

Evaluating Echelon Stock (R, nQ) Policies in Serial Production / Inventory Systems with Stochastic Demand

Fangruo Chen • Yu-Sheng Zheng

Graduate School of Business, Columbia University, New York, New York 10027

Operations and Information Management Department, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104

This paper studies echelon stock (R, nQ) policies in serial production / inventory systems with stochastic demand. We provide a recursive procedure to compute the steady state echelon inventory levels of the systems, which can be used to evaluate the long-run average holding and backorder costs as well as other performance measures. The procedure is based upon an observation of a relationship between the inventory status of adjacent stages in a serial system. We also derive exact formulas for replenishment frequencies and setup costs. Our results apply to both continuous-review systems with compound Poisson demand and periodic-review systems with independent, identically distributed demands. A preliminary numerical study was conducted to explore the cost effectiveness of echelon stock (R, nQ) policies. For two-stage systems with simple Poisson demand, we compared among the minimum costs of echelon stock (R, nQ) policies, a lower bound on the minimum achievable costs, and the minimum costs of installation stock (R, nQ) policies. Finally, we present a modification of an existing approximate evaluation procedure.

(Multi-echelon Inventory; Facility-in-Series; Batch Ordering; Reorder-Point Order-Quantity System)

1. Introduction

Consider a multi-stage, serial, production / inventory system where each stage represents an inventory stocking point. An inventory transfer between adjacent stages, which represents a production operation or a stock transition, takes a positive leadtime and incurs a setup cost. Customer demand arises only at the lowest stage and is stochastic. Demand that cannot be satisfied immediately is fully backlogged. The system data are stationary and the planning horizon is infinite.

Optimal control of the above system (under a standard cost structure of linear holding and backorder costs in addition to the setup costs) is difficult (Clark and Scarf 1962). Instead of attempting to characterize an optimal policy for the above system, we study in this paper a class of practical control policies which will be referred to as echelon stock (R, nQ) policies.

A stage's echelon stock (Clark and Scarf 1960) is equal to its outstanding orders + its on-hand stock + inventories at or in transit to all its downstream stages – customer backorders. Under an echelon stock (R, nQ) policy, each stage replenishes its stock according to a stage-specific echelon stock (R, nQ) policy: when its echelon stock y falls to or below its reorder point R , an order of nQ units is placed where Q is the stage's base order quantity and n is a minimum integer with $y + nQ > R$. Note that (R, nQ) policies are a model of the widely used reorder-point, order-quantity policies in practice. For continuous-review systems where each customer demands a single unit, (R, nQ) policies reduce to the better known special case of (R, Q) policies, under which the base order quantity becomes *the order quantity*. The main objective of this paper is to evaluate the performance of echelon stock (R, nQ) policies in serial systems.

We derive exact steady state characteristics for periodic-review systems with independent, identically distributed demands and continuous-review systems with compound Poisson demand. These characteristics can be used to compute system performance measures such as average on-hand inventories at all stages, average inventories in transit, and the average customer backorders. These results are based on an observation of a simple relationship between the inventory status at adjacent stages in a system.

We also derive replenishment, or physical inventory transfer, frequencies. In the inventory theory literature, a replenishment is rarely distinguished from an order placement; the setup cost, which is incurred for various activities associated with the order-replenishment, is typically charged to the order. These activities may include processing transactions, setting up machines, or sending trucks. Such modeling is adequate for single-stage systems where the supplier has ample stock (thus every order triggers a replenishment). It is no longer so for multi-stage systems where a supplier stage may be out of stock and each order does not necessarily trigger a replenishment. In this paper, we distinguish between ordering and replenishment frequencies.

The above performance measures can be used to select echelon stock (R, nQ) policies that minimize long-run average costs. Chen and Zheng (1992b) have developed an optimization algorithm for two-stage systems with simple Poisson demand. This algorithm is used in this paper to conduct a preliminary numerical study to explore the cost effectiveness of echelon stock (R, nQ) policies. For two-stage systems, the study compares the minimum costs of echelon stock (R, nQ) policies with a lower bound on the minimum achievable costs (by any policy). The difference is small.

Closely related to this paper are two recent publications. One is De Bodt and Graves (1985) who study echelon stock (R, Q) policies in serial systems. To our knowledge, they are the first to study simple and easily implementable echelon stock policies in multi-stage inventory systems with setup costs at all stages. They provide approximate performance measures under a 'nestedness' assumption: whenever a stage receives a shipment, a batch must be immediately sent down to its downstream stage. For systems with simple Poisson demand, our results serve as a benchmark for their ap-

proximation. In this paper, we present a modification of their approximation procedure with a numerical test.

Another is Badinelli (1992) who studies *installation stock* (R, Q) policies in serial systems with simple Poisson demand. (An installation stock (R, nQ) policy works in exactly the same way as an echelon stock (R, nQ) policy except that for the former a stage's reorder point is referred to its installation stock: inventory on hand and on order minus backlogged orders from its downstream stage. An installation stock (R, nQ) policy also reduces to an installation stock (R, Q) policy when demand is simple Poisson.) Under the assumption that the installation stock at each stage is nonnegative, he provides exact long-run average holding and backorder costs. As shown by Axsater and Rosling (1993) for serial systems, installation stock policies are merely a special case of echelon stock policies. Therefore, the systems considered in this paper are more general than Badinelli's. Yet, our results are simpler.

Installation stock (R, Q) policies have been more extensively studied in the context of one-warehouse multi-retailer systems (Deuermeier and Schwarz 1981, Moinzadeh and Lee 1986, Lee and Moinzadeh 1987a, b, Svoronos and Zipkin 1988, and Axsater 1991, 1993). One major reason why installation stock policies have received much attention is perhaps their modest informational requirement: installation stock is local stock information, whereas echelon stock requires a certain degree of information centralization. But this advantage is diminishing. Due to advanced information technologies, e.g., Electronic Data Interchange (EDI), centralized stock information is now readily available in many companies. In other words, the technologies for implementing echelon stock policies are in place. But how much improvement can we expect from using centralized stock information? In this paper, we attempt to answer this question by conducting a numerical comparison between echelon and installation stock (R, nQ) policies.

An important special case of the serial systems considered here is one where there are *no* setup costs at all but the highest stage. For such a system optimal policies are known. The same is true for assembly systems without setup costs. These results are established by Clark and Scarf (1960), Federgruen and Zipkin (1984), Schmidt and Nahmias (1985), and Rosling (1989) for

periodic-review systems, and recently extended by Chen and Zheng (1992a) to continuous-review systems. The echelon stock (R, nQ) policies studied in this paper are a generalization of these optimal policies to accommodate setup costs.

The rest of the paper is organized into three sections. Section 2 establishes an exact evaluation procedure for periodic-review systems. Section 3 extends the procedure to continuous-review systems, reports a numerical study, and provides a modification of De Bodt and Graves's approximation. Section 4 concludes the paper.

2. Periodic-review Systems

In this section, we derive performance measures for echelon stock (R, nQ) policies in periodic-review, serial systems with independent, identically distributed demands. We first introduce notation and assumptions. Then we make a key observation which is used to derive long-run average holding and backorder costs. Finally, we derive replenishment frequencies.

2.1. Preliminaries

Consider a periodic-review, serial system with N stages where $N (\geq 2)$ is a positive integer. The stages of the system are denoted by $1, \dots, N$ with stage 1 receiving stock from stage 2, 2 from 3, etc., and stage N from an outside supplier (stage $N + 1$) with infinite stock. Customer demand arises only at stage 1. Demand takes integer values and is assumed to be independent and identically distributed across different periods. Unsatisfied demand is fully backlogged. An inventory transfer from stage $i + 1$ to stage i takes L_i periods, where L_i is a stage-specific constant.

