Inventory rationing in an (s, Q) inventory model with lost sales and two demand classes

Philip Melchiors* Rommert Dekker† Marcel Kleijn†

SEPTEMBER 1998
ECONOMETRIC INSTITUTE REPORT 9837/A

Abstract

Whenever demand for a single item can be categorized into classes of different priority, an inventory rationing policy should be considered. In this paper we analyse a continuous review (s, Q) model with lost sales and two demand classes. A so-called critical level policy is applied to ration the inventory among the two demand classes. With this policy, low–priority demand is rejected in anticipation of future high–priority demand whenever the inventory level is at or below a prespecified critical level. For Poisson demand and deterministic lead times, we present an exact formulation of the average inventory cost. A simple optimization procedure is presented, and in a numerical study we compare the optimal rationing policy with a policy where no distinction between the demand classes is made. The benefit of the rationing policy is investigated for various cases and the results show that significant cost reductions can be obtained.

Keywords: Inventory, rationing, lost sales, two demand classes.

1 Introduction

In most of the literature on inventory models it is assumed that all demand for a single item is equally important. However, in practice, the demand for an item can often be categorized into classes of different priority. Consider, for example, the spare parts inventory in the airline industry. Most airlines have a contractual agreement with a company that supplies them with spare parts whenever an aircraft is grounded at the airport due to failure of

*University of Aarhus, Department of Operations Research, Ny Munkegade Build. 530, 8000 Aarhus C, Denmark. e–mail : philip@imf.au.dk, Corresponding author.

†Erasmus University Rotterdam, Econometric Institute, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.

some equipment. In the contractual agreement it is stated that the company promises e.g. a 98% service level to the airline. Besides these key customers, the spare parts inventory company may also satisfy demand from other airlines. These airlines are usually considered to be of lower priority, and their orders will only be satisfied if the inventory level is high enough, so that the 98% service level for the key customers is not endangered. Another situation where demand for a single item may have different priorities occurs in multi-echelon inventory systems with emergency orders (see e.g. Chiang & Gutierrez [1]). In many of these systems emergency orders are placed at the lowest echelon whenever the stock level is low or when customer demand is backordered. As a result, at the higher echelon two types of demand are faced: emergency orders and normal replenishment orders. Whenever the higher echelon has insufficient stock to meet both types of demand, priority will be given to the emergency orders. Finally, we mention an example that can be found in an assembly–to–order system, where a component may be used for several end–products. If these end–products yield different profits to the firm, then demand for this component may be categorized into classes of different priority.

A simple way of operating inventory systems with two demand classes is to use a rationing policy that reserves part of the stock for high priority demand by rejecting low–priority demand when stock is below a certain critical level. Henceforth, we refer to such a policy as a *critical level policy* and we will restrict ourselves to policies where the critical level is independent of the remaining lead time. However, such information, if available, could lead to improved policies. For example, if the inventory manager knows that a replenishment order will arrive soon, it may be optimal to satisfy low–priority demand even though the inventory level is below the critical level. A disadvantage of operating a policy that takes into account information about remaining lead times is that it is much more difficult to implement in practice.

In this paper we will consider a critical level policy in the context of a continuous review (s, Q) inventory model with lost sales. In some practical situations, a customer demand is handled in another way, e.g. through another supplier, if it cannot be delivered from stock on hand. Hence, at the inventory system, this demand may be viewed as a lost sales. The stockout cost in this case represents the additional cost for expediting the customer order. To the best of our knowledge, this model has not been analyzed in the literature so far. However, some closely related models have appeared and an overview of them is presented in Section 2. Sections 3 and 4 of this paper deal with the derivation of the average inventory cost in a continuous review (s, Q) inventory model with lost sales and two demand classes. For Poisson demand and fixed lead times we derive an expression for the average inventory cost. In Section 5 the optimization of the policy parameters is discussed, and Section 6 illustrates the model by means of some numerical examples. The main conclusions are presented in the last section of this paper.

2 Related literature

In this paper two areas in inventory control theory are combined, i.e. continuous review (s, Q) inventory models with lost sales, and inventory rationing. In this section an overview of related literature in both areas is presented.

The (s, Q) model with lost sales was first discussed by Hadley & Whitin [7]. They derived an exact formulation of the average inventory cost for an (s, Q) policy with Poisson demand and constant deterministic lead times, under the assumption that at most one order is outstanding. They also presented an easy approximation of the average cost and developed an iterative procedure to optimize the policy parameters, which has become the standard textbook approach (see e.g. Silver [12] and Tersine [13]). More recently, Johansen & Thorstenson [8] formulated and solved the same model as a semi–Markov decision model.

Inventory rationing was first introduced by Veinott [16], who proposed a critical–level policy for a periodic review model with n demand classes and zero lead time in a backorder environment. This model was also analyzed by Topkis [15], and for two demand classes by Kaplan [9] and Evans [5]. The first contribution considering multiple demand classes in a continuous review inventory model was made by Nahmias & Demmy [11]. They analysed an (s, Q) inventory model with two demand classes, Poisson demand, backordering, a fixed lead time and a critical level policy, under the assumption that there is at most one outstanding order. This assumption implies that whenever a replenishment order is triggered, the net inventory and the inventory position are identical. Their main contribution was the derivation of approximate expressions for the fill rates. In their analysis they used the notion of the *hitting time* of the critical level, i.e. the time that the inventory level reaches the critical level. Conditioning on this hitting time, it is possible to derive approximate expressions for the cost and service levels. Observe that the model presented by Nahmias & Demmy [11] is the one that is closest related to the model we present in this paper.

Dekker, Kleijn & De Rooij [4] considered a lot–for–lot inventory model with the same characteristics, but without the assumption of at most one outstanding order. They discussed a case study on the inventory control of slow moving spare parts in a large petrochemical plant, where parts were installed in equipment of different criticality. Their main result was the derivation of (approximate) expressions for the fill rates for both demand classes. The results of Nahmias & Demmy [11] were generalized by Moon & Kang [10]. They considered an (s, Q) model with compound Poisson demand, and derived (approximate) expressions for the fill rates of the two demand classes.

Rationing policies in a lost sales environment have not received much attention. Cohen, Kleindorfer & Lee [2] consider a periodic review (s, S) inventory system where all demands in each period are collected, and by the end of each period the inventory is used to satisfy high–priority demand first, and the remaining inventory is then made available for low–priority demand. Hence, they did not consider a critical–level policy. Recently, Ha [6] analyzed a lot–for–lot lost sales model with n demand classes and Poisson demand. He assumed exponentially distributed lead times and modelled the system as a single–product $M/M/1/S$ queue (Tijms [14]) with state–dependent service times. This enabled him to prove optimality of the lot–for–lot critical–level policy. Dekker, Hill & Kleijn [3] analyzed the same model with a general lead time distribution. They modelled the system as an $M/M/S/S$ queue (Tijms [14]) and developed efficient methods to determine the optimal policy. Since they restricted themselves to policies which are independent of the remaining lead time, the optimality of the critical level policy could not be guaranteed for generally distributed lead times.

