ELSEVIER

Int. J. Production Economics 72 (2001) 41–58

international journal of production economics

www.elsevier.com/locate/dsw

An inventory system with two supply modes and capacity constraints

Dimitrios Vlachos, George Tagaras*

Department of Mechanical Engineering, Aristoteles University of Thessaloniki, 54006 Thessaloniki, Greece

Received 6 September 1999; accepted 20 June 2000

Abstract

This paper analyses a periodic review inventory system with a main and an emergency supply mode, where policies of the base-stock type are used at both supply channels. Contrary to previously published models, the capacity of the emergency channel is taken into account. We examine two alternative ordering policies for that channel: an "early-ordering" policy that almost eliminates early stockouts in a replenishment cycle and a "late-ordering" policy that delays the emergency order decision until more demand information has been accumulated. Approximate cost models are developed and properties of their optimal solutions are derived. Simulation results indicate that these solutions are near optimal. Conclusions are drawn about the relative effectiveness of the two emergency ordering policies. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Inventory; Periodic review; Multiple supply modes; Capacity constraints; Reverse logistics

1. Introduction

Inventory systems with two supply modes become increasingly common in practice. One obvious case is when multiple suppliers are used on a regular basis and the total order quantity is divided into smaller orders to these suppliers. Another case is when there exists a main supply mode, used regularly for stock replenishment, as well as a secondary supply mode, used on exceptional occasions to supplement the stock supply. Inventory replenishment through the latter supply mode is often called emergency replenishment and this option is typically utilised when a stockout is very likely to occur at the receiving location. Emergency replenishments are characterised by low volumes, shorter lead times and higher acquisition costs per unit. The higher cost is usually a consequence of the increased transportation cost, due to the faster but more expensive means of transportation employed for emergency replenishments. This paper focuses on exactly this case of periodic review inventory systems with a primary supply channel used for regular orders and a secondary supply channel used for more expensive but faster emergency orders.

Previous research on inventory systems with regular and emergency replenishments includes the early papers of Barankin [1], Daniel [2] and Neuts [3], who study periodic review inventory systems with regular order lead time of one period and

* Corresponding author. Tel.: + 30-31-996062; fax: + 30-31-999018.

E-mail address: tagaras@ccf.auth.gr (G. Tagaras).

PII: S0925-5273(00)00083-9

instantaneous emergency replenishments. Bulinskaya [4], Fukuda [5] and Veinott [6] allow for longer lead times but always differing by one period. Whittemore and Saunders [7] derive the optimal policy for regular and emergency lead times that can be any multiple of the review period. Rosenshine and Obee [8] examine a standing order inventory system where a regular order of constant size is received every period and an emergency order of fixed size may be placed once per period and arrives immediately. Gross and Soriano [9] and Chiang and Gutierrez [10] analyse a different periodic review inventory system with two supply modes, where at each review epoch a decision is made about which of the two supply modes to use. The work of Chiang and Gutierrez [10] is the first one that considers lead times that can be shorter than the review period. In a sequel paper, Chiang and Gutierrez [11] allow multiple emergency orders to be placed at any (discrete) time within a review period, including the time of the regular order. Tagaras and Vlachos [12] also study a dual supply mode periodic review system where lead times can be shorter than the review period, but they allow at most one emergency order per cycle. Continuous review inventory models with emergency orders have been studied by Moinzadeh and Nahmias [13], Moinzadeh and Schmidt [14] and recently by Johansen and Thorstenson [15].

The motivation behind this paper, which constitutes a further examination of periodic review inventory systems with two supply modes, is twofold. Firstly, in all the previous models the capacity of both replenishment channels has been implicitly or explicitly assumed to be practically infinite, although this may not necessarily be always true in practice, especially for the emergency replenishment channel that must respond with short notice. Secondly, the practical relevance of models for inventory systems with two supply modes has been enhanced lately by the increasing role and importance of reverse logistics networks (see [16] for an overview). Specifically, consider a market for remanufactured items (automotive engines, gearboxes, tyres) where the demand for such items is realised and satisfied at a stocking location that is periodically supplied by the remanufacturing unit. When a costly stockout is very likely to occur, the stocking location may decide to satisfy its customers by replenishing its inventory with new and more expensive items, which are readily available and will arrive faster than the next regular order or even an emergency order from the possibly distant remanufacturer. Dekker et al. [16] report such a situation at Volkswagen, where newly manufactured parts are sold at the price of a remanufactured part. In the case of tyres, the shorter lead time of the emergency order for new tyres may be a consequence of the proximity of a retailer or a local warehouse. At the same time, the available stock (capacity) of that retailer or warehouse may not be sufficiently large to completely satisfy the emergency order.

With the above points in mind we set out to study a periodic review inventory system with regular and emergency replenishments, with a capacity constraint on the size of emergency replenishments. Our intention is to propose operating policies that are cost-effective, yet simple enough to be widely applicable both in conventional and reverse logistics networks. Section 2 presents a detailed description of the system under consideration. Section 3 contains the formulation of approximate expected cost models for the systems under consideration. Optimisation procedures for those models and properties of their optimal solutions are derived and presented in Section 4. Numerical results and comparisons are provided in Section 5, along with validation of the approximate models by simulation. The final section presents a summary and the main conclusions of this paper.

2. System description and assumptions

We consider a stocking location using an inventory system with two supply modes for managing the inventory of a single item. Regular replenishment orders are placed periodically through the primary supply channel, following a base stock policy; at each review instance, the size of the regular order is such that the inventory position is raised up to S (a decision variable). The review period, P, is fixed and it is assumed to be determined by considerations exogenous to our model, such as the need for co-ordinating the regular replenishments with those of other items. The orders arrive after a constant regular replenishment lead time, denoted by L. Both P and L are expressed as integer multiples of a suitably chosen time unit. The times of arrival of two consecutive regular orders define a replenishment cycle of length P time units.

At some point in the replenishment cycle, which depends on the emergency ordering policy and will be specified accordingly below, the stocking location may place an emergency order through the secondary supply channel. This order arrives after the emergency replenishment lead time L_e, which is also assumed to be constant and an integer multiple of the time unit, but shorter than the regular lead time and the review period.

The demand in a time unit is assumed to be a continuous non-negative random variable, with density function $g(y)$ and distribution function $G(y)$. The mean and the standard deviation of the demand per time unit are denoted by μ and σ, respectively. It is also assumed that demand is independently distributed in disjoint time intervals. Any demand that is not immediately satisfied is backordered and filled when a new regular or emergency order arrives at the stocking location. More specifically, the sequence of events at the beginning of each time unit is as follows: possible arrival of an order (regular or emergency), then materialisation of new demand and after that possible placement of a new order (regular or emergency).

To completely specify the operation of the entire system, the exact form of the ordering policy for the emergency supply channel needs to be determined. As in a typical inventory system, there are two issues to be addressed: *when* to place the order (timing) and *how much* to order (size). There are obviously many ways to answer these questions, leading to a large variety of ordering policies. However, the purpose of this research is to propose and study policies that are intuitively appealing and easy to implement in practice. Therefore, we restrict our attention to two such policies that differ in the timing of orders but both of which are of the familiar base stock type.

Before proceeding with the description and analysis of the two emergency ordering policies, it is helpful to closely examine a typical replenishment cycle. Let the P time units of this time interval be numbered 1–P, where time unit 1 represents the time unit at the beginning of which a regular order arrives and can be used to satisfy the accumulated backorders (if they exist) and new demand of the time unit. Then, time unit P stands for the last time unit before the arrival of the next regular order. The likelihood of a stockout obviously increases as the end of the replenishment cycle approaches and it is highest at time unit P. When emergency replenishment is possible, the reason for placing an emergency order is to reduce the risk or the size of a stockout exactly at that high-risk part of the cycle. Given that only one emergency order may be placed per cycle, its timing has to be decided taking into account, in addition to its lead time and cost, the following trade-off: the advantage of placing the order early in the cycle (e.g., at time $t < P - L_e$) is that "early" stockouts (e.g., in time unit $t + L_e$) are almost eliminated; on the other hand, the advantage of placing the order later in the cycle (but not later than time $t = P - L_e$) is that the size of this order is specified with more information about the stock level close to the high-risk end of the cycle.

With the above in mind, the two emergency ordering policies that we propose and examine in this paper will be called "early-ordering" and "late-ordering" policy, respectively. To simplify the analysis we assume that stockouts are non-negligible only in the last two time units of the replenishment cycle (Tagaras and Vlachos [12] show that this assumption provides a satisfactory approximation in systems with low to moderate demand variability, not too long review periods and lead times and large service levels). Therefore, we differentiate between the two policies as follows: in the *late-ordering* policy the emergency order is placed as late as possible in the cycle, i.e., at time $P - L_e$, so that it arrives at the beginning of the last, critical time unit P; in the *early-ordering* policy the emergency order is placed at time $P - L_e - 1$, so that it arrives at the beginning of time unit $P - 1$. Figs. 1 and 2 illustrate the sequence of events (placements and arrivals of regular and emergency orders) under the *late-ordering* and *early-ordering* policy, respectively.

