European Journal of Operational Research 203 (2010) 526–531

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Short Communication

An order expediting policy for continuous review systems with manufacturing lead-time

Chi Chiang *

Department of Management Science, National Chiao Tung University, Hsinchu, Taiwan, ROC

ARTICLE INFO

Article history:
Received 8 January 2009
Accepted 12 August 2009
Available online 20 August 2009

Keywords:
Supply mode
Inventory
Order expediting
Order splitting
Continuous review policy

ABSTRACT

Expedited shipments are often seen in practice. When the inventory level of an item gets dangerously low after an order has been placed, material managers are often willing to expedite the order at extra fixed and/or variable costs. This paper proposes a single-item continuous-review order expediting inventory policy, which can be considered as an extension of ordinary (s, Q) models. Besides the two usual operational parameters: reorder point s and order quantity Q, it consists of a third parameter called the expedite-up-to level R. If inventory falls below R at the end of the manufacturing lead-time, the buyer can request the upstream supplier to deliver part of an outstanding order via a fast transportation mode. The amount expedited will raise inventory to R, while the remaining order is delivered via a slow (regular) supply mode. Simple procedures are developed to obtain optimal operational parameters. Computational results show that the proposed policy can save large costs for a firm if service level is high, demand variability is large, the extra cost for expediting is small, or the manufacturing lead-time is long.

1. Introduction

Expedited shipments are quite common in practice. When the inventory level of an item gets dangerously low after an order has been issued, material managers are often willing to expedite the order at extra fixed and/or variable costs. By employing a fast transportation mode (e.g., by air), the buyer can have an outstanding order arrive earlier than planned. The use of the Internet has greatly enhanced the ability of a buyer to track the status of purchase orders. Through the online updating of orders or direct email communication (or simply phone calls), the buyer may request the upstream supplier to deliver an order or part of it via a fast supply mode if necessary.

Besides order expediting, a commonly suggested course of action when faced with an urgent situation is to place a faster or emergency order in addition to an already issued regular order (e.g., Chiang, 2003). Although both options enable a firm to acquire materials in time and achieve the service level promised to customers, expediting of existing orders does not create new orders that must be monitored or managed. As Perona and Miragliotta (2004) emphasized, by restricting the placement of fast orders, a firm may reduce its operational complexity. However, to our knowledge, research on the issue of order expediting in inventory systems is limited. Allen and D'Esopo (1968) proposed an (s, Q) policy with a third operational parameter called the expediting level. They suggested that when inventory drops to this level, an outstanding order is shipped after a short period. Bookbinder and Cakanyildirim (1999) considered a deterministic-demand continuous review model where lead-time is made endogenous through an expediting factor. Also, Lawson and Porteus (2000) and Arslan et al. (2001) respectively investigated multi-echelon and make-to-order inventory systems with expediting, and Bregman (2009) suggested a heuristic for solving the dynamic probabilistic project expediting problem.

In this paper, we propose a new single-item inventory policy with order expediting. We consider continuous review systems where lead-time consists of two components: manufacturing lead-time and delivery lead-time. The former is the time needed for the supplier to manufacture the quantity ordered which includes the set-up time, job processing time and waiting time; the latter is the time needed to deliver the finished products. The former can be either constant or random depending on how congestion at the manufacturing facility affects the job waiting time; the latter is a deterministic interval, though it can take on two different values corresponding respectively to the regular and fast transportation modes. Order expediting can occur due to the use of a fast mode and/or the effort to reduce the job waiting time. In this paper, we assume the manufacturing lead-time to be constant, and focus on the use of a fast mode to expedite supply. Note that Cakanyildirim et al. (2000) studied a deterministic-demand inventory model where lead-time is also made up of two periods: one for materials handling, waiting and set-up and the other for the manufacturing of a lot size. The time needed for delivery of a finished lot was not considered in their model.