For $t < t'$, let $[t, t']$, $[t, t')$, and (t, t') be respectively the interval of periods t, \dots, t' , the interval of periods $t, \dots, t' - 1$, and the interval of periods $t + 1, \dots, t' - 1$. Denote the total demands in $[t, t']$, $[t, t')$, and (t, t') by $D[t, t']$, $D[t, t')$, and $D(t, t')$ respectively. Therefore, $D[t, t]$ is the demand at period t , while $D[t, t) = D(t, t) = 0$. For notational simplicity, we sometimes use $D[T]$, $D(T)$, and $D(T)$ to denote the total demand over $T + 1$ periods, T periods, and $T - 1$ periods respectively. Denote the probability density functions of $D[L_i]$, $D[L_i)$, and $D(L_i)$ by $f^i[\cdot]$, $f^i[\cdot)$, and $f^i(\cdot)$ respectively.

An echelon stock (R, nQ) policy is specified by N pairs of integers, (R_i, Q_i) for $i = 1, \dots, N$, where R_i is

the reorder point and Q_i the base order quantity at stage i . We restrict our attention to *integer-ratio policies* where Q_{i+1} is a positive integer multiple of Q_i . Integer-ratio policies have proved to be cost-effective in similar inventory systems with deterministic demand (Roundy 1985, 1986 and Atkins and Sun 1991). This is mainly because the cost of the EOQ model is insensitive to the choice of the order quantity. Since similar insensitivity results hold for single-location (R, nQ) systems (Zheng 1992 and Zheng and Chen 1992), we suspect that integer-ratio policies are also cost-effective for our serial systems with stochastic demand. Furthermore, the integer-ratio constraint on the choice of base order quantities facilitates quantity coordination between adjacent stages.

Due to the integer-ratio restriction, each shipment leaving stage $i + 1$ (to stage i) is an integer multiple of Q_i . Therefore, if the initial on-hand inventory at stage $i + 1$ is not an integer multiple of Q_i , then there are units of inventory which will remain at stage $i + 1$ forever. This is clearly not cost effective. To remove the residual units in the system, first send the residual units at stage N to stage $N - 1$; and once they arrive at stage $N - 1$, ship the residual units at stage $N - 1$ to stage $N - 2$; and so on. For the rest of the paper, we assume that the initial on-hand inventory at stage $i + 1$ is an integer multiple of Q_i for $i = 1, \dots, N - 1$.

For clarity, we assume that replenishment activities—ordering, shipping, and receiving—in a period occur at the beginning of the period. Demand occurs during the period. Inventory holding and backorder costs are assessed on the period-ending inventory and backorder levels. For period t , define

| | |
|-------------|--|
| $I_i(t)$ | echelon i inventory = inventory at stage i plus inventories in transit to or at stage $j, j = 1, \dots, i - 1$, after demand, |
| $B(t)$ | backorder level at stage 1 after demand, |
| $IL_i(t)$ | echelon i inventory level = $I_i(t) - B(t)$, |
| $IL_i^-(t)$ | echelon i inventory level after ordering/receiving and before demand = inventory at stage i plus inventories in transit to or at stage $j, j = 1, \dots, i - 1$, minus backorder level at stage 1 after ordering/receiving but before demand, |
| $NIP_i(t)$ | echelon stock or nominal echelon inventory |

position at stage $i = IL_i^-(t) +$ outstanding orders of stage i after ordering and before demand, and

$IP_i(t)$ echelon inventory position at stage $i = IL_i^-(t) +$ inventories in transit to stage i after ordering/receiving and before demand.

When the time index t is suppressed, the notation represents the corresponding steady state variables.

Above we have introduced two different notions of inventory position, the nominal echelon inventory position and the inventory position. The nominal echelon inventory position (echelon stock) has been used to define echelon stock (R, nQ) policies: stage i orders (from stage $i + 1$) when its nominal echelon inventory position is at or below R_i . Although convenient, this is not the only way to define the policies. An alternative definition would be: stage $i + 1$ ships to stage i based on its observation of the echelon inventory position at stage i , i.e., if the echelon inventory position at stage i is at or below R_i and stage $i + 1$ has positive on-hand inventory, then a shipment in units of Q_i is made to stage i to raise its echelon inventory position to above R_i if possible, or as close as possible to R_i otherwise. It is easy to verify that these definitions lead to exactly the same material flow in the system. They also have the same information requirement: each stage of the system needs the customer demand information at stage 1 in order to keep track of its own nominal echelon inventory position or its downstream stage's echelon inventory position. The only difference is that the former definition requires order placement between internal stages, while the latter does not.

We proceed to derive a relationship between $IP_i(t)$ and $IL_{i+1}^-(t)$ for $i = 1, \dots, N - 1$. Note that stock is shipped into, or out of, stage $i + 1$ in units of Q_i . Since the initial on-hand inventory at stage $i + 1$ is also in units of Q_i , we see that the on-hand inventory at stage $i + 1$ is always a nonnegative integer multiple of Q_i . From definition, we know that $IL_{i+1}^-(t) - IP_i(t)$ is the on-hand inventory at stage $i + 1$. Therefore,

$$IL_{i+1}^-(t) - IP_i(t) = mQ_i, \quad m \geq 0 \text{ integer.} \quad (1)$$

Under the echelon stock policy, we know that $IP_i(t)$ is above R_i whenever stage $i + 1$ has positive on-hand inventory. Now suppose $IL_{i+1}^-(t) \leq R_i$. From (1) we have $IP_i(t) \leq R_i$ since $m \geq 0$. Thus the on-hand stock

at stage $i + 1$ cannot be positive, i.e. $m = 0$ or $IP_i(t) = IL_{i+1}^-(t)$. Now suppose that $IL_{i+1}^-(t) > R_i$. In this case, the echelon (R, nQ) policy dictates that $IP_i(t)$ must belong to the set $\{R_i + 1, \dots, R_i + Q_i\}$. Note that there is a unique point in the set that satisfies (1). In summary, $IL_{i+1}^-(t)$ uniquely determines $IP_i(t)$.

LEMMA 1 (KEY OBSERVATION). *Under the echelon stock (R, nQ) policy, we have*

$IP_i(t) = O_i[IL_{i+1}^-(t)]$ for $i = 1, \dots, N - 1$ where

$$O_i[x] = \begin{cases} x, & \text{if } x \leq R_i, \\ x - nQ_i, & \text{otherwise} \end{cases}$$

with n being the largest integer so that $x - nQ_i > R_i$. When t is suppressed, we have the relationships at steady state.

We next provide equations that relate IP_i to IL_i and IL_i^- . Since all and only those shipments released at stage $i + 1$ by period t would have arrived at stage i by period $t + L_i$, we have:

$$IL_i(t + L_i) = IP_i(t) - D[t, t + L_i].$$

Since $IL_i^-(t + L_i)$ is assessed before subtracting the demand at period $t + L_i$, we have

$$IL_i^-(t + L_i) = IP_i(t) - D[t, t + L_i].$$

At steady state

$$IL_i = IP_i - D[L_i] \quad \text{and} \quad IL_i^- = IP_i - D[L_i] \quad (2)$$

where IP_i is independent of $D[L_i]$ and $D[L_i]$.

2.2. Inventory Positions and Levels

By combining (2) and Lemma 1 in the previous section, we have a recursive procedure for computing the steady state distributions of the inventory positions and levels at all stages. First consider stage N . Since the outside supplier has infinite stock, echelon N inventory position behaves as if the whole serial system were a single-location system under the control of the (R_N, nQ_N) policy. Thus, IP_N , the steady state echelon N inventory position, is uniformly distributed over $\{R_N + 1, \dots, R_N + Q_N\}$, cf. Hadley and Whitin (1961). Therefore the distribution of IL_N^- can be obtained from (2).

Now proceed to stage $N - 1$. Notice that the distribution of IP_{N-1} can be obtained from that of IL_N^- by using Lemma 1; and the distribution of IL_{N-1}^- from that of IP_{N-1} by using (2). This process is continued until

the steady state distributions of the echelon inventory positions and levels at all stages have been obtained.

