Order-Based Backorders and Their Implications in Multi-Item Inventory Systems

Jing-Sheng Song
Graduate School of Management, University of California, Irvine, California 92697
jssong@uci.edu

In a multi-item inventory system, such as an assemble-to-order manufacturing system or an online-retailing system, a customer order typically consists of several different items in different amounts. The average order-based backorders are the average number of customer orders that are not yet completely filled. While this is an important measure of customer satisfaction, it has not been widely studied in the operations management literature. This is largely because its evaluation involves the joint distribution of inventory levels of different items and other intricate relations, which is computationally dreadful. Taking a novel approach, this paper develops a tractable way of evaluating this measure exactly. We also develop easy-to-compute bounds, which require the evaluation of item-based backorders only. Numerical experiments indicate that the average of the lower and upper bounds is very effective.

The exact results show surprisingly simple structures, which shed light on how system parameters affect the performance. Using these results, we study several examples to gain managerial insights. Questions addressed include: What are the implications of item-based inventory planning decisions on the order-based performance? What is the impact of introducing common components on inventory and service trade-offs? Would order-delivery performance be improved if we restrict the number of choices in product configurations?
(*Multi-Item Systems; Backorders; Demand Correlation; Performance Evaluation; Approximation; Assemble-to-Order; Component Commonality; Product Structure*)

1. Introduction

This paper is concerned with multi-item inventory systems in which a customer order typically consists of several different items in different amounts, and different orders may have overlapping subsets of items. A customer order is filled immediately if all the items requested are available in inventory. That is, as long as the inventories are available, the order assembly time is negligible. One example is the assemble-to-order manufacturing systems. Here, components are kept in stock, but final products, which may share common components, are assembled only after customer orders are realized, like in Dell Computer. Another example is mail-order systems, including online retailing, such as in Lands End and Amazon.com. In the move toward agile systems where customization is delayed as long as possible, this kind of system is becoming more and more important and prevalent.

A key performance measure for such systems is the average order-based backorders, which is the average number of customer orders that are not yet completely filled. This quantity is proportional to the average customer waiting time, so it directly measures customer dissatisfaction due to delivery delays. As customer satisfaction becomes a foremost priority in

0025-1909/02/4804/0499$5.00
1526-5501 electronic ISSN

today's corporations, it is of great interest to managers to be able to evaluate this measure, so as to see how inventory policies or other system parameters affect performance. The purpose of this paper is to present tractable analytical techniques for doing so. We also discuss various insights through several examples.

Although the importance of order-based backorders was recognized in the literature as early as the 1960s, the evaluation of this measure has not been widely studied, largely because it is analytically intricate and computationally dreadful. Most of the early works make restrictive assumptions. Some essentially assume independent demand and some consider only a one-time demand (a static problem). See, for example, Alt (1962), Hausman (1969), Kaplan (1971), Miller (1971), and Muckstadt (1973, 1979). More recently, Cheung and Hausman (1995) consider a multivariate Poisson demand model. The replenishment lead times of each item are i.i.d. random variables. However, complete cannibalization is assumed in order to derive the average customer backorders. Song et al. (1999) present a recursive procedure to compute the order-based backorders for a system in which demand forms a multivariate Poisson process, and there is a dedicated facility with exponentially distributed processing times to process replenishment orders for any given item. The procedure relies on the computation of the steady-state joint distribution of the outstanding orders. Zhang (1995a) studies a discrete time, continuous demand model with constant lead times, and discusses the computation and approximation of expected order-based backorders when there is only one type of order (i.e., one final product). Zhang (1995b) derives a lower bound for the average order-based backorders for a multivariate normal demand model with multiple types of orders. Several other recent works also study order-fulfillment performances in assemble-to-order systems with multicomponents and multiproducts; see, for example, Agrawal and Cohen (2001), Glasserman and Wang (1998), Hausman et al. (1998), and Song (1998). However, these works focus on order fill rates. The calculation of order-based backorders presents additional challenges.

The current paper shows how to evaluate the order-based backorders in a continuous review, base-stock inventory system with constant lead times and multiple types of orders. Instead of the traditional method that uses the joint distribution of the item-backorder vector, we take a novel approach to the exact analysis, which leads to a closed-form expression. Our approach consists of two steps. The first step deals with the case of equal lead times across the items. The key here is to relate the customer waiting-time distribution with the immediate order fill rate in a revised system, applying a related result shown in Song (1998). Using this relationship, we can obtain the expected customer waiting time for the equal lead-time case by integrating the waiting-time tail distribution. Subsequently, the expected number of backorders with equal lead times can be obtained by applying Little's law. The second step of the analysis deals with the general case with unequal lead times. The key here is to reveal certain cyclical behavior of the backorder process and then to focus on the detailed analysis within a cycle using probabilistic arguments. Section 3.1 presents the analysis for the two-item, unit demand system, while §3.2 generalizes the analysis to the J-item, unit demand system. Section 8 extends the results to systems with batch demands.

For a system as complex as the one studied here, the exact result shows surprisingly simple structure. This helps to shed light on the impact of system parameters on both the computational complexity and the service performance. For example, it shows that a dominating factor for the computation time is not how many items of which an order consists, but rather the number of order types that have overlapping items with this order. Thus, the easiest case is a pure assembly system, where there is only one order type. It also shows that the base-stock levels of different items affect the order-based backorders jointly only through the common part of the component lead times.

Although the exact result enjoys tremendous computational advantage over simulation, it can still be computationally demanding for systems with many overlapping demand types (or, equivalently, systems with many common components among products). To overcome this difficulty, in §4 we develop easy-to-compute bounds. These bounds are simple operations of the order-based demand rates and the item-based backorders; no joint distribution evaluation is involved. In our numerical studies, the average of the lower and upper bounds performs extremely well. Therefore, we recommend it be used in practice as a quick estimate.

In §§5–7 we provide a few examples to illustrate potential applications of the evaluation procedure and to gain various managerial insights. Section 5 studies the implication of item-based forecasts and inventory-planning decisions on the order-based performance. Section 6 focuses on a two-product example and investigates whether creating a common component between the two products would improve the inventory-service trade-off. Section 7 uses a personal computer example to explore the effect of product variety on total customer backorders. Section 9 concludes the paper.

2. Model Description and Preliminaries

We consider a continuous review, J-item inventory system. There are multiple classes of demands, each of which requires a fixed kit of items and arrives according to an independent Poisson process. For any subset of items K of $\{1, \ldots, J\}$, we say a demand is of type K if it requires one unit of item i in K and 0 units outside K. (We generalize the unit demand size to random batch sizes in §8.) Let $\mathcal{S}$ be the set of all demand types. When K contains a single item, say i, we abbreviate type K by type i. Similarly, we say an order is of type $i_1 i_2 \cdots i_k$ if $K = \{i_1, i_2, \ldots, i_k\}$. Notice the demand process for each item is the superposition of a finite number of Poisson processes. So, it again forms a Poisson process whose rate is the sum of that of the individual demand processes.

Throughout the paper, we use subscripts to indicate item type and superscripts order type. For any item $i \in \{1, \ldots, J\}$ and any subset $K \subset \{1, \ldots, J\}$, let

λ^K = arrival rate of demand type K;

λ = overall demand rate $= \sum_{K \in \mathcal{S}} \lambda^K$;

q^K = probability a demand is of type $K = \lambda^K / \lambda$;

$S(i)$ = the set of all demand types that consume item i $= \{K : i \in K \text{ and } q^K > 0\}$;

λ_i = aggregate demand rate of item $i = \sum_{K \in S(i)} \lambda^K$.

For any item i, the replenishment lead time is a constant L_i, and a base-stock policy with base-stock level s_i is followed to control its inventory. We assume complete backlogging for demands that cannot be filled immediately. When an order arrives and we have some of its items in stock but not all, we assume that we can either ship the in-stock items or put them aside as committed inventory. However, a customer request is considered backlogged until it is satisfied completely. Demands are filled on a First-Come-First-Served (FCFS) basis. When there are backorders, they are also filled on a FCFS basis.

Let $t \geq 0$ be the continuous time variable, and for each t denote

$IN_i(t)$ = net inventory of item i;

$N^K(t)$ = number of type-K demands by time t;

$D_i(t)$ = cumulative demand for item i by time $t = \sum_{K \in S(i)} N^K(t)$;

$B^K(t)$ = type-K backorder at t = number of type-K orders that are not yet completely satisfied by t;

$B_i(t)$ = number of backorders for item i at t.

Let D_i stand for the steady-state limit of $D_i(t - L_i, t) = D_i(t) - D_i(t - L_i)$, the lead-time demand of item i. Then, D_i has the same distribution as $D_i(L_i)$, a Poisson distribution parameter $\lambda_i L_i$. Let IN_i be the steady-state limit of $IN_i(t)$, and define B^K and B_i similarly.

The performance measure of primary interest in this paper is $\overline{B}^K = E[B^K]$ for any demand type K. Knowing how to evaluate this measure, we can readily obtain

$$\overline{B} = \sum_K \overline{B}^K = \text{the total average order-based backorders.}$$

Because this is the average number of customers being backlogged regardless of types, by Little's law we have:

$$\overline{W} = \text{average waiting time of an arbitrary customer}$$
$$= \overline{B}/\lambda.$$

As a function of the base-stock levels, both $\overline{B}$ and $\overline{W}$ can be used to demonstrate the quantitative trade-offs between inventory investment and customer service, and therefore to aid managerial decision making.

To highlight the dependence on the problem data, we sometimes use $\overline{B}^K(\mathbf{s} \mid \mathbf{L}) = \overline{B}^K(s_1, \ldots, s_J \mid L_1, \ldots, L_J)$ to denote the expected type-K backorders in a system with base-stock levels s_i and lead times L_i. The boldfaced letters are used to abbreviate vectors. When $L_i = L$ for all i, we write $\overline{B}^K(\mathbf{s} \mid \mathbf{L})$ as $\overline{B}^K(\mathbf{s} \mid L)$. Without loss of generality, we index the items so that $L_1 \leq L_2 \leq \cdots \leq L_J$. We also assume nonnegative base-stock levels, i.e., $s_i \geq 0$ for all i. As a consequence, the maximum duration of a type-K backorder is $\max_{i\in K} L_i$.