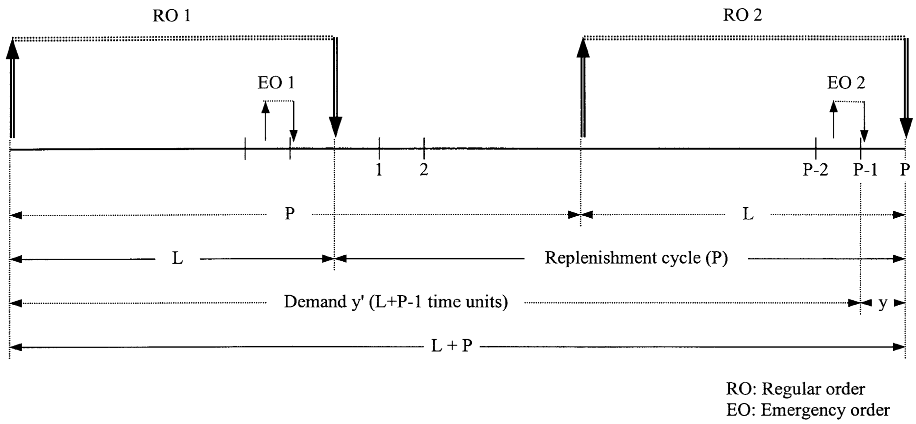

Fig. 1. Sequence of events in a typical time interval under the *late-ordering* policy.

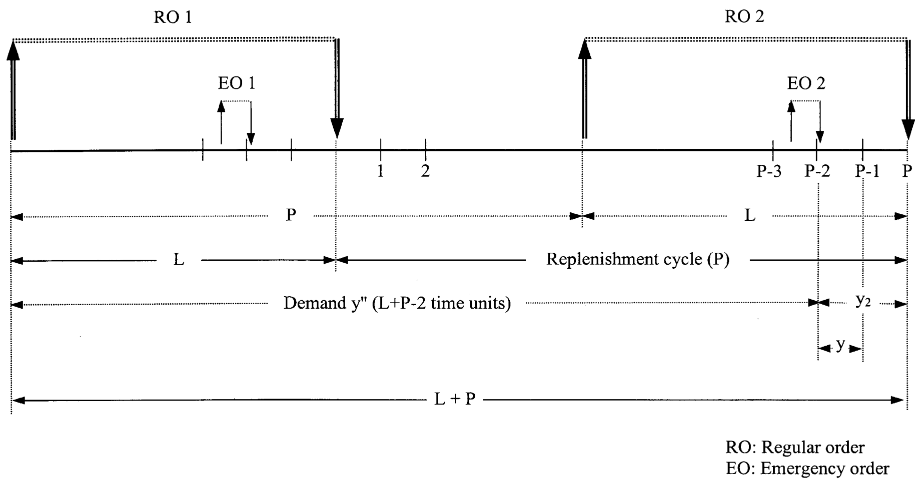

Fig. 2. Sequence of events in a typical time interval under the *early-ordering* policy.

Having determined the timing of an emergency order, the next and final step is to determine the size of that order, if such an order is needed at all. The proposed rule is similar in two cases (*late-ordering* and *early-ordering*) and is characterised by a single nonnegative parameter r (a decision variable), which represents the desired level of net stock at the time that the emergency ordering decision is made. The emergency order quantity, Q_e, is determined at time $P - L_e$ (*late-ordering*) or $P - L_e - 1$ (*early-ordering*) by comparing the net stock at that time (NS_{P-L_e} or NS_{P-L_e-1}, respectively) with the desired net stock, r, under the constraint that Q_e cannot exceed the capacity K of the emergency channel. The specific operating rule takes the following form in the two cases:

late-ordering policy:

$$Q_e = \min\{(r - NS_{P-L_e})^+, K\};$$

early-ordering policy:

$$Q_e = \min\{(r - NS_{P-L_e-1})^+, K\},$$

where $(x)^+ = \max\{0, x\}$.

There are two reasons for proposing the above rule. The first one is its simplicity and its affinity with the widely applicable in practice base stock policy. The second reason is that its form, barring the effect of the capacity constraint, is analogous to that of the rule that Chiang and Gutierrez [11] proved to be optimal in a similar context with multiple emergency orders per cycle.

The objective, which should eventually determine the preferable form of the emergency ordering policy, is to minimise the total relevant costs of the system, namely the regular and emergency acquisition costs, holding and backorder (penalty) costs. A regular order is placed at every review instance and consequently its fixed cost is not relevant in our context. Since all demand is eventually satisfied, the variable cost of a regular order is also irrelevant. The fixed cost of an emergency order is assumed to be negligible. The incremental variable cost of an emergency order (over that of a regular order) is denoted by c_e per unit. Holding cost is charged at a rate of c_h per unit and time unit, while penalty cost is charged at a rate of c_p per unit backordered and time unit. Thus, the general form of the expected cost per replenishment cycle of P time units for a given regular and emergency ordering policy with parameters S and r, is

$$C_P = c_h \sum_{i=1}^{P} E(\mathrm{OH}_i) + c_p \sum_{i=1}^{P} E(\mathrm{BO}_i) + c_e E(Q_e), \quad (1)$$

where $\mathrm{OH}_i = (\mathrm{NS}_i)^+$ and $\mathrm{BO}_i = (-\mathrm{NS}_i)^+$ denote the ending on hand inventory and backorder level, respectively, at time unit i of the cycle and $E(X)$ stands for the expected value of X.

3. Approximate cost models

In order to select the appropriate values of the decision variables S and r and to compare the two emergency ordering policies on economic grounds, suitable cost models have to be developed. Unfortunately, it is practically impossible to obtain an exact analytical expression for C_P due to the complex interrelationships among demand, regular and emergency orders in consecutive replenishment cycles. Therefore, we decided to derive approximate expressions for $E(Q_e)$, $E(\mathrm{OH}_i)$ and $E(\mathrm{BO}_i)$, $i = 1, 2, \ldots, P$, ignoring the emergency order quantities received in the time interval of length L just preceding the replenishment cycle under consideration. This simplifying assumption, suggested and used by Tagaras and Vlachos [12] in a similar model, is innocuous when the emergency orders are infrequent and small, as is usually the case. Especially, when the regular lead time L does not exceed the review period P, only one emergency order is ignored, that of the immediately previous cycle. To further simplify the notation and the presentation of the approximate cost models, we also assume $L_e = 1$, implying that the emergency order (if any) will be placed at time $P-1$ (*late-ordering)* or $P-2$ (*early ordering*) and will arrive in time unit P or $P-1$, respectively. The analysis for larger values of L_e is similar, only more tedious in terms of mathematical manipulations.

3.1. Late-ordering policy

To evaluate C_P from (1), one needs to compute $E(Q_e)$, $E(\mathrm{OH}_i)$ for $i = 1, 2, \ldots, P$ and $E(\mathrm{BO}_{P-1})$, $E(\mathrm{BO}_P)$, since by assumption $E(\mathrm{BO}_i) = 0$ for $i = 1, 2, \ldots, P-2$. To this end, we examine the time interval of length $L+P$ time units, which starts at the time (review instance) of placement of the regular order that is received at the beginning of the replenishment cycle under consideration and ends just before the arrival of the next regular order. In other words, this time interval includes and ends with the replenishment cycle. It is convenient to divide the total demand in the time interval of length $L+P$ in two parts, namely the demand of the first $L+P-1$ time units, denoted by y', and the demand of the last time unit, denoted by y. Fig. 1 explains the structure of an interval of length $L+P$ defined as above. The emergency order of the current replenishment cycle is marked EO2, while the emergency order of the previous replenishment cycle is marked EO1.

Let $f(y')$ and $F(y')$ denote the density and distribution function respectively of y'. By the independence assumption for demand in disjoint time intervals it follows that $f(y')$ is the $(L+P-1)$-fold convolution of the time unit demand density function $g(y)$. In the ensuing analysis it is assumed that all density and distribution functions of demand are continuous with infinite support and the distribution functions are also differentiable.