PROPOSITION 1. Use $p^i(\cdot)$, $w^i(\cdot)$, and $q^i(\cdot)$ to denote the density functions of IP_i , IL_i^- , and IL_i respectively. These functions can be determined by the following recursive formulas. For $i = 2, \dots, N$:

$$w^i(y) = \begin{cases} \sum_{x=0}^{R_i+Q_i-y} f^i[x]p^i(x+y) & \text{if } y \leq R_i + Q_i, \\ 0 & \text{otherwise,} \end{cases}$$

$$p^{i-1}(y) = \begin{cases} w^i(y) & \text{if } y \leq R_{i-1}, \\ \sum_{n:n \geq 0, y+nQ_{i-1} \leq R_i+Q_i} w^i(y+nQ_{i-1}) & \text{if } R_{i-1} < y \leq R_{i-1} + Q_{i-1}, \\ 0 & \text{otherwise,} \end{cases}$$

where

$$p^N(y) = \begin{cases} 1/Q_N & \text{if } y = R_N + 1, \dots, R_N + Q_N, \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore,

$$q^i(y) = \sum_{x=0}^{R_i+Q_i-y} f^i[x]p^i(x+y) \quad \text{for } i = 1, \dots, N.$$

In the above recursive formulas for $p^i(y)$ and $w^i(y)$, we have only provided upper bounds on y . Its lower bounds depend on the range of the leadtime demand $D[L_i]$. Let u_i be the maximum value of $D[L_i]$. (If the leadtime demand can be arbitrarily large, then a truncation of the leadtime demand distribution is necessary.) Let A_i be a lower bound on y for $p^i(y)$, i.e. $p^i(y) = 0$ for all $y < A_i$. We have $A_N = R_N + 1$. Define

$$a_i = A_i - u_i. \tag{3}$$

From (2), we see that a_i is a lower bound on y for $w^i(y)$. From Lemma 1, we have

$$A_{i-1} = \min\{a_i, R_{i-1} + 1\}. \tag{4}$$

Therefore, the lower bounds on y for $p^i(y)$ and $w^i(y)$ can be obtained recursively by using (3) and (4).

It is worth pausing at this point to note the differences between a common approach in the literature and ours above. In multi-echelon systems, an outstanding order of a stage can be either in transit or backordered at an

upstream stage. In the inventory literature, an 'inventory position' usually means a *nominal (echelon) inventory position*, which includes all outstanding orders, in transit or backordered at an upper stage. As far as a stage's nominal inventory position is concerned, the stage's supplier can be treated as if it had infinite stock. Thus the distribution of a stage's nominal inventory position is the same as that of a single-stage system, which is easy to obtain. A stage's inventory level is equal to its nominal inventory position minus its leadtime demand. The leadtime here consists of both the transportation time and a random delay ('retard'). To obtain a stage's inventory level, one needs to characterize the retard first. This idea has been widely used in the multi-echelon inventory literature. It has primarily been applied to establish various approximations for installation stock policies. A key feature of our approach is the introduction of the *inventory position* which only includes *in-transit* orders. Under this definition, the relationship between a stage's inventory position and its inventory level is simpler because the leadtime is now the transportation time (see (2)). But the distribution of the inventory position is no longer the same as that of a single-stage system. For our serial systems, the distribution is derived from the operation details of the systems (Lemma 1). Note that this approach leads to an exact evaluation.

With the above density functions, we can easily compute various long-run average performance measures, e.g., $E(IP_i)$, $E(IL_i^-)$, $E(IL_i)$ at each stage and $E(B)$ ($=E(IL_1^-)$). These results also provide the average on-hand inventory at stage $i + 1$, $E(IL_{i+1}^-) - E(IP_i)$, and the average inventory in transit from stage $i + 1$ to stage i , $E(IP_i) - E(IL_i^-)$.

Define fill rate (FR) to be the percentage of demand filled upon arrival. Let D be the demand in a period at steady state. Note that the on-hand inventory before the demand is $(IL_1^-)^+$. Thus the demand filled in that period is $\min\{D, (IL_1^-)^+\}$. Since D and $(IL_1^-)^+$ are independent, we have for $x \geq 0$,

$$\begin{aligned} & \Pr(\min\{D, (IL_1^-)^+\} \geq x) \\ &= \Pr(D \geq x) \Pr((IL_1^-)^+ \geq x) \\ &= [1 - F(x)] \sum_{y=x}^{R_1+Q_1} w^1(y) \end{aligned}$$

where $F(x) = \Pr(D < x)$. Consequently,

$$FR = \frac{\sum_{x=0}^{R_1+Q_1} [1 - F(x)] \sum_{y=x}^{R_1+Q_1} w^1(y)}{\mu}$$

where $\mu = E(D)$.

2.3. Replenishment Frequencies

Since an order does not necessarily trigger a replenishment in our serial systems, we distinguish between order frequencies and replenishment frequencies. The former is easy to obtain: the order frequency at stage i is the same as that of a single-stage (R_i, nQ_i) system (see Zheng and Chen 1992 for an order-frequency formula). Here we focus on the latter. This is appropriate since replenishment activities, e.g. sending trucks or setting up machines, are often the major concerns in most applications.

Let t be a period when the system is at steady state. Denote by D_t the demand at period t . Let $f(\cdot)$ be the probability density function of D_t . For $i = 1, \dots, N$, define

$$g_i(u, v) = \Pr(IP_i(t) - D_t = u, IP_i(t+1) = v).$$

Let $A = \{(u, v) | R_N + 1 \leq v \leq R_N + Q_N \text{ and } u = v - kQ_N, k \geq 0 \text{ integer}\}$. Note that $g_N(u, v) = 0$ for $(u, v) \notin A$. For any $(u, v) \in A$, we have

$$\begin{aligned} g_N(u, v) &= \Pr(IP_N(t) - D_t = u, IP_N(t+1) = v) \\ &= \Pr(IP_N(t) - D_t = u) \\ &= \frac{1}{Q_N} \sum_{m=R_N+1}^{R_N+Q_N} \Pr(D_t = m - u | IP_N(t) = m) \\ &= \frac{1}{Q_N} \sum_{m=R_N+1}^{R_N+Q_N} f(m - u) \end{aligned}$$

where the second equality follows since $(u, v) \in A$ and the outside supplier has ample stock, and the last equality follows since $IP_N(t)$ is independent of D_t .

We next derive a formula for $g_{i-1}(\cdot, \cdot)$ given $g_i(\cdot, \cdot)$. In the following, we assume $L_i > 0$. (The case with zero leadtimes can be handled similarly.) Recall that

$$\begin{aligned} IL_i^-(t + L_i) &= IP_i(t) - D[t, t + L_i] \\ &= IL_i^-(t + L_i) = IP_i(t) - D_t - D[t + 1, t + L_i]. \end{aligned}$$

Similarly, we have

$$\begin{aligned} IL_i^-(t + L_i + 1) \\ &= IP_i(t + 1) - D[t + 1, t + L_i] - D_{t+L_i}. \end{aligned}$$

From Lemma 1, we have

$$\begin{aligned} IP_{i-1}(t + L_i) &= O_{i-1}[IL_i^-(t + L_i)] \quad \text{and} \\ IP_{i-1}(t + L_i + 1) &= O_{i-1}[IL_i^-(t + L_i + 1)]. \end{aligned}$$

Therefore,

$$\begin{aligned} IP_{i-1}(t + L_i) \\ &= O_{i-1}[IP_i(t) - D_t - D[t + 1, t + L_i]] \quad \text{and} \\ IP_{i-1}(t + L_i + 1) \\ &= O_{i-1}[IP_i(t + 1) - D[t + 1, t + L_i] - D_{t+L_i}]. \end{aligned}$$

Consequently,

$$\begin{aligned} g_{i-1}(u, v) \\ &= \Pr(IP_{i-1}(t + L_i) - D_{t+L_i} = u, \\ &\quad IP_{i-1}(t + L_i + 1) = v) \\ &= \Pr(O_{i-1}[IP_i(t) - D_t - D[t + 1, t + L_i]] \\ &\quad - D_{t+L_i} = u, \quad O_{i-1}[IP_i(t + 1) \\ &\quad - D[t + 1, t + L_i] - D_{t+L_i}] = v) \\ &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} f^i(m) f(n) \Pr(O_{i-1}[IP_i(t) - D_t - m] \\ &\quad - n = u, \quad O_{i-1}[IP_i(t + 1) - m - n] \\ &\quad = v | D[t + 1, t + L_i] = m, \quad D_{t+L_i} = n) \\ &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} f^i(m) f(n) \Pr(O_{i-1}[IP_i(t) - D_t - m] \\ &\quad - n = u, \quad O_{i-1}[IP_i(t + 1) - m - n] = v) \\ &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} f^i(m) f(n) \sum_{\substack{x: O_{i-1}[x-m]-n=u, \\ y: O_{i-1}[y-m-n]=v}} g_i(x, y) \end{aligned}$$

where the fourth equality follows since $IP_i(t)$, D_t , and $IP_i(t + 1)$ are independent of $D[t + 1, t + L_i]$ and D_{t+L_i} .