3 Notation and preliminaries

In this section we introduce the notation that will be used throughout this paper. We consider an inventory model with two demand classes, each with unit Poisson demand with arrival rate λ_1 for high–priority demand and λ_2 for low–priority demand. The cost of not satisfying a demand from demand class j is denoted by π_j, $j = 1, 2$, with $\pi_1 > \pi_2 > 0$. All demand not satisfied immediately is assumed to be lost. The fixed ordering cost is K and there is a fixed lead time of L time units. The unit holding cost per time unit is denoted by $h > 0$.

The (s, Q) policy extended with a critical level is denoted as a (c, s, Q) policy, which operates as follows: whenever the inventory level drops to the reorder level s, a replenishment order of size Q is placed which arrives after L time units. Demand from both classes is satisfied whenever the inventory level exceeds the critical level c, otherwise only high–priority demand (class 1) is satisfied from stock on hand and low–priority demand is lost. Following Hadley & Whitin [7] and Nahmias & Demmy [11], we will restrict ourselves to policies in which there is at most one outstanding order. In a lost sales environment, the condition that $s < Q$ is sufficient to enforce that at most one order is outstanding. In principle, the critical level c is unbounded, but for the model to be tractable we need to require that $c < Q$. In order to be able to derive an expression for the average cost, we need some additional notation. Let $X(t)$ denote the physical inventory level at time t, and let $\{X(t), t \geq 0\}$ be the corresponding stochastic process. The restriction $Q > s$ ensures that $\{X(t), t \geq 0\}$ is a regenerative process with regenerative epochs each time the inventory level reaches the reorder level s and a replenishment order is placed. Define a cycle as the time between two consecutive regenerative epochs. Then our process can be split into independent and identically distributed renewal cycles. Using the renewal–reward theorem (see e.g. Tijms [14]) we know that the average cost per time unit equals the expected cost during a cycle divided by the expected length of a cycle. In case the inventory policy satisfies the condition $c < s$, we let H be a random variable denoting the *hitting time* of the critical level, i.e. the time from placing a replenishment order (or the time when the inventory level 'hits' the reorder level s) until the time where the inventory level 'hits' the critical level c. Since the total demand from both classes follows a Poisson distribution with parameter $\lambda := \lambda_1 + \lambda_2$, it readily follows that H is Erlang-$(s-c)$ distributed with parameter λ. Furthermore, we define R as the random variable denoting the inventory level just before a replenishment order arrives. Figure 1 illustrates the inventory process

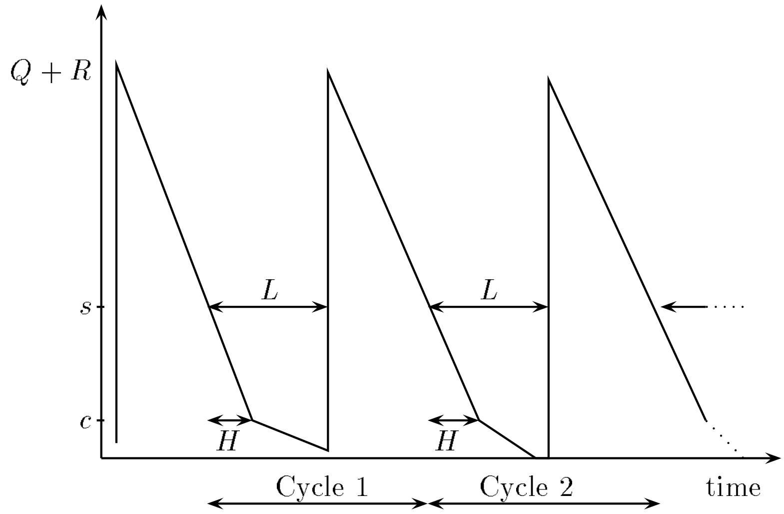

Figure 1: The inventory process with $c < s$.

over two cycles for a (c, s, Q) policy with $c < s$.
In Figure 1 R is positive in the first cycle, while in cycle 2 R is zero. Observe that the inventory level just after a replenishment order of size Q arrives is equal to $Q + R$.
Let $D_j(t)$, $j = 1, 2$, be a random variable denoting the demand from demand class j during t time units, and let $D(t) := D_1(t) + D_2(t)$ be the total demand from both classes during t time units. We can find the distribution of R as

$$\mathbb{P}(R = i) = \begin{cases} \mathbb{P}(D_1(L-H) \geq c) & \text{for } i = 0 \\ \mathbb{P}(D_1(L-H) = c - i) & \text{for } 0 < i \leq c \\ \mathbb{P}(D(L) = s - i) & \text{for } c < i \leq s \end{cases} \tag{1}$$

For a (c, s, Q) policy with $c \geq s$ we reach the critical level before we place an order and the hitting time is therefore not defined. The distribution of R is then simply

$$\mathbb{P}(R = i) = \begin{cases} \mathbb{P}(D_1(L) \geq s) & \text{for } i = 0 \\ \mathbb{P}(D_1(L) = s - i) & \text{for } 0 < i \leq s \end{cases} \tag{2}$$

We are now able to calculate the expected cost of a (c, s, Q) policy. This will be done in the next section.

4 Deriving the average cost

In this section an expression for the average cost in a (c, s, Q) inventory system will be derived. The total cost is composed of inventory holding and shortage costs, and ordering costs. The approach we follow is to derive first the expected cost during a cycle and then calculate the expected cycle length. Using the renewal–reward theorem we obtain an expression for the average cost. We divide the analysis into two parts: first, we consider policies with $c < s$, thereafter we discuss the situation where $c \geq s$.

4.1 Average cost for $c < s$

We first consider the case where $c < s$, so that the inventory level hits the critical level after a replenishment order is placed. In this case we introduce the hitting time H and we may find the expected number of stockouts per cycle $B_j^{c<s}$ for demand class j, $j = 1, 2$, by conditioning on this hitting time, i.e.

$$\begin{aligned} B_1^{c<s} &= \mathbb{E}_{D_1(L-H),H}[D_1(L-H) - c]^+ \\ B_2^{c<s} &= \mathbb{E}_{D_2(L-H),H}[D_2(L-H)] \end{aligned}$$

Since the distributions of D_1, D_2 and H have been determined it is not difficult to calculate B_1 and B_2. In Appendix 1 we give the results.
The total expected shortage cost per cycle is given by

$$\mathbb{E}[SC^{c<s}] = \pi_1 B_1^{c<s} + \pi_2 B_2^{c<s}$$

The holding cost incurred during a cycle is the sum of the holding cost incurred on each inventory level visited during the cycle. The cost incurred on one inventory level i is simply the number of units i times the unit holding cost per time unit, h, times the expected time spent on the level. It is a well known fact that (see e.g. Tijms [14], p. 24)

> "given the occurrence of n arrivals in $(0, t)$, the n arrival epochs are statistically indistinguishable from n independent observations taken from the uniform distribution on $(0, t)$"

The expected time spent on each inventory level reached during a period of length t with n demands is therefore $t/(n + 1)$ if the time interval does not end with a demand (e.g. when a replenishment arrives), and t/n if the time interval does end with a demand. We will split the holding cost up in two parts. The holding cost $HC_1^{c<s}$ incurred during the lead time, and the holding cost $HC_2^{c<s}$ incurred in the remaining part of the cycle. The holding cost incurred in the lead time depends on whether we hit the critical level during

the lead time, and if so, whether the inventory is depleted during the remaining lead time. If the total lead time demand $D(L)$ is less than $s-c$ then the holding cost incurred is

$$h\mathbb{E}_{D(L)}\left[\sum_{i=s-D(L)}^{s} i \cdot \frac{L}{D(L)+1}\right] \tag{3}$$