Denote by $g(\cdot \mid \mu)$ and $G(\cdot \mid \mu)$ the probability mass function (pmf) and cumulative distribution function (cdf) of the Poisson distribution with parameter μ, respectively. Let $G^0 = 1 - G$. For any real numbers u and v, let $u \vee v = \max\{u, v\}$, $u \wedge v = \min(u, v)$ and $u^+ = \max\{u, 0\}$. Also, denote by $\mathbf{e}_i$ the ith standard unit vector and $\mathbf{1}$ the vector of ones.

3. The Exact Approach

It is well known that for each i

$$IN_i = s_i - D_i.$$

From this, the average type-i backorder is easy to evaluate. In particular, for each i, the item-based backorder equals $B_i(s_i \mid L_i) = [-IN_i]^+ = (D_i - s_i)^+$, so the average item-$i$ backorder is

$$\overline{B}_i(s_i \mid L_i) = E[(D_i - s_i)^+] = \sum_{\ell=s_i}^{\infty} G^0(\ell|\lambda_i L_i)$$
$$= \lambda_i L_i - \sum_{\ell=0}^{s_i-1} G^0(\ell|\lambda_i L_i). \tag{1}$$

Since a request for item i is due to a type-i order with probability λ^i/λ_i, the average type-i backorder equals

$$\overline{B}^i(s_i \mid L_i) = \frac{\lambda^i}{\lambda_i}\overline{B}_i(s_i \mid L_i). \tag{2}$$

The evaluation of type-K backorders when K contains more than one item, however, is much harder. Let $B_i^K(t)$ be the number of backorders for item i at time t that are due to demand type K, where $K \in S(i)$. Since $B^K(t) = \max_{i\in K} B_i^K(t)$ for any t, we have

$$\overline{B}^K = E\Big[\max_{i\in K} B_i^K\Big], \tag{3}$$

where B_i^K is the steady-state limit of $B_i^K(t)$. Clearly, a direct approach for evaluating $\overline{B}^K$ is through (3). Observe that, given $B_i = n$, B_i^K is a binomial random variable with n trials and success probability λ^K/λ_i in each trial. In principle, then, one can compute (3) by using this fact and the joint distribution of $(B_i, i \in K)$. However, because the B_i are correlated random variables, computing its joint distribution alone is a formidable task, not to mention the conditional binomial distributions and the max operation within the expectation. In what follows we present an alternative approach that is much simpler.

The first step of the new approach relies on the following fact. Suppose $L_i = L$ for all i. Let W^K be the steady-state waiting time of a type-K backorder, and $F^{K,w}$ the type-K order fill rate with time window w, that is, the probability of satisfying a type-K order within time w. Then, for $w \leq L$, we have

$$P[W^K \leq w] = F^{K,w}(\mathbf{s}|L) = F^{K,0}(\mathbf{s}|L-w)$$
$$= P[D_i(L-w) < s_i, i \in K], \tag{4}$$

where the second equality is due to Proposition 1.1 in Song (1998).

3.1. The Two-Item System

To illustrate the idea, we first study the two-item, unit demand system. Here, there are only three possible types of demand: A Type-1 customer requires one unit of Item 1 only; a Type-2 requires one unit of Item 2 only; and a Type-12 customer asks for one unit of each item.

Observe that for any t,

$$D_i(t) = N^i(t) + N^{12}(t), \quad i = 1, 2.$$

Let $N(t)$ be the total number of arrivals (regardless of customer types) during t. Then

$$N(t) = N^1(t) + N^2(t) + N^{12}(t)$$

has the Poisson distribution with parameter λt. Since $N^K(t)$ are independent, it follows that given $N(t)=n$, $(N^1(t), N^2(t), N^{12}(t))$ has a multinomial distribution with parameters n and $q^K = \lambda^K/\lambda$, which is independent of t. Moreover,

$$\begin{aligned}
&P[D_1(t)<s_1, D_2(t)<s_2] \\
&\quad = \sum_{n=0}^{s_1\vee s_2-1} P[N(t)=n] \\
&\qquad \times P[N^1(t)+N^{12}(t)<s_1, N^2(t)+N^{12}(t)<s_2 \mid N(t)=n] \\
&\quad = \sum_{n=0}^{s_1\vee s_2-1} P[N(t)=n] \\
&\qquad \times \sum_{\substack{\ell+m<s_1\\ \ell+j<s_2\\ \ell+m+j=n}} P[N^{12}(t)=\ell, N^1(t)=m, N^2(t)=j \mid N(t)=n] \\
&\quad = \sum_{\ell=0}^{s_1\wedge s_2-1}\sum_{m=0}^{s_1-\ell-1}\sum_{j=0}^{s_2-\ell-1} \frac{(\ell+m+j)!}{\ell!m!j!}(q^{12})^\ell (q^1)^m (q^2)^j \\
&\qquad \times \frac{(\lambda t)^{\ell+m+j}}{(\ell+m+j)!}e^{-\lambda t}. \qquad (5)
\end{aligned}$$

First, suppose $L_1 = L_2 = L$. This implies $P[W^{12} > L] = 0$. Using this fact and applying (4) and (5) we obtain

$$\begin{aligned}
E[W^{12}|L] &= \int_0^\infty P[W^{12} > w]\,dw \\
&= \int_0^L (1-P[W^{12}\le w])\,dw \\
&= L - \sum_{\ell=0}^{s_1\wedge s_2-1}\sum_{m=0}^{s_1-\ell-1}\sum_{j=0}^{s_2-\ell-1}\frac{(\ell+m+j)!}{\ell!m!j!} \\
&\quad \times (q^{12})^\ell (q^1)^m (q^2)^j \int_0^L \frac{(\lambda(L-w))^{\ell+m+j}}{(\ell+m+j)!} \\
&\quad \times e^{-\lambda(L-w)}\,dw \\
&= L - \frac{1}{\lambda}\sum_{\ell=0}^{s_1\wedge s_2-1}\sum_{m=0}^{s_1-\ell-1}\sum_{j=0}^{s_2-\ell-1}\frac{(\ell+m+j)!}{\ell!m!j!} \\
&\quad \times (q^{12})^\ell (q^1)^m (q^2)^j G^0(\ell+m+j \mid \lambda L).
\end{aligned}$$

The last equation follows because

$$\int_0^L \frac{(\lambda(L-w))^n}{n!}e^{-\lambda(L-w)}\,dw = \frac{1}{\lambda}G^0(n\mid\lambda L). \qquad (6)$$

According to Little's law, $\overline{B}^{12} = \lambda^{12}E[W^{12}]$, hence

$$\begin{aligned}
\overline{B}^{12}(\mathbf{s}|L) &= \lambda^{12}L - q^{12}\sum_{\ell=0}^{s_1\wedge s_2-1}\sum_{m=0}^{s_1-\ell-1}\sum_{j=0}^{s_2-\ell-1}\frac{(\ell+m+j)!}{\ell!\,m!\,j!} \\
&\quad \times (q^{12})^\ell (q^1)^m (q^2)^j G^0(\ell+m+j\mid\lambda L). \qquad (7)
\end{aligned}$$

It is interesting to note that in standard single-item inventory theory the average backorder is obtained first through the lead-time demand distribution, and then the average backorder-waiting time is derived from Little's law. Here, in the multi-item system, we find that the reversed procedure is more convenient and effective.

Observe that (7) is similar in form to (1), except that in (7) the second term becomes a weighted sum with multinomial weights. The complexity of the computation obviously depends on how many different demand types there are. If $q^{12}=0$, i.e., there is no Type-12 demand, then (7) reduces to 0. If $q^{12}=1$ (which implies $q^1=q^2=0$), then there is only Type-12 demand, and the system is essentially a single-item system. In this case, (7) reduces to (1) with $s_1\wedge s_2$ replacing s_i, i.e.,

$$\overline{B}^{12}(\mathbf{s}|L) = \lambda L - \sum_{\ell=0}^{s_1\wedge s_2-1} G^0(\ell\mid\lambda L). \qquad (8)$$

Thus, for these special cases, the computational effort needed to evaluate the order-based backorders is the same as in the single-item system.

Equation (8) indicates that in a pure assembly system with equal lead times the average number of backorders of the final product depends on the base-stock levels only through their minimum, so it does not make sense to have different base-stock levels in such systems.

The second step of the new approach focuses on solving the case $L_1 < L_2$. Set $L = L_1$ and $\Delta = L_2 - L_1$ so that $L_2 = L+\Delta$. The following property is important in the analysis, which is shown in the appendix:

Lemma 1. *The process $\{B^{12}(t), t\ge 0\}$ is cyclical in the sense that its statistic behavior repeats in all intervals $[m(L+\Delta), (m+1)(L+\Delta))$, $m = 0,1,2,\ldots$. As a consequence, $\overline{B}^{12}$ is the expected number of Type-12 backorders occurring during a cycle $[0, L+\Delta)$ that remain to be backlogged at $(L+\Delta)^-$.*

Based on the above observation, and following a detailed probabilistic analysis within a cycle in the appendix, we obtain

Proposition 1. *The expected number of Type-12 backorders has the following expression*:

$$\overline{B}^{12}(\mathbf{s} \mid L, L+\Delta) = \frac{\lambda^{12}}{\lambda_2}\overline{B}_2(s_2 \mid \Delta) + \lambda^{12} L\, G^0(s_2 - 1 \mid \lambda_2\Delta) + \sum_{y=0}^{s_2-1} g(y \mid \lambda_2\Delta)\overline{B}^{12}(\mathbf{s} - y\mathbf{e}_2 \mid L), \quad (9)$$

where $\overline{B}^{12}(\cdot \mid L)$ *is given by* (7). *The first term of* (9) *is the expected number of Type-*12 *backorders started in interval* $[0, \Delta)$, *which is independent of* s_1. *The rest is the expected number of Type-*12 *backorders started in interval* $[\Delta, L+\Delta)$.

Because both $G^0(s_2|\lambda_2\Delta)$ and $\overline{B}_2(s_2|\lambda_2\Delta)$ equal 0 when $\Delta = 0$, it is easy to check that in this case (9) reduces to (7), as expected. The computational complexity is dominated by the last term, which is a convolution of single-variate Poisson pmf and the Type-12 backorders under equal lead time. Again, it depends on the magnitude of s_is as well as the number of different demand types (positive q^Ks).

