1\right)$ with $\theta = 1, 2, \ldots, K, K+1$, it follows from (5) that

$$\xi_i^{(\mathbf{d}_{\triangle,\theta})} = \begin{cases} \xi_0 = 1, & i = 0; \\ \alpha^i, & i = 1, 2, \ldots, \theta - 1; \\ \left(\frac{\alpha}{\beta}\right)^{\theta-1} \beta^i, & i = \theta, \theta + 1, \ldots, N. \end{cases}$$

and

$$h^{(\mathbf{d}_{\triangle,\theta})} = 1 + \sum_{i=1}^{N} \xi_i^{(\mathbf{d}_{\triangle,\theta})}$$

$$= 1 + \frac{\alpha\left(1 - \alpha^{\theta-1}\right)}{1 - \alpha} + \left(\frac{\alpha}{\beta}\right)^{\theta-1} \frac{\beta^\theta \left(1 - \beta^{N-\theta+1}\right)}{1 - \beta}.$$

It follows from (6) that

$$\pi^{(\mathbf{d}_{\triangle,\theta})}(i) = \begin{cases} \frac{1}{h^{(\mathbf{d}_{\triangle,\theta})}}, & i = 0; \\ \frac{1}{h^{(\mathbf{d}_{\triangle,\theta})}} \alpha^i, & i = 1, 2, \ldots, \theta - 1; \\ \frac{1}{h^{(\mathbf{d}_{\triangle,\theta})}} \left(\frac{\alpha}{\beta}\right)^{\theta-1} \beta^i, & i = \theta, \theta + 1, \ldots, N. \end{cases}$$

On the other hand, it follows from (8) to (11) that for $i = 0$

$$f(0) = -C_{2,1}\mu_1 - C_{2,2}\mu_2 - C_3\lambda;$$

for $i = 1, 2, \ldots, \theta - 1$,

$$f^{(\mathbf{d}_{\triangle,\theta})}(i) = R\mu_1 - C_1 i - C_{2,2}\mu_2 - C_3\lambda;$$

for $i = \theta, \theta + 1, \ldots, K$,

$$f^{(\mathbf{d}_{\triangle,\theta})}(i) = R(\mu_1 + \mu_2) - C_1 i - C_3\lambda - P\mu_2;$$

and for $i = K + 1, K + 2, \ldots, N$,

$$f(i) = R(\mu_1 + \mu_2) - C_1 i - C_3\lambda 1_{\{i<N\}} - C_4\lambda 1_{\{i=N\}}.$$

Note that

$$\eta^{\mathbf{d}_{\triangle,\theta}} = \pi^{(\mathbf{d}_{\triangle,\theta})}(0) f(0) + \sum_{i=1}^{\theta-1} \pi^{(\mathbf{d}_{\triangle,\theta})}(i) f^{(\mathbf{d}_{\triangle,\theta})}(i)$$

$$+ \sum_{i=\theta}^{K} \pi^{(\mathbf{d}_{\triangle,\theta})}(i) f^{(\mathbf{d}_{\triangle,\theta})}(i) + \sum_{i=K+1}^{N} \pi^{(\mathbf{d}_{\triangle,\theta})}(i) f(i),$$

we obtain an explicit expression for the long-run average profit of the stock-rationing queue under the static rationing policy $\mathbf{d}_{\triangle,\theta}$ as follows:

$$\eta^{\mathbf{d}_{\triangle,\theta}} = \frac{1}{h^{(\mathbf{d}_{\triangle,\theta})}} \left\{ -(C_{2,1}\mu_1 + C_{2,2}\mu_2 + C_3\lambda) + \sum_{i=1}^{\theta-1} \alpha^i (R\mu_1 - C_1 i - C_{2,2}\mu_2 - C_3\lambda) \right.$$

$$+ \sum_{i=\theta}^{K} \left(\frac{\alpha}{\beta}\right)^{\theta-1} \beta^i [R(\mu_1+\mu_2) - C_1 i - C_3\lambda - P\mu_2]$$

$$+ \left. \sum_{i=K+1}^{N} \left(\frac{\alpha}{\beta}\right)^{\theta-1} \beta^i [R(\mu_1+\mu_2) - C_1 i - C_3\lambda 1_{\{i<N\}} - C_4\lambda 1_{\{i=N\}}] \right\}$$

