A New and Simple Policy for the Continuous Review Lost Sales Inventory Model

Martin I. Reiman
Bell Labs, Lucent Technologies
Murray Hill, NJ

March 18, 2004

Abstract

We consider the single item, single location, continuous review inventory model with lost sales. Demand is assumed to arrive as a Poisson process and lead times are assumed to be fixed. The optimal policy for this system is not known. We introduce a new and simple policy: Every τ time units order an item (τ is a positive number). This policy is compared to the standard base-stock policy and is shown (numerically) to outperform it for certain parameter values. We then carry out an asymptotic analysis for both policies as the lost sales penalty and lead time grow large. This analysis, motivated by a conjecture based on our numerical results, provides reasonably explicit expressions for various quantities of interest, such as optimal costs and optimal safety stock levels. The optimal safety stock for both policies is shown to grow as the square-root of the lead time, and may be positive or negative.

1. Introduction

Karlin and Scarf (1958) described two periodic review inventory models with fixed positive lead times: Model I with lost sales, and Model II with backordering. In the backordering case they proved the optimality of a base-stock policy, whereas for the lost sales model they found that the optimal policy does not typically have a simple form. In general, the optimal policy for the lost sales case may depend on the entire 'pipeline' of items ordered but not yet delivered. Such a policy would be difficult to both compute and implement. Despite its practical importance (as Karlin and Scarf and others have noted, the lost sales model also applies to systems that employ expediting) the lost sales model has received relatively little attention in the research literature.

Our focus in this paper is a continuous review version of Model I. Customer demand arrives as a Poisson process. If the inventory level is positive when a demand arrives, it is satisfied, otherwise the demand is lost. Inventory replenishment orders arrive into inventory after a fixed lead time. Items in inventory incur (linear) holding costs, and lost demand incurs a loss penalty. The goal is to minimize the long run average cost of operation.

As in the periodic review case the continuous review lost sales model has no simple optimal policy (cf. Hill 1999), and most related work has focused on simpler policies, most notably base-stock. As recognized by Karush (1957), a base-stock policy in the continuous review lost sales model leads to an Erlang loss system, so that for a fixed base-stock level the cost can be determined using the celebrated Erlang B formula. (An interesting historical note: Karush credits Philip Morse for this insight.) Furthermore, the convexity properties of the Erlang B formula give rise to a simple search procedure for the optimal base-stock level.

Of course, the policy found by the above search is optimal among base-stock policies, but typically not among all policies. Recognizing the apparent difficulty of determining the optimal policy or analyzing the performance of policies that are more complicated than base-stock, we take a different tack in this paper. We introduce an alternate policy that is easy to implement, analyze, and optimize, and show that it sometimes (depending on the system parameters) achieves a lower cost than the optimal base-stock policy. The policy we consider has the following simple form: Place a replenishment order every τ time units. This policy, which we call constant order interval, or COI, leads to a D/M/1 queue for performance evaluation, and the optimal τ can easily be found.

This policy is 'open-loop' — the ordering 'decisions' do not depend on the demand process. This is in contrast to the base-stock policy, where orders are triggered by (successfully met) demands.

We are not necessarily advocating the use of this policy in a real system: an open-loop policy can lead to disaster in a context where parameters (such as demand arrival rate) change or have been incorrectly estimated. On the other hand, if one has a system with parameters for which the best constant order interval policy beats the best base-stock policy, then perhaps some modified version of COI, which incorporates a certain level of oversight, should be considered.

As indicated above, the performance analysis of the two policies that we consider follow from standard results in queueing theory. The optimization within each policy type is straightforward, but the expressions for the resulting minimum costs are not amenable to analytical comparison. We first carry out a numerical comparison of the policies. This comparison (carried out in Section 3) shows that neither policy majorizes the other: there are parameter values for which the constant order interval policy is better and others for which base-stock is better. The numerical results also give rise to a natural conjecture on the relative minimal costs as the lost sales cost grows large. This conjecture motivates an asymptotic analysis, which we carry out in Section 4. The asymptotic analysis is for large lost sales cost and lead time, and verifies the conjecture based on the numerical results. The asymptotic analysis also provides more explicit expressions for various quantities of interest, which in turn provide additional insights. For example, we find that the 'safety stock' for both policies grows as the square root of the lead time, and may be positive or negative. Furthermore, if the parameters are such that the optimal COI policy has positive safety stock, the optimal base-stock policy also has positive safety stock. (The converse is not true.)

We remark briefly on some related references. The recent text of Zipkin (1999), and the surveys of Porteus (1990) and Lee and Nahmius (1993) contain discussion of the lost sales model and related previous work. In addition to the previously mentioned paper of Karush (1957), a key reference for us is Smith (1977), who also considers the base-stock policy for the continuous review lost sales model. He provides a characterization of the optimal base stock level that proves to be very useful. Smith also carries out an asymptotic analysis under large lost sales cost and lead time. As described in more detail in Section 4, our asymptotic analysis is a refinement of his.

There are also papers that consider sub-optimal policies that are more complicated than base-stock. In the discounted periodic review case Morton (1969) provided monotonicity results for the value function as well as upper and lower bounds on the optimal order quantity and cost. Using this upper bound, Morton (1971) presented a myopic approximation for the optimal order quantity. For continuous review Hill (1999) showed that given a base-stock policy with base stock level $S \geq 2$, there is some $\delta > 0$ such that the 'modified base-stock' policy that also uses base-stock level S but waits at least δ time units between orders will have lower cost than the base-stock policy. (Thus no base-stock policy with $S \geq 2$ can be optimal.) Based on Hill (1999), but for the periodic review case, Johansen (2001) introduced a modified base-stock policy specified by the pair (S, t). Here again S is the base-stock level, and t is the minimum number of periods that must elapse between orders. All of these policies are more difficult to analyze, optimize and implement than the two simple policies considered in this paper.

The rest of the paper is organized as follows. In Section 2 the mathematical model is more precisely specified and the analysis (and optimization) of the two policies is presented. A numerical comparison of the two policies is provided in Section 3. Section 4 contains an asymptotic analysis under large loss penalty (and large lead time) for both policies.

2. The Mathematical Model

Customer demand arrives as a Poisson process with rate λ. Let $I(t)$ denote the number of items in inventory at epoch t. If $I(t-) > 0$ at the moment of a demand arrival, that demand is met, and $I(t+) = I(t-) - 1$. If $I(t-) = 0$, the demand is lost and a loss cost of $b > 0$ is incurred by the system. The replenishment lead time is $L > 0$. Thus, if a replenishment order is placed at epoch t, it will arrive at $t + L$. A linear holding cost of h per unit of inventory per unit of time is also charged. The goal is to minimize the long run average cost per unit of time.

2.1. Base-Stock Policy

Let $O(t)$ denote the number of orders that have been placed with the supplier, but not yet received, at epoch t. Consider the following base-stock policy: for some integer $s > 0$, place orders to keep $N(t) = O(t) + I(t)$ equal to s. Note that, aside from a possible initial transient period if $N(0) > s$, this policy will yield $N(t) = s$ for all t. If $N(0) > s$, then no orders are placed until $N(t)$ falls below s. If $N(0) < s$, an order of size $s - N(0)$ is placed at epoch 0. Otherwise, an order of size 1 is placed whenever a demand is met, and no orders are placed otherwise.

The steady-state behavior of the above system (which is all that we need for determination of long run average costs) is equivalent to the following Erlang B system. There are s servers. Customers arrive as a Poisson process of rate λ. When an arriving customer finds an idle server it enters service, otherwise the customer is lost. Service times are all equal to L. (Idle servers correspond to items in inventory. Customer arrivals correspond to demand arrivals. Busy servers correspond to orders in the pipeline.) Let O_{BS} denote the steady-state number of busy servers in this system. Then

$$p_n \equiv P\{O_{BS} = n\} = \frac{\frac{(\lambda L)^n}{n!}}{\sum_{l=0}^{s} \frac{(\lambda L)^l}{l!}}, \quad 0 \le n \le s\,.$$

The fraction of customers lost is given by p_s. To avoid confusion when we vary s, λ, and L we let

$$B(n,a) = \frac{a^n/n!}{\sum_{l=0}^{n} a^l/l!}$$

for $n > 0$ and $a > 0$. This is the Erlang loss function. (The Erlang loss function gives the loss probability for any lead time distribution with finite mean, see e.g. Wolff 1989.) The average number of idle servers, $\overline{I}_{BS}$, can be found using Little's Theorem. The average number of busy servers, $\overline{O}_{BS}$, is given by (using Little's Theorem)

$$\overline{O}_{BS} = \lambda L(1 - B(s, \lambda L))$$

so that

$$\overline{I}_{BS} = s - \lambda L(1 - B(s, \lambda L))\,.$$

Let $C_{BS}(s)$ denote the long run average cost using base-stock level s. Then

$$C_{BS}(s) = h\left[s - \lambda L(1 - B(s, \lambda L))\right] + \lambda b B(s, \lambda L)\,. \tag{1}$$

With a held fixed, $B(\cdot, a)$ is a decreasing function of its first argument: $B(n+1, a) < B(n, a)$ for $n \ge 0$. Karush (1957) showed that $B(\cdot, a)$ satisfies the convexity property

$$B(n,a) - B(n+1,a) < B(n-1,a) - B(n,a)\,. \tag{2}$$

(The benefit, in terms of reduced blocking, of adding an additional server, is decreasing in n.)