For the pure assembly system, i.e., $q^1 = q^2 = 0$, applying (1) and (8) to (9) yields

$$\overline{B}^{12}(\mathbf{s} \mid L, L+\Delta) = \lambda(L+\Delta) - \sum_{y=0}^{s_2-1}\left\{G^0(y \mid \lambda\Delta) + g(y \mid \lambda\Delta) \times \sum_{\ell=0}^{s_1\wedge(s_2-y)-1} G^0(\ell \mid \lambda L)\right\}. \quad (10)$$

Unlike in (8), the dependence of $\overline{B}^{12}$ on s_1 and s_2 is more intricate here. On the other hand, it is clear that increasing s_1 above s_2 will not decrease $\overline{B}^{12}$, so there is no reason to have $s_1 > s_2$.

3.2. The General J-Item System

We now consider the evaluation of $\overline{B}^K$ for a fixed-demand type K in the general J-item system. The challenge here is to find the right combinatorial argument and workable notation.

Because $\overline{B}^K$ depends on $\{D_i, i = 1, \ldots, J\}$ only through $\{D_i, i \in K\}$, it is more convenient to work on a transformed smaller system that consists of only the items in K. The overall demand rate to the new system is

$$\tilde{\lambda} = \sum_{A:\ A\cap K\neq\varnothing} \lambda^A.$$

For convenience, let $K_1, K_2, \ldots, K_p$ be all the demand types for the new system. Then, a demand to the new system is of type-K_α with probability

$$\tilde{q}^\alpha = \left(\sum_{A:\ K\cap A = K_\alpha} \lambda^A\right)\Big/\tilde{\lambda}.$$

Similar to the definition of $S(i)$ in the original system, for any $i \in K$, we define $\widetilde{S}(i)$ to be the set of all the demand types in the new system that require i. That is,

$$\widetilde{S}(i) = \{K_\alpha : i \in K_\alpha, \alpha = 1, \ldots, p\}.$$

For any $t \geq 0$, let $\widetilde{N}^\alpha(t)$ be the total number of arrivals of type-K_α demand in the new system by $t, \alpha = 1, \ldots, p$. Clearly, $\widetilde{N}^\alpha(t)$ has a Poisson distribution with rate $\tilde{q}^\alpha\tilde{\lambda}t$. Also, the total number of arrivals by t in the new system,

$$\widetilde{N}(t) = N^1(t) + \cdots + N^p(t),$$

has a Poisson distribution with rate $\tilde{\lambda}t$. Given $\widetilde{N}(t) = n$, $(\widetilde{N}^1(t), \ldots, \widetilde{N}^p(t))$ has a multinomial distribution with parameters n and $\tilde{q}^\alpha, \alpha = 1, \ldots, p$. Similar to (5), letting $\widetilde{S} = \max_{i\in K}\{s_i\}$, we obtain

$$\begin{aligned} P[D_i(t) < s_i, i \in K] &= \sum_{n=0}^{\tilde{s}-1} P[\widetilde{N}(t) = n] \times \sum_{j\in\mathcal{J}_K^n(\mathbf{s}-\mathbf{1})} P[\widetilde{N}^1(t) = j_1, \ldots, \widetilde{N}^p(t) = j_p | \widetilde{N}(t) = n] \\ &= \sum_{j\in\mathcal{J}_K(\mathbf{s}-\mathbf{1})} \frac{n!}{j_1!\cdots j_p!}(\tilde{q}^1)^{j_1}\cdots(\tilde{q}^p)^{j_p}\frac{(\tilde{\lambda}t)^n}{n!}e^{-\tilde{\lambda}t}. \end{aligned} \quad (11)$$

Here, $n = j_1 + \cdots + j_p$, and for any vector $\mathbf{y}$,

$$\mathcal{J}_K^n(\mathbf{y}) = \left\{\mathbf{j} = (j_1, \ldots, j_p): \sum_{\alpha:K_\alpha\in\widetilde{S}(i)} j_\alpha \leq y_i, i \in K, \sum_{\alpha=1}^p j_\alpha = n\right\},$$

and

$$\mathcal{J}_K(\mathbf{y}) = \left\{\mathbf{j} = (j_1, \ldots, j_p): \sum_{\alpha:K_\alpha\in\widetilde{S}(i)} j_\alpha \leq y_i,\ i \in K\right\}. \quad (12)$$

Suppose $L_i = L$ for all i. Then, applying (4), (11), and (6), similarly to the proof of (7) we obtain

$$\begin{aligned} E[W^K \mid L] &= \int_0^L P[W^K > w]\,dw \\ &= L - \int_0^L P[W^K \le w]\,dw \\ &= L - \frac{1}{\tilde{\lambda}} \sum_{\mathbf{j} \in \mathcal{J}_K(\mathbf{s}-\mathbf{1})} \frac{n!}{j_1! \cdots j_p!} (\tilde{q}^1)^{j_1} \cdots (\tilde{q}^p)^{j_p} \\ &\quad \times G^0(n \mid \tilde{\lambda} L), \end{aligned}$$

where $n = j_1 + \cdots + j_p$. Using Little's formula leads to

$$\begin{aligned} \overline{B}^K(\mathbf{s}|L) &= \lambda^K E[W^K|L] \\ &= \lambda^K L - \frac{\lambda^K}{\tilde{\lambda}} \sum_{\mathbf{j} \in \mathcal{J}_K(\mathbf{s}-\mathbf{1})} \frac{(j_1 + \cdots j_p)!}{j_1! \cdots j_p!} (\tilde{q}^1)^{j_1} \cdots (\tilde{q}^p)^{j_p} \\ &\quad \times G^0(j_1 + \cdots + j_p \mid \tilde{\lambda} L), \end{aligned} \tag{13}$$

which is a generalization of (7).

In the general case of unequal lead times, let $L_0 = 0$ and $\Delta_i = L_i - L_{i-1}$, then $L_i = \Delta_1 + \cdots + \Delta_i$. Without loss of generality, suppose $K = \{1, 2, \ldots, k\}$. For any $j \le k$ define a k-vector

$$\mathbf{L}^{(j)} = (L_1, \ldots, L_{j-1}, L_j, L_j, \ldots, L_j).$$

That is, $\mathbf{L}^{(j)}$ is a lead-time vector that has equal entries for components j through k. So, $\mathbf{L}^{(k)} = \mathbf{L}$, and $\mathbf{L}^{(1)}$ is a lead-time vector with equal entries of value L_1. Using similar arguments as in the two-item case, we can show that

$$\begin{aligned} \overline{B}^K(\mathbf{s} \mid \mathbf{L}) &= \frac{\lambda^K}{\lambda_k} \overline{B}_k(s_k \mid \Delta_k) + \lambda^K L_{k-1} G^0(s_k - 1 \mid \lambda_k \Delta_k) \\ &\quad + \sum_{y=0}^{s_k - 1} g(y \mid \lambda_k \Delta_k) \overline{B}^K(\mathbf{s} - y\mathbf{e}_k \mid \mathbf{L}^{(k-1)}). \end{aligned} \tag{14}$$

Proposition 2. *The expected number of type-K backorders in a general J-item system can be evaluated by applying* (14) *recursively, until the lead-time vector is reduced to* $\mathbf{L}^{(1)}$*, at which point Formula* (13) *applies. As a result, we can express* $\overline{B}^K(\mathbf{s} \mid \mathbf{L})$ *as the sum of convolutions of single-variate Poisson distributions and* $\overline{B}^K(\cdot \mid \mathbf{L}^{(1)}) = \overline{B}^K(\cdot \mid L_1)$.

Note that the procedure reduces the general problem with different lead times to a problem with equal lead times in $k-1$ steps. The computational time of (13) and (14) depends on the following factors: k (the size of K), the size of $s_i, i \in K$, the structure of the region $\mathcal{J}_K$, and the number of different lead times (i.e., positive Δ_is). Among these, the complexity of $\mathcal{J}_K$, which is determined by the number of demand types interacting with K (i.e., sharing common components), is the most dominating factor. This is due to the multinomial terms in (13).

The simplest case is when none of the items in K is required by any other demand types. That is, the subsystem $\{1, \ldots, k\}$ is a pure assembly system. (When $k = J$, the entire system is a pure assembly system.) In this case, (13) reduces to

$$\overline{B}^K(\mathbf{s} \mid \mathbf{L}) = \lambda^K L - \sum_{\ell=0}^{\min_{i \in K} s_i} G^0(\ell \mid \lambda^K L). \tag{15}$$

This is a generalization of (8), which depends on the number of components k only through $\min\{s_1, \ldots, s_k\}$; the rest of the computation is exactly the same as in the single-item system. Combining (14) and (15), we can see that the computational effort is of the order of up to k convolutions of single-variate Poisson distributions, which depends on the magnitude of s_is. The number of convolutions is determined by the number of positive Δ_is.

4. Bounds and Approximations

Relying on the joint distribution of the inventory levels, the order-based backorders are considerably more complicated to compute than item-based ones. It is therefore interesting and important to know whether the item-based backorders, which are computed through the marginal distributions, can be used to provide useful information about the order-based backorders. We discuss this issue in this section.

Using (3), we can establish lower and upper bounds on $\overline{B}^K$ that require simple operations of the order-based demand rates and the item-based backorders. All the expectations involved only use marginal distributions. Specifically,

$$\max_{i \in K} \overline{B}_i^K \le \overline{B}^K \le \sum_{i \in K} \overline{B}_i^K.$$

But $\overline{B}_i^K = (\lambda^K/\lambda_i)\overline{B}_i$, so

$$LB^K \overset{\text{def}}{=} \lambda^K \max_{i \in K} \frac{\overline{B}_i}{\lambda_i} \le \overline{B}^K \le \lambda^K \sum_{i \in K} \frac{\overline{B}_i}{\lambda_i} \overset{\text{def}}{=} UB^K.$$

Summing up all the lower bounds and upper bounds, respectively, we obtain a lower bound and an upper bound on the total average order-based backorders $\overline{B} = \sum_K \overline{B}^K$. That is,

$$LB \stackrel{\text{def}}{=} \sum_K LB^K \le \overline{B} \le \sum_K UB^K \stackrel{\text{def}}{=} UB.$$

It can be verified that

$$\begin{aligned} UB &= \sum_{i=1}^{J} \overline{B}_i \\ &= \text{the total average item-based backorders} := \overline{B}_I. \end{aligned}$$

So,

$$\overline{B} \le \overline{B}_I.$$

That is, the total item-based backorders always dominate the total order-based backorders.