$$= \frac{1}{h^{(\mathbf{d}_{\triangle,\theta})}} \left\{ -\gamma_1 + \gamma_2 \frac{\alpha(1-\alpha^{\theta-1})}{1-\alpha} - C_1 \left[\frac{\alpha(1-\alpha^{\theta-1})}{(1-\alpha)^2} - \frac{(\theta-1)\alpha^\theta}{1-\alpha} \right] \right.$$

$$- \left(\frac{\alpha}{\beta}\right)^{\theta-1} C_1 \left[\frac{(\theta-1)\beta^\theta - N\beta^{N+1}}{1-\beta} + \frac{\beta^\theta(1-\beta^{N-\theta+1})}{(1-\beta)^2} \right]$$

$$+ \left. \left(\frac{\alpha}{\beta}\right)^{\theta-1} \gamma_4 \frac{\beta^\theta(1-\beta^{K-\theta+1})}{1-\beta} + \left(\frac{\alpha}{\beta}\right)^{\theta-1} \gamma_3 \frac{\beta^{K+1}(1-\beta^{N-K})}{1-\beta} \right\}.$$

Let

$$\mathbf{d}_{\triangle,\theta}^* = \arg\max_{\mathbf{d}_{\triangle,\theta}\in\mathcal{D}^{\triangle}} \left\{ \eta^{\mathbf{d}_{\triangle,\theta}} \right\}$$

and

$$\mathbf{d}_{\triangle,\theta*} = \arg\max_{1\leq\theta\leq K+1} \left\{ \eta^{\mathbf{d}_{\triangle,\theta}} \right\}.$$

Then $\mathbf{d}_{\triangle,\theta}^* = \mathbf{d}_{\triangle,\theta*}$. Hence we call $\mathbf{d}_{\triangle,\theta}^*$ (or $\mathbf{d}_{\triangle,\theta*}$) the optimal static rationing policy in the static rationing policy set $\mathcal{D}^{\triangle}$. Since $\mathcal{D}^{\triangle} \subset \mathcal{D}$, the partially ordered set $\mathcal{D}$ shows that $\mathcal{D}^{\triangle}$ is also a partially ordered set. Based on this, it is easy to see from the two partially ordered sets $\mathcal{D}$ and $\mathcal{D}^{\triangle}$ that

$$\eta^{\mathbf{d}_{\triangle,\theta}^*} \leq \eta^{\mathbf{d}^*}, \text{ or } \mathbf{d}_{\triangle,\theta}^* \preceq \mathbf{d}^*,$$

where $\mathbf{d}^*$ is the optimal dynamic rationing policy in the set $\mathcal{D}$.

If $\eta^{\mathbf{d}_{\triangle,\theta}^*} = \eta^{\mathbf{d}^*}$, then the optimal static rationing policy $\mathbf{d}_{\triangle,\theta}^*$ is also optimal in the policy set $\mathcal{D}$, thus the optimal dynamic rationing policy is of threshold type. If $\eta^{\mathbf{d}_{\triangle,\theta}^*} < \eta^{\mathbf{d}^*}$, then the optimal static rationing policy $\mathbf{d}_{\triangle,\theta}^*$ is not optimal in the static

rationing policy subset $\mathcal{D}^\Delta$, i.e., it is also suboptimal in the policy set $\mathcal{D}$, thus the optimal dynamic rationing policy is not of threshold type.

Now, we set up some conditions under which the optimal static rationing policy $\mathbf{d}^*_{\triangle,\theta}$ is suboptimal in the dynamic rationing policy set $\mathcal{D}$.

In the static rationing policy subset $\mathcal{D}^\Delta$, it is easy to see that there must exist a minimal positive integer $\theta^* \in \{1, 2, \ldots, K, K+1\}$ such that

$$\mathbf{d}^*_{\triangle,\theta} = \mathbf{d}_{\triangle,\theta^*} = \left(0; \underbrace{0, 0, \ldots, 0}_{\theta^*-1 \text{ zeros}}, 1, 1, \ldots, 1; 1, 1, \ldots, 1 \right).$$

By using the optimal static rationing policy $\mathbf{d}^*_{\triangle,\theta}$ (or $\mathbf{d}_{\triangle,\theta^*}$), the following theorem determines the sign of the function $G^{(\mathbf{d}_{\triangle,\theta})}(\theta) + b$ in the three different points: $\theta = \theta^*-1, \theta^*, \theta^*+1$. This may be useful for us to understand how to use Proposition 1 to give the optimal long-run average profit of this system.