* Tel.: +886 35712121 57166.
E-mail address: cchiang@mail.nctu.edu.tw

0377-2217/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2009.08.012

The proposed policy can be considered as an extension of ordinary (s, Q) models. As an (s, Q) model operates, the proposed policy places an order for Q units whenever inventory drops to s; in addition, it expedites part of an outstanding order if inventory falls below R at the end of the manufacturing lead-time. The amount expedited will raise inventory to R, which is thus called the *expedite-up-to* level, while the remaining order is delivered via a regular supply mode. Consequently, an outstanding order is split into two deliveries. The use of order splitting to reduce the shortage probability and/or inventory costs has received much attention in the inventory literature (see, e.g., Sculli and Wu, 1981; Kelle and Silver, 1990; Kelle and Miller, 2001; Lau and Zhao, 1993; Chiang and Benton, 1994; Chiang and Chiang, 1996). Kiesmuller et al. (2005) investigated the use of both order splitting and expediting in a periodic review setting, and showed that an intelligent choice of supply modes could save considerable costs for a firm.

The proposed policy minimizes the expected total cost per unit time, subject to a service level constraint. Both the no-shortage probability and fill rate service measures are used in this research. We develop simple procedures to compute optimal operational parameters. Computational results indicate that the proposed policy yields significant cost savings, especially if service level is high, demand variability is large, the extra (fixed and/or variable) costs for expediting are small, or the manufacturing lead-time is long.

Notice that Chiang (2002) devised a different order expediting policy, where when inventory falls to an expediting level at the end of manufacturing, an entire order is sent via a fast supply mode. Simple thought reveals that if expediting of orders entails large costs, it is economical to expedite only part of an order. Also, Chiang's model used only the no-shortage probability service measure. Moreover, expediting supply may not be always possible. Krishnamoorthy and Raju (1998a,b) introduced the use of "local purchase" of the item which runs out of stock. They considered three different types of local purchases: for the first two cases, a retailer purchases from some other nearby shop for one or s units of the item when a demand arises during the stock out period; for the third case, the retailer makes a local purchase of S units and also cancels the outstanding order.

The rest of this paper proceeds as follows. In Section 2, we review ordinary (s, Q) models. In Section 3, we propose a continuous-review order expediting policy. Section 4 presents some computational results. Section 5 concludes this research.

2. Review of (s, Q) models

In an (s, Q) model, when the inventory position (i.e., inventory on hand + inventory on order – backlog) drops to s, an order for Q units is placed and received after a lead-time L. The literature on (s, Q) models is huge and can be divided into two groups: one with an exact cost formulation and the other with an approximate cost formulation. This paper will employ both of the exact and approximate cost expressions. Note that fuzzy theory can also be used to develop an (s, Q) model (e.g., Handfield et al., 2009).

Let A be the fixed cost of ordering and/or set-up for the manufacturing of an order, c be the unit item cost, r be the inventory holding cost rate per unit time, D be the average demand per unit time, Y_T be the demand during a time interval of length T, and $f_T(\cdot)$ be its probability density function. Demand not filled immediately is backlogged and assumed to be continuous for convenience of notation. Under constant lead-time but otherwise fairly general conditions, the exact average inventory level is

$$I^{\text{exact}} = (1/Q)\int_s^{s+Q}\left[\int_0^y (y-\xi)f_L(\xi)d\xi\right]dy \tag{1}$$

(e.g., Zheng, 1992), where the inventory position in steady state is uniformly distributed on $(s, s+Q]$. On the other hand, the approximate average inventory level is

$$I^a = s - DL + .5Q \tag{2}$$

(Hadley and Whitin, 1963, pp. 162–167); I^a is accurate when the expected number of shortages per cycle is quite small, or the shortage probability is sufficiently low (e.g., Lau and Lau, 2002). It is assumed that Q is large enough for the probability of more than one outstanding order, i.e., $Pr(Y_L > Q)$, to be negligible or approximately zero (Hadley and Whitin, 1963, pp. 162 and 201). This assumption is often used in the literature of continuous review models (e.g., Kelle and Silver, 1990; Bookbinder and Cakanyildirim, 1999). The average total cost per unit time is expressed by

$$TC(s, Q) = AD/Q + rcI, \tag{3}$$

where I is given by (1) or (2). Notice that the shortage cost is not included, because it is rarely out-of-pocket and quantifying it is difficult in practice. Instead, management often specifies a desired service level. Service level is usually defined as the probability of no shortage during an order cycle, denoted by α, or the fraction of demand filled directly from stock (i.e., the fill rate), denoted by β (see, e.g., Silver et al., 1998 for other service measures). In the former case, one minimizes (3) subject to

$$\Pr(Y_L > s) \leqslant 1 - \alpha. \tag{4}$$

In the latter case, one minimizes (3) subject to

$$\int_s^{\infty}(\xi - s)f_L(\xi)d\xi \leqslant (1-\beta)Q. \tag{5}$$

To find the optimal solution s^* and Q^*, we observe that as s becomes lower, I is smaller, whether I is given by (1) or (2). Thus if a α service level is applied, s^* is obtained by finding the smallest (integer, for discrete demand) value of s satisfying (4), and Q^* is obtained by the EOQ formula

$$Q = (2DA/rc)^{0.5} \tag{6}$$

if I is given by (2), or obtained by directly minimizing (3) if I is given by (1).