A natural approximation for $\overline{B}^K$ is to take the average of LB^K and UB^K, denoted by AB^K. ("AB" stands for "Average of Bounds.") That is,

$$\begin{aligned} AB^K &= \frac{\lambda^K}{2} \max_{i\in K} \frac{\overline{B}_i}{\lambda_i} + \frac{\lambda^K}{2} \sum_{i\in K} \frac{\overline{B}_i}{\lambda_i} \\ &= \lambda^K \max_{i\in K} \left\{ \frac{\overline{B}_i}{\lambda_i} + \frac{1}{2} \sum_{j\in K\setminus\{i\}} \frac{\overline{B}_j}{\lambda_j} \right\}. \qquad (16) \end{aligned}$$

Consequently,

$$AB = \sum_K AB^K = \frac{1}{2}(LB + UB) \qquad (17)$$

is an approximation of $\overline{B}$.

In particular, in the two-item system, we have

$$\begin{aligned} AB^{12} &= \lambda^{12} \max\left\{ \frac{\overline{B}_1}{\lambda_1} + \frac{\overline{B}_2}{2\lambda_2}, \frac{\overline{B}_2}{\lambda_2} + \frac{\overline{B}_1}{2\lambda_1} \right\} \\ &= \max\left\{ \overline{B}_1^{12} + \frac{1}{2}\overline{B}_2^{12}, \overline{B}_2^{12} + \frac{1}{2}\overline{B}_1^{12} \right\}. \qquad (18) \end{aligned}$$

Applying (18) to (17) yields

$$\begin{aligned} AB &= \overline{B}^1 + \overline{B}^2 + \max\left\{ \overline{B}_1^{12} + \frac{1}{2}\overline{B}_2^{12}, \overline{B}_2^{12} + \frac{1}{2}\overline{B}_1^{12} \right\} \\ &= \max\left\{ \overline{B}_1 + \overline{B}_2 - \frac{1}{2}\overline{B}_2^{12}, \overline{B}_1 + \overline{B}_2 - \frac{1}{2}\overline{B}_1^{12} \right\} \\ &= \overline{B}_1 + \overline{B}_2 - \frac{\lambda^{12}}{2} \min\left\{ \frac{\overline{B}_1}{\lambda_1}, \frac{\overline{B}_2}{\lambda_2} \right\}. \qquad (19) \end{aligned}$$

Thus, to approximate the total order-based backorders, AB takes some "correction" from the total item-based backorders $\overline{B}_1 + \overline{B}_2$.

We tested the effectiveness of the bounds and the approximation through some numerical examples. In a two-item system with $L_1 = 1, L_2 = 2$, and $\lambda = 20$, four vectors (q^{12}, q^1, q^2) were chosen, representing different levels of demand correlation. They are $(0.2, 0.4, 0.4), (0.5, 0.25, 0.25), (0.8, 0.1, 0.1)$, and $(0.5, 0.33, 0.17)$. With each system configuration, the base-stock levels were set following the form of

$$s_i = \lambda_i L_i + z_i \sqrt{\lambda_i L_i}, \qquad (20)$$

with z_i varying among $0, 0.67$, and 1.64. Since s_i are integers in our model, we assigned to s_i the integer part of the right-hand-side value in (20). We observe that the approximation AB provides reliable information on $\overline{B}$, especially for higher values of z_i. In particular, in this set of 36 examples, the percentage error of AB, defined by $100|AB - \overline{B}|/\overline{B}$, is 2.82. The percentage errors of LB and $\overline{B}_I$ (defined similarly), on the other hand, are 11.27 and 10.37, respectively. This observation remains true for other experiments with different values of λ and L_i. (Section 7 provides more evidence in a six-item, six-product system.) Therefore, we recommend AB be used in practice as a quick estimate of $\overline{B}$.

Before concluding this section, it is worth explaining the rationale of using (20). First, recall that the lead-time demand D_i has the Poisson distribution with mean $\lambda_i L_i$. Because a Poisson distribution with a large mean can be approximated well by a normal distribution, we can treat D_i as a normal random variable with mean $\lambda_i L_i$ and standard deviation $\sqrt{\lambda_i L_i}$, assuming $\lambda_i L_i$ is not too small. (See the accurate performance of this approximation when $\lambda_i L_i = 5$ and 10 in Figure 6.4.1 of Zipkin 2000). Standard inventory-planning models suggest that the base-stock level for item i takes the form of (20), where z_i is called the safety factor. This policy is optimal if we minimize the average inventory-holding cost of item i subject to an item-i fill rate constraint. In particular, $z_i = 0, 0.39, 0.67, 1.04$, and 1.64 correspond to item-i's fill rates 50%, 65%, 75%, 85%, and 95%, respectively. This policy is also optimal if we minimize the average inventory and backorder cost. Assuming h_i to be the unit-holding cost rate and b_i the unit backorder cost rate, then $\Phi(z_i) = b_i/(b_i + h_i)$, where $\Phi(\cdot)$ is the standard normal cdf. See, for example, §§6.4.2 and 6.5.2 in Zipkin (2000). Finally, this total cost minimization model is equivalent to a problem minimizing the average inventory-holding cost subject to an upper bound on the average number of backorders. The unit backorder cost rate in the former model corresponds to the Lagrangian multiplier of the constraint in the latter.

Given that the form of the optimal policy is unknown for the system studied here, (20) represents reasonable and plausible policy parameters—those widely used in practice employing the standard item-based inventory-planning tools. We follow (20) to set the base-stock levels in all the numerical examples in this paper.

5. The Value of Order-Based Demand Information

In practice, managers often make inventory decisions based on the item-based information, such as the demand rates λ_i and the lead times L_i. There are two reasons. First, the item-based information is much easier to obtain; its forecast can be made independent of other items. Second, the standard inventory-planning models are mostly item based; they assume the demands for different items are independent. We now investigate the implication of such a practice if the demands for different items can occur at the same time and hence are correlated. This would help reveal the value of identifying different demand types and estimating demand correlation. To keep notation simple, we focus on the two-item system. Here, the Type-12 demand couples the demand processes of both items.

Assume $\lambda_i, i = 1, 2$ are fixed, which are the item demand rates forecasts. Suppose that we control the item inventories using the base-stock levels given in (20), but the real demand environment is multiclass as captured in our model. Indeed, the proportion of Type-12 demand q^{12} is a measure of demand correlation. Note that the relation

$$\lambda_i = (q^i + q^{12})\lambda, \quad i = 1, 2$$

leads to

$$\lambda = (\lambda_1 + \lambda_2)/(1 + q^{12}). \tag{21}$$

Because λ_1 and λ_2 are fixed, the higher q^{12} is, the lower is the overall demand rate λ. Also, as q^{12} changes, to keep λ_i fixed, we must have

$$q^i = \frac{\lambda_i}{\lambda_1 + \lambda_2} - \left(1 - \frac{\lambda_i}{\lambda_1 + \lambda_2}\right) q^{12}, \quad i = 1, 2$$

and

$$q^{12} \le \min\left\{\frac{\lambda_1}{\lambda_2}, \frac{\lambda_2}{\lambda_1}\right\}.$$

Under any fixed inventory policy, the average item backorders $\overline{B}_1, \overline{B}_2$ and the average item backorder waiting times $\overline{W}_1 = \overline{B}_1/\lambda_1$ and $\overline{W}_2 = \overline{B}_2/\lambda_2$ are all constant as q^{12} changes. The question is: How would the total average order-based service measure behave? Is it sensitive to q^{12}? If it is not, then it is not very important to identify the demand correlation, in particular, to estimate q^{12}. In other words, it is sufficient to forecast λ_i alone. Otherwise, it is worthwhile investigating the demand correlation. In that case, a natural question would be: Can safety factors z_i be chosen cleverly so that the order-based service measure is relatively insensitive to q^{12}?

These questions can be addressed by examining the average *customer* waiting time $\overline{W} = \overline{B}/\lambda$ as q^{12} increases. Given that the overall demand rate λ decreases as q^{12} increases, it makes sense to study the effect on the average waiting time only. Numerical results (not reported here) show that $\overline{W}$ increases in q^{12} regardless of the level of the safety factors. Increasing the safety factor of one item improves the overall order-based waiting time. However, there does not appear to be a simple way of adjusting the item-based inventory planning to keep the order-based performance at a constant level. This suggests that acquiring order-based demand information is important.

This can also be seen from the approximation derived in the last section. Let $\widehat{W} = AB/\lambda$ be the approximate average waiting time. Then, according to (19) and (21) and assuming $\overline{W}_1 \le \overline{W}_2$, we have

$$\widehat{W} = (1+q^{12})(\overline{B}_1+\overline{B}_2)/(\lambda_1+\lambda_2) - \frac{q^{12}}{2}\min\{\overline{W}_1, \overline{W}_2\}$$
$$= (1+q^{12})(\lambda_1\overline{W}_1+\lambda_2\overline{W}_2)/(\lambda_1+\lambda_2) - \frac{q^{12}}{2}\overline{W}_1$$
$$= \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\overline{W}_1 + \frac{\lambda_2}{\lambda_1+\lambda_2}\overline{W}_2\right) + \left[\frac{\lambda_2}{\lambda_1+\lambda_2}(\overline{W}_2-\overline{W}_1)+\frac{1}{2}\overline{W}_1\right]q^{12}. \quad (22)$$

(A symmetric equation can be obtained by assuming $\overline{W}_2 \le \overline{W}_1$.) Since $\overline{W}_2 - \overline{W}_1 \ge 0$, the coefficient of q^{12} is positive. Thus, approximately, the average customer waiting time increases linearly in q^{12}.

6. A Two-Product Example: The Effect of Component Commonality

In this section we consider an example of two symmetric products, each of which consists of two components. Product 1 is made from Components 1 and 3, and Product 2 is made from Components 2 and 4. Also, both products have the same demand rate μ. Components 1 and 2 are product specific and have the same lead times L. Components 3 and 4 have longer lead times $L' = L + \Delta$. Our objective is to examine the changes in the inventory-service trade-offs after combining Components 3 and 4 into one common component, Component 5. Assume the lead time for Component 5 remains L'.