Theorem 13 *In the stock-rationing queue, the static rationing policies* $\mathbf{d}_{\triangle,\theta^*-1}$, $\mathbf{d}_{\triangle,`*}$ *and* $\mathbf{d}_{\triangle,\theta^*+1}$ *satisfy the following conditions:*

$$G^{(\mathbf{d}_{\triangle,\theta^*-1})}\left(\theta^*-1\right) + b \le 0, \; G^{(\mathbf{d}_{\triangle,\theta^*})}\left(\theta^*-1\right) + b \le 0,$$

and

$$G^{(\mathbf{d}_{\triangle,\theta^*})}\left(\theta^*\right) + b \ge 0, \; G^{(\mathbf{d}_{\triangle,\theta^*+1})}\left(\theta^*\right) + b \ge 0.$$

Proof We consider three static rationing policies with an interrelated structure as follows:

$$\mathbf{d}_{\triangle,\theta^*-1} = \left(0; \underbrace{0, 0, \ldots, 0}_{\theta^*-2 \text{ zeros}}, 1, 1, 1, 1, \ldots, 1; 1, 1, \ldots, 1 \right),$$

$$\mathbf{d}_{\triangle,\theta^*} = \left(0; \underbrace{0, 0, \ldots, 0, 0}_{\theta^*-1 \text{ zeros}}, 1, 1, 1, \ldots, 1; 1, 1, \ldots, 1 \right),$$

$$\mathbf{d}_{\triangle,\theta^*+1} = \left(0; \underbrace{0, 0, \ldots, 0, 0, 0}_{\theta^* \text{ zeros}}, 1, \ldots, 1; 1, 1, \ldots, 1 \right).$$

Note that $\mathbf{d}_{\triangle,\theta^*}$ is the optimal static rationing policy, and $\mathbf{d}^*_{\triangle,\theta} = \mathbf{d}_{\triangle,\theta^*}$. It is clear that $\mathbf{d}_{\triangle,\theta^*} \succeq \mathbf{d}_{\triangle,\theta^*-1}$ and $\mathbf{d}_{\triangle,\theta^*} \succeq \mathbf{d}_{\triangle,\theta^*+1}$. Thus it follows from (39) that

$$\eta^{\mathbf{d}_{\triangle,\theta^*+1}} - \eta^{\mathbf{d}_{\triangle,\theta^*}} = -\mu_2 \pi^{(\mathbf{d}_{\triangle,\theta^*+1})}\left(\theta^*\right) \left[G^{(\mathbf{d}_{\triangle,\theta^*})}\left(\theta^*\right) + b \right],$$

which, together with $\eta^{\mathbf{d}_{\triangle,\theta^*+1}} - \eta^{\mathbf{d}_{\triangle,\theta^*}} \leq 0$, leads to

$$G^{(\mathbf{d}_{\triangle,\theta^*})}\left(\theta^*\right) + b \geq 0.$$

On the other hand, it follows from (39) that

$$\eta^{\mathbf{d}_{\triangle,\theta^*}} - \eta^{\mathbf{d}_{\triangle,\theta^*+1}} = \mu_2 \pi^{(\mathbf{d}_{\triangle,\theta^*})}\left(\theta^*\right)\left[G^{(\mathbf{d}_{\triangle,\theta^*+1})}\left(\theta^*\right) + b\right],$$

this gives

$$G^{(\mathbf{d}_{\triangle,\theta^*+1})}\left(\theta^*\right) + b \geq 0$$

Similarly, by using $\eta^{\mathbf{d}_{\triangle,\theta^*}} \geq \eta^{\mathbf{d}_{\triangle,\theta^*-1}}$ and

$$\eta^{\mathbf{d}_{\triangle,\theta^*}} - \eta^{\mathbf{d}_{\triangle,\theta^*-1}} = -\mu_2 \pi^{(\mathbf{d}_{\triangle,\theta^*})}\left(\theta^* - 1\right)\left[G^{(\mathbf{d}_{\triangle,\theta^*-1})}\left(\theta^* - 1\right) + b\right],$$

this gives

$$G^{(\mathbf{d}_{\triangle,\theta^*-1})}\left(\theta^* - 1\right) + b \leq 0;$$

and

$$\eta^{\mathbf{d}_{\triangle,\theta^*-1}} - \eta^{\mathbf{d}_{\triangle,\theta^*}} = \mu_2 \pi^{(\mathbf{d}_{\triangle,\theta^*-1})}\left(\theta^* - 1\right)\left[G^{(\mathbf{d}_{\triangle,\theta^*})}\left(\theta^* - 1\right) + b\right],$$

we obtain

$$G^{(\mathbf{d}_{\triangle,\theta^*})}\left(\theta^* - 1\right) + b \leq 0.$$

This completes the proof.   □

9 Numerical experiments

In this section, by observing several different penalty costs, we conduct numerical experiments to demonstrate our theoretical results and to gain insights on the optimal dynamic and static rationing policies in the stock-rationing queue.