If a β service level is applied, a Lagrangian formulation including (3) and (5) could be used and by solving the first-order condition two equations involving s and Q are formed (e.g., Chiang and Chiang, 1996; Platt et al., 1997). However, the usual method of iterative substitution for finding s^* and Q^* does not guarantee convergence of solutions. In this paper, we suggest that one finds the smallest value of s satisfying (5) for a given Q, i.e., s is set as a function of Q, and then minimizes (3) over Q (e.g., using a Fortran program) whether I is given by (1) or (2) (see, e.g., Tijms and Groenevelt, 1984 for a similar approach that determines s given Q in service-constrained inventory systems).

3. An order expediting inventory control policy

Assume that the lead-time L consists of two periods: the supplier's manufacturing lead-time M and delivery lead-time N. While M is constant, N can be shortened to G (i.e., $G < N$) if a fast supply mode (e.g., by air) is used. Assume without loss of generality that the supplier arranges the order's delivery at the end of M. (If the supplier has to arrange delivery at a time earlier than the end of M, the remaining manufacturing lead-time becomes part of G or N.) If inventory is dangerously low at the end of M, the buyer can request the upstream supplier to deliver part of an order via a fast mode. A natural question arises: how much of the outstanding order is expedited? The amount expedited cannot be too large in order to save procurement costs, while it cannot be too small because the on-hand inventory at the end of M plus the amount expedited are expected to satisfy backorders (if any) and meet demand during the upcoming time interval N. Thus, it seems that there is an inventory level (which is less than s), denoted by R, such that if inventory falls to $H < R$ at the end of M, the amount expedited equals $R - H$ while the amount not expedited is $Q - R + H$ (see Fig. 1). We call R the *expedite-up-to* level.

Hence, we propose a new continuous review policy that consists of three decision variables: s, Q, and R, or equivalently, Δ, Q, and R where $\Delta \equiv s - R > 0$. Let A' be the fixed cost of expediting and c' be the *incremental* (out-of-pocket) unit expediting cost. We next derive the average total cost per unit time for the proposed policy. Notice that the probability of expediting is $Pr(Y_M > \Delta)$ and the average amount expedited is

$$Q^E \equiv \int_{\Delta}^{\Delta+Q} (\zeta - \Delta) f_M(\zeta) d\zeta + Q \int_{\Delta+Q}^{\infty} f_M(\zeta) d\zeta, \tag{7}$$

which is a function of Δ and Q. Since $Pr(Y_L > Q)$ is assumed to be negligible, $Pr(Y_M \geqslant \Delta + Q)$ is approximately zero due to $M < L$ and $\Delta > 0$ and thus

$$Q^E \approx \int_{\Delta}^{\infty} (\zeta - \Delta) f_M(\zeta) d\zeta, \tag{7'}$$

which depends on Δ only. Also, Chiang and Chiang (1996) showed that if an order is split into two deliveries in a continuous review system, the average cycle stock is reduced by the proportion of Q shipped in the second delivery times the average demand during the inter-arrival time between the two deliveries (see also Lau and Zhao, 1993 in a two-supplier setting). Here, part of Q may be expedited (the expected proportion of Q expedited is Q^E/Q) as opposed to delayed in Chiang and Chiang and the time between the expedited and normal shipments is $N - G$. Thus, the average cycle stock is increased by $(Q^E/Q)D(N - G)$. It follows from (3) that the average total cost per unit time is written by

$$\begin{aligned} TC(\Delta, R, Q) &= AD/Q + (A'D/Q)Pr(Y_M > \Delta) \\ &\quad + c'DQ^E/Q + rc[I + (Q^E/Q)D(N - G)] \\ &= (D/Q)\{A + A'Pr(Y_M > \Delta) \\ &\quad + [c' + rc(N - G)]Q^E\} + rcI, \end{aligned} \tag{8}$$