Let

$$s^* = \max\left\{ \begin{array}{c} \arg\min \; C_{BS}(s) \\ s \ge 0 \end{array} \right\}.$$

(When there is more than one minimizing s for $C_{BS}(s)$ the largest such s is chosen for s^*.)

A useful characterization of s^* was provided by Smith (1977) (in equation (7.4)). Translated into our notation and using our largest arg min definition, it is

$$B(s^*, \lambda L) - B(s^*+1, \lambda L) < \frac{h}{\lambda(b + Lh)} \le B(s^*-1, \lambda L) - B(s^*, \lambda L)\,. \tag{3}$$

This provides a simple way to find s^*, and also provides a simple sufficient condition for $s^* \ge 1$. Note that $B(0, \lambda L) = 1$ and $B(1, \lambda L) = \lambda L/(1 + \lambda L)$. Thus, if

$$\frac{1}{1 + \lambda L} \ge \frac{h}{\lambda(b + Lh)}\,, \tag{4}$$

then (3) assures that $s^* \geq 1$. The condition (4) is equivalent to

$$\lambda b \geq h \,. \tag{5}$$

It is straightforward to check the necessity of (5) by comparing the cost with $s = 0$ and $s = 1$: We have $C_{BS}(0) = \lambda b$ and $C_{BS}(1) = (h + \lambda^2 bL)/(1 + \lambda L)$. It is easy to verify that if $\lambda b < h$, then $C_{BS}(0) < C_{BS}(1)$. It is clear that, since $B(s, \lambda L) \to 0$ as $s \to \infty$, (3) always has a solution as long as (5) holds. The minimum cost achievable using a base-stock policy is

$$C_{BS}^* = \min_{s \geq 0} C_{BS}(s) = C_{BS}(s^*) \,.$$

The following result, whose proof is given in the appendix, provides some insight on the effect of the lead time L on the optimal cost.

Lemma 1. *With h, b, and λ held fixed, C_{BS}^* is a continuous and strictly increasing function of L.*

2.2. Constant Order Interval Policy

We propose the following policy: For $\tau > 0$ (we impose $\tau > \lambda^{-1}$ below for stability), place a replenishment order every τ time units. After the system has been in operation for L time units, the pipeline is 'filled' and orders are delivered into inventory every τ time units. Thus $\{I(t), t \geq L\}$ behaves as the number-in-system process in a D/M/1 queue. (Customer demand arrivals correspond to service completions. Demands arriving when $I(t) = 0$ do not change the state of the system.) Let ρ denote the traffic intensity of this D/M/1 queue. Then $\rho = (\lambda\tau)^{-1}$. This system will have a steady-state distribution if and only if $\rho < 1$, which corresponds to $\tau > \lambda^{-1}$. If $\tau \leq \lambda^{-1}$ the average inventory level will be infinite, which will correspond to an infinite long run average cost. So we impose the condition $\tau > \lambda^{-1}$. We also assume, as we did for the base-stock policy, that $\lambda b \geq h$.

By PASTA (the Poisson 'arrivals' here correspond to demand arrivals) the fraction of lost demands is $1 - \rho$. The average number-in-system (= inventory level) $\overline{I}_{COI}$ is obtained from the standard GI/M/1 analysis (c.f. Wolff 1989):

$$\overline{I}_{COI} = \frac{\rho}{1 - \alpha(\rho)} \,, \tag{6}$$

where $\alpha(\rho)$ is the unique root in $(0, 1)$ of

$$z = R(\rho, z) \,, \tag{7}$$

with

$$R(\rho, z) = e^{-(1-z)/\rho} \,. \tag{8}$$

Let $C_{COI}(\tau)$ denote the long run average cost using constant order interval τ. Then, for $\tau > \lambda^{-1}$,

$$C_{COI}(\tau) = h\overline{I}_{COI}([\lambda\tau]^{-1}) + \lambda b(1 - [\lambda\tau]^{-1}) .$$

It is easier to work with ρ than τ, and then translate: $\tau = [\lambda\rho]^{-1}$. Let

$$f(\rho) = h\overline{I}_{COI}(\rho) + \lambda b(1 - \rho) .$$

From (6),

$$\frac{d\overline{I}_{COI}}{d\rho}(\rho) = \frac{1 - \alpha(\rho) + \rho\alpha'(\rho)}{[1 - \alpha(\rho)]^2} , \tag{9}$$

where $\alpha'(\rho) = \frac{d\alpha}{d\rho}(\rho)$. Using (7) we can write

$$\frac{d\alpha}{d\rho}(\rho) = \frac{\partial R}{\partial \rho}(\rho, \alpha(\rho)) + \frac{\partial R}{\partial z}(\rho, \alpha(\rho))\frac{d\alpha}{d\rho}(\rho) . \tag{10}$$

Differentiating (8) yields

$$\frac{\partial R}{\partial \rho}(\rho, z) = \frac{1 - z}{\rho^2} R(\rho, z)$$

and

$$\frac{\partial R}{\partial z}(\rho, z) = \frac{1}{\rho} R(\rho, z) .$$

Substituting these into (10), and noting that $\alpha(\rho) = R(\rho, \alpha(\rho))$, we obtain

$$\frac{d\alpha}{d\rho}(\rho) = \frac{[1 - \alpha(\rho)]\alpha(\rho)}{\rho[\rho - \alpha(\rho)]} . \tag{11}$$

For the D/M/1 queue, $\alpha(\rho) < \rho$ (cf. Rogozin 1966, Theorem 1), so $\alpha'(\rho) > 0$. Substituting (11) into (9) yields

$$\frac{d\overline{I}_{COI}}{d\rho}(\rho) = \frac{\rho}{[1 - \alpha(\rho)][\rho - \alpha(\rho)]} . \tag{12}$$

Differentiating (12) and using (11) we determine that (after some cancellation)

$$\frac{d^2\overline{I}_{COI}}{d\rho^2}(\rho) = \frac{\alpha(\rho)}{[\rho - \alpha(\rho)]^3} . \tag{13}$$

Recall that $\alpha(\rho) < \rho$, so $\frac{d^2\overline{I}_{COI}}{d\rho^2}(\rho) > 0$, and $\overline{I}_{COI}$ is a strictly increasing, strictly convex function of ρ. By the definition of f

$$\frac{df}{d\rho}(\rho) = h\frac{d\overline{I}_{COI}}{d\rho}(\rho) - \lambda b .$$

Let ρ^* denote the unique root of

$$\frac{df}{d\rho}(\rho) = 0 . \tag{14}$$

We show that (14) always has a root (for $\lambda b \geq h$) by showing that $\alpha(\rho)/\rho \to 0$ as $\rho \to 0$ and $\alpha(\rho) \to 1$ as $\rho \to 1$. The uniqueness of the root follows from convexity of f. Because $\alpha(\rho) < \rho$,

$$R(\rho, \alpha(\rho)) \leq e^{(\rho-1)/\rho} \,.$$

Since $e^{(\rho-1)/\rho}/\rho \to 0$ as $\rho \to 0$, $\alpha(\rho)/\rho \to 0$ as $\rho \to 0$. To show that $\alpha(\rho) \to 1$ as $\rho \to 1$ we use the heavy traffic limit for the D/M/1 queue. The classical heavy traffic result of Kingman (1962) (see also e.g. Wolff 1989) for the GI/G/1 queue can be applied here. Specialized to the D/M/1 queue, this result states that

$$(1-\rho)\overline{I}_{COI}(\rho) \to \frac{1}{2} \quad \text{as} \quad \rho \to 1 \,. \tag{15}$$

Thus

$$\frac{(1-\rho)\rho}{1-\alpha(\rho)} \to \frac{1}{2} \quad \text{as} \quad \rho \to 1 \,. \tag{16}$$

From (16) we see immediately that $\alpha(\rho) \to 1$ as $\rho \to 1$.