In Examples 1–4, we take some common parameters in the stock-rationing queue with two demand classes as follows:

$$C_1 = 1, C_{2,1} = 4, C_{2,2} = 1, C_3 = 5, C_4 = 1, R = 15, N = 100.$$

In Examples 1 and 2, we analyze some difference between the optimal static and dynamic rationing policies, and use the optimal static rationing policy to show whether or not the the optimal dynamic rationing policy is of threshold type.

Fig. 3 The optimal long-run average profit $\eta^{\mathbf{d}^*}$ versus the threshold θ
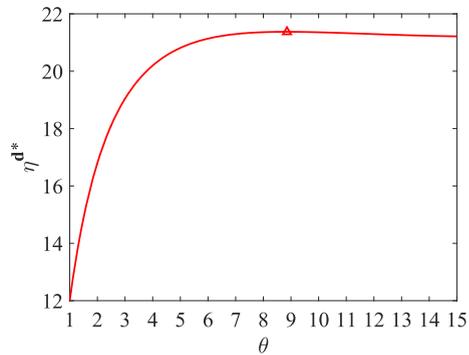
Example 1 We give some useful comparisons of the optimal long-run average profit between two different penalty costs, and further verify how the optimality depends on the penalty cost in Theorems 4 and 6 for the the optimal dynamic rationing policy. To this end, we further take the system parameters as $\lambda = 3$, $\mu_1 = 4$, $\mu_2 = 2$, $K = 15$ and $1 \leq i \leq 15$.

Case one: A higher penalty cost

We take a higher penalty cost $P = 10$. If $d_i^* = 0$ for $1 \leq i \leq 15$ such that a possible optimal dynamic rationing policy $\mathbf{d}^* = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1)$, then we obtain $\eta^{\mathbf{d}^*} = 22.3$. On the other hand, if $d_i'^* = 1$ for $1 \leq i \leq 15$ such that another possible optimal dynamic rationing policy $\mathbf{d}'^* = (0; 1, 1, \ldots, 1; 1, 1, \ldots, 1)$, then we get $\eta^{\mathbf{d}'^*} = 13$. By comparing $\eta^{\mathbf{d}^*} = 22.3$ with $\eta^{\mathbf{d}'^*} = 13$, it is easy to see that the possible optimal dynamic rationing policy should be $\mathbf{d}^* = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1)$.

Case two: A lower penalty cost

We choose a lower penalty cost $P = 0.1$. If $d_i^* = 1$ for $1 \leq i \leq 15$ such that a possible optimal dynamic rationing policy $\mathbf{d}^* = (0; 1, 1, \ldots, 1; 1, 1, \ldots, 1)$, then we obtain $\eta^{\mathbf{d}^*} = 22.9$. On the other hand, if $d_i'^* = 0$ for $1 \leq i \leq 15$ such that another possible optimal dynamic rationing policy $\mathbf{d}'^* = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1)$, then we get $\eta^{\mathbf{d}'^*} = 22.3$. Obviously, the possible optimal dynamic rationing policy should be $\mathbf{d}^* = (0; 1, 1, \ldots, 1; 1, 1, \ldots, 1)$.

Example 2 We use the numerical example to demonstrate whether or not the optimal static rationing policy is suboptimal in the policy set $\mathcal{D}$. If yes, then we show that the optimal dynamic rationing policy is not of threshold type. To this end, we take some system parameters: $\lambda = 3$, $\mu_1 = 4$, $\mu_2 = 2$, $K = 15$, $1 \leq \theta \leq 15$. These parameters are the same as those in Example 1.

In what follows our observation is to focus on the higher penalty cost $P = 10$ and the lower penalty cost $P = 0.1$, respectively.