where I is given by (1) or (2) with $s = \Delta + R$. Notice that the inventory position in steady state is still uniformly distributed on $(s, s + Q]$ if using (1) for I, since it remains unchanged with order expediting. Let

$$A^* \equiv \{A + A'Pr(Y_M > \Delta) + [c' + rc(N - G)]Q^E\}. \tag{9}$$

Then, (8) can be written by the following expression, which is of the same form as (3)

$$TC(\Delta, R, Q) = A^* D/Q + rcI \tag{10}$$

If service level is defined as the probability of no shortage during an order cycle, one minimizes (10) subject to

$$\begin{aligned} &\int_0^{\Delta} \left(\int_{\Delta+R-\zeta}^{\infty} f_N(\xi) d\xi \right) f_M(\zeta) d\zeta + \int_{\Delta}^{\Delta+R} \left(\int_{\Delta+R-\zeta}^{\infty} f_G(\xi) d\xi \right) f_M(\zeta) d\zeta \\ &\quad + \int_{\Delta+R}^{\infty} f_M(\zeta) d\zeta + \int_{\Delta}^{\infty} f_M(\zeta) d\zeta \int_R^{\infty} f_N(\xi) d\xi \leqslant 1 - \alpha. \end{aligned} \tag{11}$$

The first term of (11) is for the situation where expediting does not occur, while the second and third terms are the shortage probability before the expedited shipment arrives and the last term is the shortage probability before the remaining order arrives.

To find the optimal combination of Δ, Q, and R, we first state the following lemma.

Lemma 1. *For a certain value of Δ, as R decreases, the shortage probability during an order cycle, i.e., the left-hand side of (11), increases.*

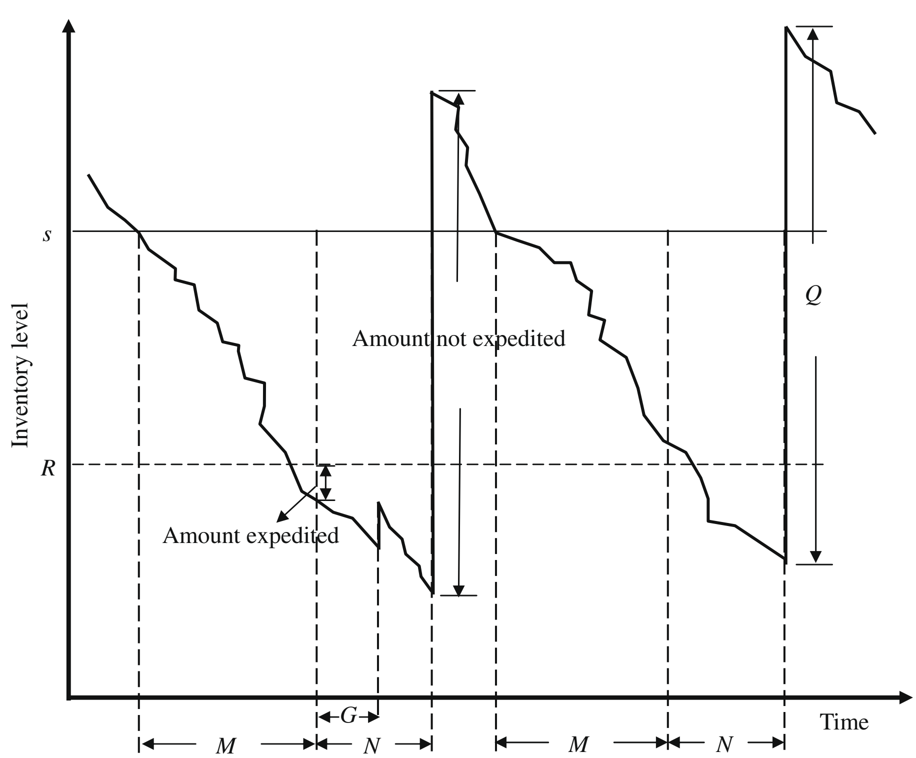

Fig. 1. An order expediting inventory control policy.