PROPOSITION 2. The $g_i(\cdot, \cdot)$'s can be computed recursively:

$$g_{i-1}(u, v) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} f^i(m) f(n) \sum_{\substack{x: O_{i-1}[x-m]-n=u, \\ y: O_{i-1}[y-m-n]=v}} g_i(x, y) \quad (5)$$

for $i = 2, \dots, N$ where

$$g_N(u, v) = \begin{cases} \frac{1}{Q_N} \sum_{m=R_N+1}^{R_N+Q_N} f(m-u) & \text{if } (u, v) \in A, \\ 0 & \text{otherwise.} \end{cases}$$

The long-run average replenishment frequency at stage i is

$$\sum_{u < v} g_i(u, v).$$

The summation ranges for x and y in (5) can be clarified as follows. From the definition of $O_{i-1}[\cdot]$, we have

$$\{x : O_{i-1}[x - m] - n = u\} = \begin{cases} \{m + n + u\} & \text{if } n + u \leq R_{i-1}, \\ \{m + n + u + lQ_{i-1}, l = 0, 1, \dots\} & \text{if } R_{i-1} < n + u \leq R_{i-1} + Q_{i-1}, \\ \emptyset & \text{otherwise.} \end{cases}$$

Similarly,

$$\{y : O_{i-1}[y - m - n] = v\} = \begin{cases} \{m + n + v\} & \text{if } v \leq R_{i-1}, \\ \{m + n + v + lQ_{i-1}, l = 0, 1, \dots\} & \text{if } R_{i-1} < v \leq R_{i-1} + Q_{i-1}, \\ \emptyset & \text{otherwise.} \end{cases}$$

These ranges can be further restricted to $x \leq R_i + Q_i$ and $y \leq R_i + Q_i$ since otherwise $g_i(x, y) = 0$. Upper bounds on m and n in (5) depend on the upper bound on a period's demand. If the demand is unbounded, a truncation is necessary.

2.4. Long-Run Average Costs

Let h_i be the unit echelon holding cost per period at stage i and p be the unit backorder cost per period (at stage 1). The steady state distributions in Proposition

1 can be used to compute the long-run average system-wide holding and backorder costs

$$\begin{aligned} & \sum_{i=1}^N h_i E(I_i) + pE(B) \\ &= \sum_{i=1}^N h_i E(IL_i) + (H_1 + p)E(B) \end{aligned}$$

where $H_1 = \sum_{i=1}^N h_i$ and the equality follows since $IL_i = I_i - B$.

Let K_i be the setup cost for each replenishment (not order) at stage i , independent of the replenishment size. From Proposition 2, we have the long-run average system-wide setup cost

$$\sum_{i=1}^N K_i \sum_{u < v} g_i(u, v).$$

Therefore, the long-run average total cost of the system is

$$\sum_{i=1}^N K_i \sum_{u < v} g_i(u, v) + \sum_{i=1}^N h_i E(IL_i) + (H_1 + p)E(B).$$

(To include fixed backorder costs, let π be the unit backorder cost, independent of backorder duration, and μ be the expected one-period demand. The long-run average fixed backorder cost is thus $\mu\pi(1 - FR)$. For a discussion on fixed backorder costs, see Chen and Zheng 1993a.)

3. Continuous-Review Systems

In this section, we assume that the demand is compound Poisson and the echelon stock (R, nQ) policy is implemented with continuous review. We first extend the results in the previous section to the continuous-review system, and then present a numerical study and an approximation.

3.1. Performance Measures

Most of the notation introduced previously can be used for the continuous-review system under the following convention: ordering, shipping and receiving activities occur before demand at any time epoch; the order triggered by a demand at time t is placed at time t^+ (immediately after time t). But we must modify the definition of leadtimes and aggregate demands. Here, L_i

can be any nonnegative real value, $D[t, t']$ is the total demand in the time interval $[t, t']$, and $D[T]$ is the total demand over a closed interval of length T . Similarly, $D[t, t']$, $D(t, t')$, $D[T]$ and $D(T)$ should be re-defined. Finally, the time index t in the state variables can be any time epoch.

It is easy to see that Lemma 1 and equation (2) are still valid. Therefore, Proposition 1 also holds. (Note that $w^i(\cdot)$ and $q^i(\cdot)$ are now identical.) The steady state characteristics can be used to compute the long-run average inventory positions and levels. Due to PASTA—Poisson Arrivals See Time Averages (Wolff 1982), the inventory level at stage 1 observed by an incoming customer is IL_1^- . Therefore, the expression for the fill rate there is also valid here, where μ is the average demand size of a customer.

We proceed to derive replenishment frequencies. Let K_i be the setup cost for each shipment from stage $i + 1$ to stage i , independent of the shipment size. A shipment from stage $i + 1$ to stage i can be either a *normal* shipment or a *direct* shipment. A normal shipment is triggered by a demand arrival: a demand arrival causes echelon i inventory position to drop to or below R_i and there is stock available at stage $i + 1$. On the other hand, a direct shipment is triggered by a delivery at stage $i + 1$: just prior to a delivery at stage $i + 1$, echelon i inventory position is at or below R_i and stage $i + 1$ is out of stock; and immediately after the delivery, a shipment (the direct shipment) is sent down to stage i . Now the delivery to stage $i + 1$ that triggers the direct shipment to stage i can be itself a direct or a normal shipment to stage $i + 1$. If the delivery is a direct shipment to stage $i + 1$, then this direct shipment (to stage $i + 1$) is again triggered by either a direct or a normal shipment to stage $i + 2$, and so on. Since the outside supplier has infinite stock, every shipment to stage N (the highest stage) is a normal shipment. Therefore, each direct shipment to stage i can be *uniquely* associated with a normal shipment to an upstream stage, say stage j : the direct shipment is a direct (non-stop) transfer of a subbatch of the normal shipment from stage j to stage i . We know that every normal shipment is naturally associated with (or triggered by) a customer. By tracing the normal-shipment origin of a direct shipment, the direct shipment can also be associated with a customer: a customer triggers a direct shipment via a normal shipment.

Take any $i = 1, \dots, N$. Suppose that a customer arrives at time t and observes $IL_{i+1}^-(t)$ and thus $IP_i(t) = O_i[IL_{i+1}^-(t)]$. Let D be the demand size of the customer. From the above discussion, the customer triggers a normal shipment to stage i if and only if

$$IL_{i+1}^-(t) - IP_i(t) > 0 \quad \text{and} \quad IP_i(t) - D \leq R_i.$$

From Lemma 1, the above inequalities are equivalent to

$$IL_{i+1}^-(t) > R_i + Q_i \quad \text{and} \quad O_i[IL_{i+1}^-(t)] - D \leq R_i.$$