Case one: A higher penalty cost

We observe how the optimal long-run average profit $\eta^{\mathbf{d}^*}$ depends on the threshold from $\theta = 1$ to $\theta = 15$. From Fig. 3, it is seen that the optimal threshold is $\theta^* = 9$ and $\eta^{\mathbf{d}_\Delta,\theta^*} = 21.4$. However, from Case one of Example 1, $\eta^{\mathbf{d}^*} = 22.3$. Thus we obtain that $\eta^{\mathbf{d}_\Delta,\theta^*} = 21.4 < \eta^{\mathbf{d}^*} = 22.3$. This shows that the optimal static rationing policy

Fig. 4 The optimal long-run average profit $\eta^{\mathbf{d}*}$ versus the threshold θ
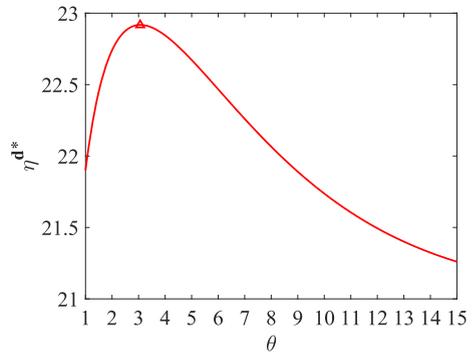
Fig. 5 $\eta^{\mathbf{d}*}$ versus λ under three different thresholds K
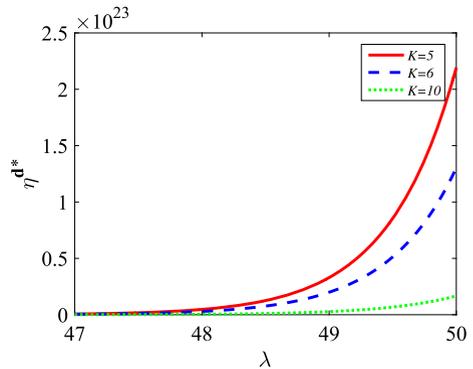
is suboptimal in the policy set $\mathcal{D}$, and the optimal dynamic rationing policy is not of threshold type. Thus, $\mathbf{d}^* = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1)$, given in Example 1, is not the optimal dynamic rationing policy yet.

Case two: A lower penalty cost

From Fig. 4, it is seen that the optimal threshold is $\theta^* = 3$ and $\eta^{\mathbf{d}_\triangle,\theta*} = 22.9$. From Case two of Example 1, we obtained $\eta^{\mathbf{d}*} = 22.9$. This gives that $\eta^{\mathbf{d}_\triangle,\theta*} = \eta^{\mathbf{d}*} = 22.9$. Therefore, the optimal static rationing policy is the same as the optimal dynamic rationing policy, and it is optimal in the policy set $\mathcal{D}$, and the optimal dynamic rationing policy is of threshold type. Thus, $\mathbf{d}^* = (0; 1, 1, \ldots, 1; 1, 1, \ldots, 1)$, given in Example 1, is the optimal dynamic rationing policy.

Example 3 We analyze how the optimal long-run average profit of the stock-rationing queue depends on the arrival rate. Our observation focuses on the higher penalty cost $P = 10$ and the lower penalty cost $P = 0.1$, respectively. To do this, we further take the system parameters: $\mu_1 = 30$, $\mu_2 = 40$ and the threshold: $K = 5, 6, 10$.

Case one: A higher penalty cost

Let $P = 10$ and $\mathbf{d}^* = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1)$. From Fig. 5, it is seen that the optimal long-run average profit $\eta^{\mathbf{d}*}$ increases as λ increases. In addition, with the threshold K increases, the optimal long-run average profit $\eta^{\mathbf{d}*}$ increases less slowly as λ increases.

Case two: A lower penalty cost

Fig. 6 $\eta^{\mathbf{d}^*}$ versus λ under three different thresholds K
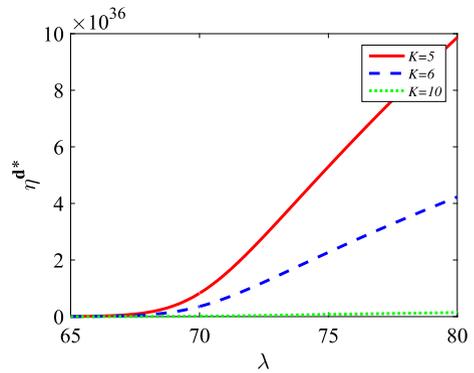
Fig. 7 The long-run average profit $\eta^{\mathbf{d}^*}$ versus the penalty cost P
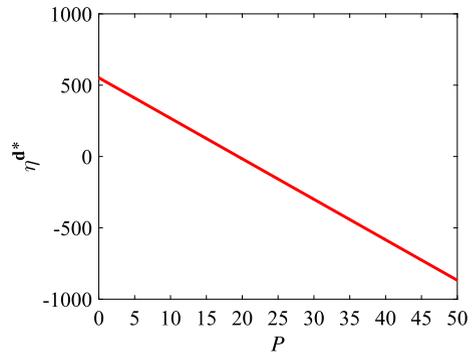
Let $P = 0.1$ and $\mathbf{d}^* = (0; 1, 1, \ldots, 1; 1, 1, \ldots, 1)$. We discuss how the optimal long-run average profit $\eta^{\mathbf{d}^*}$ depends on λ for $\lambda \in (65, 80)$. From Fig. 6, it is seen that the optimal long-run average profit $\eta^{\mathbf{d}^*}$ increases as λ increases. In addition, with the threshold K increases, the optimal long-run average profit $\eta^{\mathbf{d}^*}$ increases less slowly as λ increases.