This two-product example has been used in a number of studies on the effect of component commonality issues. See, for example, Baker et al. (1986), Eynan and Rosenblatt (1996), Hillier (1999), and the references therein. However, these works assume either there is only a single period or there are no replenishment lead times. In both cases, $L = L' = 0$. It was recognized that the analysis of models with dynamic demand and positive lead times, such as the one studied in the current paper, is considerably harder; see, e.g., Baker (1985). It is therefore of interest to see whether the findings drawn from the single-period models can be extended to a more general and realistic setting.

The inventory measure we use is the total investment in the base-stock levels, i.e., $\sum_i c_i s_i$, where c_i is the unit cost of item i. The service measure is the average total product backorders $\overline{B}$, which is proportional to the average customer waiting time. We would like to see, for the same service level, how much inventory investment can be saved by introducing a common component. The overall optimization of this problem is a difficult one, and we shall pursue it in future research. What we are interested in here is whether we can see the benefit of component commonality by simple alterations of the original inventory policy.

Clearly, for the original system, Components 1 and 2 should have the same base-stock levels, say s_1. Similarly, let s_3 be the base-stock level for Components 3 and 4. In the new system, we set the same base-stock levels for the product-specific Components 1 and 2 as in the original system, while changing the base-stock level for the common Component 5, say s_5. We plot the total average backorders of the new system as a function of s_5 and compare it against the total average backorders in the old system, which is a constant over s_5. Without loss of generality, we assume $c_3 = 1$. Further assuming $c_5 = c_3$, the difference between the inventory investments before and after introducing the common component equals $2s_3 - s_5$. (Note that this is an upper bound for inventory savings if $c_5 > c_3$.) It is thus interesting to see whether the two curves cross *before* s_5 increases to $2s_3$, implying that the new system can reach the same service level with less inventory investment.

For convenience, we refer to the original system as "System-NC," where "NC" indicates "No Commonality." Applying (10), we obtain the following common expression of the average backorders for each product:

$$\overline{B}_{\mathrm{NC}}(s_1, s_3) = \mu(L+\Delta) - \sum_{y=0}^{s_1-1}\Bigg[G^0(y \mid \mu\Delta) + g(y \mid \mu\Delta) \times \sum_{\ell=0}^{s_1\wedge(s_3-y)-1} G^0(\ell \mid \mu L)\Bigg].$$

Similarly, we refer to the new system as "System-C," where "C" indicates "with Commonality." Note that in System-C Component 5 has a demand rate 2μ, and Product 1 corresponds to the type-K demand with $K = \{1, 5\}$. Since K has only two items, we also say this is a Type-15 demand. In the subsystem consisting of Components 1 and 5, there are two demand types: Type-15 and Type-5. The overall demand rate is $\tilde{\lambda} = 2\mu$. The relative proportion of each demand type is given by $\tilde{q}^{15} = \tilde{q}^{5} = \mu/(2\mu) = 1/2$. Now, applying (9), (7) (replacing Index 2 by Index 5), and (1) yields the following expression for the average backorders for Product 1:

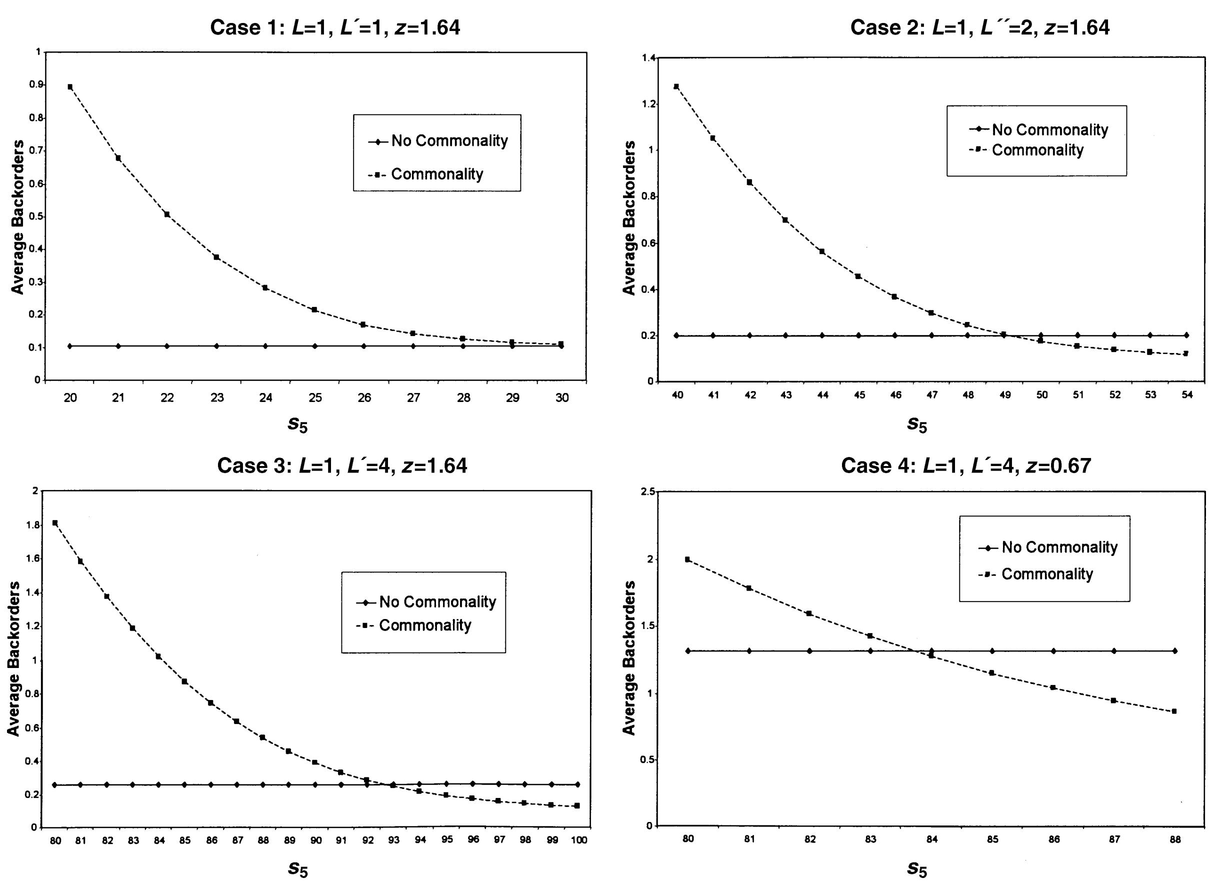
$$\begin{aligned}\overline{B}_{\mathrm{C}}(s_1, s_5) &= \frac{1}{2}\overline{B}_5(s_5 \mid \Delta) + \mu L\, G^0(s_5 - 1 \mid 2\mu\Delta) \\ &\quad + \sum_{y=0}^{s_5-1} g(y \mid 2\mu\Delta)\overline{B}^{15}(s_1, s_5 - y \mid L) \\ &= \mu(L+\Delta) - \frac{1}{2}\sum_{y=0}^{s_5-1}\Bigg[G^0(y \mid 2\mu\Delta) + g(y \mid 2\mu\Delta) \\ &\quad \times \sum_{\ell=0}^{s_1 \wedge (s_5-y)-1}\ \sum_{j=0}^{s_5-y-\ell-1}\frac{(\ell+j)!}{\ell!\,j!}\left(\frac{1}{2}\right)^{\ell+j} G^0(\ell+j \mid 2\mu L)\Bigg].\end{aligned}$$

Due to symmetry, the above formula is also the average backorder for Product 2.

Let s_1 and s_3 be set as in (20). Also, let both items have the same safety factor, $z_1 = z_3 = z$, where z varies from 0 to 1.64. For a given value of z, L, and Δ, s_1 and s_3 are fixed, therefore $\overline{B}_{\mathrm{NC}}(s_1, s_3)$ is a constant. Figure 1 shows the changes of $\overline{B}_{\mathrm{C}}(s_1, s_5)$ as s_5 increases from $2\mu L'$ (corresponding to the safety factor $z_3 = 0$) to $2s_3$. (Note that $2s_3 > 2\mu L'$.)

Case 1 in Figure 5 shows a surprising result: If $L = L'$, then there is no inventory saving in System-C. Even at the point $s_5 = 2s_3$, we still have $\overline{B}_{\mathrm{C}}$ slightly larger than $\overline{B}_{\mathrm{NC}}$. This is perhaps due to the fact that

Figure 1 **Effect of Commonality**



we use the same base-stock level for Component 1 in both systems. It is likely that the optimal policy for System-C is to increase the stock level for the product-specific component (Component 1) but to keep less combined stock level for the common component, i.e., $s_5 < 2s_3$. However, in that case the overall inventory savings can still be hard to see if Component 3, and hence the common Component 5, have much higher costs than the product-specific components.

Nonetheless, in the other cases in which $L' > L$, we do see $\overline{B}_{\rm C}$ reaches $\overline{B}_{\rm NC}$ before s_5 reaches $2s_3$, and this effect is more pronounced when L' increases. Comparing all these cases, we can see that, within the type of policies specified here, the benefit of using common components is larger when the lead time of the common component differs more from that of the product-specific component.

But exactly how much inventory savings can one expect from using a common component when $L' > L$? One conjecture is that if we use the same safety factor for both Items 3 and 5, i.e., let $z_3 = z_5 = z$, then $\overline{B}_{\rm C} \le \overline{B}_{\rm NC}$. If so, then the safety stock savings is at least $1 - z\sqrt{2\mu L'}/(2z\sqrt{\mu L'}) = 1 - 1/\sqrt{2} = 29\%$. Numerical results show (not reported here), however, that this is not always the case. In some cases within this setting, $\overline{B}_{\rm C} > \overline{B}_{\rm NC}$.

In conclusion, using a model with dynamic demands and positive lead times, we find that the benefit of using component commonality is not as easy to see as one would imagine. The benefit depends on how carefully inventory is managed. Simple alterations of the inventory policies for systems without common components may not lead to the advantage of using common components. This is specially true when the lead times for both product-specific components and common components are close. It is also possible that we do not see the benefit if the unit cost of the common component is much higher than those of the product-specific components.

Finally, we note that the analysis and formulas in this section apply to systems with any number of components and products, as long as each product is made of two components.