If the customer indeed triggers a normal shipment to stage i , it will be released at time t^+ and arrive at stage i at time $t^+ + L_i$. Just before the arrival, the echelon inventory position at stage $i - 1$ is $IP_{i-1}(t + L_i) = O_{i-1}[IL_i^-(t + L_i)]$ (Lemma 1). This arrival triggers a direct shipment to stage $i - 1$ if $IP_{i-1}(t + L_i) \leq R_{i-1}$, or $IL_i^-(t + L_i) \leq R_{i-1}$, or

$$IP_i(t) - D - D(t, t + L_i) \leq R_{i-1},$$

since

$$IL_i^-(t + L_i) = IP_i(t) - D - D(t, t + L_i).$$

Therefore, the customer triggers a direct shipment to stage $i - 1$ via a normal shipment to stage i if and only if

$$IL_{i+1}^-(t) > R_i + Q_i,$$

$$IP_i(t) - D \leq R_i, \quad IP_i(t) - D - D(t, t + L_i) \leq R_{i-1}.$$

In general, the customer triggers a direct shipment to stage $m (< i)$ via a normal shipment to stage i if and only if

$$IL_{i+1}^-(t) > R_i + Q_i,$$

$$IP_i(t) - D \leq R_i,$$

$$IP_i(t) - D - \sum_{j=k}^i D_j \leq R_{k-1} \quad \text{for } k = i, i - 1, \dots, m + 1,$$

where

$$D_i \stackrel{\text{def}}{=} D(t, t + L_i) \quad \text{and}$$

$$D_j \stackrel{\text{def}}{=} D[t + L_i + \dots + L_{j+1}, t + L_i + \dots + L_j],$$

$$j = 1, \dots, i - 1.$$

Now suppose t is a time epoch when the system is at steady state. Suppress the time index t to represent steady state variables and define

$$\alpha_{ii} = \Pr(IL_{i+1}^- > R_i + Q_i, O_i[IL_{i+1}^-] - D \leq R_i)$$

and

$$\alpha_{im} = \Pr\left(IL_{i+1}^- > R_i + Q_i, O_i[IL_{i+1}^-] - D \leq R_i, \right. \\ \left. O_i[IL_{i+1}^-] - D - \sum_{j=k}^i D_j \leq R_{k-1}, k = i, \dots, m + 1 \right), \\ m = 1, \dots, i - 1.$$

When IL_{N+1}^- appears in the above definitions, substitute ∞ for IL_{N+1}^- and IP_N for $O_N[IL_{N+1}^-]$. Note that IL_{i+1}^- , D , and D_j , $j = 1, \dots, i$, are independent of one another.

The customer triggers a normal shipment to stage i with probability α_{ii} , and a direct shipment to stage m via a normal shipment to stage i with probability α_{im} , $m = 1, \dots, i - 1$. Therefore, the total expected setup cost "incurred" by this customer is

$$\sum_{i=1}^N \sum_{m=1}^i \alpha_{im} K_m.$$

PROPOSITION 3. Let λ be the customer arrival rate. The long-run average system-wide setup cost is

$$\lambda \sum_{i=1}^N \sum_{m=1}^i \alpha_{im} K_m.$$

The replenishment frequency at stage m is $\lambda \sum_{i=m}^N \alpha_{im}$, $m = 1, \dots, N$.

When demand is simple Poisson, the replenishment frequencies and the setup costs can be computed easily through a recursive procedure (see Appendix).

3.2. Numerical Examples

We present a set of numerical examples to illustrate the cost effectiveness of echelon stock (R, nQ) policies. We only consider two-stage systems with simple Poisson demand, for which an exact optimization algorithm has been developed (Chen and Zheng 1992b). For each example, we identified 1) an optimal echelon stock (R, nQ) policy, 2) an optimal installation stock (R, nQ) policy, and 3) a lower bound on the minimum achievable cost (by any policy). (We used the so-called integrated lower bound recently developed in Chen and Zheng 1992a.) The results are summarized in Table 1. In the table, $C_e(C_i)$ denotes the minimum cost of echelon (installation) stock (R, Q) policies; LB is the lower bound; and

$$\epsilon_1 = \frac{C_i - C_e}{C_e} \times 100, \quad \epsilon_2 = \frac{C_e - \text{LB}}{\text{LB}} \times 100.$$

Therefore, ϵ_1 is the percentage cost increase of installation stock (R, nQ) policies over their echelon stock counterpart, and ϵ_2 is an upper bound on the gap between echelon stock (R, nQ) policies and the true (but unknown) optimal. Note that r_1 and r_2 refer to installation inventory positions. When $R_2 - R_1$ is an integer multiple of Q_1 , the echelon stock policy reduces to an installation stock policy with $r_1 = R_1$ and $r_2 = R_2 - R_1 - Q_1$.

The examples suggest that optimal echelon stock (R, nQ) policies are close to the true optimal. When $K_2 > K_1$, they are within one percent of the lower bound. The gap is larger for cases with $K_2 < K_1$. The largest gap is 4.5%. The difference in minimum costs between installation and echelon stock policies is small, indicating that the value of centralized stock information is insignificant for these examples. It is interesting to note that the ratio of Q_2 to Q_1 of the optimal echelon stock (R, nQ) policies is very close to

$$\max\left\{ \sqrt{\frac{K_2/h_2}{K_1/h_1}}, 1 \right\},$$

which is the optimal ratio suggested for closely related deterministic systems (Roundy 1985).

3.3. An Approximation

In this section, we develop an approximation for holding and backorder costs of echelon stock (R, Q) policies, which is based on the idea of De Bodt and Graves (1985).

For easy presentation, first consider a two-stage, continuous-review system with simple Poisson demand. De Bodt and Graves's approximation is based on a 'nestedness' assumption: whenever stage 2 receives a batch, it immediately ships a sub-batch to stage 1. Let $Q_2 = nQ_1$. Then an incoming batch at stage 2 consists of n sub-batches each of size Q_1 . The sub-batch that is sent down immediately is called a joint replenishment (JR), whereas the remaining $n - 1$ sub-batches are called normal replenishments (NRs). This assumption is violated when the leadtime demand of stage 2 is smaller than $R_2 - R_1$ or larger than $R_2 - R_1 + Q_1$. Therefore, the assumption itself is an approximation. Based on this assumption, they further approximate holding and

Table 1 Comparisons Between Optimal Echelon Stock (R, nQ) Policies, Optimal Installation Stock (R, nQ) Policies, and Lower Bounds

$L_1 = 1, L_2 = 2, K_1 = 10, h_1 = 0.5, h_2 = 1, p = 5$

| No. | λ | K_2 | Echelon Stock Policy | | | | | Installation Stock Policy | | | | | ϵ_1 (%) | LB | ϵ_2 (%) |
|-----|-----------|-------|----------------------|-------|-------|-------|----------|---------------------------|-------|-------|-------|----------|------------------|----------|------------------|
| | | | R_1 | Q_1 | R_2 | Q_2 | C_e | r_1 | Q_1 | r_2 | Q_2 | C_i | | | |
| 1 | 1 | 5 | 0 | 6 | 1 | 6 | 8.3828 | 1 | 6 | -6 | 6 | 8.3965 | 0.16 | 8.0216 | 4.50 |
| 2 | 1 | 100 | 0 | 8 | -1 | 16 | 17.2446 | 0 | 8 | -8 | 16 | 17.2798 | 0.20 | 17.1317 | 0.66 |
| 3 | 1 | 200 | 0 | 8 | -2 | 24 | 22.5396 | -1 | 11 | -11 | 22 | 22.6591 | 0.53 | 22.3669 | 0.77 |
| 4 | 1 | 400 | 0 | 8 | -3 | 32 | 30.0028 | 1 | 5 | -10 | 30 | 30.2177 | 0.72 | 29.8456 | 0.53 |
| 5 | 5 | 5 | 4 | 14 | 12 | 14 | 21.4394 | 12 | 13 | -13 | 13 | 21.5399 | 0.47 | 20.6433 | 3.86 |
| 6 | 5 | 100 | 4 | 18 | 8 | 36 | 41.3823 | 6 | 19 | -19 | 38 | 41.7969 | 1.00 | 41.1448 | 0.58 |
| 7 | 5 | 200 | 4 | 17 | 5 | 51 | 53.1838 | 4 | 17 | -17 | 51 | 53.2712 | 0.16 | 52.8355 | 0.66 |
| 8 | 5 | 400 | 5 | 14 | 1 | 70 | 69.9265 | 3 | 17 | -17 | 68 | 70.0236 | 0.14 | 69.5613 | 0.52 |
| 9 | 10 | 5 | 9 | 19 | 26 | 19 | 33.2192 | 9 | 18 | 0 | 18 | 33.2664 | 0.14 | 32.1126 | 3.45 |
| 10 | 10 | 100 | 8 | 25 | 20 | 50 | 61.4478 | 19 | 46 | -46 | 46 | 62.5464 | 1.79 | 61.1264 | 0.53 |
| 11 | 10 | 200 | 8 | 24 | 16 | 72 | 78.1643 | 13 | 18 | -18 | 72 | 78.9519 | 1.01 | 77.6614 | 0.65 |
| 12 | 10 | 400 | 10 | 20 | 10 | 100 | 101.8232 | 10 | 20 | -20 | 100 | 101.8232 | 0.00 | 101.3039 | 0.51 |
| 13 | 15 | 5 | 14 | 23 | 41 | 23 | 43.4355 | 16 | 24 | 0 | 24 | 43.4869 | 0.12 | 42.0567 | 3.28 |
| 14 | 15 | 100 | 13 | 31 | 33 | 62 | 78.0018 | 13 | 21 | 0 | 63 | 79.2287 | 1.57 | 77.6118 | 0.50 |
| 15 | 15 | 200 | 13 | 29 | 29 | 87 | 98.4661 | 13 | 20 | 0 | 80 | 100.1879 | 1.75 | 97.8676 | 0.61 |
| 16 | 15 | 400 | 15 | 25 | 21 | 125 | 127.4852 | 17 | 25 | -25 | 125 | 127.9855 | 0.39 | 126.8215 | 0.52 |