Example 4 Our observation is to focus on how the penalty cost P influences the long-run average profit $\eta^{\mathbf{d}}$ for any given policy $\mathbf{d}$. From (17), it is easy to see that for any given policy $\mathbf{d}$, the long-run average profit $\eta^{\mathbf{d}}$ is linear in the penalty cost P. To show this, we take the system parameters: $P \in (0, 50)$, $\mu_1 = 4$, $\mu_2 = 2$, $\lambda = 3$ and $K = 15$. In this case, we observe the special policy: $\mathbf{d}_1 = \mathbf{d}^* = (0; 0, 0, \ldots, 0; 1, 1, \ldots, 1)$. Figure 7 shows that for the special policy $\mathbf{d}_1 = \mathbf{d}^*$, the long-run average profit $\eta^{\mathbf{d}^*}$ linearly decreases as P increases.

10 Concluding remarks

In this paper, we highlight intuitive understanding on the optimal dynamic rationing policy of the stock-rationing queue with two demand classes by means of the sensitivity-based optimization. To find the optimal dynamic rationing policy, we establish a policy-based birth-death process and a more general reward function such that the

long-run average profit of the stock-rationing queue is expressed explicitly. Furthermore, we set up a policy-based Poisson equation and provide an explicit expression for its solution. Based on this, we derive a performance difference equation between any two policies such that we can find the optimal dynamic rationing policy and compute the maximal long-run average profit from three different areas of the penalty costs. Therefore, we provide an algebraic method to set up a complete algebraic solution to the optimal dynamic rationing policy. We show that the optimal dynamic policy must be of transformational threshold type, which leads to refining three simple sufficient conditions under each of which the optimal dynamic policy is of threshold type. In addition, we develop some new structural properties (e.g., set-structured monotonicity, and class property of policies) of the optimal dynamic rationing policy. Therefore, we believe that the methodology and results developed in this paper can be applicable to analyzing supply chain finance and with applications of blockchain technology, and open a series of potentially promising research.

Along such a line, there are a number of interesting directions for potential future research, for example:

- Extending to the stock-rationing queues with multiple demand classes, multiple types of products, backorders, batch order, batch production, and so on;
- analyzing non-Poisson input, such as Markovian arrival processes (MAPs); and/or non-exponential service times, e.g. the PH distributions;
- discussing how the long-run profit can be influenced by some concave or convex reward functions;
- studying individual or social optimization for stock-rationing queues from a perspective of game theory by means of the sensitivity-based optimization;
- investigating optimal dynamic rationing policy in supply chain finance and with applications of blockchain technology.

Data availibility Enquiries about data availability should be directed to the authors.

Declarations

Conflict of interest The authors have not disclosed any competing interests.

References

Alfieri A, Pastore E, Zotteri G (2017) Dynamic inventory rationing: how to allocate stock according to managerial priorities. An empirical study. Int J Prod Econ 189:14–29

Altug MS, Ceryan O (2022) Optimal dynamic allocation of rental and sales inventory for fashion apparel products. IISE Trans 54(6):603–617

Arslan H, Graves SC, Roemer TA (2007) A single-product inventory model for multiple demand classes. Manag Sci 53(9):1486–1500

Baron O, Lu T, Wang J (2019) Priority, capacity rationing, and ambulance diversion in emergency departments. SSRN, 3387439

Benjaafar S, ElHafsi M (2006) Production and inventory control of a single product assemble-to-order system with multiple customer classes. Manag Sci 52(12):1896–1912

Cao XR (2007) Stochastic learning and optimization: a sensitivity-based approach. Springer

Chew EP, Lee LH, Liu S (2013) Dynamic rationing and ordering policies for multiple demand classes. OR Spectr 35(1):127–151

Cohen MA, Kleindorfer PR, Lee HL (1988) Service constrained (s, S) inventory systems with priority demand classes and lost sales. Manag Sci 34(4):482–499