Proof. As R decreases, the four terms, except the second one, in the left-hand side of (11) increase. But the inner integral of the second term is less than one; hence, as R decreases, the second and third terms combined increase. □

The above lemma intuitively makes sense. Also, we observe that as R becomes lower, the total cost in (10) is smaller. Thus, for a certain Δ, we find the smallest (integer, for discrete demand) value of R satisfying (11) and Q can be found from (10) by the equation

$$Q = (2DA^*/rc)^{0.5} \tag{12}$$

if I is given by (2), or obtained by directly minimizing (10) if I is given by (1), as in the ordinary no-expediting model. Then, one can perform a simple search on (integer) values of Δ in order to determine the optimal solution.

On the other hand, if service is defined as the fill rate, one minimizes (10) subject to

$$\begin{aligned}&\int_0^{\Delta}\left(\int_{\Delta+R-\zeta}^{\infty}[\xi-(\Delta+R-\zeta)]f_N(\xi)d\xi\right)f_M(\zeta)d\zeta \\ &+\int_{\Delta}^{\Delta+R}\left(\int_{\Delta+R-\zeta}^{\infty}[\xi-(\Delta+R-\zeta)]f_G(\xi)d\xi\right)f_M(\zeta)d\zeta+\int_{\Delta+R}^{\infty}[\zeta \\ &-(\Delta+R)+DG]f_M(\zeta)d\zeta+\int_{\Delta}^{\infty}f_M(\zeta)d\zeta\int_R^{\infty}(\xi-R)f_N(\xi)d\xi \\ &\leqslant(1-\beta)Q.\end{aligned} \tag{13}$$

The first term in the left-hand side of (13) is the average backorder if expediting does not occur. The second and third terms are the average amount backlogged before the expedited shipment arrives, while the last term is the expected backorder before the remaining order arrives (notice the same range of integrals in (11) and (13)).

To find the optimal combination of Δ, Q, and R, we suggest that one treats (10) as a function of Δ, as in the case of α service levels, and follows the solution procedure described above for the ordinary (s, Q) models. In other words, for a certain Δ, one finds the smallest value of R satisfying (13) for a given Q, due to the following lemma, and then minimizes (10) over Q whether I is given by (1) or (2).

Lemma 2. *Given Δ and Q, as R decreases, the fraction of demand not filled directly from stock, i.e., the left-hand side of (13) divided by Q, increases.*

Proof. Omitted (similar to that of Lemma 1). □

Lemma 2 echoes Lemma 1 and makes intuitive sense as well. Again, one carries out a simple search on values of Δ to obtain the optimal solution.

4. Computational results

Consider the base case: unit time = 1 year = 250 (working) days, $D = 500$ units/year, $A = \$16$, $A' = \$4$, $r = 25\%$/year, $c = \$4.0$/unit, $c' = \$0.5$/unit, $L = .1$ years, $M = .08$ years, $N = .02$ years, $G = .004$ years, α (or β) = 99.9%. Demand is Poisson with mean DT for a period of length T. Note that the average inventory should be added by one-half if demand is Poisson and one uses I^a for I (Hadley and Whitin, 1963, pp. 186–187).

Consider first employing Chiang's model (2002) that expedited a whole outstanding order if necessary and used only the α service level; the cost saving is only 0.80%. If the proposed policy is used, we find that $\Delta = 49, R = 18, Q = 128$, and the savings is 2.92%. Noticing $Pr(Y_M \geqslant \Delta + Q)$ in (7) for the base case (which is approximately zero), we see that it is very unlikely that one has a whole order shipped via the fast mode. This result seems to be observed throughout the computation. If a β service level is applied, then $\Delta = 49, R = 11, Q = 130$, and the savings is 0.98%. More results are shown in Tables 1–3. It is clear that the proposed policy is superior to Chiang's model. Also, the higher the service level, the larger the percentage savings given by the proposed policy. This result is expected, for high service levels justify the expediting of part of an outstanding order.

Notice that we used both I^{exact} and I^a to compute the optimal solutions and their costs. The optimal operational parameters found are the same for all problems solved in Tables 1–8, which is probably because we have set service at high levels, so that it is worthwhile to expedite part of an outstanding order. For a given optimal solution, the expected total costs computed by using I^{exact} and I^a differ by no more than \$0.01 (only the exact costs are shown in the tables), except for three problems (where the expected costs differ by about \$0.02 or \$0.03). This result agrees with Lau and Lau (2002), who found that I^a is more accurate than the inventory literature implies.