backorder costs through a single-location technique used by Hadley and Whitin (1963).

Recall that the holding and backorder costs of the system can be expressed as

$$E[h_2IL_2 + h_1IL_1 + (H_1 + p)B].$$

The first term $h_2E[IL_2]$ is easy to determine, and no approximation is necessary. Below we approximate the holding and backorder costs at stage 1, i.e.,

$$E[h_1IL_1 + (H_1 + p)B],$$

by inheriting De Bodt and Grave's nestedness assumption but avoiding the second step of their approximation.

A replenishment batch at stage 2 is split into one JR and $n - 1$ NRs. First consider the JR. It is shipped directly from the outside supplier to stage 1. The transfer time is $L_1 + L_2$. The stage-2 echelon inventory position just before the JR is ordered is R_2 . Now imagine a single-location, (R_2, Q_1) system with leadtime $L_1 + L_2$ and the same demand process as in the original system. Note that the average holding and backorder costs incurred by the JR (at stage 1) are exactly the same as the average costs of the single-location system:

$$\frac{1}{Q_1} \sum_{y=R_2+1}^{R_2+Q_1} G_2(y) \quad \text{where} \quad (6)$$

$$G_2(y) \stackrel{\text{def}}{=} E[h_1(y - D[L_1 + L_2]) + (H_1 + p)(y - D[L_1 + L_2])^-].$$

Now consider the NRs. The stage 1 inventory position just before each normal replenishment is R_1 . A NR takes L_1 units of time to move from stage 2 to stage 1. Similarly, a single-location analogy leads to the average holding and backorder costs incurred by each NR:

$$\frac{1}{Q_1} \sum_{y=R_1+1}^{R_1+Q_1} G_1(y) \quad \text{where} \quad (7)$$

$$G_1(y) \stackrel{\text{def}}{=} E[h_1(y - D[L_1]) + (H_1 + p)(y - D[L_1])^-].$$

Since one out of n batches received at stage 1 is a JR and the rest are NRs, the average holding and backorder costs at stage 1 are a weighted average of (6) and (7):

$$\frac{n-1}{n} \frac{1}{Q_1} \sum_{y=R_1+1}^{R_1+Q_1} G_1(y) + \frac{1}{n} \frac{1}{Q_1} \sum_{y=R_2+1}^{R_2+Q_1} G_2(y).$$

The same logic can be used to derive approximate formulas for $E(I_1)$ and $E(B)$.

Table 2 Approximate vs. Exact Measures

| No. | R_1 | Q_1 | R_2 | Q_2 | $E^e(I_1)$ | $E^a(I_1)$ | error (%) | $E^e(B)$ | $E^a(B)$ | error (%) |
|-----|-------|-------|-------|-------|------------|------------|-----------|----------|----------|-----------|
| 1 | 0 | 6 | 1 | 6 | 1.7801 | 1.9072 | 7.1400 | 0.4155 | 0.4072 | -1.9976 |
| 2 | 0 | 8 | -1 | 16 | 2.4990 | 2.4995 | 0.0200 | 0.4997 | 0.4995 | -0.0400 |
| 3 | 0 | 8 | -2 | 24 | 2.6848 | 2.6865 | 0.0633 | 0.5201 | 0.5198 | -0.0577 |
| 4 | 0 | 8 | -3 | 32 | 2.8055 | 2.8102 | 0.1675 | 0.5611 | 0.5602 | -0.1604 |
| 5 | 4 | 14 | 12 | 14 | 4.7416 | 5.1820 | 9.2880 | 0.7019 | 0.6820 | -2.8352 |
| 6 | 4 | 18 | 8 | 36 | 6.3345 | 6.3412 | 0.1058 | 0.8416 | 0.8412 | -0.0475 |
| 7 | 4 | 17 | 5 | 51 | 6.0957 | 6.0996 | 0.0640 | 1.1002 | 1.0996 | -0.0545 |
| 8 | 5 | 14 | 1 | 70 | 5.9030 | 6.1157 | 3.6033 | 1.4532 | 1.4157 | -2.5805 |
| 9 | 9 | 19 | 26 | 19 | 6.3933 | 6.9953 | 9.4161 | 1.0213 | 0.9953 | -2.5458 |
| 10 | 8 | 25 | 20 | 50 | 8.2809 | 8.3002 | 0.2331 | 1.3014 | 1.3002 | -0.0922 |
| 11 | 8 | 24 | 16 | 72 | 8.1089 | 8.1122 | 0.0407 | 1.6128 | 1.6122 | -0.0372 |
| 12 | 10 | 20 | 10 | 100 | 8.2647 | 8.5656 | 3.6408 | 2.1201 | 2.0656 | -2.5706 |
| 13 | 14 | 23 | 41 | 23 | 8.1306 | 9.0603 | 11.4346 | 1.0979 | 1.0603 | -3.4247 |
| 14 | 13 | 31 | 33 | 62 | 10.5109 | 10.5345 | 0.2245 | 1.5358 | 1.5345 | -0.0846 |
| 15 | 13 | 29 | 29 | 87 | 10.1223 | 10.1263 | 0.0395 | 1.7934 | 1.7929 | -0.0279 |
| 16 | 15 | 25 | 21 | 125 | 10.3314 | 10.6265 | 2.8563 | 2.4769 | 2.4265 | -2.0348 |

We next use our exact measures to assess the accuracy of the above approximation, i.e. the nestedness requirement. We used the numerical examples and the optimal echelon stock (R, Q) policies in Table 1. Both the exact and the approximate expected on-hand inventories and backorders at stage 1 are reported in Table 2, where $E^e(I_1)$ and $E^e(B)$ are exact and $E^a(I_1)$ and $E^a(B)$ are approximate. From the table, we see that the above approximation is very good for most of the examples. Notice that the approximation consistently overestimates the on-hand inventories and underestimates the backorders. We expect to see the accuracy of the approximation to deteriorate as R_2 deviates from the optimal: if R_2 is increased, the JR is more likely to pause at stage 2; while if R_2 is decreased, the number of NR's is more likely to be less than $(n - 1)$.