Since the stocking location follows an order-up-to-S policy for regular orders and the previous emergency order (EO1 in Fig. 1) is negligible by assumption, the determination of $E(\mathrm{OH}_{P-1})$ and $E(\mathrm{BO}_{P-1})$ is straightforward following classical inventory theory arguments:

$$E(\mathrm{OH}_{P-1}) = \int_{y'=0}^{S} (S-y')f(y')\,\mathrm{d}y' = \int_{y'=0}^{S} F(y')\,\mathrm{d}y', \quad (2)$$

$$E(\mathrm{BO}_{P-1}) = \int_{y'=S}^{\infty} (y'-S)f(y')\,\mathrm{d}y' = \mu(L+P-1) - S + E(\mathrm{OH}_{P-1}). \quad (3)$$

By the definition of the *late-ordering* policy, an emergency order will be placed at time $P-1$ of the cycle only if $\mathrm{NS}_{P-1} < r$. Since $\mathrm{NS}_{P-1} = S - y'$, the emergency order will be placed only if $y' > S - r$ and its size will then be $Q_e = \min\{y' + r - S, K\}$. Consequently,

$$E(Q_e) = \int_{y'=S-r}^{S-r+K} (y'+r-S)f(y')\,\mathrm{d}y' + K\int_{y'=S-r+K}^{\infty} f(y')\,\mathrm{d}y' = K - \int_{y'=S-r}^{S-r+K} F(y')\,\mathrm{d}y'. \quad (4)$$

From (4) it is clear that $E(Q_e)$ is an increasing function of the emergency channel capacity K, as expected, stabilising at a certain maximum value when K becomes so large that it is practically irrelevant (unconstrained Q_e).

To evaluate $E(\mathrm{OH}_P)$ and $E(\mathrm{BO}_P)$ one needs to take into account the emergency order Q_e, which arrives in time unit P and can be used to satisfy part of the demand y of that time unit and/or existing backorders from time unit $P-1$. It is therefore necessary to consider the following possible realisations of y' and y:

- if $y' \leqslant S - r$, then $Q_e = 0$ and $\mathrm{NS}_p = S - y - y'$, i.e.,
  - if $y + y' \leqslant S$, then $\mathrm{OH}_P = S - y - y'$ and $BO_P = 0$,
  - if $y + y' > S$, then $\mathrm{OH}_P = 0$ and $BO_P = y + y' - S$;
- if $S - r < y' < S - r + K$, then $Q_e = y' - (S - r)$ and $\mathrm{NS}_p = r - y$, i.e.,
  - if $y \leqslant r$, then $\mathrm{OH}_P = r - y$ and $\mathrm{BO}_P = 0$,
  - if $y > r$, then $\mathrm{OH}_P = 0$ and $\mathrm{BO}_P = y - r$;
- if $y' > S - r + K$, then $Q_e = K$ and $\mathrm{NS}_p = S + K - y - y'$, i.e.,
  - if $y + y' \leqslant S + K$, then $\mathrm{OH}_P = S + K - y - y'$ and $\mathrm{BO}_P = 0$,
  - if $y + y' > S + K$, then $\mathrm{OH}_P = 0$ and $\mathrm{BO}_P = y + y' - S - K$.

Consequently, the expressions for $E(\mathrm{OH}_P)$ and $E(\mathrm{BO}_P)$ at the end of the cycle are

$$E(\mathrm{OH}_P) = \int_{y'=0}^{S-r}\left[\int_{y=0}^{S-y'}(S-y-y')g(y)\,\mathrm{d}y\right]f(y')\,\mathrm{d}y' + \int_{y'=S-r}^{S-r+K}\left[\int_{y=0}^{r}(r-y)g(y)\,\mathrm{d}y\right]f(y')\,\mathrm{d}y' + \int_{y'=S-r+K}^{\infty}\left[\int_{y=0}^{S+K-y'} \times (S+K-y-y')g(y)\,\mathrm{d}y\right]f(y')\,\mathrm{d}y',$$

$$E(\mathrm{BO}_P) = \int_{y'=0}^{S-r}\left[\int_{y=S-y'}^{\infty}(y+y'-S)g(y)\,\mathrm{d}y\right]f(y')\,\mathrm{d}y' + \int_{y'=S-r}^{S-r+K}\left[\int_{y=r}^{\infty}(y-r)g(y)\,\mathrm{d}y\right]f(y')\,\mathrm{d}y' + \int_{y'=S-r+K}^{\infty}\left[\int_{y=S+K-y'}^{\infty} \times (y+y'-S-K)g(y)\,\mathrm{d}y\right]f(y')\,\mathrm{d}y'$$

which, after long manipulations, simplify to the following:

$$E(\mathrm{OH}_P) = \int_{y=0}^{r} G(y)F(S+K-y)\,\mathrm{d}y + \int_{y=r}^{S} G(y)F(S-y)\,\mathrm{d}y, \quad (5)$$

$$E(\mathrm{BO}_P)=\int_{y=0}^{r}G(y)F(S+K-y)\,\mathrm{d}y+\int_{y=r}^{S}G(y)F(S-y)\,\mathrm{d}y+(L+P)\mu-S-K+\int_{y'=S-r}^{S-r+K}F(y')\,\mathrm{d}y'$$
$$=E(\mathrm{OH}_P)+(L+P)\mu-S-K+\int_{y'=S-r}^{S-r+K}F(y')\,\mathrm{d}y'. \quad (6)$$

The determination of the terms of the expected cost function, C_P, can now be completed by deriving expressions for $E(\mathrm{OH}_1),\dots, E(\mathrm{OH}_{P-2})$, or, equivalently, for $E(\mathrm{NS}_1),\dots, E(\mathrm{NS}_{P-2})$ because by assumption $E(\mathrm{BO}_1)=E(\mathrm{BO}_2)=\cdots=E(\mathrm{BO}_{P-2})=0$. Since no demand is lost, it is clear from Fig. 1 that the expected net stock immediately after the arrival of the regular order at time unit 1 of the replenishment cycle and before the demand of that time unit is materialised is $S-L\mu$ (recall that EO1 is ignored). Consequently, $E(\mathrm{NS}_1)=S-L\mu-\mu$ and more generally

$$E(\mathrm{OH}_i)=E(\mathrm{NS}_i)=S-(L+i)\mu,\quad i=1,2,\dots,P-2, \quad (7)$$

and

$$\sum_{i=1}^{P-2}E(\mathrm{OH}_i)=(P-2)[S-(L+P)\mu]+\mu\left[\frac{P(P-1)}{2}-1\right]. \quad (8)$$

Combining all the expressions above, the approximate expected cost function takes the following form:

$$C_P=c_\mathrm{h}\left[\sum_{i=1}^{P-2}E(\mathrm{OH}_i)+E(\mathrm{OH}_{P-1})+E(\mathrm{OH}_P)\right]+c_\mathrm{p}[E(\mathrm{BO}_{P-1})+E(\mathrm{BO}_P)]+c_\mathrm{e}E(Q_\mathrm{e})$$
$$=c_\mathrm{h}\mu\left[\frac{P(P-1)}{2}-1\right]+c_\mathrm{h}(P-2)[S-\mu(L+P)]+c_\mathrm{p}[\mu(2L+2P-1)-2S]$$
$$+(c_\mathrm{h}+c_\mathrm{p})\left[\int_{y'=0}^{S}F(y')\,\mathrm{d}y'+\int_{y=0}^{r}G(y)F(S+K-y)\,\mathrm{d}y+\int_{y=r}^{S}G(y)F(S-y)\,\mathrm{d}y\right]$$
$$+(c_\mathrm{e}-c_\mathrm{p})\left[K-\int_{y'=S-r}^{S-r+K}F(y')\,\mathrm{d}y'\right]. \quad (9)$$

3.2. *Early-ordering policy*

The derivation of $E(Q_\mathrm{e})$, $E(\mathrm{OH}_i)$ for $i=1,2,\dots,P$ and $E(\mathrm{BO}_{P-1}), E(\mathrm{BO}_P)$ for the *early-ordering* policy follows a similar path and it will be presented much more succinctly in this section. The main difference from the analysis of the *late-ordering* policy is that now the demand in the time interval of length $L+P$ is divided into the demand of the first $L+P-2$ time units, denoted by y'', and the demand of the last two time units, denoted by y_2 (see Fig. 2). Furthermore, the demand in time unit $P-1$, denoted by y, also needs to be taken into account separately in the derivation of expressions for $E(\mathrm{OH}_{P-1})$ and $E(\mathrm{BO}_{P-1})$.