In Tables 4 and 5, we investigate the effect of the ratios A'/A and c'/c on the performance of the proposed policy. It can be seen that

Table 1
Effect of the α service level on performance of Chiang's model, 2002; $D = 500$ units/year, 1 year = 250 days, $A = \$16$, $A' = \$4$, $r = 25\%$/year, $c = \$4.0$/unit, $c' = \$0.5$/unit, $L = .1$ years, $M = .08$ years, $N = .02$ years, $G = .004$ years.

α (%)	Expediting is not allowed			Expediting is allowed				% Savings
	s	Q	$TC(s,Q)$	s	E[a]	Q	$TC(s,E,Q)$	
95.0	62	126	138.99	Expediting is not economical				
99.0	67	126	143.99	Expediting is not economical				
99.9	73	126	149.99	71	12	127	148.79	0.80
99.99	78	126	154.99	74	16	127	152.20	1.80

[a] E is called the expediting level in Chiang's model.

Table 2
Effect of the α service level on performance of the proposed policy; $D = 500$ units/year, 1 year = 250 days, $A = \$16$, $A' = \$4$, $r = 25\%$/year, $c = \$4.0$/unit, $c' = \$0.5$/unit, $L = .1$ years, $M = .08$ years, $N = .02$ years, $G = .004$ years.

α (%)	Expediting is not allowed			Expediting is allowed						% Savings
	s	Q	$TC(s,Q)$	Δ	R	Q	Q^E	A^*	$TC(\Delta,R,Q)$	
95.0	62	126	138.99	49	11	128	.252	16.41	138.61	0.27
99.0	67	126	143.99	48	15	129	.345	16.55	142.14	1.28
99.9	73	126	149.99	49	18	128	.252	16.41	145.61	2.92
99.99	78	126	154.99	48	22	129	.345	16.55	149.14	3.77

Table 3
Effect of the β service level on performance of the proposed policy; $D = 500$ units/year, 1 year = 250 days, $A = \$16$, $A' = \$4$, $r = 25\%$/year, $c = \$4.0$/unit, $c' = \$0.5$/unit, $L = .1$ years, $M = .08$ years, $N = .02$ years, $G = .004$ years.

β (%)	Expediting is not allowed			Expediting is allowed						% Savings
	s	Q	$TC(s,Q)$	Δ	R	Q	Q^E	A^*	$TC(\Delta,R,Q)$	
99.0	54	130	131.07	Expediting is not economical						
99.9	63	126	139.99	49	11	130	.252	16.41	138.62	0.98
99.99	69	132	146.11	47	16	131	.464	16.72	142.81	2.26

Table 4
Effect of the ratio A'/A on performance of the proposed policy; $D = 500$ units/year, 1 year = 250 days, $A = \$16$, $r = 25\%$/year, $c = \$4.0$/unit, $c' = \$0.5$/unit, $L = .1$ years, $M = .08$ years, $N = .02$ years, $G = .004$ years, and α or $\beta = 99.9\%$.

A'	No-shortage probability constraint						% Savings	Fill-rate constraint						% Savings
	Δ	R	Q	Q^E	A^*	TC		Δ	R	Q	Q^E	A^*	TC	
\$16	52	17	128	.091	16.50	147.94	1.37	55	7	134	.029	16.17	139.84	0.11
8	49	18	129	.252	16.69	146.70	2.19	51	10	128	.129	16.38	139.48	0.37
4	49	18	128	.252	16.41	145.61	2.92	49	11	130	.252	16.41	138.62	0.98
2	47	19	128	.464	16.48	144.87	3.41	47	12	128	.464	16.48	137.88	1.51
0	45	20	128	.807	16.42	143.63	4.24	42	14	137	1.663	16.86	136.54	2.46

Table 5
Effect of the ratio c'/c on performance of the proposed policy; $D = 500$ units/year, 1 year = 250 days, $A = \$16$, $A' = \$4$, $r = 25\%$/year, $c = \$4.0$/unit, $L = .1$ years, $M = .08$ years, $N = .02$ years, $G = .004$ years, and α or $\beta = 99.9\%$.