We proceed to extend the above approximation to N -stage systems with simple Poisson demand under the nestedness assumption. Consider stage k . The replenishments at stage k are categorized into NRs and different types of JRs. Use $NR(k + 1, k)$ to denote a NR (from stage $k + 1$) to stage k and $JR(j, k)$ to denote a JR from stage j to stage k , i.e., a batch that has paused at stage j and then directly moves to stage k , $j = k + 2, \dots, N + 1$. (Stage $N + 1$ is the outside supplier.) Let $Q_i = n_i Q_{i-1}$. Let us examine how a replenishment batch

at stage N , Q_N , is shipped to (or through) stage k . For each Q_N , there are $n_N n_{N-1} \dots n_{k+2}$ replenishments to stage $k + 1$. Of these replenishments,

$$n_N n_{N-1} \dots n_{k+2} (n_{k+1} - 1)$$

sub-batches of size Q_k would pause at stage $k + 1$. Therefore, the total number of $NR(k + 1, k)$'s is

$$n_N n_{N-1} \dots n_{k+2} (n_{k+1} - 1).$$

Now consider the number of $JR(k + 2, k)$ s. Note that there are $n_N n_{N-1} \dots n_{k+3}$ replenishments (each of size Q_{k+2}) to stage $k + 2$. Of these replenishments,

$$n_N n_{N-1} \dots n_{k+3} (n_{k+2} - 1)$$

(each of size Q_{k+1}) would pause at stage $k + 2$. Each of these pausing batches would generate exactly one $JR(k + 2, k)$. Therefore, the total number of $JR(k + 2, k)$ s is

$$n_N n_{N-1} \dots n_{k+3} (n_{k+2} - 1).$$

The numbers of other types of JRs can be similarly identified. Table 3 summarizes these numbers.

We next derive an approximation for the expected system-wide holding and backorder cost, $\sum_{k=1}^N h_k E(IL_k)$

+ $(H_1 + p)E(B)$. Consider $E(IL_k)$. As before, we can associate each type of replenishment at stage k with a single-location system, as shown in Table 4.

Therefore for $k = 2, \dots, N$,

$$E(IL_k) = \sum_{j=k}^N \frac{r_{j+1,k}}{r_k} \left(R_j + \frac{Q_k + 1}{2} - \mu_{jk} \right)$$

where $r_k = \sum_{j=k}^N r_{j+1,k} = n_N n_{N-1} \dots n_{k+1}$ and μ_{jk} is the average demand over leadtime $L_k + \dots + L_j$. The holding and backorder cost at stage 1, $h_1 E(IL_1) + (H_1 + p)E(B)$, can be obtained similarly:

$$\sum_{j=1}^N \frac{r_{j+1,1}}{r_1} \frac{1}{Q_1} \sum_{y=R_j+1}^{R_j+Q_1} G_j(y)$$

where

$$G_j(y) \stackrel{\text{def}}{=} E[h_1(y - D[L_1 + \dots + L_j]) + (H_1 + p)(y - D[L_1 + \dots + L_j])^-].$$

4. Conclusion

This paper has developed a procedure for exact performance evaluation of echelon stock (R, nQ) policies in serial systems with stochastic demand. The procedure applies to both continuous-review systems with compound Poisson demand and periodic-review systems with independent, identically distributed demands. A preliminary numerical study suggests that echelon stock (R, nQ) policies are close to optimal. Some generalizations are possible. When demands form a renewal process, an exact evaluation procedure can still be established. This extension is conceptually simple, but algebraically tedious. We refer interested readers to Chen and Zheng (1991) for a treatment of renewal demand. Throughout the paper, we have assumed constant in-

Table 4 Single-Location Equivalents

| Type | Single-Location Analogy | |
|------------------|-------------------------|-----------------------------------|
| | Policy | Leadtime |
| NR($k + 1, k$) | (R_k, Q_k) | L_k |
| JR($k + 2, k$) | (R_{k+1}, Q_k) | $L_k + L_{k+1}$ |
| \vdots | \vdots | \vdots |
| JR(N, k) | (R_{N-1}, Q_k) | $L_k + L_{k+1} + \dots + L_{N-1}$ |
| JR($N + 1, k$) | (R_N, Q_k) | $L_k + L_{k+1} + \dots + L_N$ |

ventory transfer times between stages. For systems with stochastic transfer leadtimes that can be modeled by the framework of Zipkin (1986), our evaluation procedure still applies with minor modifications. (This was suggested by Paul Zipkin.) Research is in progress to extend the analysis to multi-echelon systems with more general network topologies, e.g. the one-warehouse multi-retailer systems (Chen and Zheng 1993b).

Optimization within the class of echelon stock (R, nQ) policies remains a major challenge. This is especially true for systems with many stages. An efficient algorithm requires not only an efficient evaluation procedure but also a cost function with a simple form. Here approximation may become useful. For example, the approximation in §3.3 may be used to find close-to-optimal control parameters. The cost effectiveness of echelon stock (R, nQ) policies represents another challenge. In this direction, a theoretical worst case analysis would be the most desirable result. Finally, the value of centralized stock information in general multi-stage stochastic production / inventory systems remains to be an intriguing question.*

* The authors wish to thank the referees for their helpful comments on an earlier draft, which have led to many improvements. In particular, the test on a modification of De Bodt and Graves's approximation was suggested by one referee. This research was supported in part by NSF grant DDM 9111183.

Table 3 Numbers of NRs and JRs

| Type | Number | Notation |
|------------------|---|--------------------------------------|
| NR($k + 1, k$) | $n_N n_{N-1} \dots n_{k+2} (n_{k+1} - 1)$ | $\stackrel{\text{def}}{=} r_{k+1,k}$ |
| JR($k + 2, k$) | $n_N n_{N-1} \dots (n_{k+2} - 1)$ | $\stackrel{\text{def}}{=} r_{k+2,k}$ |
| \vdots | \vdots | \vdots |
| JR(N, k) | $n_N - 1$ | $\stackrel{\text{def}}{=} r_{N,k}$ |
| JR($N + 1, k$) | 1 | $\stackrel{\text{def}}{=} r_{N+1,k}$ |

Appendix

Here we provide a recursive procedure to compute setup costs for continuous-review systems with simple Poisson demand.

Take any $i = 2, \dots, N$. Suppose that at time t , a batch of size q is sent from stage $i + 1$ to stage i . Just prior to the shipment, stage i 's echelon inventory position is r , i.e. $IP_i(t^-) = r$. Since each customer demands a single unit (simple Poisson), each stage orders its base

quantity every time it places an order. As a result, batches coming to a stage at different times remain separate in all subsequent shipments to downstream stages. (This is a key difference between simple-Poisson and compound-Poisson systems.) Therefore, we can define the total setup cost for sending a batch from a stage to stage 1. Let $\kappa_i[r, q]$ be the expected total setup cost for sending the batch q from stage $i + 1$ to stage 1. Since no shipment will be made to stage i if $r \geq R_i + 1$, let $\kappa_i[r, q] = 0$ for $r \geq R_i + 1$.

Due to the integer-ratio constraint on base order quantities, we know that q is an integer multiple of Q_{i-1} . Let $q = nQ_{i-1}$ where n is a positive integer. The batch arrives at stage i at time $T = t + L_i$ and will be shipped down to stage $i - 1$ through at most one direct shipment and possibly several normal shipments. Let $b_i(r, q)$ be the number of normal shipments. Define $a_i(r, q) = n - b_i(r, q)$. Thus, $a_i(r, q)$ is the number of sub-batches (each of size Q_{i-1}) included in the direct shipment. Since the echelon inventory position at stage $i - 1$ just prior to the arrival at time T is $IP_{i-1}(T^-)$, the expected total setup cost for sending the direct shipment to stage 1 is

$$\begin{aligned} E(\kappa_{i-1}[IP_{i-1}(T^-), a_i(r, q)Q_{i-1}]) \\ = E(\kappa_{i-1}[IL_i^-(T^-), a_i(r, q)Q_{i-1}]), \end{aligned}$$

where the equality follows because $IP_{i-1}(T^-)$ and $IL_i^-(T^-)$ differ only when $IL_i^-(T^-) \geq R_{i-1} + Q_{i-1}$, in which case both sides of the above equation are zero. On the other hand, since the echelon inventory position at stage $i - 1$ just prior to each normal shipment is R_{i-1} the expected total setup cost associated with each normal shipment is

$$\kappa_{i-1}[R_{i-1}, Q_{i-1}].$$

PROPOSITION 4. Suppose that at time t , a batch of size q is sent from stage $i + 1$ to stage i and $IP_i(t^-) = r$. Then for $i = 2, \dots, N$

$$\begin{aligned} \kappa_i[r, q] = K_i + E(b_i(r, q))\kappa_{i-1}[R_{i-1}, Q_{i-1}] \\ + E(\kappa_{i-1}[IL_i^-(T^-), a_i(r, q)Q_{i-1}]), \end{aligned}$$

where $T = t + L_i$ and $\kappa_i[r, q] = K_i$.