The analysis now starts with the determination of the emergency order quantity at time $P-2$. According to the *early-ordering* policy, an emergency order will be placed at that time only if $\mathrm{NS}_{P-2}<r$. Since $\mathrm{NS}_{P-2}=S-y''$, such an order will be placed only if $y''>S-r$ and its size will then be $Q_\mathrm{e}=\min\{y''+r-S,K\}$. Thus,

$$E(Q_\mathrm{e})=\int_{y''=S-r}^{S-r+K}(y''+r-S)h(y'')\,\mathrm{d}y''+K\int_{y''=S-r+K}^{\infty}h(y'')\,\mathrm{d}y''$$
$$=K-\int_{y''=S-r}^{S-r+K}H(y'')\,\mathrm{d}y''. \quad (10)$$

The derivation of expressions for $E(\mathrm{OH}_{P-1})$ and $E(\mathrm{BO}_{P-1})$ is almost identical to the analogous derivation of $E(\mathrm{OH}_P)$ and $E(\mathrm{BO}_P)$ in the *late-ordering* policy, with the modification that y'' replaces y'. For brevity, we provide below the final results:

$$E(\mathrm{OH}_{P-1}) = \int_{y=0}^{r} G(y)H(S+K-y)\,\mathrm{d}y + \int_{y=r}^{S} G(y)H(S-y)\,\mathrm{d}y, \tag{11}$$

$$E(\mathrm{BO}_{P-1}) = E(\mathrm{OH}_{P-1}) + (L+P-1)\mu - S - K + \int_{y''=S-r}^{S-r+K} H(y'')\,\mathrm{d}y''. \tag{12}$$

$E(\mathrm{OH}_P)$ and $E(\mathrm{BO}_P)$ are obtained in a similar fashion, but now y_2 replaces y. The final expressions are

$$E(\mathrm{OH}_P) = \int_{y_2=0}^{r} G_2(y_2)H(S+K-y_2)\,\mathrm{d}y_2 + \int_{y_2=r}^{S} G_2(y_2)H(S-y_2)\,\mathrm{d}y_2, \tag{13}$$

$$E(\mathrm{BO}_P) = E(\mathrm{OH}_P) + (L+P)\mu - S - K + \int_{y''=S-r}^{S-r+K} H(y'')\,\mathrm{d}y''. \tag{14}$$

Finally, the expressions for $E(\mathrm{OH}_1), \ldots, E(\mathrm{OH}_{P-2})$ remain the same as in (7). The approximate total expected cost per replenishment cycle, C_P, under the *early-ordering policy* is given by

$$\begin{aligned} C_P = {} & c_\mathrm{h}\mu\left[\frac{P(P-1)}{2} - 1\right] + c_\mathrm{h}(P-2)[S - \mu(L+P)] + c_\mathrm{p}[\mu(2L+2P-1) - 2S] \\ & + (c_\mathrm{h}+c_\mathrm{p})\left[\int_{y=0}^{r} [G(y)+G_2(y)]\,H(S+K-y)\mathrm{d}y + \int_{y=r}^{S} [G(y)+G_2(y)]\,H(S-y)\,\mathrm{d}y\right] \\ & + (c_\mathrm{e} - 2c_\mathrm{p})\left[K - \int_{y''=S-r}^{S-r+K} H(y'')\,\mathrm{d}y''\right]. \end{aligned} \tag{15}$$

4. Optimization and properties

The parameters S and r that minimise C_P of the corresponding inventory system must necessarily satisfy the first-order conditions

$$\frac{\partial C_P}{\partial S} = 0, \quad \frac{\partial C_P}{\partial r} = 0. \tag{16}$$

The analytical expressions of the first-order conditions are obviously different for the two systems. In the two subsections below, we derive expressions for determining the optimal S, r with regard to C_P and properties of the optimal solutions, under the *late-ordering* and *early-ordering* policies for emergency replenishments.

4.1. Late-ordering policy

Using (9), the first-order conditions (16) are rewritten as

$$c_\mathrm{h}(P-2) - 2c_\mathrm{p} + (c_\mathrm{h}+c_\mathrm{p})\left[F(S) + \int_{y=0}^{r} G(y)f(S+K-y)\,\mathrm{d}y + \int_{y=r}^{S} G(y)f(S-y)\,\mathrm{d}y\right] + (c_\mathrm{e}-c_\mathrm{p})[F(S-r) - F(S-r+K)] = 0, \tag{17}$$

$$[F(S-r+K) - F(S-r)] \times [c_\mathrm{e} - c_\mathrm{p} + (c_\mathrm{h}+c_\mathrm{p})\,G(r)] = 0. \tag{18}$$

Proposition 1 specifies how to determine the optimal S and r under certain conditions. The proofs of Propositions 1 and 2 that follow are given in the appendix.

Proposition 1. *If* (a) *$G(y)$ is strictly increasing in y,* (b) *the decision parameters must satisfy the relationship* $0 < r < S$ *and* (c)

$$2c_\mathrm{p} - c_\mathrm{h}(P-2) > (c_\mathrm{p}+c_\mathrm{h})\left[F(r^0) + \int_{y=0}^{r^0} F(r^0+K-y)g(y)\,\mathrm{d}y\right],$$

where r^0 is such that

$$G(r^0) = \frac{c_p - c_e}{c_p + c_h}, \tag{19}$$

then the optimal r is r^0 and the optimal S satisfies the equation

$$F(S) + \int_{x=0}^{r^0} F(S+K-x)g(x)\,dx + \int_{x=r^0}^{S} F(S-x)g(x)\,dx = \frac{2c_p - c_h(P-2)}{c_p + c_h}. \tag{20}$$

The conditions of Proposition 1 hold in all the numerical examples of the following section. Consequently, Proposition 1 means that in many practical situations C_P has a unique global minimum that is obtained as the unique solution to the system of equations (19) and (20). This solution, which hereunder will be denoted S^0, r^0, is the exact optimum of the approximate cost function C_P but it is generally a sub-optimal solution of the exact cost function. Its quality will be evaluated numerically in the next section.

The following proposition shows the effect of the emergency channel capacity K on the optimal S^0 and r^0 of C_p.

Proposition 2. *S^0 is a monotonically decreasing function of the capacity of the emergency channel K, while r^0 is independent of K.*

The dependence of S^0 on K is intuitively appealing; the lower the capacity of the emergency channel, the more the inventory system has to rely on regular replenishments and this can be achieved by increasing the base stock S. On the other hand, the independence between r^0 and K is puzzling. Note, in addition, that by expression (19) the optimal r^0 is also independent of S and L. The explanation for this remarkable stability of r^0 is that the determination of the optimal r in the context of the approximate cost model follows essentially a Newsboy-type analysis for time unit P with initial net stock NS_{P-1} independent of r and K.

4.2. Early-ordering policy

Using (15), the first-order conditions (16) become in this case

$$\begin{aligned} &c_h(P-2) - 2c_p + (c_h + c_p) \\ &\left[\int_{y=0}^{r} G(y)h(S+K-y)\,dy + \int_{y=r}^{S} G(y)h(S-y)\,dy\right] \\ &\quad + (c_h + c_p)\left[\int_{y=0}^{r} G_2(y)h(S+K-y)\,dy \right.\\ &\quad \left. + \int_{y=r}^{S} G_2(y)h(S-y)\,dy\right] \\ &\quad + (c_e - 2c_p)\,[H(S-r) - H(S-r+K)] = 0, \end{aligned} \tag{21}$$

$$\begin{aligned} &[H(S-r+K) - H(S-r)] \\ &\quad \times [c_e - 2c_p + (c_h + c_p)\,(G(r) + G_2(r))] = 0. \end{aligned} \tag{22}$$

Propositions 3 and 4 (and their proofs) are analogous to Propositions 1 and 2 of the *early-ordering* policy and are presented below without discussion, because their interpretation and the relevant comments are also similar.

Proposition 3. *If* (a) *G(y) is strictly increasing in y,* (b) *the decision parameters must satisfy the relationship* $0 < r < S$ *and* (c)

$$2c_p - c_h(P-2) > (c_p + c_h)\left[\int_{y=0}^{r^0} [g(y) + g_2(y)]H(r^0 + K - y)\,dy\right],$$

where r^0 is such that

$$G(r^0) + G_2(r^0) = \frac{2c_p - c_e}{c_p + c_h}, \tag{23}$$

then the optimal r is r^0 and the optimal S satisfies the equation

$$\begin{aligned} &\int_{x=0}^{r^0} H(S+K-x)\,[g(x) + g_2(x)]\,dx \\ &\quad + \int_{x=0}^{S-r^0} H(x)\,[g(S-x) + g_2(S-x)]\,dx \\ &\quad = \frac{2c_p - c_h(P-2)}{c_p + c_h}. \end{aligned} \tag{24}$$

Proposition 4. *S^0 is a monotonically decreasing function of the capacity of the emergency channel K, while r^0 is independent of K.*

5. Numerical results

The numerical investigation we undertook and report in this section had three objectives:

- Evaluate the accuracy of the approximate models and the quality of the resulting solutions.
- Examine the effect of the limited emergency channel capacity on the performance of the inventory systems under consideration.
- Compare the relative cost effectiveness of the two emergency ordering policies under different conditions.