c'	No-shortage probability constraint						% Savings	Fill-rate constraint						% Savings
	Δ	R	Q	Q^E	A^*	TC		Δ	R	Q	Q^E	A^*	TC	
\$2.0	49	18	130	.252	16.79	147.08	1.94	54	8	127	.043	16.14	139.55	0.32
1.0	49	18	129	.252	16.54	146.10	2.59	49	11	130	.252	16.54	139.11	0.63
0.5	49	18	128	.252	16.41	145.61	2.92	49	11	130	.252	16.41	138.62	0.98
0	47	19	128	.464	16.49	144.90	3.39	47	12	128	.464	16.48	137.90	1.50

Table 6
Effect of G on performance of the proposed policy; $D = 500$ units/year, 1 year = 250 days, $A = \$16$, $A' = \$4$, $r = 25\%$/year, $c = \$4.0$/unit, $c' = \$0.5$/unit, $L = .1$ years, $M = .08$ years, $N = .02$ years, and α or $\beta = 99.9\%$.

G	No-shortage probability constraint						% Savings	Fill-rate constraint						% Savings
	Δ	R	Q	Q^E	A^*	TC		Δ	R	Q	Q^E	A^*	TC	
.016	52	19	127	.091	16.16	148.61	0.92	Expediting is not economical						
.008	48	19	129	.345	16.55	146.13	2.57	51	10	134	.129	16.22	139.03	0.69
.004	49	18	128	.252	16.41	145.61	2.92	49	11	130	.252	16.41	138.62	0.98
.001	46	19	130	.616	16.92	145.58	2.94	49	11	128	.252	16.41	138.61	0.99

Table 7
Effect of M on performance of the proposed policy (with α service level); $D = 500$ units/year, 1 year = 250 days, $A = \$16$, $A' = \$4$, $r = 25\%$/year, $c = \$4.0$/unit, $c' = \$0.5$/unit, $N = .02$ years, $G = 0.004$ years, and $\alpha = 99.9\%$.

M	Expediting is not allowed			Expediting is allowed						% Savings
	s	Q	$TC(s,Q)$	Δ	R	Q	Q^E	A^*	$TC(\Delta,R,Q)$	
0.02	35	126	141.99	17	17	127	.028	16.07	141.27	0.51
0.04	48	126	144.99	27	18	128	.141	16.28	143.10	1.30
0.08	73	126	149.99	49	18	128	.252	16.41	145.61	2.92
0.16	121	126	157.99	95	18	128	.199	16.28	151.10	4.36

Table 8
Effect of M on performance of the proposed policy (with β service level); $D = 500$ units/year, 1 year = 250 days, $A = \$16$, $A' = \$4$, $r = 25\%$/year, $c = \$4.0$/unit, $c' = \$0.5$/unit, $N = .02$ years, $G = 0.004$ years, and $\beta = 99.9\%$.

M	Expediting is not allowed			Expediting is allowed						% Savings
	s	Q	$TC(s,Q)$	Δ	R	Q	Q^E	A^*	$TC(\Delta,R,Q)$	
0.02	27	141	134.74	18	9	131	.013	16.03	134.21	0.39
0.04	39	142	136.84	27	11	128	.141	16.28	136.10	0.54
0.08	63	126	139.99	49	11	130	.252	16.41	138.62	0.98
0.16	108	127	144.99	90	12	132	.629	16.81	142.19	1.93

the smaller the ratio A'/A or c'/c, the more cost-effective the proposed policy and the savings could be larger than 4%. This result is also expected without explanation.

Next, we examine the effect of G and M. It is seen from Table 6 that as G is shorter, the proposed policy becomes more attractive. Moreover, it appears from Tables 7 and 8 that the longer the manufacturing lead-time M, the more cost-effective the proposed policy. This result is intuitively reasonable, since with a longer M the lead-time demand becomes more volatile and expediting part of an order at the end of M is more beneficial.

Finally, we consider a high-volume item for which demand is normally distributed. We find that as demand variability is larger, the proposed policy yields a larger percentage savings (detailed computational results are available from the author upon request), which also intuitively makes sense.

5. Conclusion

This paper proposes an order expediting inventory control policy. It consists of three operational parameters: s, Q, and R, or equivalently, Δ, Q, and R where Δ equals $s - R$, such that if inventory falls below R at the end of the manufacturing lead-time, part of an outstanding order is delivered via a fast supply mode. We derive the average total cost per unit time and minimize it subject to a service level constraint. Either the fill rate or the probability-of-no-shortage-during-a-cycle service measure is utilized. We show that the average total cost is a function of Δ. Thus, one can perform a simple search on Δ to obtain optimal operational parameters. Computational results show that the proposed policy is worthwhile to use, especially if service level is high, demand variability is large, the extra expediting cost is small, or the manufacturing lead-time is long.