To carry out the above recursion, we need $E(b_i(r, q))$ and the joint distribution of $IL_i^-(T^-)$ and $a_i(r, q)$. Note that

- If $IL_i^-(T^-) \geq R_{i-1} + 1$ then $a_i(r, q) = 0$ and $b_i(r, q) = n$;
- If $IL_i^-(T^-) \leq R_{i-1} - (n - 1)Q_{i-1}$ then $a_i(r, q) = n$ and $b_i(r, q) = 0$;
- If $R_{i-1} - jQ_{i-1} < IL_i^-(T^-) \leq R_{i-1} - (j - 1)Q_{i-1}$ then $a_i(r, q) = j$ and $b_i(r, q) = n - j$, for $j = 1, \dots, n - 1$, and that

$$IL_i^-(T^-) = IP_i(t^-) - D[t, T] = r - D[t, T].$$

These relations determine the joint distribution of $IL_i^-(T^-)$ and $a_i(r, q)$. It can be easily verified that

$$E(b_i(r, q)) = \sum_{j=0}^{n-1} F^i(r - R_{i-1} + jQ_{i-1} - 1)$$

where $F^i(x) = \Pr(D[t, T] \leq x)$.

The computation of $\kappa_i[\cdot, \cdot]$ can be restricted to a finite set. Note that the size of any shipment to stage i , direct or normal, is a positive integer multiple of Q_i and is less than or equal to Q_N . Also, the echelon

inventory position at stage i is always less than or equal to $R_i + Q_i$. Therefore, $\kappa_i[r, q]$ is nonzero only if $r + q \leq R_i + Q_i$ and $q = mQ_i \leq Q_N$ where m is a positive integer. On the other hand, note that if

$$r + q \leq \min\{R_j + Q_j, j = 1, \dots, i - 1\}$$

then the batch q will be sent down to stage 1 through only direct shipments and thus $\kappa_i[r, q] = \sum_{j=1}^i K_j$. Consequently, the evaluation range for $\kappa_i[\cdot, \cdot]$ can be further confined to $r + q > \min\{R_j + Q_j, j = 1, \dots, i - 1\}$.

Since the outside supplier has ample stock, each shipment to stage N is a normal shipment. Thus the expected total setup cost incurred by each replenishment batch at stage N is $\kappa_N[R_N, Q_N]$. Since the order frequency at stage N is λ/Q_N , the average system-wide setup cost is $\lambda\kappa_N[R_N, Q_N]/Q_N$.

References

- Atkins, D. and D. Sun, "98%-Effective Lot-Sizing for Series Inventory System with Backlogging," *Oper. Res.* (forthcoming).
- Axsater, S., "Exact and Approximate Evaluation of Batch Ordering Policies for Two-Level Inventory Systems," *Oper. Res.*, 41 (1993), 777-785.
- , "Evaluation of (R, Q) Policies for Two-Level Inventory Systems with Poisson Demand," Lulea University of Technology, Sweden, 1991.
- and K. Rosling, "Installation vs. Echelon Stock Policies for Multi-Level Inventory Control," *Management Sci.*, 39 (1993), 1274-1280.
- Badinelli, R., "A Model for Continuous-Review Pull Policies in Serial Inventory Systems," *Oper. Res.*, 40 (1992), 142-156.
- Chen, F. and Y.-S. Zheng, "Evaluating the Performance of a Serial Inventory System with Stochastic Demand," Working Paper 91-06-08, Decision Sciences Department, The Wharton School, University of Pennsylvania, 1991.
- and —, "Lower Bounds for Multi-Echelon Stochastic Inventory Systems," *Management Sci.*, 40 (November 1994).
- and —, "Computing Optimal Echelon Stock (R, nQ) Policies in Serial Systems," Working Paper, Decision Sciences Department, The Wharton School, University of Pennsylvania, 1992.
- and —, "Inventory Models with General Backorder Costs," *European J. Operational Res.*, 65 (1993a), 175-186.
- and —, "One-Warehouse Multi-Retailer Systems with Centralized Stock Information," Working Paper, Decision Sciences Department, The Wharton School, University of Pennsylvania, 1993b.
- Clark, A. J. and H. Scarf, "Optimal Policies for a Multi-Echelon Inventory Problem," *Management Sci.*, 6 (1960), 475-490.
- and —, "Approximate Solutions to a Simple Multi-Echelon Inventory Problem," in *Studies in Applied Probability and Management Science*, pp. 88-110, K. J. Arrow, S. Karlin and H. Scarf (Eds.), Stanford University Press, Stanford, CA, 1962.
- De Bodt, M. A. and S. C. Graves, "Continuous Review Policies for a Multi-Echelon Inventory Problem with Stochastic Demand," *Management Sci.*, 31 (1985), 1286-1295.
- Deuermeyer, B. and L. B. Schwarz, "A Model for the Analysis of System Service Level in Warehouse/Retailer Distribution Sys-

- tems: The Identical Retailer Case," in L. B. Schwarz (Ed.), *Studies in the Management Sciences: The Multi-Level Production/Inventory Control Systems*, Vol. 16, North-Holland, Amsterdam, 1981, 163–193.
- Federgruen, A. and P. Zipkin, "Computational Issues in an Infinite Horizon, Multi-Echelon Inventory Model," *Oper. Res.*, 32 (1984), 818–836.
- Hadley, G. and T. M. Whitin, "A Family of Inventory Models," *Management Sci.*, 7 (1961), 351–371.
- and —, *Analysis of Inventory Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1963.
- Lee, H. L. and K. Moinzadeh, "Two-Parameter Approximations for Multi-Echelon Repairable Inventory Models with Batch Ordering Policy," *IIE Transactions* 19 (1987a), 140–149.
- and —, "Operating Characteristics of a Two-Echelon Inventory System for Repairable and Consumable Items Under Batch Ordering and Shipment Policy," *Naval Res. Logist. Quart.*, 34 (1987b), 365–380.
- Moinzadeh, K. and H. L. Lee, "Batch Size and Stocking Levels in Multi-Echelon Repairable Systems," *Management Sci.*, 32 (1986), 1567–1581.
- Rosling, K., "Optimal Inventory Policies for Assembly Systems Under Random Demands," *Oper. Res.*, 37 (1989), 565–579.
- Roundy, R. O., "98% Effective Integer-Ratio Lot-Sizing for One-Warehouse Multi-Retailer Systems," *Management Sci.*, 31 (1985), 1416–1430.
- , "98% Effective Lot-Sizing Rule for a Multi-Product, Multi-Stage Production-Inventory System," *Math. Oper. Res.*, 11 (1986), 699–727.
- Schmidt, C. P. and S. Nahmias, "Optimal Policy for a Two-Stage Assembly System Under Random Demand," *Oper. Res.*, 33 (1985), 1130–1145.
- Svoronos, A. and P. Zipkin, "Estimating the Performance of Multi-Level Inventory Systems," *Oper. Res.*, 36 (1988), 57–72.
- Wolff, R. W., "Poisson Arrivals See Time Averages," *Oper. Res.*, 30 (1982), 223–231.
- Zheng, Y.-S., "On Properties of Stochastic Inventory Systems," *Management Sci.*, 38 (1992), 87–103.
- and F. Chen, "Inventory Policies with Quantized Ordering," *Naval Res. Logist.*, 39 (1992), 285–305.
- Zipkin, P., "Stochastic Leadtimes in Continuous-Time Inventory Models," *Naval Res. Logist. Quart.*, 33 (1986), 763–774.

Accepted by Awi Federgruen; received August 23, 1991. This paper has been with the authors $7\frac{1}{2}$ months for 3 revisions.