Dekker R, Hill RM, Kleijn MJ, Teunter RH (2002) On the $(S − 1, S)$ lost sales inventory model with priority demand classes. Naval Res Logist 49(6):593–610

Dekker R, Kleijn MJ, De Rooij PJ (1998) A spare parts stocking policy based on equipment criticality. Int J Prod Econ 56:69–77

Deshpande V, Cohen MA, Donohue K (2003) A threshold inventory rationing policy for service-differentiated demand classes. Manag Sci 49(6):683–703

Ding Q, Kouvelis P, Milner J (2016) Inventory rationing for multiple class demand under continuous review. Prod Oper Manag 25(8):1344–1362

Elaydi SN (1999) Dynamics of first order difference equations. In: An introduction to difference equations. Springer, pp 1–48

Elhafsi M, Zhi L, Camus H, Craye E (2015) An assemble-to-order system with product and components demand with lost sales. Int J Prod Res 53(3):718–735

ElHafsi M, Fang J, Camus H (2018) Optimal control of a continuous-time W-configuration assemble-to-order system. Eur J Oper Res 267(3):917–932

Escalona P, Ordóñez F, Kauak I (2017) Critical level rationing in inventory systems with continuously distributed demand. OR Spectr 39(1):273–301

Escalona P, Ordóñez F, Marianov V (2015) Joint location-inventory problem with differentiated service levels using critical level policy. Transp Res Part E: Logist Transp Rev 83:141–157

Fadıloğlu MM, Bulut Ö (2010) A dynamic rationing policy for continuous-review inventory systems. Eur J Oper Res 202(3):675–685

Frank KC, Zhang RQ, Duenyas I (2003) Optimal policies for inventory systems with priority demand classes. Oper Res 51(6):993–1002

Gayon JP, Benjaafar S, De Véricourt F (2009) Using imperfect advance demand information in production-inventory systems with multiple customer classes. Manuf Serv Oper Manag 11(1):128–143

Gayon JP, De Vericourt F, Karaesmen F (2009) Stock rationing in an $M/E_k/1$ multi-class make-to-stock queue with backorders. IIE Trans 41(12):1096–1109

Goedhart J, Haijema R, Akkerman R (2022) Inventory rationing and replenishment for an omni-channel retailer. Comput Oper Res 140:105647

Ha AY (1997) Inventory rationing in a make-to-stock production system with several demand classes and lost sales. Manag Sci 43(8):1093–1103

Ha AY (1997) Stock-rationing policy for a make-to-stock production system with two priority classes and backordering. Naval Res Logist 44(5):457–472

Ha AY (2000) Stock rationing in an $M/E_k/1$ make-to-stock queue. Manag Sci 46(1):77–87

Haynsworth HC, Price BA (1989) A model for use in the rationing of inventory during lead time. Naval Res Logist 36(4):491–506

Huang B, Iravani SM (2008) A make-to-stock system with multiple customer classes and batch ordering. Oper Res 56(5):1312–1320

Hung HC, Chew EP, Lee LH, Liu S (2012) Dynamic inventory rationing for systems with multiple demand classes and general demand processes. Int J Prod Econ 139(1):351–358

Hung YF, Hsiao JY (2013) Inventory rationing decision models during replenishment lead time. Int J Prod Econ 144(1):290–300

Isotupa KS (2006) An (s, Q) Markovian inventory system with lost sales and two demand classes. Math Comput Model 43(7–8):687–694

Ioannidis S, Xanthopoulos AS, Sarantis I, Koulouriotis DE (2021) Joint production, inventory rationing, and order admission control of a stochastic manufacturing system with setups. Oper Res 21(2):827–855

Jain A, Moinzadeh K, Dumrongsiri A (2015) Priority allocation in a rental model with decreasing demand. Manuf Serv Oper Manag 17(2):236–248

Kranenburg AA, van Houtum GJ (2007) Cost optimization in the $(S - 1, S)$ lost sales inventory model with multiple demand classes. Oper Res Lett 35(4):493–502

Li QL (2010) Constructive computation in stochastic models with applications: the RG-factorizations. Springer

Li QL, Ma JY, Fan RN, Xia L (2019a) An overview for Markov decision processes in queues and networks. In: Stochastic models in reliability, network security and system safety, pp 44–71

Li QL, Li YM, Ma JY, Liu HL (2019b) A complete algebraic solution to the optimal dynamic rationing policy in the stock-rationing queue with two demand classes. arXiv: 1908.09295