To meet these objectives we solved a total of 144 problems, 72 under the *late-ordering* policy for emergency replenishments and 72 under the *early-ordering* policy. These problems were generated by combining a common set of 24 combinations of review period (P), regular order lead time (L), coefficient of variation of demand (σ/μ), backorder penalty cost (c_p) and emergency order cost (c_e) with three levels of emergency channel capacity (K). Specifically, three combinations of P and L were examined and two values were used for each of σ/μ, c_p and c_e. The resulting 24 combinations of these parameters are shown in Table 1. In all cases the holding cost rate was used as the cost unit ($c_h = 1$) and the emergency replenishment lead time was used as the time unit ($L_e = 1$). The demand per time unit was assumed to follow a normal distribution left-truncated at 0, i.e., with the negative part of the normal distribution redistributed proportionally to its positive part [17,18], with a mean (μ) of 100 units. The three levels of K were set equal to 20, 100 and 200 units. The case $K = 20$ represents low capacity and the system approaches a classical inventory system with a single replenishment mode. $K = 100$ represents medium capacity, equal to the average demand in a time unit ($K = \mu$). $K = 200$ represents high emergency channel capacity, equal to the average demand of two time units ($K = 2\mu$). This case approaches a system without constraint on the size of emergency orders.

Table 1
Combination of problem parameters used in the numerical investigation

Combination	P	L	σ/μ	c_p	c_e
1	7	4	0.2	50	20
2	7	4	0.2	100	20
3	7	4	0.2	50	40
4	7	4	0.2	100	40
5	7	4	0.4	50	20
6	7	4	0.4	100	20
7	7	4	0.4	50	40
8	7	4	0.4	100	40
9	7	7	0.2	50	20
10	7	7	0.2	100	20
11	7	7	0.2	50	40
12	7	7	0.2	100	40
13	7	7	0.4	50	20
14	7	7	0.4	100	20
15	7	7	0.4	50	40
16	7	7	0.4	100	40
17	14	7	0.2	50	20
18	14	7	0.2	100	20
19	14	7	0.2	50	40
20	14	7	0.2	100	40
21	14	7	0.4	50	20
22	14	7	0.4	100	20
23	14	7	0.4	50	40
24	14	7	0.4	100	40

For each of the 144 problems, S^0 and r^0 were first obtained by numerically solving the systems of equations (19) and (20) of Proposition 1, and (23) and (24) of Proposition 3 (where necessary, the convolutions of demand per time unit distributions were approximated by normal distributions with the same mean and variance). Then, a combination of simulation and search led to the determination of the "exact optimal" solutions, denoted S^*, r^*, which yield the minimum simulated cost, C_P^*. Specifically, under each combination of S, r that was examined, the system was simulated 3000 times for 500 replenishment cycles. The programme simulated the operation of the system exactly as it was described in Section 2 and derived estimates for all the terms appearing in the cost function (1). The two-sided 95% confidence intervals for the operational characteristics were in all cases tighter than 0.1% of the respective point estimates. The optimal combination S^*, r^* was identified through a grid search over integer only values of S, r, first to keep the computational time reasonable and second because the expected cost function is very flat in the vicinity of the optimum.

Tables 2 and 3 contain a representative sample of the detailed results. Specifically, Table 2 presents the solutions S^0, r^0 of the approximate cost model (rounded to the closest integers) for the 24 problems under the *late-ordering* policy with $K = 20$, along with operational characteristics and expected costs, C_P, as they are computed from the respective expressions of Section 3. Similarly, Table 3 contains the solutions of the approximate cost model under the *early-ordering* policy with $K = 100$. The rightmost columns of Tables 2 and 3 summarise the simulation results at the solutions S^0, r^0. In all 144 problems, the conditions of Propositions 1 and 3 for unique global minima were satisfied.

The accuracy in estimating the quantities of interest with the approximate model of the *late-ordering* policy can be inferred by comparing the corresponding columns of Table 2. This comparison shows that the differences are generally small. Moreover, the simulated values of $E(\mathrm{BO}_i)$, $i = 1, 2, \ldots, P-2$ are indeed very close to 0, except for the cases with $P = 14$ and high demand variability where $E(\mathrm{BO}_{P-2})$ is not negligible. Similar conclusions are drawn for the *early-ordering* policy from Table 3. It is fair to say, though, that the satisfactory accuracy of the approximate models is partly due to the fact that in all numerical examples the regular replenishment lead time is not longer than the review period: $L \leqslant P$. Although the models in Section 3 may also be used for $L > P$, their accuracy will deteriorate in those situations because more than one previous emergency replenishments are ignored. This case is addressed by Tagaras and Vlachos [12] in the unconstrained version of the *late-ordering* policy by means of a heuristic procedure.

An even more important issue is the quality of the solutions S^0, r^0, i.e., their proximity to S^*, r^* and the cost penalty for using the former rather than the latter. Tables 4 and 5 contain S^* and r^* for the 24 problems under late and early emergency ordering with $K = 20$, 100 and 200 and the expected costs C_P^* of these solutions, along with the respective solutions S^0, r^0 and their *simulated* costs. ΔC_P is the percentage cost penalty of the solution S^0, r^0 with respect to the optimum measured in terms of C_P^*.

The main conclusion from Tables 4 and 5 is that the cost penalties are uniformly very low. For the 72 problems with late emergency ordering (Table 4), the average cost penalty ΔC_P is 0.17%; in the worst case, $\Delta C_P = 0.78\%$. For the 72 problems with early emergency ordering (Table 5), the average cost penalty ΔC_P is 0.08%; in the worst case, $\Delta C_P = 0.39\%$. The cost penalties are lowest in the low-capacity problems ($K = 20$), because in those cases the simplifying assumption of the approximate model (ignoring the previous emergency order) is harmless. Another important observation is that S^* is non-increasing in K, exactly like S^0 by Propositions 2 and 4. In addition, S^0 is in most cases higher than S^*, but the differences are generally not very large, hence the low-cost penalties. At the same time, r^0 is usually but not always lower than r^*. In particular, the optimal r (r^*) is decreasing in K and increasing in L for fixed P (while according to the approximate model r^0 is independent of K and L), but the relationship is not so strong as to lead to very unsatisfactory solutions S^0, r^0. Overall, it can be argued that the approximate models of Section 4 are quite accurate and their solutions are near-optimal resulting in low-cost penalties with respect to the exact optimal solutions.

We can now pursue the second objective of the numerical investigation, that is to examine the effect of the limited emergency channel capacity on the performance of the inventory systems. The relationship between K and the optimal S and r has already been discussed. The exact effect of K on the total expected cost can be determined by comparing the simulated costs of S^0, r^0 or S^*, r^* between problems differing only in K. Without going into details, it suffices to say that when the capacity increases from its low level ($K = 20$) to its high level ($K = 200$) in the *late-ordering* system, the cost reduction in terms of C_P^* averages 3.17% over the 24 parameter combinations, with a maximum of 8.03%. The analogous cost reduction in the *early-ordering* system averages 3.32%, with a maximum of 8.57%. The cost differences increase as the review period and the regular replenishment leadtime become longer.

Table 2
S^0, r^0 and operational characteristics (approximate and by simulation) under the *late-ordering* policy ($K = 20$)

Problem	S^0	r^0	Approximate model						Simulation					
			$E(OH_{P-1})$	$E(OH_P)$	$E(BO_{P-1})$	$E(BO_P)$	$E(Q_e)$	C_P	$E(OH_{P-1})$	$E(OH_P)$	$E(BO_{P-1})$	$E(BO_P)$	$E(Q_e)$	C_P
1	1166	104	165.7	71.8	0.09	3.56	2.62	2800.5	168.4	73.8	0.06	3.14	2.35	2790.9
2	1187	116	187.0	90.7	0.03	1.65	2.04	2921.5	188.9	92.2	0.02	1.46	1.86	2911.0
3	1172	83	171.9	76.6	0.06	3.58	1.19	2837.5	173.2	77.6	0.05	3.35	1.12	2831.1
4	1192	105	191.8	94.6	0.02	1.56	1.27	2953.7	193.2	95.8	0.02	1.42	1.18	2945.7
5	1252	109	253.1	160.1	1.10	5.87	2.27	3567.1	251.2	158.2	0.98	6.18	2.20	3570.8
6	1295	132	295.4	199.1	0.42	2.41	1.72	3786.9	293.1	196.9	0.37	2.91	1.67	3822.0
7	1257	66	258.0	164.1	0.99	5.98	1.12	3600.6	255.0	161.3	0.92	6.48	1.10	3608.9
8	1299	109	299.4	202.5	0.38	2.34	1.14	3814.9	296.5	199.8	0.35	2.90	1.11	3853.5
9	1476	104	175.9	82.3	0.18	3.97	2.61	2896.5	178.4	84.0	0.11	3.39	2.32	2875.7
10	1500	116	200.0	103.7	0.06	1.86	1.95	3034.5	201.8	105.1	0.04	1.57	1.76	3012.3
11	1482	83	182.0	87.1	0.13	3.94	1.32	2935.0	183.3	88.0	0.10	3.59	1.21	2920.4
12	1504	105	204.5	107.5	0.05	1.75	1.29	3066.3	205.3	108.0	0.03	1.57	1.21	3048.1
13	1576	109	277.5	184.6	1.54	6.39	2.20	3782.8	274.3	181.3	1.34	6.61	2.13	3772.2
14	1624	132	324.6	228.3	0.62	2.68	1.62	4035.4	320.9	224.7	0.53	3.13	1.57	4054.0
15	1581	66	282.4	188.6	1.40	6.45	1.19	3816.6	278.3	184.6	1.28	6.88	1.16	3813.5
16	1628	109	328.6	232.8	0.57	2.60	1.13	4062.6	324.4	227.7	0.51	3.11	1.09	4086.3
17	2156	104	157.8	72.1	1.46	10.85	4.90	10 618.8	161.6	74.6	1.17	9.63	4.51	10 593.7
18	2192	116	193.0	100.8	0.50	5.01	3.36	11 021.4	195.5	102.7	0.41	4.50	3.14	10 993.5
19	2162	83	163.4	75.8	1.24	10.50	3.19	10 699.6	166.0	77.6	1.05	9.65	2.98	10 678.6
20	2196	105	196.8	103.7	0.44	4.76	2.57	11 080.2	198.8	105.2	0.37	4.36	2.43	11 055.9
21	2249	109	255.7	170.5	6.68	17.48	4.02	12 503.0	251.2	166.1	6.29	17.59	4.01	12 539.5
22	2319	132	322.6	230.3	2.62	7.66	2.69	13 275.1	316.1	224.1	2.51	8.19	2.71	13 307.6
23	2254	66	260.3	174.0	6.28	17.32	2.69	12 570.3	254.7	168.6	6.05	17.77	2.69	12 610.1
24	2323	109	325.5	232.6	2.52	7.53	2.11	13 323.2	319.4	226.8	2.41	8.06	2.11	13 358.4