Thus, defining $\tau^* = [\lambda\rho^*]^{-1}$ we have

$$\tau^* = \underset{\tau > 0}{\arg\min} \, \{C_{COI}(\tau)\} \,,$$

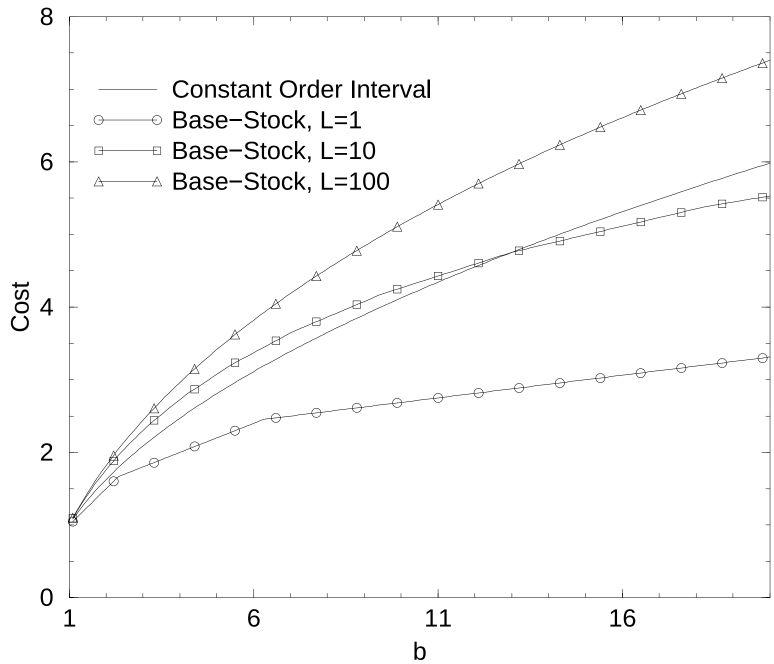
and $C^*_{COI} = C_{COI}(\tau^*) = f(\rho^*)$.

3. Comparison of the Two Policies

The preceding analysis does not provide any hint of the comparative performance of the optimal base-stock policy and the optimal constant order interval policy. So we begin this section by presenting a numerical comparison.

For all of our numerical results we set $\lambda = 1$ and $h = 1$. This is without loss of generality: Setting $\lambda = 1$ fixes the time unit, and setting $h = 1$ fixes the monetary unit. We are left with two parameters, b and L. Recall that C^*_{COI} does not depend on L. In Figure 1 we plot $C^*_{COI}(b)$ and $C^*_{BS}(b, L)$ vs. b for $1 < b \leq 20$ and $L = 1, 10, 100$. (Here and in other places below we explicitly indicate the dependence of C^*_{COI} and C^*_{BS} on some parameters. Precise use will vary and should be clear from the context.) This figure clearly indicates that it is possible to have $C^*_{COI}(b) < C^*_{BS}(b, L)$ for some (b, L): The optimal constant order interval policy sometimes provides a lower cost than the optimal base-stock policy. (The figure also indicates that, for some (b, L), $C^*_{COI}(b) > C^*_{BS}(b, L)$.)

Now that we have established that neither policy majorizes the other, it would be nice to determine, for any given (b, L), which policy provides a lower cost. Lemma 1 provides a key step

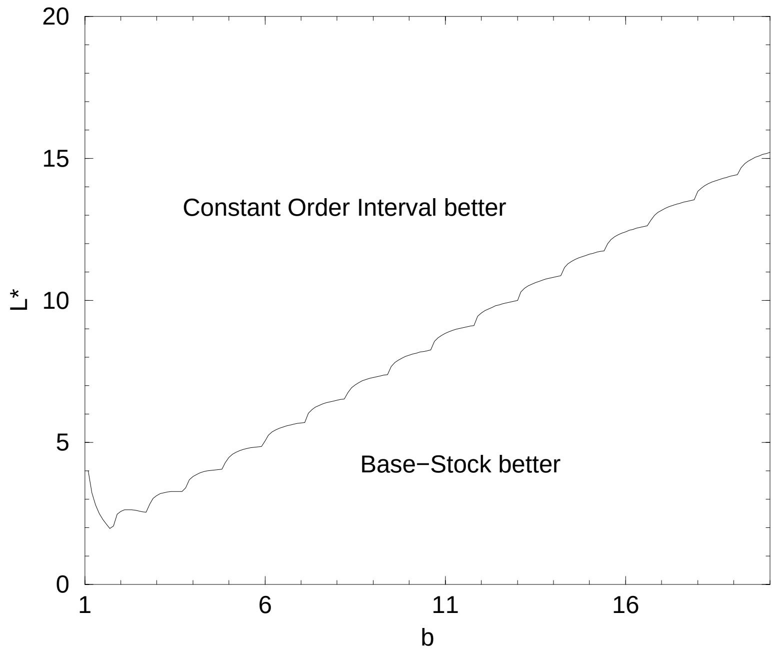


in this comparison. Because C^*_{COI} does not depend on L, by Lemma 1 for each $b > 1$ there is an $L^*(b)$ such that $C^*_{BS}(b, L) < C^*_{COI}(b)$ for $L < L^*(b)$ and $C^*_{BS}(b, L) > C^*_{COI}(b)$ for $L > L^*(b)$. In Figure 2 we plot $L^*(b)$ vs. b for $1 < b \leq 20$. There are two immediate observations one can make from the figure: $L^*(b)$ is not monotone in b, and for large b $L^*(b)$ appears to be nearly linear in b. The latter observation is shown to be true, in an asymptotic sense, in Section 4: We show there that

$$\lim_{b \to \infty} b^{-1} L^*(b) = \gamma^* , \tag{17}$$

where γ^* is approximately 0.69787.

It is worth pondering, at least momentarily, what a reasonable relationship is between b and L, especially given (17). This topic appears to be worthy of further study. Such a study is beyond the scope of this paper, so we provide some simple observations. If the lead time is purely due to transportation, *and* the holding cost during transport is borne by the decision-maker, then we must have $b \geq hL$. Indeed, if $b < hL$ then it would be cheaper to expedite all orders. The picture is a bit murkier when the lead time is caused by production, and the 'holding' costs during production are borne by the producer. We can imagine an item going through a series of production steps, each of which adds some value. We can also imagine that the holding cost (certainly the financial piece) is proportional to the current total value $V(t)$. Then the holding cost incurred by the producer is proportional to $\int_0^L V(s)ds$. If value is added in a reasonably linear fashion (i.e. not all in the



interval $[L - \epsilon, L]$) then the expediting cost should grow proportionally with L. (Again, if it does not, the producer would always use the expediting mode.) Another view of expediting cost — the cost to the producer of the disruption to other production caused by expediting, would require a deeper analysis of the entire production process and all demands placed on it.

4. The Large Loss Penalty Limit

In this section we examine the behavior of ρ^* and s^* as $b \to \infty$, with a particular focus on determining the behavior of L^* as $b \to \infty$. As a byproduct we also obtain long lead time asymptotics ($L \to \infty$) for s^*.

4.1. The Constant Order Interval Policy

We begin with an analysis of the constant order interval policy. We consider the limit $b \to \infty$ with λ and h held fixed. Thus we explicitly indicate the dependence of various quantities on b. (As indicated earlier, L plays no role in this policy.)

Before stating and proving Theorem 1, which provides the asymptotic behavior of the COI policy as $b \to \infty$, we first provide a heuristic derivation of the results. Intuitively, as $b \to \infty$ we want to have the lost sales fraction shrink to zero. This corresponds to $\rho^* \to 1$: heavy traffic. (Recall that we introduced the heavy traffic limit for the D/M/1 queue in Section 2.2.) Anticipating the behavior of $\rho^*(b)$ we assume that

$$\sqrt{b}(1-\rho^*(b)) \to r^*, \qquad 0 < r^* < \infty\,. \tag{18}$$

One might reasonably guess this behavior as follows. With $\rho^*(b)$ satisfying (18) the average inventory holding cost rate, $h\overline{I}_{COI}(\rho^*(b))$, satisfies

$$b^{-1/2} h\overline{I}_{COI}(\rho^*(b)) \to \frac{h}{2r^*} \quad \text{as} \quad b \to \infty\,. \tag{19}$$

The lost sales cost rate, $\lambda b(1-\rho^*(b))$, satisfies

$$b^{-1/2}\lambda b(1-\rho^*(b)) \to \lambda r^* \quad \text{as} \quad b \to \infty\,. \tag{20}$$

Thus, with $\rho^*(b)$ satisfying (18) the holding and lost sales costs are of the same order of magnitude.