If we hit the critical level (i.e. $D(L) \geq s-c$), we divide the holding cost in the holding cost before and after the hitting time H. The expected holding cost incurred before the hitting time is

$$h\mathbb{E}_{H}\left[\sum_{i=c+1}^{s} i \cdot \frac{H}{s-c}\right] \tag{4}$$

If after the lead time the inventory is not depleted, that is $D_1(L-H) < c$, the expected holding cost incurred is

$$h\mathbb{E}_{D_1(L-H),H}\left[\sum_{i=c-D_1(L-H)}^{c} i \cdot \frac{L-H}{D_1(L-H)+1}\right] \tag{5}$$

If the inventory is depleted during the remaining part of the lead time, the expected arrival time of the last demand satisfied is $\frac{c}{D_1(L-H)+1}(L-H)$ time units after the hitting time. The expected holding cost incurred is therefore

$$h\mathbb{E}_{D_1(L-H),H}\left[\sum_{i=1}^{c} i \cdot \frac{(L-H)c/(D_1(L-H)+1)}{c}\right] \tag{6}$$

We have now described the holding cost incurred during the lead time as a function of the random variables H, $D(L)$ and $D_1(L-H)$. Since their distributions are known, we can determine the expected holding cost incurred during the lead time. In Appendix 1 a complete derivation of the expected inventory cost $HC_1^{c<s}$, suitable for implementation, is presented.

We will now find the expected holding cost $HC_2^{c<s}$ incurred in the remaining part of the cycle. Observe that after a replenishment order arrives the inventory level is $R+Q$. The expected holding cost incurred while the inventory level drops to s is the sum of the holding cost incurred on each level, and since we have unit demand the expected time spent on each level is $1/\lambda$. Taking expectations with respect to R yields

$$HC_2^{c<s} = h\mathbb{E}_R\Big[\sum_{i=s+1}^{Q+R} i \cdot \frac{1}{\lambda}\Big]$$

The total expected holding cost is

$$\mathbb{E}[HC^{c<s}] = HC_1^{c<s} + HC_2^{c<s}$$

All we need now to derive an expression for the average inventory cost is the expected length of a cycle, which is the lead time plus the expected length of the period where the inventory is reduced from $R+Q$ to s.

$$\mathbb{E}[LoC^{c<s}] = L + \frac{Q + \mathbb{E}[R] - s}{\lambda}$$

Hence, the average cost of a (c, s, Q) policy with $c < s$ is given by

$$TC^{c<s}(c,s,Q) = \frac{\mathbb{E}[SC^{c<s}] + \mathbb{E}[HC^{c<s}] + K}{\mathbb{E}[LoC^{c<s}]}$$

4.2 Average cost for $c \geq s$

In the model developed above only rationing policies with $c < s$ were considered. In this section we will find the expected cost of a policy with $c \geq s$. For such policies we will start rejecting demand from demand class 2 before we place an order. The analysis is similar to the one in the previous section and we will adopt the same notation. The expected number of stockouts for demand class 1 and 2 is

$$\begin{aligned} B_1^{c\geq s} &= \mathbb{E}_{D_1(L)}[D_1(L) - s]^+ \\ B_2^{c\geq s} &= \mathbb{E}_{D_2(L+\tau),\tau}[D_2(L+\tau)] \end{aligned}$$

where $\tau := \inf\{t \geq 0 : D_1(t) \geq c - s\}$ denotes the time between the start of rejecting low–priority demand and placing a replenishment order. By observing that $\mathbb{E}\tau = (c-s)/\lambda_1$ we obtain by the memoryless property of the Poisson process that $B_2^{c\geq s} = \lambda_2(L + (c-s)/\lambda_1)$. The calculation of $B_1^{c\geq s}$ is straightforward.
The expected total stockout cost per cycle is

$$\mathbb{E}[SC^{c\geq s}] = \pi_1 B_1^{c\geq s} + \pi_2 B_2^{c\geq s}$$

To calculate the expected holding cost during a cycle, we divide the holding cost in three parts as shown in Figure 2.
The expected holding cost incurred is found by the same principles used in the previous section.

$$\begin{aligned} HC_3^{c\geq s} &= h\mathbb{E}_R\Big[\sum_{i=c+1}^{Q+R} i \cdot \frac{1}{\lambda}\Big] \\ HC_2^{c\geq s} &= h \sum_{i=s+1}^{c} i \cdot \frac{1}{\lambda_1} \end{aligned}$$

By conditioning on whether the inventory is depleted or not and using the same reasoning as in (6) we obtain

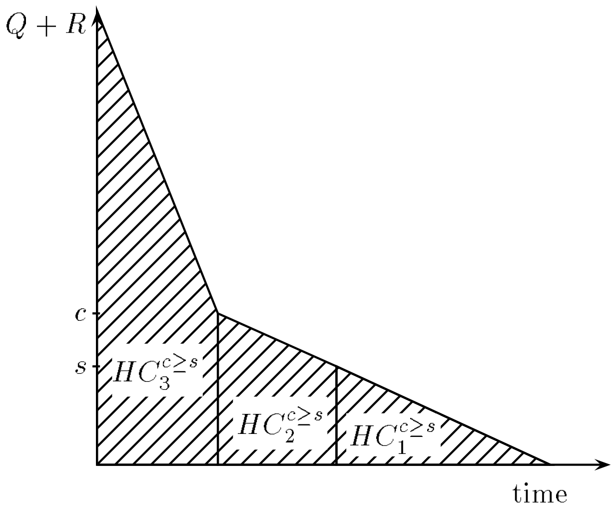

Figure 2: The inventory process with $c \geq s$.

$$\begin{aligned} HC_1^{c\geq s} &= h\mathbb{E}_{D_1(L)}\Bigg[\mathbf{1}_{\{D_1(L)<s\}} \sum_{i=s-D_1(L)}^{s} i \cdot \frac{L}{D_1(L)+1} \\ &\quad +\mathbf{1}_{\{D_1(L)\geq s\}} \sum_{i=1}^{s} i \cdot \frac{Ls/(D_1(L)+1)}{s}\Bigg] \end{aligned}$$

The total expected holding cost is now given by

$$\mathbb{E}[HC^{c\geq s}] = HC_1^{c\geq s} + HC_2^{c\geq s} + HC_3^{c\geq s}$$

It is easy to calculate the total expected holding cost (see Appendix 2).
The expected length of a cycle is

$$\mathbb{E}[LoC^{c\geq s}] = L + \frac{Q+\mathbb{E}[R]-c}{\lambda} + \frac{c-s}{\lambda_1}$$

Hence, the total cost of a (c, s, Q) policy with $c \geq s$ is given by

$$TC^{c\geq s}(c,s,Q) = \frac{\mathbb{E}[SC^{c\geq s}] + \mathbb{E}[HC^{c\geq s}] + K}{\mathbb{E}[LoC^{c\geq s}]}$$

We have now concluded the analytic derivations of the expected cost of the rationing policies. Using the results presented in Section 4.1 and 4.2, we obtain that the average cost $TC(c, s, Q)$ of a (c, s, Q) inventory policy is given by

$$TC(c,s,Q) = \begin{cases} TC^{c<s}(c,s,Q) & \text{if } c < s \\ TC^{c\geq s}(c,s,Q) & \text{if } c \geq s \end{cases}$$

5 Optimization

Due to the complexity of the average cost formula it has not been possible to derive an explicit expression for the optimal policy. The optimization procedure is therefore based on enumeration and bounding.