7. A Personal Computer Example: The Impact of Product Structure

In this section we consider an assemble-to-order personal computer setting. Suppose there are six items that play a key role in differentiating major demand types. These items are:

(1) built-in zip drive;
(2) standard hard drive;
(3) high-profile hard drive;
(4) DVD-Rom drive;
(5) standard processor; and
(6) high-profile processor.

There are six major demand types resulting from different choices and combinations of these items; their compositions are, respectively, $\{2,5\}$, $\{3,5\}$, $\{1,2,5\}$, $\{1,3,6\}$, $\{1,3,4,5\}$, and $\{1,3,4,6\}$. (See Figure 2.) The percentages of the individual types of demands relative to the total demand are $q^{25} = 0.10$, $q^{35} = 0.40$, $q^{125} = 0.15$, $q^{136} = 0.10$, $q^{1345} = 0.20$, $q^{1346} = 0.05$. The overall demand rate λ is 8. The item lead-time vector is $\mathbf{L} = (L_1, L_2, L_3, L_4, L_5, L_6) = (1, 1, 1, 1, 2, 2)$.

For simplicity and ease of presentation, we focus on base-stock policies with equal safety factors across items. That is, the base-stock levels are chosen according to (20) with $z_i = z$ for all i. This kind of policy is commonly used in practice; see, for example, Agrawal and Cohen (2001) and Hausman et al. (1998).

Figure 3 plots $\overline{B}$, AB, LB, and $\overline{B}_I$ as the common safety factor z increases from 0 to 1.64 gradually. It is

Figure 2 The PC Example: Product Structure

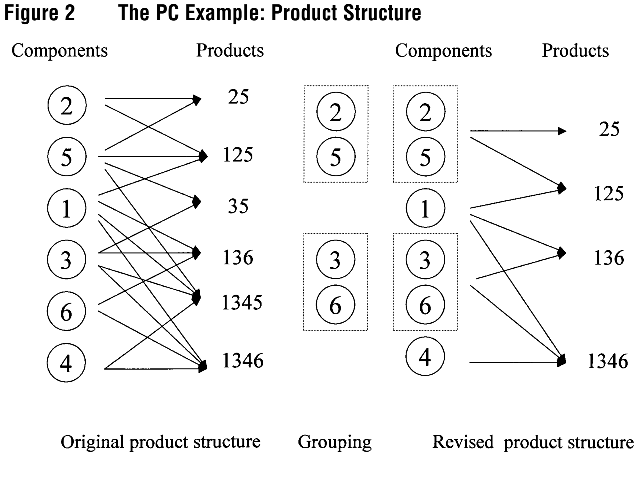
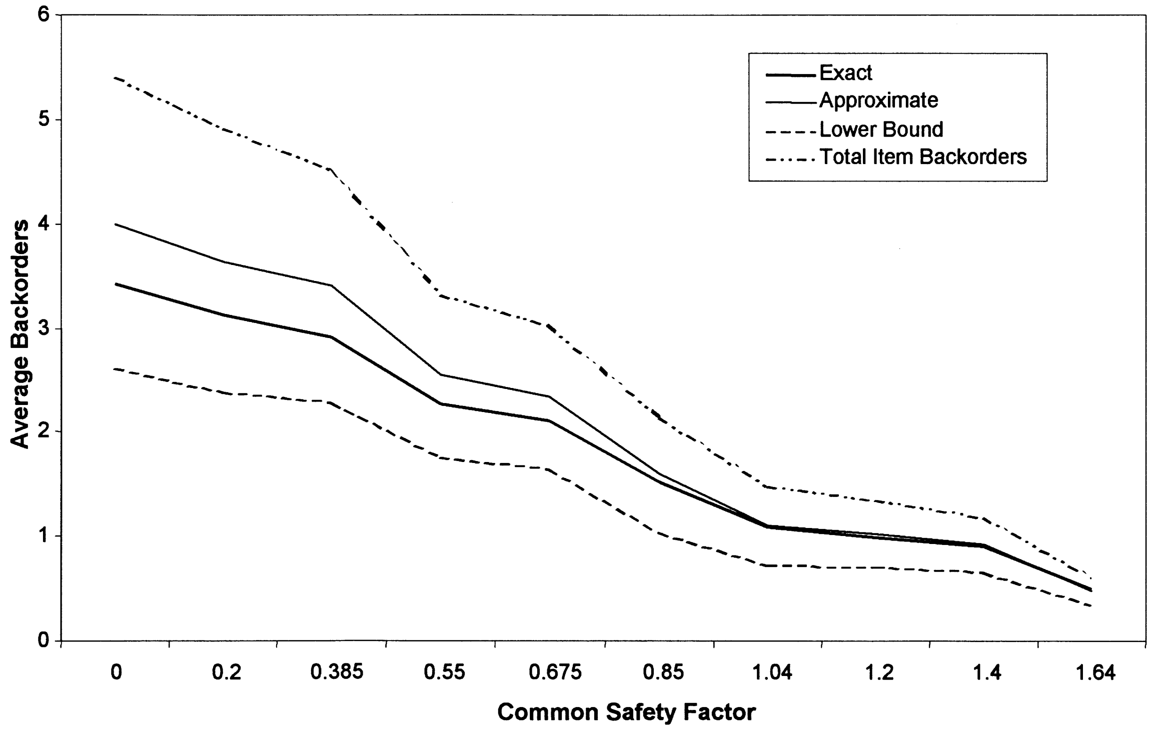


Figure 3 Personal Computer Example: Exact Result vs. Approximations



easy to observe that the approximation AB performs extremely well for higher values of the safety factor, corresponding to item fill rates 85% and above. The same observation applies to other experiments with different values of λ and L_i.

We now discuss the impact of product structure. In the current setting, the PC manufacturer/distributer allows the customers to specify what they want in terms of the combination of standard or high-profile hard drives and processors. Suppose the company wants to simplify the product feature offerings by grouping the standard and high-profile options together. That is, it offers the standard option $\{2, 5\}$ and the high-profile option $\{3, 6\}$ only and eliminates the demand types that consist of $\{3, 5\}$ and $\{2, 6\}$. In other words, customers now can choose from the combinations of the following:

(1) built-in zip drive;

(4) DVD-Rom drive;

(2, 5) standard option for processor and hard drive; and

(3, 6) high profile option for processor and hard drive.

(See Figure 2.) Here, options (2, 5) and (3, 6) are exclusive, that is, choose either one but not both. The question is: What is the operational implication of the new strategy? Would this significantly improve the average customer waiting time, or equivalently the total order-based backorders $\overline{B}$?

We need to determine the market demand for the new product line. First, we assume that the total demand rate λ is not affected. Second, we assume that half of the original demand Type 35 will divert into Type 25, and half into Type 136. Third, the original Types 1345 and 1346 will converge to Type 1346. This is the simplest such scenario, though of course not the only plausible one. Thus, there are four major demand types in the new strategy. They are $\{2, 5\}, \{1, 2, 5\}, \{1, 3, 6\}$, and $\{1, 3, 4, 6\}$. The corresponding percentages are

$$q^{25} = 0.30, q^{125} = 0.15, q^{136} = 0.30, \text{ and } q^{1346} = 0.25.$$

Figure 4 compares the average backorders in the original and revised systems. Here, we use the same common safety factor z in both systems. (Note that by changing the product structure, the item demand rates have been changed in the revised system. Therefore, the base-stock levels in the revised system are set according to the revised item demand rates.) It is surprising to observe that simpler product structure does *not* necessarily lead to fewer total order-based backorders for the same level of common safety factor. One plausible explanation is that the type of policy we employed here, though commonly used in practice, is suboptimal for this system. The true optimal policy should take into account the demand correlation across items, which likely results in different ways of setting item safety factors. For example, Item 1 is shared by more products than any other item, so its safety factor perhaps should be higher too. However, exactly how much higher, and how to relate the safety factors of all items, should be addressed in future research.

It is tempting to conjecture that in the revised system the total inventory investment is less. This is harder to measure, however, because it depends on the unit cost of each item. In the literature, the sum of the base-stock levels $\sum_i s_i$ is sometimes used as a measure of inventory investment. This is viable if the item costs are similar. Applying this measure here, we find many counterexamples to the conjecture. In particular, Figure 5 presents the same set of data as in Figure 4 in an alternative way. Instead of viewing the average backorder as a function of the common safety factor, we now view it as a function of the total base stock ($\sum_i s_i$) corresponding to the safety factor. As the figure shows, $\overline{B}$ is not necessarily *smaller* under the revised product structure, even with the same number of base-stock units. Thus, the presumed value of simplified product structure cannot be easily achieved by simple adjustments of the item-based inventory policies. It depends critically on how carefully inventory is managed by taking into account demand correlation.

Figure 4 **Effect of Product Structure: Total Backorders vs. Common Safety Factor**