Table 3
S^0, r^0 and operational characteristics (approximate and by simulation) under the *early-ordering* policy ($K = 100$)

Problem	S^0	r^0	Approximate model						Simulation					
			$E(OH_{P-1})$	$E(OH_P)$	$E(BO_{P-1})$	$E(BO_P)$	$E(Q_e)$	C_P	$E(OH_{P-1})$	$E(OH_P)$	$E(BO_{P-1})$	$E(BO_P)$	$E(Q_e)$	C_P
1	1156	205	162.1	64.8	0.00	2.69	6.55	2770.1	166.9	69.3	0.00	2.37	5.45	2771.0
2	1169	222	176.0	77.1	0.00	1.10	7.39	2854.0	181.1	82.1	0.00	0.96	6.06	2855.6
3	1171	174	172.1	75.7	0.00	3.62	1.33	2836.0	173.5	76.9	0.00	3.45	1.24	2833.9
4	1187	206	189.2	90.5	0.00	1.28	2.46	2939.7	191.4	92.6	0.00	1.19	2.20	2937.3
5	1224	209	232.6	139.9	0.00	5.85	9.07	3465.4	236.1	140.8	0.16	5.26	7.97	3458.2
6	1258	243	266.5	169.6	0.00	2.57	9.07	3664.5	270.0	171.8	0.06	2.25	7.97	3652.1
7	1245	160	247.3	154.5	0.00	6.67	2.84	3574.0	246.5	152.4	0.18	6.45	2.65	3564.1
8	1280	212	283.2	186.5	0.00	2.67	3.87	3791.8	283.2	185.3	0.04	2.52	3.56	3772.9
9	1461	205	168.8	71.6	0.00	2.80	7.92	2843.3	173.1	75.5	0.00	2.36	6.06	2823.2
10	1475	222	183.6	84.8	0.00	1.17	8.51	2931.5	188.0	88.9	0.00	0.96	6.47	2909.5
11	1479	174	181.2	85.0	0.00	3.78	1.89	2927.7	182.3	85.9	0.00	3.54	1.65	2914.8
12	1496	206	198.9	100.3	0.00	1.32	3.11	3034.9	201.1	102.3	0.00	1.18	2.57	3017.3
13	1542	209	251.5	158.1	0.00	6.30	9.80	3630.6	253.5	158.2	0.25	5.33	8.23	3595.0
14	1581	243	289.7	193.0	0.00	2.83	9.21	3855.2	291.4	193.2	0.10	2.28	7.75	3807.4
15	1565	160	268.2	175.5	0.00	6.96	3.58	3759.7	266.4	172.3	0.27	6.54	3.17	3733.1
16	1602	212	305.9	209.4	0.00	2.89	4.50	3994.3	304.8	206.9	0.08	2.59	3.93	3950.0
17	2117	205	142.4	50.0	0.22	7.77	24.85	10 297.2	154.6	60.2	0.10	5.73	18.77	10 315.0
18	2144	222	166.0	69.5	0.09	3.50	21.58	10 557.9	177.1	79.5	0.04	2.49	16.53	10 569.6
19	2149	174	157.8	66.6	0.07	8.89	8.62	10 606.9	163.5	71.3	0.05	7.85	7.26	10 597.2
20	2171	206	181.6	84.8	0.03	3.19	10.54	10 862.6	188.3	91.0	0.02	2.68	8.66	10 852.7
21	2197	209	221.7	135.4	2.93	16.66	21.75	11 789.2	229.4	140.2	2.56	13.78	18.93	12 021.0
22	2260	243	278.0	184.9	0.79	7.65	17.23	12 571.1	283.5	188.3	1.05	6.26	15.22	12 643.4
23	2223	160	236.4	151.2	1.98	16.81	11.41	12 259.3	237.8	150.7	2.17	15.43	10.32	12 318.4
24	2279	212	290.4	197.3	0.42	7.34	10.98	12 851.2	291.8	197.0	0.86	6.52	9.95	12 908.0

Table 4
Percentage cost penalties ΔC_P of the approximate solutions with respect to the "true" optima (simulation) under the *late-ordering* policy

Problem	$K = 20$							$K = 100$							$K = 200$						
	S^*	r^*	C_P^*	S^0	r^0	C_P	ΔC_P [a]	S^*	r^*	C_P^*	S^0	r^0	C_P	ΔC_P [a]	S^*	r^*	C_P^*	S^0	r^0	C_P	ΔC_P [a]
1	1161	107	2788.8	1166	104	2790.9	0.08	1140	104	2726.5	1152	104	2730.6	0.15	1140	104	2724.5	1150	104	2727.7	0.12
2	1181	119	2909.2	1187	116	2911.0	0.06	1154	116	2790.4	1163	116	2795.6	0.18	1151	116	2785.0	1160	116	2789.2	0.15
3	1169	89	2830.1	1172	83	2831.1	0.04	1166	85	2821.4	1170	83	2822.7	0.05	1166	85	2821.4	1170	83	2822.7	0.05
4	1188	108	2944.5	1192	105	2945.7	0.04	1177	105	2897.5	1182	105	2899.5	0.07	1176	105	2896.6	1182	105	2898.8	0.07
5	1248	117	3569.9	1252	109	3570.8	0.03	1222	111	3460.7	1235	109	3466.6	0.17	1216	111	3438.0	1228	109	3443.0	0.15
6	1291	138	3820.7	1295	132	3822.0	0.03	1254	135	3636.4	1267	132	3643.6	0.20	1240	132	3587.0	1256	132	3594.5	0.21
7	1255	80	3608.0	1257	66	3608.9	0.02	1248	74	3584.2	1254	66	3586.5	0.06	1244	73	3582.9	1253	66	3584.4	0.04
8	1296	116	3852.6	1299	109	3853.5	0.02	1277	112	3759.5	1284	109	3762.8	0.09	1271	112	3742.0	1280	109	3745.3	0.09
9	1467	109	2871.5	1476	104	2875.7	0.15	1441	107	2770.8	1460	104	2786.5	0.57	1439	107	2766.4	1458	104	2781.2	0.53
10	1494	120	3007.9	1500	116	3012.3	0.15	1451	118	2838.6	1473	116	2857.4	0.66	1447	118	2828.0	1469	116	2845.3	0.61
11	1477	91	2917.5	1482	83	2920.4	0.10	1470	91	2894.7	1480	83	2901.9	0.25	1468	89	2894.5	1479	83	2900.7	0.21
12	1499	109	3045.5	1504	105	3048.1	0.08	1480	107	2970.0	1492	105	2977.9	0.27	1480	107	2967.7	1491	105	2975.2	0.25
13	1571	122	3770.2	1576	109	3772.2	0.05	1533	119	3606.3	1558	109	3625.7	0.54	1521	119	3561.4	1550	109	3585.4	0.67
14	1619	146	4051.3	1624	132	4054.0	0.06	1570	142	3805.6	1596	132	3830.8	0.66	1549	139	3725.9	1581	132	3755.1	0.78
15	1578	86	3811.3	1581	66	3813.5	0.06	1563	84	3757.3	1577	66	3767.7	0.28	1559	84	3749.3	1576	66	3760.7	0.30
16	1625	119	4084.4	1628	109	4086.3	0.05	1594	118	3946.5	1612	109	3959.4	0.33	1586	118	3913.9	1606	109	3929.0	0.39
17	2148	107	10 589.3	2156	104	10 593.7	0.04	2114	104	10 342.1	2134	104	10 367.1	0.24	2107	103	10 310.1	2126	104	10 330.3	0.20
18	2186	119	10 989.3	2192	116	10 993.5	0.04	2142	116	10 582.7	2159	116	10 607.8	0.24	2129	115	10 505.3	2147	116	10 527.9	0.21
19	2157	88	10 675.6	2162	83	10 678.6	0.03	2146	85	10 622.0	2158	83	10 633.1	0.10	2146	83	10 622.7	2157	83	10 632.2	0.09
20	2191	108	11 052.9	2196	105	11 055.9	0.03	2165	105	10 836.0	2178	105	10 850.4	0.13	2161	105	10 810.8	2173	105	10 822.7	0.11
21	2255	118	12 535.3	2249	109	12 539.5	0.03	2220	115	12 234.3	2229	109	12 239.2	0.04	2201	109	12 118.6	2213	109	12 122.9	0.04
22	2321	145	13 305.5	2319	132	13 307.6	0.02	2278	139	12 855.8	2289	132	12 863.7	0.06	2255	135	12 648.8	2267	132	12 656.4	0.06
23	2261	83	12 604.4	2254	66	12 610.1	0.05	2243	77	12 504.9	2249	66	12 507.7	0.02	2241	72	12 487.3	2246	66	12 489.4	0.02
24	2326	123	13 356.6	2323	109	13 358.4	0.01	2296	115	13 075.3	2304	109	13 080.0	0.04	2284	112	12 970.0	2292	109	12 974.3	0.03