Assume that the order size Q is given, and denote the associated optimal values of c and s by $c^*(Q)$ and $s^*(Q)$. To obtain an upper bound on the value of $s^*(Q)$ we conjecture that $s^*(Q)$ will be less than or equal to the optimal reorder level for the model without a critical level (which is equivalent to our model with $c = 0$).

Although we cannot give a formal proof of this, we give an intuitive explanation. When the critical level is positive, the average demand rate during the lead time will decrease. Since the main purpose of using a reorder level is to cover the lead time demand, and the optimal reorder is increasing with the total lead time demand, we conjecture that the optimal reorder level $s^*(Q)$ is decreasing with the critical level c. We note that all numerical experiments supported this conjecture.

The inventory model without a critical level is identical to a simple lost sales (s, Q) model with demand rate $\lambda := \lambda_1 + \lambda_2$ and lost sales cost $\pi := (\lambda_1\pi_1 + \lambda_2\pi_2)/\lambda$. Instead of determining the optimal reorder point for this model we will use the reorder point following from the Hadley–Whitin heuristic which is an upper bound on $s^*(Q)$, as shown in the following lemma.

Lemma 1. *Let $\tilde{s}(Q)$ be the reorder point obtained by using the Hadley-Whitin heuristic and let $\bar{s}(Q)$ be the optimal reorder point in the simple lost sales model. For a fixed value of Q it follows that $\tilde{s}(Q) \geq \bar{s}(Q)$. Furthermore, $\tilde{s}(Q)$ is found as the solution to*

$$\tilde{s}(Q) = min \; \{s \geq 0 : \mathbb{P}(D(L) \geq s + 1) \leq \frac{h}{h + \pi\lambda/Q}\}$$

Proof. The average cost of the simple (s, Q) policy is given by (see Hadley & Whitin [7])

$$TC_0(s, Q) = \frac{[K\lambda/Q + h[(Q+1)/2 + s - \lambda L] + (h + \frac{\pi\lambda}{Q})\mathbb{E}[D(L) - s]^+]}{(Q + \mathbb{E}[D(L) - s]^+)/\lambda}$$

Since Q is fixed we can write this as

$$TC_0(s, Q) = \frac{f(s)}{Q/\lambda + g(s)} \text{ where } g(s) \geq 0$$

Following the heuristic we approximate the average cycle length by Q/λ, as if there were no stockouts, and obtain

$$\tilde{TC}_0(s, Q) = \frac{f(s)}{Q/\lambda}$$

The reorder level $\tilde{s}(Q)$ that minimizes $\tilde{TC}_0(s,Q)$ surely minimizes $f(s)$ too. Thus for any $y > 0$ we have

$$\begin{aligned} TC_0(\tilde{s}(Q)+y,Q) &\geq \frac{f(\tilde{s}(Q))}{Q/\lambda + g(\tilde{s}(Q))+y)} \\ &\geq \frac{f(\tilde{s}(Q))}{Q/\lambda + g(\tilde{s}(Q))} \\ &= TC_0(\tilde{s}(Q),Q) \end{aligned}$$

establishing the upper bound. Observe that the second inequality is a result of $g(s)$ decreasing. The reorder level $\tilde{s}(Q)$ is found as (see Hadley & Whitin [7])

$$\tilde{s}(Q) = \min\{s \geq 0 : \mathbb{P}(D(L) \geq s+1) \leq \frac{h}{h+\pi\lambda/Q}\}$$

□

By Lemma 1 and our previous conjecture, we obtain that $\tilde{s}(Q)$ is an upper bound on the optimal reorder level $s^*(Q)$. In our computational experiments we experienced that it is possible to end up in local minima when searching for $s^*(Q)$. Hence, we suggest enumeration over all values between 0 and $\tilde{s}(Q)$.
We also suggest enumeration to determine the optimal critical level. Given the reorder level s, we evaluate all critical levels between 0 and $s-1$ using the average cost function $TC^{c<s}(c,s,Q)$. Let c' be the value which gives the minimum cost, i.e.

$$c' = \text{argmin } \{TC^{c<s}(c,s,Q) : 0 \leq c < s\}$$

Regarding $c \geq s$ it can easily be proved that for fixed values of s and Q, the average cost function is either convex or concave in c, depending on the underlying model and the values of s and Q. The critical level that minimizes the average cost function for $c \geq s$ is denoted by

$$c'' = \text{argmin } \{TC^{c\geq s}(c,s,Q) : s \leq c < Q\}$$

If the average cost function is convex, c'' is found in the global minimum, which can be found explicitly. If the average cost function is concave, let $c'' = s$ if $TC^{c\geq s}(s,s,Q) < TC^{c\geq s}(Q-1,s,Q)$. Otherwise let $c'' = Q-1$. Finally, let the optimal critical level given the reorder level s and the order size Q be given by

$$c = \begin{cases} c' & \text{if } TC^{c<s}(c',s,Q) < TC^{c\geq s}(c'',s,Q) \\ c'' & \text{if } TC^{c<s}(c',s,Q) \geq TC^{c\geq s}(c'',s,Q) \end{cases}$$

In many practical situations, the order size Q is prespecified. However, if one also wants to determine the optimal value of Q, one can use a local search algorithm with the economic order quantity as a starting solution. Numerical experiments have indicated that the average cost function is unimodal in Q.

6 Numerical results

In this section we will investigate the performance of the rationing policy discussed in the previous chapters. As a performance measure we will use the cost reduction CR of using the optimal (c, s, Q) rationing policy compared to the best possible (s, Q) policy. Hence, CR is defined as

$$CR := \frac{\min\{TC_0(s,Q) : s \geq 0, Q > s\} - \min\{TC(c,s,Q) : Q > c \geq 0, Q > s \geq 0\}}{\min\{TC_0(s,Q) : s \geq 0, Q > s\}}$$

To determine $\min\{TC_0(s,Q) : s \geq 0, Q > s\}$ we used an enumeration approach similar to the one described in the previous section. Alternatively, one may use the method described in Johansen & Thorstenson [8].
From computational experiments it appeared that the critical level may influence the optimal reorder level s and replenishment order size Q in two ways: if $c < s$ the main effect of the critical level is a reduction of the optimal reorder level s, whereas if $c \geq s$ the main effect lies in the reduction of the optimal order size Q. In this section we will consider examples that lead to both types of rationing policies, and at the end of the section, try to describe what determines the structure of the optimal policy.

Example 1

In the first example, we consider an inventory system with the following characteristics: $L = 1$, $h = 1$, $K = 100$, $\lambda_1 = 1$, $\lambda_2 = 10$, $\pi_1 = 1000$, and $\pi_2 = 10$. In Table 1 we have calculated the optimal critical level policy and the optimal non-rationing (s, Q) policy for Example 1. Observe that all costs are average cost per time unit.