Recall that $\rho^*(b)$ is the unique solution of

$$\frac{\rho^*(b)}{[1-\alpha(\rho^*(b))][\rho^*(b)-\alpha(\rho^*(b))]} = \frac{\lambda b}{h}\,. \tag{21}$$

Assuming (18), and using (16) we obtain

$$\frac{\rho^*(b)}{b[1-\alpha(\rho^*(b))][\rho^*(b)-\alpha(\rho^*(b))]} \to \frac{1}{2(r^*)^2}\,,$$

so that, by (21), $r^* = \sqrt{\frac{h}{2\lambda}}$. Combining this with (19) and (20) yields

$$b^{-1/2} C^*_{COI}(b) \to \sqrt{2h\lambda}\,.$$

Parts (i) and (ii) of the following theorem provide a rigorous justification for the above heuristically derived results. Part (iii) of the theorem shows the asymptotic optimality of a natural approximation to $\rho^*(b)$ based on (18). Solving (18) for $\rho^*(b)$ we obtain

$$\rho^*(b) \sim 1 - \frac{r^*}{\sqrt{b}}\,.$$

Using the value of r^* obtained above, it is natural to propose the use of

$$\hat{\rho}(b) = 1 - \sqrt{\frac{h}{2\lambda b}}\,,$$

or $\hat{\tau}(b) = [\lambda\hat{\rho}(b)]^{-1}$ in a system with parameters h, λ and b. Part (iii) of Theorem 1 shows that, to leading order (the optimal cost and cost under $\hat{\rho}(b)$ both grow to infinity), the cost under $\hat{\rho}(b)$ is the same as the cost under $\rho^*(b)$.

Theorem 1. *With h and λ held fixed, the following hold:*

(i) $\rho^*(b) \to 1$ *as* $b \to \infty$ *with*

$$\sqrt{b}(1 - \rho^*(b)) \to \sqrt{\frac{h}{2\lambda}} \quad as \quad b \to \infty\,. \tag{22}$$

(ii)

$$b^{-1/2} C^*_{COI}(b) \to \sqrt{2h\lambda} \quad as \quad b \to \infty\,. \tag{23}$$

(iii) Let $\hat{\rho}(b) = 1 - \sqrt{\frac{h}{2\lambda b}}$ *and* $\hat{\tau}(b) = [\lambda \hat{\rho}(b)]^{-1}$.

Then

$$\frac{C_{COI}(\hat{\tau}(b))}{C^*_{COI}(b)} \to 1 \quad as \quad b \to \infty\,.$$

Proof. We first show that $\rho^*(b) \to 1$ as $b \to \infty$. As $b \to \infty$, the left-hand side of (21) must grow to ∞ as well. The numerator is clearly bounded, so we must have the denominator converge to 0 as $b \to \infty$. This can only occur if $\rho^*(b) \to 1$ as $b \to \infty$.

We rewrite (16) as

$$1 - \alpha(\rho) = 2(1 - \rho) + \delta(\rho)(1 - \rho) \tag{24}$$

where $\delta(\rho) \to 0$ as $\rho \to 1$. Using (24) we can write

$$1 - \alpha(\rho^*(b)) = 2(1 - \rho^*(b)) + \delta_2(b)(1 - \rho^*(b)) \tag{25}$$

with $\delta_2(b) \to 0$ as $b \to \infty$. (We have $\delta_2(b) = \delta(\rho^*(b))$; $\delta(\rho^*(b)) \to 0$ as $b \to \infty$ since $\rho^*(b) \to 1$ as $b \to \infty$.) Let $D(b)$ denote the denominator of the left-hand side of (21). Substituting (25) into $D(b)$ yields

$$D(b) = (2 + \delta_2(b))(1 + \delta_2(b))(1 - \rho^*(b))^2,$$

so that

$$[1 - \rho^*(b)]^{-2} D(b) \to 2 \quad \text{as} \quad b \to \infty\,.$$

Combining this with (21) yields

$$\sqrt{b}(1 - \rho^*(b)) \to \sqrt{\frac{h}{2\lambda}} \quad \text{as} \quad b \to \infty\,,$$

and completes the proof of (i).

We have

$$C^*_{COI}(b) = f(\rho^*(b)) = \frac{h\rho^*(b)}{1 - \alpha(\rho^*(b))} + \lambda b(1 - \rho^*(b))\,.$$

Using (22) and (25),

$$\sqrt{b}[1-\alpha(\rho^*(b))] \to \sqrt{2h/\lambda} \quad \text{as} \quad b \to \infty \,. \tag{26}$$

Thus (using (22) for the second term of $f(\rho^*(b))$) we obtain

$$b^{-1/2} C^*_{COI}(b) \to \sqrt{\frac{h\lambda}{2}} + \sqrt{\frac{h\lambda}{2}} \quad \text{as} \quad b \to \infty \,,$$

which is (ii).

With $\hat{\rho}(b) = 1 - \sqrt{\frac{h}{2\lambda b}}$ and $\hat{\tau}(b) = [\lambda \hat{\rho}(b)]^{-1}$,

$$C_{COI}(\hat{\tau}(\rho)) = f(\hat{\rho}(b)) = \frac{h\hat{\rho}(b)}{1-\alpha(\hat{\rho}(b))} + \lambda b(1-\hat{\rho}(b)) \,.$$

Using (24) we can write

$$1-\alpha(\hat{\rho}(b)) = [2+\delta(\hat{\rho}(b))]\sqrt{\frac{h}{2\lambda}} \,,$$

so that

$$b^{-1/2} C_{COI}(\hat{\tau}(b)) \to \sqrt{2h\lambda} \,. \tag{27}$$

Combining (23) and (27) yields

$$\frac{C_{COI}(\hat{\tau}(b))}{C^*_{COI}(b)} \to 1 \quad \text{as} \quad b \to \infty \,,$$

which is (iii). ∎

4.2. The Base-Stock Policy

We now undertake an asymptotic analysis of the base-stock policy as $b \to \infty$. As in the analysis of the constant order interval policy we hold λ and h fixed. In contrast to the constant order interval policy, however, the performance of the base-stock policy, as well as the choice of s^*, depend on L. Motivated by the desire to prove the asymptotic linearity of L^* in b, we let $L_b = \gamma b$ for $0 < \gamma < \infty$. (Other scalings for L_b vs. b are commented on below, in Section 4.3.)

As in the previous section, we begin with a heuristic derivation of the asymptotic behavior of $s^*(b)$ and $C^*_{BS}(b)$ as $b \to \infty$. As for the COI policy, intuitively, as $b \to \infty$ we want to have the lost sales fraction shrink to zero. The fraction of lost sales is $B(s^*(b), \lambda b \gamma)$. Although it seems clear that we want $s^*(b) \to \infty$ as $b \to \infty$, it is not immediately clear how $s^*(b)$ should relate to b as both grow large.

A key role in our analysis is played by an asymptotic result for $B(n,a)$ due to Jagerman (1974), which states that, if $\beta_a \to \beta$ as $a \to \infty$ with $-\infty < \beta < \infty$, then

$$\lim_{a\to\infty} \sqrt{a} B\left(\lfloor a + \beta_a \sqrt{a} \rfloor, a\right) = \psi(-\beta) \,, \tag{28}$$

where

$$\psi(x) = \frac{\phi(x)}{1-\Phi(x)}\,, \qquad -\infty < x < \infty\,,$$

$\phi(x) = e^{-x^2/2}/\sqrt{2\pi}$, and $\Phi(x) = \int_{-\infty}^{x} \phi(z)dz$.

The limit (28) provides us with the key to determining how $s(b)$ needs to grow to have the holding and lost sales costs be of the same order. In particular, if $s(b) = \lfloor \lambda b\gamma + \beta\sqrt{\lambda b\gamma} \rfloor$ for some $\beta \in (-\infty, \infty)$, then the average inventory holding cost rate $h[s(b) - \lambda b\gamma(1 - B(s(b), \lambda b\gamma))]$ satisfies

$$b^{-1/2} h\left[s(b) - \lambda b\gamma(1 - B(s(b), \lambda b\gamma))\right] \to h\sqrt{\lambda\gamma}(\beta + \psi(-\beta)) \quad \text{as} \quad b \to \infty\,. \tag{29}$$

The lost sales cost rate $\lambda b B(s(b), \lambda b\gamma)$ satisfies

$$b^{-1/2}\lambda b B(s(b), \lambda b\gamma) \to \sqrt{\frac{\lambda}{\gamma}}\psi(-\beta)\,. \tag{30}$$

Combining (29) and (30) we obtain

$$b^{-1/2} C_{BS}(s(b)) \to h\beta\sqrt{\lambda\gamma} + \sqrt{\frac{\lambda}{\gamma}}(h\gamma + 1)\psi(-\beta) \equiv g(\beta) \quad \text{as} \quad b \to \infty\,. \tag{31}$$

It was shown by Morrison and Ramakrishnan (2003) that ψ is strictly increasing and strictly convex. Thus g is also convex. A simple calculation yields

$$\psi'(x) \equiv \frac{d\psi}{dx}(x) = \psi(x)[\psi(x) - x], \quad -\infty < x < \infty\,.$$

Differentiating (31),

$$g'(\beta) = h\sqrt{\lambda\gamma} - \sqrt{\frac{\lambda}{\gamma}}(h\gamma + 1)\psi'(-\beta)\,,$$

so that $g'(\beta) = 0$ if and only if

$$\psi'(-\beta) = \frac{h\gamma}{1 + h\gamma}\,. \tag{32}$$

The following lemma, which is proved in the appendix, verifies that (32) has a unique solution.