References

Allen, S.G., D'Esopo, D.A., 1968. An ordering policy for stock items when delivery can be expedited. Operations Research 16, 880–883.

Arslan, H., Ayhan, H., Olsen, T.L., 2001. Analytic models for when and how to expedite in make-to-order systems. IIE Transactions 33 (11), 1019–1029.

Bookbinder, J.H., Cakanyildirim, M., 1999. Random lead times and expedited orders in (Q, r) inventory systems. European Journal of Operational Research 115 (2), 300–313.

Bregman, R.L., 2009. A heuristic procedure for solving the dynamic probabilistic project expediting problem. European Journal of Operational Research 192 (1), 125–137.

Cakanyildirim, M., Bookbinder, J.H., Gerchak, Y., 2000. Continuous review inventory models where random lead time depends on lot size and reserved capacity. International Journal of Production Economics 68 (3), 217–228.

Chiang, C., 2002. Ordering policies for inventory systems with expediting. Journal of the Operational Research Society 53 (11), 1291–1295.

Chiang, C., 2003. Optimal replenishment for a periodic review inventory system with two supply modes. European Journal of Operational Research 149 (1), 229–244.

Chiang, C., Benton, W.C., 1994. Sole sourcing versus dual sourcing under stochastic demands and lead times. Naval Research Logistics 41 (5), 609–624.

Chiang, C., Chiang, W.C., 1996. Reducing inventory costs by order splitting in the sole sourcing environment. Journal of the Operational Research Society 47 (3), 446–456.

Hadley, G., Whitin, T.M., 1963. Analysis of Inventory Systems. Prentice-Hall, Englewood Cliffs, NJ.

Handfield, R., Warsing, D., Wu, X., 2009. (Q, r) inventory policies in a fuzzy uncertain supply chain environment. European Journal of Operational Research 197 (2), 609–619.

Kelle, P., Miller, P.A., 2001. Stockout risk and order splitting. International Journal of Production Economics 71 (1–3), 407–415.

Kelle, P., Silver, E.A., 1990. Safety stock reduction by order splitting. Naval Research Logistics 37 (5), 725–743.

Kiesmuller, G.P., de Kok, A.G., Fransoo, J.C., 2005. Transportation mode selection with positive manufacturing lead time. Transportation Research Part E 41 (6), 511–530.

Krishnamoorthy, A., Raju, N., 1998a. (s, S) inventory with lead time-the N-policy. International Journal of Information and Management Sciences 9 (4), 45–52.

Krishnamoorthy, A., Raju, N., 1998b. N-policy for (s, S) perishable inventory system with positive lead time. Korean Journal of Computational and Applied Mathematics 5 (1), 253–261.

Lau, A.H., Lau, H., 2002. A comparison of different methods for estimating the average inventory level in a (Q, R) system with backorders. International Journal of Production Economics 79 (3), 303–316.

Lau, H., Zhao, L., 1993. Optimal ordering policies with two suppliers when lead times and demands are all stochastic. European Journal of Operational Research 68 (1), 120–133.

Lawson, D.G., Porteus, E.L., 2000. Multistage inventory management with expediting. Operations Research 48 (6), 878–893.

Perona, M., Miragliotta, G., 2004. Complexity management and supply chain performance assessment: A field study and a conceptual framework. International Journal of Production Economics 90 (1), 103–115.

Platt, D.E., Robinson, L.W., Freund, R.B., 1997. Tractable (Q, R) heuristic models for constrained service levels. Management Science 43 (7), 951–965.

Sculli, D., Wu, S.Y., 1981. Stock control with two suppliers and normal lead times. Journal of the Operational Research Society 32 (11), 1003–1009.

Silver, E.A., Pyke, D.F., Peterson, R., 1998. Inventory Management and Production Planning and Scheduling. John Wiley and Sons, New York.

Tijms, H.C., Groenevelt, H., 1984. Simple approximations for the reorder point in periodic and continuous review (s, S) inventory systems with service level constraints. European Journal of Operational Research 17 (2), 175–190.

Zheng, Y.S., 1992. On properties of stochastic inventory systems. Management Science 38 (1), 87–103.