Policy	$(c, s, Q) = (2, 14, 48)$	$(s, Q) = (17, 48)$
Total cost	52.49	54.96
Holding cost	27.87	30.52
Shortage cost	2.09	1.55
Ordering cost	22.54	22.88
Cycle length	4.44	4.37
Number of Stockouts Class 1	0.000058	0.00032
Number of Stockouts Class 2	0.041	0.0032

Table 1: Comparison of the optimal (c, s, Q) policy and the optimal (s, Q) policy.

We see that a cost reduction of 4.5% is obtained when a critical level policy is applied. As expected, the effect of the rationing policy is a reduced reorder level, leading to a lower average holding cost. The average stockout cost increases, but not enough to compensate the decrease in the holding cost. The expected ordering cost decreases due to the increase in the expected length of a cycle. We can also observe that the rationing policy has a dramatic effect on the expected number of stockouts for demand class 2, which increases by 1200%, whereas the number of stockouts for demand class 1 is reduced with 80%.

We have performed some variations of this example to show how the optimal policy is influenced by changes in the parameter values, and to investigate under which conditions the gain of rationing is most significant. The results are presented in Table 2.

π_1	100	500	1000	5000	10000	100000
(c,s,Q)	(1,12,49)	(1,14,48)	(2,14,48)	(2,15,48)	(3,15,48)	(4,16,48)
(s,Q)	(13,49)	(16,48)	(17,48)	(19,49)	(20,49)	(23,49)
CR	0.0062	0.0305	0.0449	0.0700	0.0819	0.1086
π_2	1	5	10	25	50	100
(c,s,Q)	(16,4,24)	(3,11,48)	(2,14,48)	(1,16,48)	(1,17,48)	(0,18,48)
(s,Q)	(17,48)	(17,48)	(17,48)	(17,49)	(17,49)	(18,48)
CR	0.4973	0.0698	0.0449	0.0217	0.0106	0.0000
K	25	50	100	200	500	1000
(c,s,Q)	(2,15,25)	(2,15,35)	(2,14,48)	(2,13,67)	(2,11,106)	(24, 3,132)
(s,Q)	(18,25)	(18,34)	(17,48)	(16,68)	(15,107)	(15,150)
CR	0.0625	0.0548	0.0449	0.0374	0.0296	0.1268

Table 2: Cost reductions for variations of Example 1.

In Table 2 we see what happens when we change the stockout cost of demand class 1. For small values of π_1 the cost reduction is negligible, whereas for larger values of π_1 the cost reduction is quite significant. The opposite is true if we change the value of π_2. Hence, it seems reasonable to conclude that the greater the difference between π_1 and π_2, the greater the cost reduction obtained by applying a critical level policy. When $\pi_2 = 1$ an interesting phenomenon occurs, i.e. the structure of the policy changes. From Table 2 one can observe that for $\pi_2 \geq 5$ the optimal policies satisfy $c < s$, whereas for $\pi_2 = 1$ we obtain an optimal critical level policy with $c \geq s$. A similar observation can be made with respect to the fixed order cost K. If we increase K the cost reduction decreases because the average ordering cost constitutes a larger part of the total cost. Moreover, the optimal order size Q increases and the reorder level s decreases. For all $K \leq 500$ the optimal critical level is equal to 2. However, for $K = 1000$ the structure of the optimal policy changes and we get an optimal critical level of $c = 24$.
In Figure 3 we study in more detail the change in structure of the policy, or 'bang–bang'

effect, for varying values of π_2 and K. We have calculated the optimal policy with respect to two different restrictions. The solid lines represent policies where $c \geq s$ and the dashed lines policies with $c < s$. For small values of K the restriction $c \geq s$ leads to policies

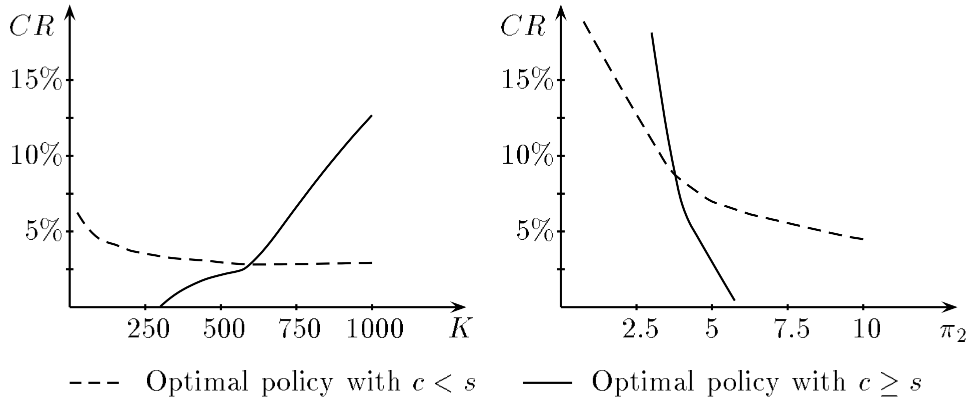

Figure 3: Cost reductions for policies with $c < s$ and policies with $c \geq s$ (Example 1).

that perform worse than the optimal (s, Q) policy. But as K increases, the optimal order size is increased too, and the cost of carrying inventory increases. When the fixed order cost exceeds a certain value (approximately 600 for this example) the cost of carrying inventory gets so high that it is optimal to reject some low–priority demand in order to reduce demand and thereby the holding cost. Another 'bang–bang' effect is found when we change the cost of rejecting low–priority demand. As seen in Table 2 the cost reduction increase as π_2 decrease, and Figure 3 illustrates that as the cost of rejecting low–priority demand gets low enough the cost reduction of using policies with $c \geq s$ increases rapidly, and the structure of the optimal policy changes. In Example 2 we will discuss policies with $c \geq s$ in more detail.

To investigate the effect of changing the demand rates we have calculated the cost reduction for 400 combinations of λ_1 and λ_2 with all other parameters fixed (see Figure 4). The classification is chosen in order to equalize the size of the areas. No rigorous conclusion is possible, but it seems clear that the cost reduction is strongly connected to the demand ratio λ_1/λ_2. If this ratio is greater than one, i.e. most demand is considered to be of high priority, the cost reduction is very small. This is also the case if the demand ratio is smaller than 1/20. In between it appears that the largest cost reduction is obtained for demand ratios between 1/4 and 1/10. This observation differs from the observations made by Ha [6], who concluded that the greatest cost reduction, in an $(S-1, S)$ model, is obtained for demand ratios around one.

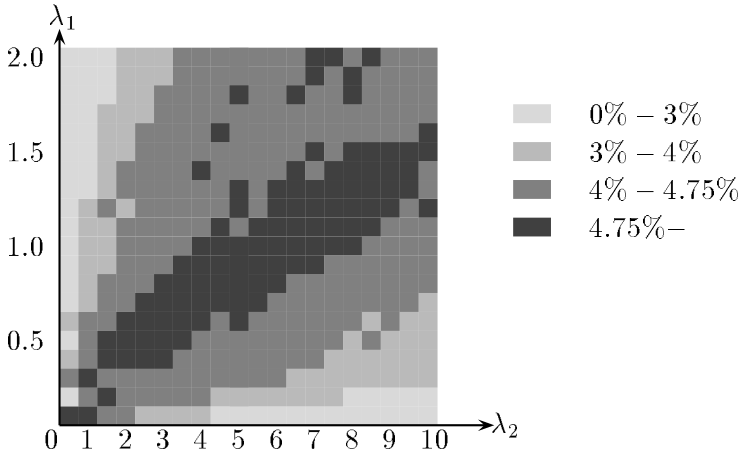

Figure 4: Cost reduction for Example 1 obtained by rationing for different values of λ_1 and λ_2.