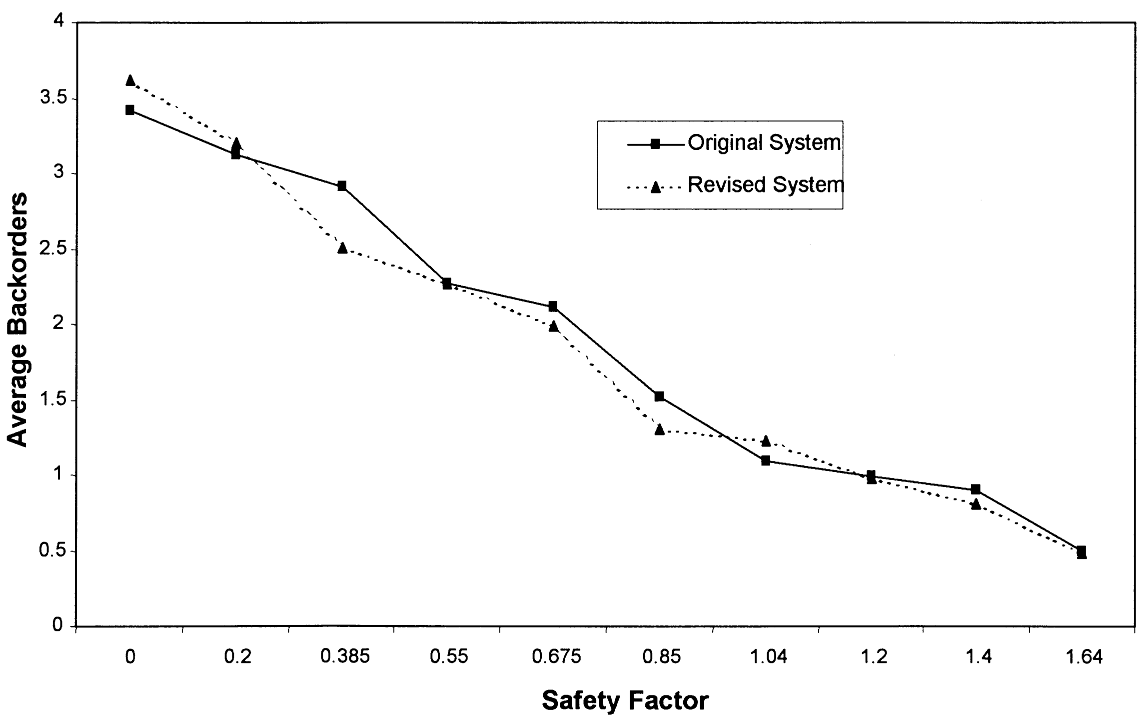
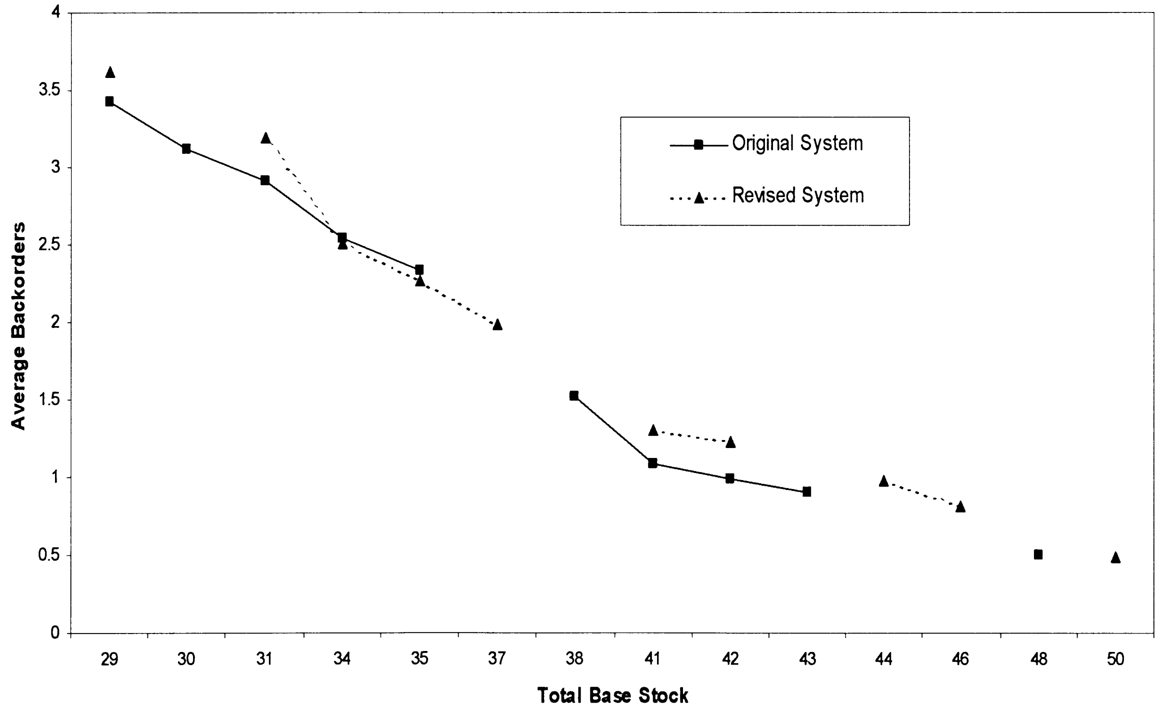


Figure 5 Effect of Product Structure: Total Backorders vs. Total Base Stock



8. Batch Demands

Although the paper focuses on the unit demand systems, the analysis can be extended to systems with batch demands. For any fixed $K \subseteq \{1, \ldots, J\}$, assume that the number of units of item i requested by a type-K demand, Z_i^K, is a positive integer random variable, $i \in K$. Let ψ^K and Ψ^K be the joint pmf and cdf of the batch size vector $\mathbf{Z}^K = (Z_i^K)_{i \in K}$, respectively. Define

$$Z_i = \text{demand batch size for item } i = Z_i^K \text{ with probability } \lambda^K/\lambda_i;$$

$$\Psi_i = \text{cdf of batch size } Z_i = \sum_{K \in S(i)} (\lambda^K/\lambda_i)\Psi_i^K.$$

Let D_i be the steady-state lead-time demand of item i. Then, D_i has a compound Poisson distribution with jump parameter $\lambda_i = \sum_{K \in S(i)} \lambda^K$ and batch size Z_i. Denote by $\gamma(\cdot \mid \mu, U)$ and $\Gamma(\cdot \mid \mu, U)$ the pmf and cdf, respectively, of the compound Poisson distribution with jump parameter μ and batch size distribution U, and denote $\Gamma^0 = 1 - \Gamma$. Then the average item-i backorder is given by

$$\overline{B}_i(s_i \mid L_i) = E[(D_i - s_i)^+] = \sum_{k=s_i}^{\infty} \Gamma^0(k|\lambda_i L_i, \Psi_i).$$

To compute $\overline{B}^K$ again, it is more convenient to work with a transformed system that contains items in K only, as in §3.2. In addition to the transformed quantities introduced in §3.2, let Y^α be the demand size of type K_α in the new system and let ϕ^α be its probability mass function. Then, with probability q^A, $\phi^\alpha = \psi^A$ if $A \cap K = K_\alpha$. Let $\mathcal{D}(t) = (D_i(t) : i \in K)$ be the cumulative demand vector by t, and define for every $i \in K$:

$$V_i^\alpha = \begin{cases} Y_i^\alpha, & \text{if } i \in K_\alpha; \\ 0, & \text{otherwise.} \end{cases}$$

Denote $V^\alpha = (V_i^\alpha : i \in K)$. Then

$$\mathcal{D}(t) = \sum_{m_1=1}^{N^1(t)} V^1(m_1) + \cdots + \sum_{m_p=1}^{N^p(t)} V^p(m_p),$$

where $V^\alpha(m_\alpha)$ are independent copies of V^α. Conditioning on $N^\alpha(t) = j_\alpha$, $\alpha = 1, \ldots, p$, $\mathcal{D}(t)$ has the probability mass function

$$\phi^{(\mathbf{j})} = [\phi^1]^{j_1} * [\phi^2]^{j_2} * \cdots * [\phi^p]^{j_p},$$

where $*$ means convolution and $[\phi^\alpha]^j$ stands for j-fold convolution of ϕ^α. Using a similar approach as in §3.2, we can show that

$$\begin{aligned}\overline{B}^K(\mathbf{s}|L) = \lambda^K L - \frac{\lambda^K}{\tilde{\lambda}} \sum_{\mathbf{z}\le\mathbf{s}} \psi^K(\mathbf{z}) \\ \times \sum_{\mathbf{j}\in\mathcal{J}_K(\mathbf{s}-\mathbf{z})} \frac{n!}{j_1!\cdots j_p!}(q^{-1})^{j_1}\cdots(q^{-p})^{j_p} \\ \times \sum_{\mathbf{x}\le\mathbf{s}-\mathbf{z}} \phi^{(\mathbf{j})}(\mathbf{x}) G^0(n \mid \tilde{\lambda} L)\end{aligned}$$

with $n = j_1 + \cdots + j_p$. Suppose $K = \{1, 2, \ldots, k\}$, we also have

$$\begin{aligned}\overline{B}^K(\mathbf{s} \mid \mathbf{L}) = \frac{\lambda^K}{\lambda_k}\overline{B}_k(s_k \mid \Delta_k) + \lambda^K L_{k-1} G^0(s_k - 1 \mid \lambda_k \Delta_k) \\ + \sum_{y=0}^{s_k-1} \gamma(y \mid \lambda_k\Delta_k, \Psi_k)\overline{B}^K(\mathbf{s} - y\mathbf{e}_k \mid \mathbf{L}').\end{aligned}$$

Here, $\mathbf{L}' = (L_1, L_2, \ldots, L_{k-1}, L_{k-1})$. Applying the same technique to $\overline{B}^K(\mathbf{s} - y\mathbf{e}_k|\mathbf{L}')$, and continuing in the same fashion, we can express $\overline{B}^K(\mathbf{s}|\mathbf{L})$ as the sum of convolutions of one-dimensional compound Poisson distributions and $\overline{B}^K(\cdot|L)$.

Clearly, the exact result for the batch demand case is much more computationally demanding, due to the complication of convolutions of the batch size distribution. Therefore, it is of practical use only if the latter has a simple form. In general, we propose the following approximation, building on the lower-bound idea used in the unit demand case. First, the average number of backorders of item i that are due to demand type K is approximately

$$\frac{\lambda^K E[Z_i^K]}{\sum_{A\in S(i)} \lambda^A E[Z_i^A]}\overline{B}_i.$$

This quantity divided by $E[Z_i^K]$ approximates the average number of type-K backorders that have item i in short. Since the type-K backorder is the maximum of the latter among all items, applying Jensen's inequality we obtain an approximation

$$\overline{B}^K \approx \lambda^K \max_{i\in K}\left\{\overline{B}_i \Big/ \sum_{A\in S(i)} \lambda^A E[Z_i^A]\right\}.$$

The approximation is quite effective; we refer the reader to Lu et al. (2001), in which a more general model with independent, identically distributed leadtimes is studied, for more details.

9. Concluding Remarks

In this paper we studied continuous review, multi-item inventory systems with multiple demand classes. Each demand class requires a specific subset of the items in stock. The model also applies to the assemble-to-order manufacturing systems. The primary concern is how to evaluate the average order-based backorders or the average number of customers whose orders are not yet filled, for a given base-stock policy. Both exact and approximation approaches were developed. As shown in Song (2000), these methods can be easily adapted to more general systems with batch ordering policies.

The results developed here greatly enhance the analytical and computational tractability of performance analysis. As such, they lay a foundation for further development of performance optimization techniques, including characterization of optimal policies for such systems.