Ma JY, Xia L, Li QL (2019) Optimal energy-efficient policies for data centers through sensitivity-based optimization. Discret Event Dyn Syst 29(4):567–606

Ma JY, Li QL, Xia L (2021) Optimal asynchronous dynamic policies in energy-efficient data centers. Systems 10(2):27

Melchiors P (2003) Restricted time-remembering policies for the inventory rationing problem. Int J Prod Econ 81(1):461–468

Melchiors P, Dekker R, Kleijn MJ (2000) Inventory rationing in an (s, Q) inventory model with lost sales and two demand classes. J Oper Res Soc 51(1):111–122

Möllering K (2007) Inventory rationing: a new modeling approach using Markov chain theory. Springer

Möllering KT, Thonemann UW (2008) An optimal critical level policy for inventory systems with two demand classes. Naval Res Logist 55(7):632–642

Möllering KT, Thonemann UW (2010) An optimal constant level rationing policy under service level constraints. OR Spectr 32(2):319–341

Nadar E, Akan M, Scheller-Wolf A (2014) Optimal structural results for assemble-to-order generalized M-systems. Oper Res 62(3):571–579

Nahmias S, Demmy WS (1981) Operating characteristics of an inventory system with rationing. Manag Sci 27(11):1236–1245

Pang Z, Shen H, Cheng TCE (2014) Inventory rationing in a make-to-stock system with batch production and lost sales. Prod Oper Manag 23(7):1243–1257

Papastavrou E, Andreou P, Efstathiou G (2014) Rationing of nursing care and nurse-patient outcomes: a systematic review of quantitative studies. Int J Health Plan Manag 29(1):3–25

Puterman ML (2014) Markov decision processes: discrete stochastic dynamic programming. Wiley

Schulte B, Pibernik R (2017) Profitability of service-level-based price differentiation with inventory rationing. Prod Oper Manag 26(5):903–923

Shen X, Yu Y (2019) Capacity allocation with multiple suppliers and multiple demand classes. Prod Oper Manag 28(11):2792–2807

Sobel MJ, Zhang RQ (2001) Inventory policies for systems with stochastic and deterministic demand. Oper Res 49(1):157–162

Tan T, Güllü R, Erkip N (2009) Using imperfect advance demand information in ordering and rationing decisions. Int J Prod Econ 121(2):665–677

Tempelmeier H (2006) Supply chain inventory optimization with two customer classes in discrete time. Eur J Oper Res 174(1):600–621

Teunter RH, Haneveld WKK (2008) Dynamic inventory rationing strategies for inventory systems with two demand classes, Poisson demand and backordering. Eur J Oper Res 190(1):156–178

Topkis DM (1968) Optimal ordering and rationing policies in a nonstationary dynamic inventory model with n demand classes. Manag Sci 15(3):160–176

Turgay Z, Karaesmen F, Örmeci EL (2015) A dynamic inventory rationing problem with uncertain demand and production rates. Ann Oper Res 231(1):207–228

Van Foreest ND, Wijngaard J (2014) On optimal policies for production-inventory systems with compound Poisson demand and setup costs. Math Oper Res 39(2):517–532

van Wijk ACC, Adan IJ, van Houtum GJ (2019) Optimal lateral transshipment policies for a two location inventory problem with multiple demand classes. Eur J Oper Res 272(2):481–495

Veinott Jr AF (1965) Optimal policy in a dynamic, single product, nonstationary inventory model with several demand classes. Oper Res 13(5):761–778

Wang D, Tang O (2014) Dynamic inventory rationing with mixed backorders and lost sales. Int J Prod Econ 149:56–67

Wang D, Tang O, Huo J (2013) A heuristic for rationing inventory in two demand classes with backlog costs and a service constraint. Comput Oper Res 40(12):2826–2835

Wang Y, Zhang SH, Sun L (2013) Anticipated rationing policy for two demand classes under service level constraints. Comput Ind Eng 65(2):331–340

Wang R, Qin Y, Sun H (2021) Research on location selection strategy for airlines spare parts central warehouse based on METRIC. Comput Intell Neurosci 2021:1–16

Xia L, Zhang ZG, Li QL (2021) A c/μ-rule for job assignment in heterogeneous group-server queues. Prod Oper Manag 31(3):1191–1215

You PS (2003) Dynamic rationing policies for product with incremental upgrading demands. Eur J Oper Res 144(1):128–137

Zhao H, Deshpande V, Ryan JK (2005) Inventory sharing and rationing in decentralized dealer networks. Manag Sci 51(4):531–547