Example 2

In this example, we will turn our attention to rationing policies with $c \geq s$. There is no way to guarantee that the optimal rationing policy has $c \geq s$. However, in order to favour the policies with $c \geq s$ we can increase the fixed order cost and the unit holding cost, or lower the stockout cost of demand class 2. For the second example the following parameter values are used: $L = 1$, $h = 2$, $K = 200$, $\lambda_1 = 1$, $\lambda_2 = 5$, $\pi_1 = 500$, and $\pi_2 = 6$. The optimal critical level policy, the optimal (s, Q) policy, and the corresponding costs for Example 2 are reported in Table 3.Again, all costs are average cost per time unit.

Policy	$(c, s, Q) = (12, 3, 28)$	$(s, Q) = (9, 36)$
Total cost	60.76	78.68
Holding cost	21.41	43.13
Shortage cost	23.97	2.36
Ordering cost	15.38	33.18
Cycle length	13.00	6.03
Number of stockouts class 1	0.0018	0.0045
Number of stockouts class 2	3.85	0.13

Table 3: Comparison of the optimal (c, s, Q) policy and the optimal (s, Q) policy.

We see that a considerable cost reduction of 22.8% is obtained. It is very interesting to see that the cost allocation in the optimal critical level policy is very different from the allocation in the standard (s, Q) policy. The average holding and ordering cost are both

reduced with about 50% while the average shortage cost has increased with a factor 10. This is caused by the fact that the expected cycle length is doubled because we reject on average 3.85 demands from demand class 2 per time unit. The expected holding and ordering cost per cycle is more or less unchanged, so the reduction per time unit is mainly due to the longer expected cycle length. The expected number of stockouts per cycle for demand from demand class 1 hardly changes. However, since the cycle length is doubled, the average stockout cost per time unit for demand class 1 is reduced.
Also for Example 2 we have analysed the effect of variations in the parameter values on the optimal policy and the cost reduction. The results are shown in Table 4.

π_1	40	100	500	1000	5000	10000
(c, s, Q)	(10, 0,28)	(11, 1,28)	(12, 3,28)	(12, 3,28)	(13, 4,28)	(14, 5,28)
(s, Q)	(0,35)	(5,36)	(9,36)	(10,36)	(12,36)	(13,35)
CR	0.1896	0.2046	0.2278	0.2344	0.2470	0.2470
π_2	2	4	6	10	25	50
(c, s, Q)	(15, 3,19)	(14, 3,24)	(12, 3,28)	(6, 3,34)	(1, 8,36)	(1, 9,36)
(s, Q)	(9,36)	(9,36)	(9,36)	(9,36)	(9,36)	(9,36)
CR	0.4439	0.3307	0.2278	0.0664	0.0203	0.0073
K	50	100	200	400	1000	1500
(c, s, Q)	(2,7,18)	(3, 7,23)	(12, 3,28)	(18, 3,34)	(30, 2,47)	(37, 2,54)
(s, Q)	(10,18)	(9,26)	(9,36)	(8,50)	(7,79)	(7,96)
CR	0.0978	0.1364	0.2278	0.3108	0.4038	0.4353

Table 4: Cost reduction for variations of Example 2.

By changing the value of the parameters π_1, π_2 and K, we see that high cost reductions can be obtained by applying a critical level policy. The cost reduction CR increases when π_1 increases, but the optimal policies remain more or less the same. As π_2 gets very small, the advantage of using the rationing policy increases rapidly. Note that while the optimal (s, Q) policy does not change at all, the (c, s, Q) policy is sensitive to changes in π_2. The 'bang–bang' effect that occurs is similar to the one observed in Example 1 (see Figure 3). If the fixed order cost K increases, the optimal replenishment order size Q increases as well, both for the (s, Q) policy and the (c, s, Q) policy. However, for the latter policy this increase is limited due to the increasing level of c, thus part of the holding cost is replaced by additional stockout cost.
In Figure 5 we observe that there is no clear relation between the cost reduction CR and the demand ratio λ_1/λ_2, as was the case for Example 1. It seems like CR decreases as

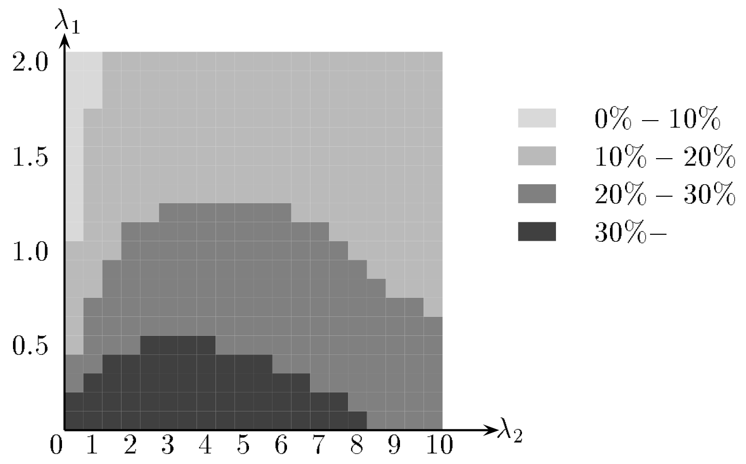

Figure 5: Cost reduction for Example 2 obtained by rationing for different values of λ_1 and λ_2.

λ_1 increases. When the share of high–priority demand in the total demand increases, the influence of the rationing policy declines, which explains the dependence of the cost reduction with λ_1. The dependence with λ_2 is more complicated. The largest cost reductions are found for values of λ_2 between 2 and 6. It is obvious that when the demand rate approaches zero there will be no gain of rationing. On the other hand, if the demand rate gets very high, the cost of rejecting demand will increase so that it is not profitable to exchange holding cost for stockout cost. Finally, we observe that for all combinations of λ_1 and λ_2, the cost reduction is substantial.

We will conclude this section by investigating what determines the structure of the optimal policy. We have previously seen that the values of the fixed order cost and the stockout cost for demand class 2 have great influence on the structure of the optimal policy. To obtain further insight we have determined the structure of the optimal policy for a number of different parameter values. The effect of changing the demand rate or the stockout cost of demand class 1 turned out to be negligible. More interesting is the effect of the parameter values connected with demand class 2 with respect to the structure of the optimal policy. Figure 6 shows how the structure of the optimal policy depends on the values of K, λ_2 and π_2. Observe that the area above the curve corresponds to optimal policies satisfying $c \geq s$, whereas the area below is associated with optimal policies satisfying $c < s$.
The effect of K and π_2 on the structure of the optimal policy are as expected: if K is small and π_2 is large then the optimal policy will satisfy $c < s$. We also see that the demand rate of demand class 2 significantly influences the structure of the optimal policy. If the demand rate for class 2 is relatively low, then the optimal policy is more likely to satisfy $c \geq s$. This can be explained by observing that a policy with $c \geq s$ implies that most