The results can also be used as an analytical tool to study broader strategic and operational issues. Sections 5 to 7 provided preliminary examples; many insights have been derived with regard to multi-item inventory-planning decisions and product structure issues. Much more effort is needed, however, to resolve related modeling issues such as performance optimization, and to identify the key determinants of optimal inventory policies. We intend to study these issues in future research. Indeed, some progress has been made; see, for example, Song and Yao (2000) and Lu et al.

Acknowledgments

The author thanks the editor-in-chief, Hau Lee, the associate editor, and the referees for their suggestions to improve the paper's exposition. The author also thanks Savas Dayanik, the late Morton Klein, and Paul Zipkin for their helpful comments on earlier drafts. The seminar participants at Duke University, Northwestern University, INSEAD, and Ecole Centrale Paris provided useful feedback as well. The research was supported in part by NSF Grants DMI-9700037, DMI-9896339, and DMI-0084922.

Appendix

Proof of Lemma 1. Because this is a base-stock system, each item demand triggers a replenishment order for that item. Also, since we assume nonnegative base-stock levels and the lead times are constant, each item backorder must be filled by a replenishment triggered by an earlier demand. In other words, a backorder for item i occurred at time t must be filled by an order placed earlier than t, i.e., it must be filled by time $t+L_i$. This implies that all the existing Type-12 backorders at time 0^- will have been cleared by $L+\Delta$. As a consequence, $\{B^{12}(t), t \geq 0\}$ demonstrates a cyclical behavior. In particular,

$$B^{12}(t) = C(t) + R(t), \quad t \in [0, L+\Delta),$$

where

$C(t)$ = cumulative new Type-12 backorders incurred since time 0;
$R(t)$ = remaining Type-12 backorders incurred before time 0.

$C(t)$ is increasing and $R(t)$ is decreasing in the interval $[0, L+\Delta)$. Also, $C(0)=0$, $R(L+\Delta-)=0$, $R(L+\Delta)=C(L+\Delta-)$, and $C(L+\Delta)=0$. This pattern repeats in all intervals $[m(L+\Delta), (m+1)(L+\Delta))$, $m=0,1,2,\ldots$. Also, the statistical behavior of the two processes $C(t)$ and $R(t)$ is identical in each of these intervals. Hence, $\overline{B}^{12}$ is the expected number of Type-12 backorders occurred during a cycle $[0, L+\Delta)$ that remain to be backlogged at $(L+\Delta)^-$. □

Proof of Proposition 1. According to Lemma 1, $\overline{B}^{12}$ is the expected number of Type-12 backorders occurring during a cycle $[0, L+\Delta)$ that remain to be backlogged at $(L+\Delta)^-$. We now calculate this expected value by conditioning on the time interval in which the backorders started. First, define

$IO_2(\Delta)$ = outstanding orders of Item 2 at time Δ;
$IO_2^{\text{old}}(\Delta)$ = orders of Item 2 that were placed before the cycle but remain outstanding at time Δ;
$IO_2^{\text{new}}(\Delta)$ = orders of Item 2 that were placed during $[0, \Delta)$;
$IP_i(\Delta)$ = inventory position of Item 2 at time $\Delta = IN_i(\Delta) + IO_i(\Delta)$, $i=1,2$.

Obviously,

$$IO_2(\Delta) = IO_2^{\text{old}}(\Delta) + IO_2^{\text{new}}(\Delta).$$

Also, since $L_2 = L+\Delta$, $IO_2^{\text{old}}(\Delta)$ will have arrived to the system by time $L+\Delta$ while $IO_2^{\text{new}}(\Delta)$ will still remain outstanding at that time. According to the law of flow conservation, we have

$$IN_2(L+\Delta) = IN_2(\Delta) + IO_2^{\text{old}}(\Delta) - D_2(\Delta, L+\Delta). \tag{A1}$$

By the definition of base-stock policy, we have $IP_2(\Delta) = s_2$. Also, $IO_2^{\text{new}}(\Delta) = D_2(\Delta)$. So

$$s_2 = IN_2(\Delta) + IO_2(\Delta) = IN_2(\Delta) + IO_2^{\text{old}}(\Delta) + D_2(\Delta),$$

which leads to

$$IN_2(\Delta) + IO_2^{\text{old}}(\Delta) = s_2 - D_2(\Delta). \tag{A2}$$

In other words, at Δ the Item 2 base-stock coverage until $L+\Delta$ is $s_2 - D_2(\Delta)$. (A2) is the key for the rest of the proof.

Suppose the quantity in (A2) is positive, then it means that any backorders occurred before Δ will be filled by $L+\Delta$. If the quantity is negative, then we must have $IO_2^{\text{old}}(\Delta) = 0$ and $IN_2(\Delta) < 0$. But, backorders occurred before 0 should be filled by orders that were triggered before 0. So, $IO_2^{\text{old}}(\Delta) = 0$ implies that the backorders of Item 2 at Δ, $[-IN_2(\Delta)]^+$ all occurred during $[0, \Delta)$. Combining these arguments, we conclude that the expected number of Item 2 backorders that started in interval $[0, \Delta)$ equals

$$E[(D_2(\Delta) - s_2)^+] = \overline{B}_2(s_2 \mid \Delta).$$

Notice that for Type-12 demands that arrived during $[0, \Delta)$, only those backlogged for Item 2 may remain backlogged at $(L+\Delta)^-$, because all the backlogs for Item 1 that occurred in $[0, \Delta)$ will have been filled by then. Thus, the number of Type-12 backorders started in interval $[0, \Delta)$ have expected value

$$\frac{\lambda^{12}}{\lambda_2} \overline{B}_2(s_2 \mid \Delta). \tag{A3}$$

Now consider Type-12 backorders in $[\Delta, L+\Delta)$. Applying (A2) to (A1) yields

$$IN_2(L+\Delta) = s_2 - D_2(\Delta) - D_2(\Delta, L+\Delta). \tag{A4}$$

By the definition of IP_1 and the law of flow conservation, we also have

$$IN_1(L+\Delta) = IP_1(\Delta) - D_1(\Delta, L+\Delta) = s_1 - D_1(\Delta, L+\Delta). \tag{A5}$$

Recall that $D_i(\Delta, L+\Delta)$ has the same distribution of $D_i(L)$. Using (A4) and (A5), the number of Type-12 backorders started in $[\Delta, L+\Delta)$ have expected value

$$\begin{aligned} &E[\max\{[-IN_1(L+\Delta)]^+, [-IN_2(L+\Delta)]^+\}] \\ &\quad = E[\max\{[D_1(L) - s_1]^+, [D_2(L) - (s_2 - D_2(\Delta))]^+\}] \\ &\quad = \sum_{y=0}^{s_2-1} P(D_2(\Delta) = y) \overline{B}^{12}(\mathbf{s} - y\mathbf{e}_2 \mid L) \\ &\qquad + P(D_2(\Delta) \geq s_2) E[D^{12}(L)]. \end{aligned} \tag{A6}$$

The second term in the last equation follows because if $D_2(\Delta) \geq s_2$, then from (A2) $IN_2(\Delta) \leq 0$, and therefore all the Type-12 demands during $[\Delta, L+\Delta)$ will be backlogged, with expected value $E[D^{12}(L)]$.

Combining (A3) and (A6) yields (9). □

References

Agrawal, N., M. Cohen. 2001. Optimal material control and performance evaluation in an assembly environment with component commonality. *Naval Res. Logist.* Forthcoming.

Alt, F. 1962. Safety levels in military inventory management. *Oper. Res.* **10** 786–794.

Baker, K. 1985. Safety stocks and component commonality. *J. Oper. Management* **6** 13–22.

——, M. Magazine, H. Nuttle. 1986. The effect of commonality of safety stock in a simple inventory model. *Management Sci.* **32** 982–988.

Cheung, K., W. Hausman. 1995. Multiple failures in a multi-item spare inventory model. *IIE Trans.* **27** 171–180.

Glasserman, P., Y. Wang. 1998. Leadtime-inventory tradeoffs in assemble-to-order systems. *Oper. Res.* **46** 858–871.

Eynan, A., M. Rosenblatt. 1996. Component commonality effects on inventory costs. *IIE Trans.* **28** 93–104.

Hausman, W. 1969. Minimizing customer line items backordered in inventory control. *Management Sci.* **15** B628–B634.

——, H. L. Lee, A. X. Zhang. 1998. Joint demand fulfillment probability in a multi-item inventory system with independent order-up-to policies. *Eur. J. Oper. Res.* **109** 646–659.

Hillier, M. 1999. Component commonality in a multi-period inventory model with service level constraints. *Internat. J. Production Res.* **37** 2665–2683.

Kaplan, A. 1971. Computing expected values of customer requests backordered. *Management Sci.* **17** 647–651.

Lu, Y., J.-S. Song, D. Yao. 2001. Optimal inventory-backorder tradeoffs in an assemble-to-order system. Working paper, Graduate School of Management, University of California, Irvine, CA.

Miller, B. 1971. A multi-item inventory model with joint backorder criterion. *Oper. Res.* **19** 1467–1476.

Muckstadt, J. 1973. A model for a multi-item, multi-echelon, multi-indenture inventory system. *Management Sci.* **20** 472–481.

——. 1979. A three-echelon, multi-item model for recoverable items. *Naval Res. Logist. Quart.* **26** 199–222.

Song, J.-S. 1998. On the order-fill-rate in a multi-item, base-stock inventory system. *Oper. Res.* **46** 831–845.

——. 2000. A note on assemble-to-order systems with batch ordering. *Management Sci.* **46** 739–743.

——, D. Yao. 2000. Performance analysis and optimization in assemble-to-order systems with random leadtimes. *Oper. Res.* Forthcoming.

——, S. Xu, B. Liu. 1999. Order-fulfillment performance measures in an assemble-to-order system with stochastic leadtimes. *Oper. Res.* **47** 131–149.

Zhang, A. 1995a. Decentralized order-up-to policies in an assemble-to-order system with a single product. Working paper, School of Business Administration, University of Southern California, Los Angeles, CA.

——. 1995b. Performance evaluation of an assemble-to-order system with dependent demands and decentralized order-up-to policies. Working paper, School of Business Administration, University of Southern California, Los Angeles, CA.

Zipkin, P. 2000. *Foundations of Inventory Management.* McGraw-Hill, New York.

Accepted by Hau L. Lee; received November 29, 2001. This paper was with the author 6 days for 1 revision.