Lemma 2. *For any $y \in (0,1)$ there exists a unique $w^*(y)$ such that*

$$\psi'(w^*(y)) = y\,.$$

Let

$$\beta^* = -w^*\left(\frac{h\gamma}{1 + h\gamma}\right)\,. \tag{33}$$

The above derivation indicates that

$$\frac{s^*(b) - \lambda b\gamma}{\sqrt{\lambda b\gamma}} \to \beta^* \quad \text{as} \quad b \to \infty$$

and

$$b^{-1/2}C^*_{BS}(b) \to g(\beta^*) \quad \text{as} \quad b \to \infty\,.$$

These two limits, as well as the asymptotic optimality of the base-stock level

$$\hat{s}(b) = \lfloor \lambda b\gamma + \beta^* \sqrt{\lambda b\gamma} \rfloor$$

are shown in Theorem 2.

Theorem 2. *With h, λ, and γ held fixed, $0 < h, \lambda, \gamma < \infty$, and $L_b = \gamma b$, the following hold:*

(i) $s^*(b) \to \infty$ *with*

$$\frac{s^*(b) - \lambda b\gamma}{\sqrt{\lambda b\gamma}} \to \beta^* \quad \textit{as} \quad b \to \infty\,. \tag{34}$$

(ii)

$$b^{-1/2}C^*_{BS}(b) \to h\beta^* \sqrt{\lambda\gamma} + \sqrt{\frac{\lambda}{\gamma}}(h\gamma + 1)\psi(-\beta^*) \quad \textit{as} \quad b \to \infty\,. \tag{35}$$

(iii) Let

$$\hat{s}_b = \lfloor \lambda b\gamma + \beta^* \sqrt{\lambda b\gamma} \rfloor\,. \tag{36}$$

Then

$$\frac{C_{BS}(\hat{s}_b)}{C^*_{BS}(b)} \to 1 \quad \textit{as} \quad b \to \infty\,.$$

Proof. We use the following standard identity for $B(\cdot,\cdot)$: For $1 \le s < \infty$ and $0 < a < \infty$,

$$B(s-1, a) - B(s, a) = B(s-1, a)\left[1 - \frac{a}{s + aB(s-1, a)}\right]\,. \tag{37}$$

We can write (28) as

$$B\left(\lfloor \lambda b\gamma + \beta_a \sqrt{\lambda b\gamma} \rfloor, \lambda b\gamma\right) = \frac{\psi(-\beta) + \delta_3(b)}{\sqrt{\lambda b\gamma}}\,, \tag{38}$$

where $\delta_3(b) \to 0$ as $b \to \infty$. Let $\Delta > 0$. With $\hat{s}_b$ defined as in (36) and using (38) on the right-hand side of (37), we have

$$\begin{aligned} &B\left(\hat{s}_b - 1 + \lfloor \Delta\sqrt{\lambda b\gamma} \rfloor,\ \lambda b\gamma\right) - B\left(\hat{s}_b + \lfloor \Delta\sqrt{\lambda b\gamma} \rfloor,\ \lambda b\gamma\right) \\ &\quad = \frac{\psi(-\beta^* - \Delta) + \delta_3(b)}{\sqrt{\lambda b\gamma}}\left[\frac{\beta^*\sqrt{\lambda b\gamma} + \Delta\sqrt{\lambda b\gamma} + \sqrt{\lambda b\gamma}(\psi(-\beta^* - \Delta) + \delta_4(b))}{\lambda b\gamma + \beta^*\sqrt{\lambda b\gamma} + \Delta\sqrt{\lambda b\gamma} + \sqrt{\lambda b\gamma}(\psi(-\beta^* - \Delta) + \delta_4(b))}\right] \\ &\quad \equiv \frac{R_b(\Delta)}{\lambda b\gamma}\,, \end{aligned} \tag{39}$$

where $\delta_4(b)$ includes $\delta_3(b)$ as well as round-off error due to $\lfloor\cdot\rfloor$; $\delta_4 \to 0$ as $b \to \infty$. As $b \to \infty$, $R_b(\Delta) \to R(\Delta)$ with

$$R(\Delta) = \psi(-\beta^* - \Delta)\left[\beta^* + \Delta + \psi(-\beta^* - \Delta)\right] = \psi'(-\beta^* - \Delta)\,.$$

By the definition of β^* and the convexity of ψ, $R(\Delta) < \frac{h\gamma}{1+h\gamma}$. Define $C(\Delta) > 0$ by $C(\Delta) = \frac{h\gamma}{1+h\gamma} - R(\Delta)$. Using (39) we can write

$$b\left[B\left(\hat{s}_b - 1 + \lfloor\Delta\sqrt{\lambda b\gamma}\rfloor,\ \lambda b\gamma\right) - B\left(\hat{s}_b + \lfloor\Delta\sqrt{\lambda b\gamma}\rfloor,\ \lambda b\gamma\right)\right] = \frac{h}{\lambda(1+h\gamma)} - \frac{C(\Delta)}{\lambda\gamma} + \delta_5(b)\,, \tag{40}$$

where $\delta_5(b) \to 0$ as $b \to \infty$. For b large enough the right-hand side of (40) is strictly less than $\frac{h}{\lambda(1+h\gamma)}$. From (3),

$$b\left[B(s^*(b) - 1,\ \lambda b\gamma) - B(s^*(b),\ \lambda b\gamma)\right] \geq \frac{h}{\lambda(1+h\gamma)}\,,$$

so that

$$\hat{s}_b + \lfloor\Delta\sqrt{\lambda b\gamma}\rfloor \geq s^*(b)\,. \tag{41}$$

Using a similar argument on $B(\hat{s}_b - 1 - \lfloor\Delta\sqrt{\lambda b\gamma}\rfloor,\ \lambda b\gamma) - B(\hat{s}_b - \lfloor\Delta\sqrt{\lambda b\gamma}\rfloor,\ \lambda b\gamma)$ (and using $R(-\Delta) > \frac{h\gamma}{1+h\gamma}$) yields

$$\hat{s}_b - \lfloor\Delta\sqrt{\lambda b\gamma}\rfloor \leq s^*(b)\,. \tag{42}$$

Combining (41) and (42) we obtain that, for any $\Delta > 0$,

$$\hat{s}_b - \lfloor\Delta\sqrt{\lambda b\gamma}\rfloor \leq s^*(b) \leq \hat{s}_b + \lfloor\Delta\sqrt{\lambda b\gamma}\rfloor\,. \tag{43}$$

Since $\Delta > 0$ was arbitrary this implies that

$$\frac{s^*(b) - \hat{s}_b}{\sqrt{\lambda b\gamma}} \to 0 \quad \text{as} \quad b \to \infty\,,$$

which, by the definition of $\hat{s}_b$, implies (i).

We prove (ii) and (iii) together. Using (1) we can write

$$\begin{aligned} C_{BS}(\hat{s}_b) \;&=\; h\left[\lfloor\lambda b\gamma + \beta^*\sqrt{\lambda b\gamma}\rfloor - \lambda b\gamma\left(1 - B(\lfloor\lambda b\gamma + \beta^*\sqrt{\lambda b\gamma}\rfloor, \lambda b\gamma)\right)\right] \\ &\quad + \lambda b B(\lfloor\lambda b\gamma + \beta^*\sqrt{\lambda b\gamma}\rfloor,\ \lambda b\gamma)\,. \end{aligned} \tag{44}$$

Combining (44) with (38) we then have

$$b^{-1/2}C_{BS}(\hat{s}_b) = h\beta^*\sqrt{\lambda\gamma} + \sqrt{\frac{\lambda}{\gamma}}(h\gamma+1)(\psi(-\beta^*) + \delta_3(b)) + \delta_6(b)\,, \tag{45}$$

where $\delta_6(b)$ contains round-off error due to $\lfloor\cdot\rfloor$; $\delta_6(b) \to 0$ as $b \to \infty$. From (45) it is immediate that

$$b^{-1/2} C_{BS}(\hat{s}_b) \to h\beta^* \sqrt{\lambda\gamma} + \sqrt{\frac{\lambda}{\gamma}}(h\gamma+1)\psi(-\beta^*) \equiv \overline{C}^*_{BS} \quad \text{as} \quad b \to \infty \,. \tag{46}$$