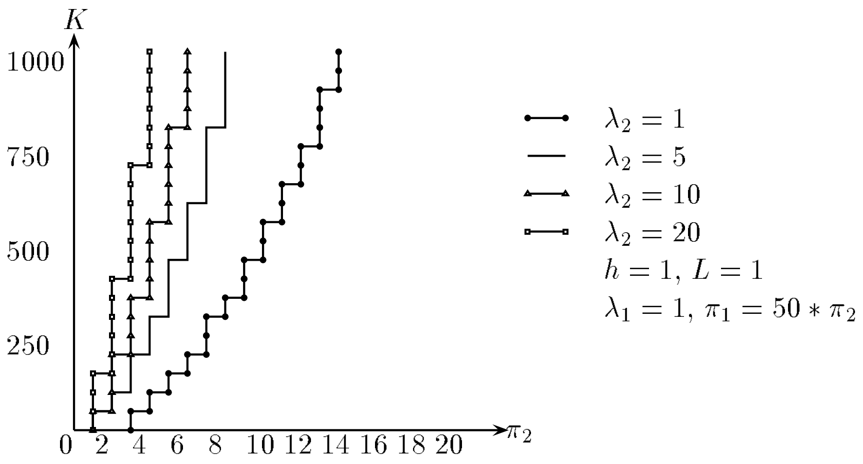

Figure 6: The structure of the optimal policy.

demand from class 2 is lost which does not lead to high lost sales cost if λ_2 is small.

7 Conclusions and further research

In this paper we have discussed an (s, Q) inventory model with lost sales and two demand classes. We have introduced a lead time independent rationing policy, i.e. the (c, s, Q) policy. This so–called critical level policy reserves part of the inventory for high–priority demand, i.e. if the inventory level is at or below the critical level c, low–priority demand is rejected in anticipation of future high–priority demand. We have derived an exact expression for the average cost of (c, s, Q) policies, that satisfy $0 \leq c < s < Q$ or $0 \leq s \leq c < Q$. We have shown that this rationing policy can have two different effects on the optimal reorder level and replenishment order size, depending on whether the critical level is below or above the reorder level. In the first case, the critical level policy reduces the safety stock needed. Significant cost reductions can be obtained if the stockout cost of high–priority demand is considerably larger than the stockout cost of low–priority demand. If the critical level is at or above the reorder point, then in general the rationing policy will reduce the average holding cost, by rejecting a great part of the low–priority demand. This is in particular advantageous if the cost of rejecting low–priority demand is small (compared to the holding cost rate) or if the fixed order cost is high.

Although the lead time independent (c, s, Q) policy is easy to understand and implement in practice, it may be cost effective to consider a lead time dependent policy. If the inventory level is below the critical level, and a low–priority customer arrives, it may be optimal to

deliver this demand anyway if the inventory manager knows that a replenishment order will arrive soon. It would be interesting to compare a (c,s,Q) policy with a lead time dependent policy and see how much the average cost can be reduced.
Another possible extension of our model is to consider more than two demand classes. However, if we have n demand classes we would require a different approach to calculate the exact cost, because conditioning on $n-1$ hitting times seems computationally cumbersome.

Appendix 1: Derivation of average cost for $c < s$

To simplify the notation let

$$\begin{aligned} p_j(i) &= \mathbb{P}(D_j(L) = i) \text{ for } j = 1,2 \text{ and } i = 0,1,2,\ldots \\ p(i) &= \mathbb{P}(D(L) = i) \text{ for } i = 0,1,2,\ldots \end{aligned}$$

Moreover let $f_H(t)$ denote the pdf of the hitting time H. It is easy to find an expression for $B_2^{c<s}$, i.e.

$$\begin{aligned} B_2^{c<s} &= \mathbb{E}_{D_2(L-H)}[D_2(L-H)] \\ &= \int_0^L f_H(t)\lambda_2(L-t)dt \\ &= \int_0^L e^{-\lambda t}\lambda^{s-c}\frac{t^{s-c-1}}{(s-c-1)!}(L-t)\lambda_2\, dt \\ &= \frac{\lambda_2}{\lambda}(\lambda L - s + c)\Big[1 - \sum_{j=0}^{s-c-1} p(j)\Big] + \frac{\lambda_2}{\lambda}\frac{e^{-L\lambda}(\lambda L)^{s-c}}{(s-c-1)!} \end{aligned}$$

which is equivalent to the result obtained by Nahmias & Demmy [11].
To find $B_1^{c<s}$ we need the distribution of $D_1(L-H)$. By using the binomial expansion for $(L-t)^i$ we obtain

$$\begin{aligned} \mathbb{P}(D_1(L-H) = i) &= \int_0^L f_H(t) \cdot P(D_1(L-t) = i)dt \\ &= \int_0^L e^{-\lambda t}\lambda^{s-c}\frac{t^{s-c-1}}{(s-c-1)!}e^{-\lambda_1(L-t)}\lambda_1^i\frac{(L-t)^i}{i!}dt \\ &= \sum_{k=0}^{i}\frac{e^{-\lambda_1 L}L^{i-k}(-1)^k}{(i-k)!k!}\frac{(k+s-c-1)!}{(s-c-1)!}\frac{\lambda^{s-c}\lambda_1^i}{\lambda_2^{k+s-c}}\Big[1 - \sum_{j=0}^{k+s-c-1} p_2(j)\Big] \\ &= \sum_{k=0}^{i}(-1)^k A(i,k)\Big[1 - \sum_{j=0}^{k+s-c-1} p_2(j)\Big] \end{aligned}$$

with

$$
\begin{aligned}
A(i,0) &:= \frac{e^{-\lambda_1 L} L^i \lambda^{s-c} \lambda_1^i}{i! \lambda_2^{s-c}} \\
A(i,k) &:= A(k-1) \cdot \frac{(k+s-c-1)\cdot(i-k+1)}{L \cdot k \cdot \lambda_2}
\end{aligned} \tag{7}
$$

Hence, we find

$$
\begin{aligned}
B_1^{c<s} &= \mathbb{E}_{D_1(L-H)}[D_1(L-H)-c]^+ \\
&= \sum_{i=c+1}^{\infty} \mathbb{P}(D_1(L-H)=i)\ (i-c)
\end{aligned}
$$

Nahmias & Demmy [11] suggest that the integral is solved using numerical integration. However, this is a slow procedure whereas the expression developed above is exact and fast.