[a] In %.

Table 5
Percentage cost penalties ΔC_P of the approximate solutions with respect to the "true" optima (simulation) under the *early-ordering* policy

Problem	$K = 20$							$K = 100$							$K = 200$						
	S^*	r^*	C_P^*	S^0	r^0	C_P	ΔC_P[a]	S^*	r^*	C_P^*	S^0	r^0	C_P	ΔC_P[a]	S^*	r^*	C_P^*	S^0	r^0	C_P	ΔC_P[a]
1	1164	208	2805.4	1166	205	2806.0	0.02	1152	204	2769.4	1156	205	2771.0	0.06	1152	204	2768.9	1155	205	2770.1	0.04
2	1185	225	2929.2	1188	222	2930.1	0.03	1163	221	2853.7	1169	222	2855.6	0.07	1162	221	2851.5	1167	222	2853.0	0.05
3	1170	178	2837.1	1172	174	2837.4	0.01	1169	175	2833.6	1171	174	2833.9	0.01	1169	175	2833.6	1171	174	2833.9	0.01
4	1191	208	2961.4	1193	206	2961.8	0.01	1184	205	2936.7	1187	206	2937.3	0.02	1184	205	2936.4	1187	206	2937.0	0.02
5	1248	215	3569.2	1249	209	3569.6	0.01	1216	211	3456.2	1224	209	3458.2	0.06	1204	211	3431.8	1213	209	3432.6	0.02
6	1290	249	3825.5	1293	243	3826.2	0.02	1247	244	3649.0	1258	243	3652.1	0.09	1231	244	3598.7	1240	243	3600.7	0.06
7	1252	171	3601.3	1254	160	3601.5	0.01	1240	165	3564.0	1245	160	3564.1	0.00	1240	163	3559.1	1243	160	3559.3	0.00
8	1296	217	3856.2	1297	212	3856.5	0.01	1275	214	3771.8	1280	212	3772.9	0.03	1272	211	3757.6	1274	212	3758.0	0.01
9	1470	210	2884.9	1476	205	2887.9	0.11	1447	207	2816.1	1461	205	2823.2	0.25	1447	207	2814.0	1459	205	2820.0	0.21
10	1494	227	3024.0	1501	222	3027.9	0.13	1461	224	2901.0	1475	222	2909.5	0.29	1458	223	2895.0	1472	222	2902.6	0.26
11	1478	185	2924.7	1482	174	2927.0	0.08	1472	182	2911.8	1479	174	2914.8	0.10	1472	182	2911.7	1479	174	2914.7	0.10
12	1501	211	3060.9	1505	206	3062.9	0.06	1488	209	3012.9	1496	206	3017.3	0.14	1487	209	3011.9	1495	206	3015.6	0.12
13	1569	221	3763.7	1572	209	3765.3	0.04	1524	220	3584.6	1542	209	3595.0	0.29	1506	216	3532.9	1524	209	3540.9	0.23
14	1617	258	4048.7	1621	243	4050.8	0.05	1559	252	3792.7	1581	243	3807.4	0.39	1529	247	3701.6	1554	243	3713.2	0.32
15	1573	183	3801.5	1577	160	3802.8	0.03	1556	174	3728.3	1565	160	3733.1	0.13	1553	171	3715.3	1561	160	3719.6	0.12
16	1621	225	4082.7	1625	212	4084.0	0.03	1590	221	3943.8	1602	212	3950.0	0.16	1580	218	3911.5	1593	212	3917.3	0.15
17	2146	207	10 580.0	2153	205	10 583.8	0.04	2098	204	10 296.5	2117	205	10 315.0	0.18	2082	202	10 246.3	2097	205	10 257.2	0.11
18	2184	226	10 989.2	2190	222	10 992.7	0.03	2125	220	10 550.1	2144	222	10 569.6	0.19	2096	220	10 431.2	2113	222	10 443.0	0.11
19	2155	181	10 663.4	2159	174	10 665.4	0.02	2141	175	10 592.2	2149	174	10 597.2	0.05	2140	173	10 590.4	2148	174	10 595.1	0.04
20	2190	208	11 055.9	2195	206	11 058.7	0.03	2160	206	10 844.4	2171	206	10 852.7	0.08	2155	205	10 820.3	2165	206	10 826.9	0.06
21	2251	220	12 487.3	2243	209	12 492.8	0.04	2195	215	12 020.0	2197	209	12 021.0	0.01	2159	210	11 780.0	2150	209	11 782.9	0.03
22	2317	257	13 264.1	2314	244	13 266.3	0.02	2253	249	12 640.5	2260	243	12 643.4	0.02	2207	243	12 269.3	2202	243	12 270.4	0.01
23	2257	177	12 554.8	2248	160	12 561.8	0.06	2225	169	12 316.4	2223	160	12 318.4	0.02	2213	165	12 228.4	2206	160	12 230.3	0.02
24	2323	223	13 318.5	2318	212	13 320.9	0.02	2277	217	12 906.9	2279	212	12 908.0	0.01	2252	213	12 715.3	2249	212	12 715.7	0.00

[a] In %.

The third and final objective of the numerical investigation was to compare the relative cost effectiveness of the two emergency ordering policies under different conditions. This comparison is performed by examining the corresponding costs in Tables 4 and 5 side by side. The interesting observation here is that the results are mixed. The early-ordering policy outperforms the late-ordering policy in 38 of the 72 cases and the reverse happens in the remaining 34 cases. Specifically, the cost performance of the early-ordering policy, relative to that of the late-ordering policy, improves as the variability of demand increases, the shortage cost decreases and the regular leadtime and the review period become longer. The cost differences between the two policies, in either direction, are not large but they become more pronounced as the capacity of the emergency channel increases. The maximum cost advantage of the early-ordering policy is about 3% of C_P^* in problem 22 with $K = 200$, whereas the maximum cost advantage of the late-ordering policy approaches 2.4% in problem 2 with $K = 200$.

6. Conclusion

We examined periodic review inventory systems with regular and emergency supply modes with capacity limitations on the emergency channel and two alternative options for the timing of ordering and receiving emergency replenishments. The approximate cost models that we developed were shown to provide near optimal solutions that can be derived very efficiently. The main conclusions of the extensive numerical investigation regarding the effect of the limited emergency channel capacity and the relative cost effectiveness of the two emergency ordering options can be summarised as follows:

- The constraint on the capacity of the emergency channel has a significant effect on the system performance under both emergency ordering policies, especially when the review period and the regular replenishment leadtime are long, or equivalently if the emergency leadtime is short and consequently emergency replenishments are more effective. This result is certainly intuitive and expected, but the proposed cost models permit the quantification of this effect.
- Placing the emergency order early in the replenishment cycle becomes relatively more advantageous than placing the order later in the cycle as the demand becomes more variable, the shortage cost decreases and the regular leadtime and review period become longer. The differences in the economic performance of the two emergency ordering policies are small, but they become relatively more substantial when the emergency channel capacity is large.