Using (43) and the fact that $B(n,a)$ is decreasing in n, we obtain, for $\Delta > 0$,

$$B\left(\hat{s}_b + \lfloor \Delta\sqrt{\lambda b\gamma}\rfloor,\ \lambda b\gamma\right) \leq B(s^*(b),\ \lambda b\gamma) \leq B\left(\hat{s} - \lfloor \Delta\sqrt{\lambda b\gamma}\rfloor,\ \lambda b\gamma\right) . \tag{47}$$

Let

$$L_\Delta(b) = b^{-1/2} h\left(\hat{s}_b - \lfloor \Delta\sqrt{\lambda b\gamma}\rfloor\right) - h\lambda\gamma\sqrt{b} + \lambda\sqrt{b}(h\gamma+1)B\left(\hat{s}_b + \lfloor \Delta\sqrt{\lambda b\gamma}\rfloor,\ \lambda b\gamma\right) \tag{48}$$

and

$$U_\Delta(b) = b^{-1/2} h\left(\hat{s}_b + \lfloor \Delta\sqrt{\lambda b\gamma}\rfloor\right) - h\lambda\gamma\sqrt{b} + \lambda\sqrt{b}(h\gamma+1)B\left(\hat{s}_b - \lfloor \Delta\sqrt{\lambda b\gamma}\rfloor,\ \lambda b\gamma\right) . \tag{49}$$

Combining (43) and (47) yields

$$L_\Delta(b) \leq b^{-1/2} C^*_{BS}(b) \leq U_\Delta(b) \,. \tag{50}$$

In the same manner that (44) led to (46), from (48) and (49) we obtain

$$\lim_{b\to\infty} L_\Delta(b) = \overline{C}^*_{BS} - h\Delta\sqrt{\lambda\gamma} + \sqrt{\frac{\lambda}{\gamma}}(h\gamma+1)\left[\psi(-\beta^*-\Delta) - \psi(-\beta^*)\right] \equiv \overline{L}_\Delta$$

and

$$\lim_{b\to\infty} U_\Delta(b) = \overline{C}^*_{BS} + h\Delta\sqrt{\lambda\gamma} + \sqrt{\frac{\lambda}{\gamma}}(h\gamma+1)\left[\psi(-\beta^*+\Delta) - \psi(-\beta^*)\right] \equiv \overline{U}_\Delta \,.$$

By the continuity of ψ, $\lim_{\Delta\to 0} \overline{L}_\Delta = \lim_{\Delta\to 0} \overline{U}_\Delta = \overline{C}^*_{BS}$. Since $\Delta > 0$ was arbitrary,

$$b^{-1/2} C^*_{BS}(b) \to \overline{C}^*_{BS} \quad \text{as} \quad b \to \infty \,, \tag{51}$$

which proves (ii).

Finally, combining (46) and (51) yields

$$\frac{C_{BS}(\hat{s}_b)}{C^*_{BS}(b)} \to 1 \quad \text{as} \quad b \to \infty \,,$$

proving (iii). ∎

Smith (1977) considered the limit (translated to our notation, as are all of Smith's results below) $L \to \infty$ with $b_L = \gamma^{-1} L$. This is equivalent to the limiting regime considered here. With $s(b) = \lfloor \lambda b\gamma + x\sqrt{\lambda b\gamma}\rfloor$ he used the approximation

$$e^{-\lambda b\gamma} \sum_{n=0}^{s(b)} \frac{(\lambda b\gamma)^n}{n!} \approx 1 \,.$$

The result of Jagerman (1974) that we use is based on the more accurate limit

$$e^{-\lambda b\gamma} \sum_{n=0}^{s(b)} \frac{(\lambda b\gamma)^n}{n!} \to \Phi(x) \quad \text{as} \quad b \to \infty\,.$$

One consequence of Smith's less accurate approximation is that he obtains $\beta^* > 0$ for all values of h. This contrasts with our results, since the solution to (33) can be negative. (See the discussion on safety stock in Section 4.3.)

4.3. Comparison of the Limits

The essence of Theorems 1 and 2, from the point of view of comparing the two policies is the following:

$$b^{-1/2} C^*_{COI}(b) \to \sqrt{2h\lambda} \quad \text{as} \quad b \to \infty$$

and

$$b^{-1/2} C^*_{BS}(b) \to h\beta^* \sqrt{\lambda\gamma} + \sqrt{\frac{\lambda}{\gamma}}(h\gamma + 1)\psi(-\beta^*) \quad \text{as} \quad b \to \infty\,.$$

Let

$$\theta_b(h, \lambda, \gamma) = \frac{C^*_{BS}(b)}{C^*_{COI}(b)}\,. \tag{52}$$

Then Theorems 1 and 2 imply that

$$\theta_b(h, \lambda, \gamma) \to \overline{\theta}(h\gamma) \quad \text{as} \quad b \to \infty\,, \tag{53}$$

where

$$\overline{\theta}(x) = \beta^*(x)\sqrt{\frac{x}{2}} + \frac{(1+x)\psi(-\beta^*(x))}{\sqrt{2x}}\,, \tag{54}$$

and $\beta^*(x)$ is defined by (33) with $h\gamma = x$. Interestingly, the λ terms cancel out in the ratio, and h, γ appear only via their product. By (33) and the convexity of ψ, β^* is decreasing in x. The following lemma, showing that $\overline{\theta}$ is an increasing function, is proved in the appendix.

Lemma 3. *For $x > 0$, $\overline{\theta'}(x) > 0$.*

We are now ready to prove the asymptotic linearity of L^*_b in b as $b \to \infty$. Let x^* be the unique point where

$$\overline{\theta}(x^*) = 1\,. \tag{55}$$

The uniqueness of x^* satisfying (55) follows from Lemma 3. The existence of x^* follows from numerical calculation: $\overline{\theta}(0.69786) < 1$ and $\overline{\theta}(0.69788) > 1$. Thus $0.69786 < x^* < 0.69788$. By Lemma 3, $\overline{\theta}(x) < 1$ for $x < x^*$ and $\overline{\theta}(x) > 1$ for $x > x^*$.

Theorem 3. *With h and λ fixed,*

$$\frac{L_b^*}{b} \to \frac{x^*}{h} \equiv \gamma^* \quad as \quad b \to \infty\,. \tag{56}$$

Proof. We prove that $\limsup_{b\to\infty} \frac{L_b^*}{b} \le \gamma^*$ and that $\liminf_{b\to\infty} \frac{L_b^*}{b} \ge \gamma^*$.

By (53) and the definition of γ^*,

$$\theta_b(h, \lambda, \gamma^*) \to 1 \quad \text{as} \quad b \to \infty\,.$$

In addition, by Lemma 3, for any $\Delta > 0$,

$$\begin{aligned}\lim_{b\to\infty} \theta_b(h, \lambda, \gamma^* - \Delta) &< 1 \\ \lim_{b\to\infty} \theta_b(h, \lambda, \gamma^* + \Delta) &> 1\,.\end{aligned}$$

Thus, given any $\Delta > 0$, for all b large enough,

$$\gamma^* - \Delta \le \frac{L_b^*}{b} \le \gamma^* + \Delta\,,$$

which implies

$$\limsup_{b\to\infty} \frac{L_b^*}{b} \le \gamma^* + \Delta$$

and

$$\liminf_{b\to\infty} \frac{L_b^*}{b} \ge \gamma^* - \Delta\,.$$

Since $\Delta > 0$ was arbitrary the theorem is proved. ∎

In the above asymptotic analysis we have assumed that $L_b = \gamma b$ with $0 < \gamma < \infty$ and let $b \to \infty$. Although a detailed analysis is beyond the scope of this paper, we provide a quick overview of the cases $\frac{L_b}{b} \to 0$ and $\frac{L_b}{b} \to \infty$ as $b \to \infty$. When $\frac{L_b}{b} \to 0$ as $b \to \infty$ it can be shown that

$$\frac{s^*(b) - \lambda L_b}{\sqrt{\lambda L_b}} \to \infty \quad \text{as} \quad b \to \infty\,.$$

A heuristic calculation leads us to conjecture that, with h and λ held fixed,

$$\frac{C_{BS}^*(b, L_b)}{C_{COI}^*(b)} \to 0 \quad \text{as} \quad b \to \infty\,.$$

When $\frac{L_b}{b} \to \infty$ it can be shown that

$$\frac{s^*(b) - \lambda L_b}{\sqrt{\lambda L_b}} \to -\infty \quad \text{as} \quad b \to \infty\,.$$

A heuristic calculation for this case leads us to conjecture that, with h and λ held fixed,

$$\frac{C^*_{BS}(b, L_b)}{C^*_{COI}(b)} \to \infty \quad \text{as} \quad b \to \infty\,.$$