In Section 4.1 the holding cost incurred during the lead time was found as a function of the random variables H, $D(L)$ and $D_1(L-H)$. We will now find the expected holding cost $HC_1^{c<s}$ incurred during the lead time, by conditioning on these variables. For $D(L) < s-c$ we apply (3), for $D(L) \geq s-c$ (which is equivalent with $0 < H < L$) we apply (4) and either (5) or (6).

$$
\begin{aligned}
HC1^{c<s} &= h \sum_{j=0}^{s-c-1} p(j) \left(\sum_{i=s-j}^{s} i \cdot \frac{L}{j+1} \right) \\
&\quad + h \int_0^L f_H(t) \left[\sum_{i=c+1}^{s} i \cdot \frac{t}{s-c} \right. \\
&\quad + \sum_{j=0}^{c} \mathbb{P}(D_1(L-t)=j) \left(\sum_{i=c-j}^{c} i \cdot \frac{L-t}{j+1} \right) \\
&\quad \left. + \sum_{j=c+1}^{\infty} \mathbb{P}(D_1(L-t)=j) \left(\sum_{i=1}^{c} i \cdot \frac{(L-t)c/(j+1)}{c} \right) \right] dt
\end{aligned}
$$

Exploiting the properties of $\sum_{i=a}^{b} i$ yields

$$\begin{aligned}
HC_1^{c<s} &= h\sum_{j=0}^{s-c-1} p(j)(s-j/2)L + h\int_0^L f_H(t)\frac{c+1+s}{2}t\ dt \\
&\quad +h\int_0^L f_H(t)\sum_{i=0}^{c}\mathbb{P}(D_1(L-t)=i)(c-i/2)(L-t)dt \\
&\quad +h\int_0^L f_H(t)\sum_{i=c+1}^{\infty}\mathbb{P}(D_1(L-t)=i)\frac{c+1}{2}\cdot(L-t)\frac{c}{i+1}dt \\
&= h\sum_{j=0}^{s-c-1} p(j)(s-j/2)L \\
&\quad +h\frac{s^2-c^2+s-c}{2\lambda}\Big[1-\sum_{i=0}^{s-c}p(i)\Big] \\
&\quad +h\sum_{j=0}^{c}(c-j/2)\sum_{k=0}^{j+1}(-1)^k B(k,j)\Big[1-\sum_{i=0}^{s-c-1+k}p_2(i)\Big] \\
&\quad +h\sum_{j=c+1}^{\infty}\frac{c+1}{2}\frac{c}{j+1}\sum_{k=0}^{j+1}(-1)^k B(k,j)\Big[1-\sum_{i=0}^{s-c-1+k}p_2(i)\Big]
\end{aligned}$$

with

$$\begin{aligned}
B(j,0) &:= e^{-\lambda_1 L}\frac{L^{j+1}\lambda^{s-c}\lambda_1^j}{j!\lambda_2^{s-c}} \qquad (8)\\
B(j,k) &:= B(j,k-1)\cdot\frac{(s-c-1+k)\ \cdot(j+2-k)}{L\cdot\lambda_2\cdot k}
\end{aligned}$$

The expected holding cost incurred in the remaining part of the cycle is easily found, i.e.

$$\begin{aligned}
HC_2^{c<s} &= h\mathbb{E}_R\Big[\sum_{i=s+1}^{Q+R} i\cdot\frac{1}{\lambda}\Big] \\
&= h\mathbb{E}_R\Big[\frac{Q+R+s+1}{2}\cdot\frac{Q+R-s}{\lambda}\Big] \\
&= h\frac{Q^2+2Q\mathbb{E}[R]+\mathbb{E}[R^2]+Q+\mathbb{E}[R]-s-s^2}{2\lambda}
\end{aligned}$$

The first two moments of the random variable R are easily found from (1). The total expected holding cost is

$$\mathbb{E}[HC^{c<s}] = HC_1^{c<s} + HC_2^{c<s}$$

The expressions developed in this appendix are valid for all combinations of parameters. However during implementation, numerical problems can arise when evaluating $\mathbb{P}(D_1(L-H))$ and $H_1^{c<s}$. If $\lambda_2 < \lambda_1$ the terms in(7) and (8) gets very big, causing representation problems, and the integrals should be solved using numerical integration.

Appendix 2: Derivation of the expected holding cost for $c \geq s$

The expected holding cost is divided in three parts according to Figure 2.

$$
\begin{aligned}
HC_3^{c\geq s} &= h\mathbb{E}_R\Big[\sum_{i=c+1}^{Q+R} i\cdot\frac{1}{\lambda}\Big] \\
&= h\frac{Q^2+2QE[R]+E[R^2]+Q+E[R]-c-c^2}{2\lambda} \\
HC_2^{c\geq s} &= h\sum_{i=s+1}^{c} i\cdot\frac{1}{\lambda_1} \\
&= h\frac{c+s+1}{2}\frac{c-s}{\lambda_1}
\end{aligned}
$$

For $HC_1^{c\geq s}$ we condition on whether the inventory is depleted or not.

$$
\begin{aligned}
HC_1^{c\geq s} &= h\mathbb{E}_{D_1(L)}\left[\mathbf{1}_{\{D_1(L)<s\}}\sum_{i=s-D_1(L)}^{s} i\cdot\frac{L}{D_1(L)+1}\right. \\
&\qquad \left. +\mathbf{1}_{\{D_1(L)<s\}}\sum_{i=1}^{s} i\cdot\frac{Ls/(D_1(L)+1)}{s}\right] \\
&= h\sum_{i=0}^{s-1}p_1(i)L(s-i/2)+h\sum_{i=s}^{\infty}p_1(i)\frac{s+1}{i+1}\frac{s}{2}L
\end{aligned}
$$

Hence, we have

$$
\mathbb{E}[HC^{c\geq s}] = HC_1^{c\geq s}+HC_2^{c\geq s}+HC_3^{c\geq s}
$$

References

[1] C. Chiang and G. J. Gutierrez. A periodic review inventory system with two supply modes. *European Journal of Operational Research*, 94(3):527–547, 1996.

[2] M.A. Cohen, P.R. Kleindorfer, and H.L. Lee. Service constrained (s, S) inventory systems with priority demand classes and lost sales. *Management Science*, 34:482–499, 1988.

[3] R. Dekker, R.M. Hill, and M.J. Kleijn. On the $(S-1, S)$ lost sales inventory model with priority demand classes. Technical Report 9743/A, Econometric Institute, Erasmus University Rotterdam, The Netherlands, 1997.

[4] R. Dekker, M.J. Kleijn, and P.J. De Rooij. A spare parts stocking system based on equipment criticality. *to appear in the International Journal of Production Economics*, 1996.

[5] R.V. Evans. Sales and restocking policies in a single item inventory system. *Management Science*, 14:463–472, 1968.

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

[7] G. Hadley and T.M. Whitin. *Analysis of Inventory Systems.* Prentice–Hall, Englewood Cliffs, NJ, 1963.

[8] S.G. Johansen and A. Thorstenson. Optimal and approximate (Q, r) inventory policies with lost sales and gamma-distributed lead time. *International Journal of Production Economies*, 30–31:179–194, 1993.

[9] A. Kaplan. Stock rationing. *Management Science*, 15:260–267, 1969.

[10] I. Moon and S. Kang. Rationing policies for some inventory systems. *Journal of the Operational Research Society*, 49:509–518, 1998.

[11] S. Nahmias and S. Demmy. Operating characteristics of an inventory system with rationing. *Management Science*, 27:1236–1245, 1981.

[12] E.A. Silver and R. Peterson. *Decision Systems for Inventory Management and Production Planning.* John Wiley & Sons, New York, 1985.

[13] R.J. Tersine. *Principles of Inventory and Materials Management.* North-Holland, New York, 1988.

[14] H.C. Tijms. *Stochastic Models: An Algorithmic Approach.* Wiley, New York, 1994.

[15] D.M. Topkis. Optimal ordering and rationing policies in a nonstationary dynamic inventory model with n demand classes. *Management Science*, 15:160–176, 1968.

[16] A.F. Veinott Jr. Optimal policy in a dynamic, single product, nonstationary inventory model with several demand classes. *Operations Research*, 13:761–778, 1965.