The emergency ordering policies that have been examined in this paper do not exhaust the class of possible policies for inventory systems with two supply modes. It is clear that more complex policies, such as deciding about the (variable) timing of a single emergency order using current inventory information or allowing multiple emergency orders per cycle will generally be more effective in terms of holding, stockout and emergency order cost. However, the proposed policies have the advantage of operational simplicity, which is a decisive factor for practical implementation. In addition, they are well suited for an environment where orderly management of inventory is especially desirable, since even the emergency ordering decisions are taken periodically at well-specified times. Despite these advantages, future research is needed to evaluate the savings that may be expected to result from the use of more flexible policies.

Acknowledgements

The research presented in this paper has been supported by the European Commission as part of the TMR network REVLOG – Reverse Logistics and its effects on industry (Contract No. ERB FMR XCT 97-0151). The authors would like to thank two anonymous referees for their useful comments and suggestions that contributed to a significantly improved presentation of the material in this paper.

Appendix A. Proofs of propositions

Proof of Proposition 1. The first partial derivatives of C_P with respect to S and r are

$$\begin{aligned}\frac{\partial C_P}{\partial S} &= c_h(P-2) - 2c_p + (c_h + c_p)\Bigg[F(S) + \int_{y=0}^{r} G(y) f(S+K-y)\,dy + \int_{y=r}^{S} G(y) f(S-y)\,dy\Bigg] \\ &\quad + (c_e - c_p)\,[F(S-r) - F(S-r+K)] \\ &= c_h(P-2) - 2c_p + (c_h + c_p)\Bigg[F(S) + \int_{y=0}^{r} F(S+K-y) g(y)\,dy + \int_{y=r}^{S} F(S-y) g(y)\,dy\Bigg] \\ &\quad + [(c_e - c_p) + G(r)(c_h + c_p)] \times [F(S-r) - F(S-r+K)], \end{aligned} \tag{A.1}$$

$$\frac{\partial C_P}{\partial r} = [F(S-r+K) - F(S-r)] \times [c_e - c_p + (c_h + c_p)G(r)]. \tag{A.2}$$

Since $F(S-r) < F(S-r+K)$ for all finite S, r and $K > 0$, the first derivative of C_P with respect to r vanishes only for r such that (19) holds. By condition (a) there is a unique r that satisfies (19). Substituting (19) into (A.1), $\partial C_P/\partial S = 0$ yields (20). For $0 < r < S$, per condition (b), the left-hand side of (20) is a positive increasing function of S with lowest value $F(r) + \int_{y=0}^{r} F(r+K-y) g(y)\,dy$ (at $S = r$). By condition (c), the right-hand side of (20) exceeds the lowest value and consequently there is a unique S that satisfies (20). Therefore, there is a unique combination of S and r that satisfies the first-order necessary conditions for optimality.

The second derivatives are

$$\frac{\partial^2 C_P}{\partial S \partial r} = \frac{\partial^2 C_P}{\partial r \partial S} = [f(S-r+K) - f(S-r)] \times [c_e - c_p + (c_h + c_p)G(r)] = A, \tag{A.3}$$

$$\begin{aligned}\frac{\partial^2 C_P}{\partial r^2} &= [f(S-r) - f(S-r+K)] \times [c_e - c_p + (c_h + c_p)G(r)] \\ &\quad + [F(S-r+K) - F(S-r)]\,(c_h + c_p)\,g(r) \\ &= -A + [F(S-r+K) - F(S-r)] \times (c_h + c_p)\,g(r) = B - A, \end{aligned} \tag{A.4}$$

where $B > 0$ since $g(r) > 0$ by condition (a) and $F(S-r+K) > F(S-r)$ for $K > 0$.

$$\begin{aligned}\frac{\partial^2 C_P}{\partial S^2} &= (c_h + c_p)\Bigg[f(S) + \frac{\partial}{\partial S}\int_{y=0}^{r} F(S+K-y) g(y)\,dy + \frac{\partial}{\partial S}\int_{y=r}^{S} F(S-y) g(y)\,dy\Bigg] \\ &\quad + [c_e - c_p + (c_h + c_p)G(r)] \times [f(S-r) - f(S-r+K)] \\ &= (c_h + c_p)\Bigg[f(S) + \int_{y=0}^{r} f(S+K-y) g(y)\,dy + \int_{y=r}^{S} f(S-y) g(y)\,dy\Bigg] - A = C - A \end{aligned}$$

with $C > 0$.

Evaluation of the second derivatives at the combination of S and r that satisfies the first-order conditions confirms that these S and r also satisfy the second-order sufficient conditions for a minimum. Thus, this combination of S and r is the global minimum of C_P.

Proof of Proposition 2. Differentiating (20) that characterises the optimal S, with respect to K gives

$$\frac{\partial A(S)}{\partial S}\frac{\partial S}{\partial K} + \frac{\partial A(S)}{\partial K} = 0, \tag{A.5}$$

where $A(S)$ is the left-hand side of (20).

The partial derivative of $A(S)$ with respect to S is

$$\frac{\partial A(S)}{\partial S} = \frac{\partial}{\partial S}\Bigg[F(S) + \int_{x=0}^{r} F(S+K-x) g(x)\,dx + \int_{x=r}^{S} F(S-x) g(x)\,dx\Bigg]$$

$$= f(S) + \int_{x=0}^{r} f(S + K - x)g(x)\,dx + \int_{x=r}^{S} f(S - x)g(x)\,dx > 0. \tag{A.6}$$

The partial derivative of $A(S)$ with respect to K is

$$\frac{\partial A(S)}{\partial K} = \frac{\partial}{\partial K}\left[F(S) + \int_{x=0}^{r} F(S + K - x)g(x)\,dx + \int_{x=r}^{S} F(S - x)g(x)\,dx\right] = \int_{x=0}^{r} f(S + K - x)g(x)\,dx > 0. \tag{A.7}$$

Finally, since $\partial A(S)/\partial S > 0$ and $\partial A(S)/\partial K > 0$, (A.5) yields $\partial S/\partial K < 0$, which proves that the optimal S is decreasing in K. The independence between the optimal r and K is obvious from (20).

References

[1] E.W. Barankin, A delivery-lag inventory model with an emergency provision, Naval Research Logistics Quarterly 8 (1961) 285–311.

[2] K.H. Daniel, A delivery-lag inventory model with emergency, in: H.E. Scarf, D.M. Gilford, M.W. Shelly (Eds.), Multistage Inventory Models and Techniques, Stanford University Press, Stanford, CA, 1962.

[3] M.F. Neuts, An inventory model with optional lag time, SIAM Journal of Applied Mathematics 12 (1964) 179–185.

[4] E.V. Bulinskaya, Some results concerning optimum inventory policies, Theory of Probability and Applications 9 (1964) 389–403.

[5] Y. Fukuda, Optimal policies for the inventory problem with negotiable leadtime, Management Science 10 (1964) 690–708.

[6] A.F. Veinott Jr., The status of mathematical inventory theory, Management Science 12 (1966) 745–777.

[7] A.S. Whittemore, S. Saunders, Optimal inventory under stochastic demand with two supply options, SIAM Journal of Applied Mathematics 32 (1977) 293–305.

[8] M. Rosenshine, D. Obee, Analysis of a standing order inventory system with emergency orders, Operations Research 24 (1976) 1143–1155.

[9] D. Gross, A. Soriano, On the economic application of airlift to product distribution and its impact on inventory levels, Naval Research Logistics Quarterly 19 (1972) 501–507.

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

[11] C. Chiang, G.J. Gutierrez, Optimal control policies for a periodic review inventory system with emergency orders, Naval Research Logistics 45 (1998) 187–204.

[12] G. Tagaras, D. Vlachos, A periodic review inventory system with emergency replenishments, Management Science 47 (3) 2001.

[13] K. Moinzadeh, S. Nahmias, A continuous review model for an inventory system with two supply modes, Management Science 34 (1988) 761–773.

[14] K. Moinzadeh, C.P. Schmidt, An (S − 1, S) inventory system with emergency orders, Operations Research 39 (1991) 308–321.

[15] S.G. Johansen, A. Thorstenson, An inventory model with Poisson demands and emergency orders, International Journal of Production Economics 56–57 (1998) 275–289.

[16] R. Dekker, J. Bloemhof-Ruwaard, M. Fleischmann, J. van Nunen, E. van der Laan, L. van Wassenhove, Operational research in reverse logistics: Some recent contributions, International Journal of Logistics: Research and Applications 1 (1998) 141–155.

[17] A.M. Mood, F.A. Graybill, D.C. Boes, Introduction to the Theory of Statistics, McGraw-Hill, New York, 1974.

[18] T.P. Ryan, Statistical Methods for Quality Improvement, Wiley, New York, 1989.