A key measure of how much an inventory system hedges against uncertain demand is safety stock. The safety stock is defined as the difference between the average inventory position and the expected number of demands over a lead time. For a base-stock policy the inventory position is always equal to the base-stock level. Thus, if we let $Z_{BS}(s)$ denote the safety stock of the base-stock policy with base-stock level s,

$$Z_{BS}(s) = s - \lambda L\,.$$

For the COI policy the average inventory position is the sum of the average inventory level and the number of items in the pipeline. Letting $Z_{COI}(\rho)$ denote the safety stock of the COI policy with order interval $\tau = (\lambda\rho)^{-1}$,

$$Z_{COI}(\rho) = \frac{\rho}{1 - \alpha(\rho)} + (\rho - 1)\lambda L\,.$$

Using the results of Theorems 1 and 2 we can examine the asymptotic behavior (assuming $L_b = \gamma b$ with $0 < \gamma < \infty$) of the safety stock of these two policies under the respective optimal s^* and ρ^*. Let

$$\widehat{Z}_{BS}(b) = b^{-1/2} Z_{BS}(s^*(b))$$

and

$$\widehat{Z}_{COI}(b) = b^{-1/2} Z_{COI}(\rho^*(b))\,.$$

From Theorem 2 we immediately obtain

$$\widehat{Z}_{BS}(b) \to \beta^* \sqrt{\lambda\gamma} \equiv \widehat{Z}_{BS}(\infty) \quad \text{as} \quad b \to \infty\,. \tag{57}$$

Similarly, Theorem 1 yields

$$\widehat{Z}_{COI}(b) \to \sqrt{\frac{\lambda}{2h}}(1 - \gamma h) \equiv \widehat{Z}_{COI}(\infty) \quad \text{as} \quad b \to \infty\,. \tag{58}$$

(Although it appears that $\widehat{Z}_{BS}(\infty)$ does not depend on h, recall that β^* depends on h.)

In the unlimited backlogging version of the inventory model we are considering, the long lead time limit leads to safety stock that is proportional to the square-root of the lead time and is always *positive*. For the lost sales model, on the other hand, although the safety stock is again proportional to the square-root of the lead time, it may be positive or negative, depending on the values of h and γ. (The value of λ affects the magnitude of $\widehat{Z}_{BS}(\infty)$ and $\widehat{Z}_{COI}(\infty)$ but not their signs.)

The sign of $\widehat{Z}_{BS}(\infty)$ will be precisely the sign of β^*. Recall from (33) that β^* depends on γ and h through their product γh, and that β^* is decreasing in γh. If we find x_0 such that $\beta^*(x_0) = 0$ then $\beta^* < 0$ for $\gamma h > x_0$ and $\beta^* > 0$ for $\gamma h < x_0$. From (33) we can write

$$\psi^2(0) = \frac{x_0}{1 + x_0}$$

so that $x_0 = \psi^2(0)/(1 - \psi^2(0))$. Note that $\psi(0) = \sqrt{\frac{2}{\pi}}$ so that $x_0 = 2/(\pi - 2) \approx 1.752$.

From (58) it is immediate that $\widehat{Z}_{COI}(\infty) < 0$ for $\gamma h > 1$ and $\widehat{Z}_{COI}(\infty) > 0$ for $\gamma h < 1$. Thus, when $\gamma h < 1$ both policies have positive safety stock. When $1 < \gamma h < x_0$, the COI policy has negative safety stock while the base stock policy has positive safety stock. For $\gamma h > x_0$ both policies have negative safety stock.

Appendix

In this appendix we provide proofs of Lemmas 1, 2 and 3.

Lemma 1. *With h, b, and λ held fixed, C^*_{BS} is a continuous and strictly increasing function of L.*

Proof. Throughout the proof we assume that h, b, and λ are fixed. We also assume that $\lambda b \geq h$, so that $s^* \geq 1$. We use the following result from measure theory: A continuous function whose right hand derivative is non-negative everywhere and positive almost everywhere is strictly increasing. (This follows, for example, from Roydin 1968, Problem 3, p.98 and Theorem 2, p.96.)

Recall that, for $s \geq 0$,

$$C_{BS}(s) = h\left[s - \lambda L(1 - B(s, \lambda L))\right] + \lambda b B(s, \lambda L)\,.$$

For fixed s, this is differentiable in L, with

$$\frac{\partial C_{BS}}{\partial L}(s) = -h\lambda(1 - B(s, \lambda L)) + h\lambda^2 L \frac{\partial B}{\partial a}(s, \lambda L) + \lambda^2 b \frac{\partial B}{\partial a}(s, \lambda L)\,. \tag{A1}$$

Straightforward differentiation of $B(s, a)$ with respect to its second argument yields

$$\frac{\partial B}{\partial a}(s, a) = p_{s-1} - B(s, a)[1 - B(s, a)]\,. \tag{A2}$$

It is known that (cf. problem 5-11 on p. 301 in Wolff 1989)

$$p_{s-1} = B(s-1, a)[1 - B(s, a)], \qquad s \geq 1\,. \tag{A3}$$

Substituting (A3) into (A2) we obtain

$$\frac{\partial B}{\partial a}(s, a) = [B(s-1, a) - B(s, a)][1 - B(s, a)]\,. \tag{A4}$$

Assume (momentarily) that the right-hand inequality in (3) is strict. Then $s^*(L)$ will not change for sufficiently small changes in L. Thus

$$\begin{aligned}\frac{\partial C^*_{BS}}{\partial L} &= \frac{\partial C_{BS}}{\partial L}(s^*(L)) \\ &= \left\{\lambda^2(hL + b)[B(s^*(L) - 1, \lambda L) - B(s^*(L), \lambda L)] - h\lambda\right\}[1 - B(s^*(L), \lambda L)]\,. \end{aligned} \tag{A5}$$

By (3),

$$B(s^*(L) - 1, \lambda L) - B(s^*(L), \lambda L) > \frac{h}{\lambda(b + Lh)}\,,$$

so that

$$\frac{\partial C^*_{BS}}{\partial L} > 0\,.$$

We now deal with the case where the right-hand inequality in (3) is not strict. In this case

$$B(s^*(L)-1,\lambda L)-B(s^*(L),\lambda L)=\frac{h}{\lambda(b+Lh)}, \tag{A6}$$

and using (A5) we obtain $\frac{\partial C_{BS}}{\partial L}(s^*(L))=0$. Note also that, at L, $s^*(L)$ and $s^*(L)-1$ have the same total cost. Since $C_{BS}(s^*(L))$ and $C_{BS}(s^*(L)-1)$ are both continuous in L, C^*_{BS} is continuous at L.

In order to complete the proof for this case we show that, for δ sufficiently small, $s^*(L+\delta)=s^*(L)$, which immediately implies that $\frac{\partial C_{BS}}{\partial L}(s^*(L))$ is the right hand derivative of C^*_{BS}. This will be true if

$$\frac{\partial}{\partial L}\left(\frac{h}{\lambda(b+Lh)}\right)<\frac{\partial}{\partial L}\left(B(s^*(L)-1,\lambda L)-B(s^*(L),\lambda L)\right)\,. \tag{A7}$$

The left-hand side is immediate:

$$\frac{\partial}{\partial L}\left(\frac{h}{\lambda(b+Lh)}\right)=-\frac{h^2}{\lambda(b+Lh)^2}\,.$$

For the right-hand side, if $s^*(L)\geq 2$ we use (A4) to write

$$\begin{aligned}&\frac{\partial}{\partial L}\left[B(s^*(L)-1,\lambda L)-B(s^*(L),\lambda L)\right]\\ &\quad=\lambda\left\{\left[B(s^*(L)-2,\lambda L)-B(s^*(L)-1,\lambda L)\right]\left[1-B(s^*(L)-1,\lambda L)\right]\right.\\ &\qquad\left.-\left[B(s^*(L)-1,\lambda L)-B(s^*(L),\lambda L)\right]\left[1-B(s^*(L),\lambda L)\right]\right\}\,.\end{aligned}$$

Recall that

$$B(s^*(L)-2,\lambda L)-B(s^*(L)-1,\lambda L)>B(s^*(L)-1,\lambda L)-B(s^*(L),\lambda L)=\frac{h}{\lambda(b+Lh)}\,,$$

and define $\eta>0$ by

$$\eta=B(s^*(L)-2,\lambda L)-B(s^*(L)-1,\lambda L)-\frac{h}{\lambda(b+Lh)}\,.$$

We can thus write

$$\frac{\partial}{\partial L}\left[B(s^*(L)-1,\lambda L)-B(s^*(L),\lambda L)\right]=\lambda\eta\left[1-B(s^*(L)-1,\lambda L)\right]-\frac{h^2}{\lambda(b+Lh)^2},$$

so that (A7) holds.

If $s^*(L)=1$ and (A6) holds, then $\lambda b=h$. In this case

$$\frac{\partial}{\partial L}(B(0,\lambda L)-B(1,\lambda L))=-\frac{\lambda}{(1+\lambda L)^2}=-\frac{h^2}{\lambda(b+Lh)^2}\,,$$

and (A7) again holds. ∎

Lemma 2. *For any $y \in (0,1)$ there exists a unique w^* such that*

$$\psi'(w^*) = y\,. \tag{A8}$$

Proof. The uniqueness of w^* follows immediately from convexity of ψ, but existence requires a bit more effort. In order to show that a finite solution w^* to (A8) exists for every $y \in (0,1)$ we show that $\psi'(x) \to 0$ as $x \to -\infty$ and $\psi'(x) \to 1$ as $x \to \infty$.

We first deal with $x \to -\infty$. We can write

$$\psi(x) = \frac{e^{-x^2/2}}{\int_x^\infty e^{-t^2/2}dt}$$

and

$$\psi'(x) = \frac{e^{-x^2/2}}{\int_x^\infty e^{-t^2/2}dt}\left[\frac{e^{-x^2/2}}{\int_x^\infty e^{-t^2/2}dt} - x\right]. \tag{A9}$$

As $x \to -\infty$,

$$\int_x^\infty e^{-t^2/2}dt \to \sqrt{2\pi}\,,$$

$e^{-x^2/2} \to 0$ and $xe^{-x^2/2} \to 0$. Thus $\psi'(x) \to 0$ as $x \to -\infty$.

We now deal with $x \to \infty$. Our starting point is relation 586. on p. 136 of Dwight (1961):

$$\frac{1}{\sqrt{2\pi}}\int_{-x}^{x} e^{-t^2/2}dt = 1 - \sqrt{\frac{2}{\pi}}\frac{e^{-x^2/2}}{x}\left[1 - \frac{1}{x^2} + \frac{3}{x^4} + \tilde{\epsilon}(x)\right], \tag{A10}$$

with $|\tilde{\epsilon}(x)| \le 3/x^4$. (A similar relation, which would require a bit of manipulation to meet our purposes is (7.1.13) on p. 298 of Abramowitz and Stegun 1972.) We can rewrite (A10) as

$$\int_x^\infty e^{-t^2/2}dt = \frac{e^{-x^2/2}}{x}\left[1 - \frac{1}{x^2} + \frac{3}{x^4} + \tilde{\epsilon}(x)\right]$$

so that

$$\frac{e^{-x^2/2}}{\int_x^\infty e^{-t^2/2}dt} = \frac{x}{1 - \frac{1}{x^2} + \epsilon(x)}\,,$$

where $\epsilon(x) = \frac{3}{x^4} + \tilde{\epsilon}(x)$ and $x^2\epsilon(x) \to 0$ as $x \to \infty$. Thus

$$\psi(x) = x + \frac{1}{x} + \epsilon'(x)\,, \tag{A11}$$

where $x\epsilon'(x) \to 0$ as $x \to \infty$. Using (A11) we can write

$$\psi'(x) = \left(x + \frac{1}{x} + \epsilon'(x)\right)\left(\frac{1}{x} + \epsilon'(x)\right) \to 1 \quad \text{as} \quad x \to \infty\,,$$

completing the proof. ∎

Lemma 3. *For $x \geq 0$, $\overline{\theta}'(x) > 0$.*

Proof. We re-write (54) as

$$\overline{\theta}(x) = \sqrt{\frac{x}{2}}\,[\psi(-\beta^*(x)) + \beta^*(x)] + \frac{\psi(-\beta^*(x))}{\sqrt{2x}}\,. \tag{A12}$$

Differentate (A12):

$$\begin{aligned}\frac{d\overline{\theta}}{dx}(x) &= \frac{d\beta^*}{dx}\left\{\sqrt{\frac{x}{2}}\left[1 - \psi'(-\beta^*(x))\right] - \frac{1}{\sqrt{2x}}\psi'(-\beta^*(x))\right\} \\ &\quad + \frac{1}{2\sqrt{2x}}[\beta^*(x) + \psi(-\beta^*(x)] - \frac{1}{2x\sqrt{2x}}\psi(-\beta^*(x))\,. \end{aligned}\tag{A13}$$

From (33) we can substitute $\psi'(-\beta^*(x)) = \frac{x}{1+x}$. Performing this substitution, the term multiplying $\frac{d\beta^*}{dx}$ is seen to be zero. We thus have

$$\frac{d\overline{\theta}}{dx}(x) = \frac{1}{2\sqrt{2x}}\left[\beta^*(x) + \psi(-\beta^*(x)) - \frac{1}{x}\psi(-\beta^*(x))\right]\,. \tag{A14}$$

Although the $\frac{d\beta^*}{dx}$ term does not appear in (A14), we obtain a useful relation from $\frac{d\beta^*}{dx}$, so we calculate it. Implicit differentiation of (33) yields

$$\frac{d\beta^*}{dx}\left\{\psi(-\beta^*(x)) - \psi'(-\beta^*(x))[2\psi(-\beta^*(x)) + \beta^*(x)]\right\} = \frac{1}{(1+x)^2}\,.$$

Substituting in $\psi'(-\beta^*(x)) = \frac{x}{1+x}$, we obtain

$$\frac{d\beta^*}{dx} = \frac{-1}{x(1+x)[\beta^*(x) + \psi(-\beta^*(x)) - \frac{1}{x}\psi(-\beta^*(x))]}\,. \tag{A15}$$

Recall that β^* is decreasing in x, so $\frac{d\beta^*}{dx} < 0$, which implies that

$$\beta^*(x) + \psi(-\beta^*(x)) - \frac{1}{x}\psi(-\beta^*(x)) > 0\,. \tag{A16}$$

Combining (A16) with (A14) yields $\frac{d\overline{\theta}}{dx} > 0$, completing the proof. ∎

References

Abramowitz, M., I. A. Stegun, eds. 1972. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 12th printing. National Bureau of Standards, Washington D.C.

Dwight, H. B. 1961. *Tables of Integrals and Other Mathematical Data*, 4th Edition. MacMillan, New York.

Hill, R. M. 1999. On the suboptimality of $(S-1, S)$ lost sales inventory policies. *Int. J. Prod. Econ.* **59** 387–393.

Jagerman, D. L. 1974. Some properties of the Erlang loss function. *Bell System Tech. J.* **53** 525–551.

Johansen, S. G. 2001. Pure and modified base stock policies for the lost sales inventory system with negligible set-up costs and constant lead times. *Int. J. Prod. Econ.* **71** 391–399.

Karlin, S., H. Scarf. 1958. Inventory models of the Arrow-Harris-Marschak type with time lag. K. Arrow, S. Karlin, H. Scarf eds. *Studies in the Mathematical Theory of Inventory and Production*, Stanford University Press, Stanford, CA, 155–178.

Karush, W. 1957. A queueing model for an inventory problem. *Oper. Res.* **5** 693–703.

Kingman, J. F. C. 1962. On queues in heavy traffic. *J. Royal Statist. Soc.* **B24** 383–392.

Lee, H. L., S. Nahmius. 1993. Single-product, single-location models. S. C. Graves, A. H. G. Rinnooy Khan, P. H. Zipkin eds. *Logistics of Production and Inventory*, North-Holland, Amsterdam, 3–55.

Morrison, J. A., K. G. Ramakrishnan. 2003. Asymptotic solution to an inverse problem for a shared unbuffered resource. *SIAM J. Appl. Math.* **63** 222–240.

Morton, T. E. 1969. Bounds on the solution of the lagged optimal inventory equation with no demand backlogging and proportional costs. *SIAM Review* **11** 572–596.

Morton, T. E. 1971. The near-mypoic nature of the lagged-proportional-cost inventory problem with lost sales. *Oper. Res.* **19** 1708–1716.

Porteus, E. L. 1990. Stochastic inventory theory. D. P. Heyman, M. J. Sobel eds. *Stochastic*

Models, North-Holland, Amsterdam, 605–652.

Rogozin, B. A. 1966. Some extremal problems in the theory of mass service. *Theor. Probab. Appl.* **11** 144–151.

Roydin, H. L. 1968. *Real Analysis*, 2nd Edition. Macmillan, London.

Smith, S. A. 1977. Optimal inventories for an $(S-1, S)$ system with no backorders. *Man. Sci.* **23** 522–528.

Wolff, R. W. 1989. *Stochastic Modeling and the Theory of Queues.* Prentice-Hall, Englewood Cliffs, NJ.

Zipkin, P. H. 1999. *Foundations of Inventory Management.* McGraw-Hill, Boston.