

A Single-Unit Decomposition Approach to Multiechelon Inventory Systems

Alp Muharremoglu

Graduate School of Business, Columbia University, New York, New York 10027, alp2101@columbia.edu

John N. Tsitsiklis

Laboratory for Information and Decision Sciences, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, jnt@mit.edu

We show the optimality of state-dependent echelon base-stock policies in uncapacitated serial inventory systems with Markov-modulated demand and Markov-modulated stochastic lead times in the absence of order crossing. Our results cover finite-time horizon problems as well as infinite-time horizon formulations, with either a discounted or an average cost criterion. We employ a novel approach, based on a decomposition of the problem into a series of single-unit single-customer problems that are essentially decoupled. Besides providing a simple proof technique, this approach also gives rise to efficient algorithms for the calculation of the base-stock levels.

Subject classifications: inventory/production: multiechelon; inventory/production: policies: review/lead times.

Area of review: Manufacturing, Service, and Supply Chain Operations.

History: Received August 2001; revisions received March 2003, July 2006, September 2007; accepted October 2007.

1. Introduction

This paper deals with serial (multiechelon) inventory systems of the following type. There are M stages. Stage 1 receives stock from stage 2, stage 2 from stage 3, etc., and stage M receives stock from an outside supplier with ample stock. Demands originate at stage 1, and unfilled demand is backlogged. There are holding, ordering, and backorder costs, and a central controller has the objective of minimizing these costs in the appropriate time frame.

In their seminal paper, Clark and Scarf (1960) characterize optimal policies for an uncapacitated serial inventory system. They consider finite-horizon problems and prove that echelon base-stock policies are optimal when the demands are independent identically distributed (i.i.d.) and the lead times between stages are deterministic. Their proof technique involves a decomposition of the multiechelon problem into a series of single-stage problems. This general approach guided much of the subsequent literature, and many extensions were obtained using the same stage-by-stage decomposition. In particular, Federgruen and Zipkin (1984) extend the results to the stationary infinite-horizon setting, and Chen and Zheng (1994) provide an alternative proof that is also valid in continuous time. Rosling (1989) shows that a general assembly system can be converted to an equivalent serial system. All of these papers assume i.i.d. demands and deterministic lead times.

In this paper, we study a system with Markov-modulated demands and stochastic lead times. We assume that demands and lead times are stochastic and are affected (modulated) by an exogenous Markov process. Such a model can capture many phenomena such as seasonalities,

exchange rate variations, fluctuating market conditions, demand forecasts, etc. This type of demand model, where the distribution of demand depends on the state of a modulating Markov chain, is certainly not new to the inventory control literature. Song and Zipkin (1993), Beyer and Sethi (1997), Sethi and Cheng (1997), and Cheng and Sethi (1999) all investigate a single-stage system with such a demand model and prove the optimality of state-dependent (s, S) policies under different time horizon assumptions, with and without backlogging assumptions. Song and Zipkin (1992) and (1996a) evaluate the performance of base-stock policies in serial inventory systems, with state-independent and state-dependent policies, respectively. More recently, Chen and Song (2001) show the optimality of state-dependent echelon base-stock policies for serial systems with Markov-modulated demand and deterministic lead times, under an infinite-horizon average cost criterion. Markov-modulated demand is also considered by Parker and Kapuscinski (2004) in a capacitated setting.

The study of stochastic lead times in inventory control dates back to the early days of the literature. Hadley and Whitin (1963) investigate the subject for a single-stage problem, and suggest that two seemingly contradictory assumptions are needed—namely, that orders do not cross each other and that they are independent. Kaplan (1970) provides a simple model of stochastic lead times that prevents order crossing, while keeping the probability that an outstanding order arrives in the next time period independent of the current status of other outstanding orders. He shows that the deterministic lead-time results carry over to his model of stochastic lead times. Nahmias (1979) and Ehrhardt (1984) streamlined Kaplan's results. Zipkin

(1986) investigated stochastic lead times in continuous-time single-stage inventory models. Song and Zipkin (1996b) study a single-stage system with Markov modulated lead times. Svoronos and Zipkin (1991) evaluate one-for-one replenishment policies in the serial system setting. However, we are not aware of any optimality results for serial systems under any type of stochastic lead times. Our stochastic lead-time model incorporates the same two important features of Kaplan's stochastic lead-time model, i.e., the absence of order crossing and the independence from the current status of other outstanding orders. In our model, just like in Kaplan's, an exogenous random variable determines which outstanding orders are going to arrive at a given stage. However, we additionally allow the stochastic lead times to depend on the state of a modulating Markov chain, and we also allow for dependencies between the lead-time random variables corresponding to different stages in the system.

The standard approach in multiechelon inventory theory is a decomposition into a series of single-stage problems. This approach, and especially the simplified and streamlined proof technique introduced by Veinott (1966), indeed does lead to many of the results and extensions discussed above. Nevertheless, our approach relies on a decomposition into a series of unit-customer pairs. Consider a single unit and a single customer. Assume that the distribution of time until the customer arrives to the system is known and that the goal is to move the unit through the system in a way that optimizes the holding versus backorder cost trade-off. Because only a single unit and a single customer are present, this problem is much simpler than the original one. We show that under the assumptions of this paper, the original problem is equivalent to a series of decoupled single-unit single-customer problems. This approach allows us to handle several extensions to the standard model in an intuitive manner, and provides an alternative to inductive arguments based on dynamic programming equations. The primary contribution of this paper is the formal proof of the decomposition of the serial inventory problem into essentially decoupled subproblems, each consisting of a single unit and a single customer.

A decomposition of the type employed here has been introduced in a series of papers by Axsäter, although without bringing it to bear on the full-fledged dynamic programming formulation of the inventory control problem. Axsäter (1990) observes that in a distribution system with a single depot and multiple retailers that follow base-stock policies, any particular unit ordered by retailer i is used to fill a particular future demand. He then matches this unit with that demand and evaluates the expected cost for this unit and "its demand." Using this approach, he develops an efficient method to evaluate the cost of a given base-stock policy for a two-echelon distribution system in continuous time with Poisson demand under the infinite-horizon average cost criterion. In Axsäter (1993a), he extends this result to batch-ordering policies and in Axsäter (1993b), he investigates

the system with periodic review, using the virtual allocation rule suggested by Graves (1996) and a base-stock policy. As we show in this paper, Axsäter's insight, when used properly, leads to the decomposition of the problem into single-unit single-customer problems, and provides a powerful technique for developing optimality results and algorithms for multiechelon systems.

A related work is the masters thesis by Achy-Brou (2001) (supervised by the second author, concurrently with this work), who studies the single-unit single-customer subproblem for the case of i.i.d. demands and deterministic lead times and a discounted cost criterion. This work formulates the subproblem as a dynamic program, describes and implements the associated dynamic programming algorithm, analyzes structural properties of the solution, and discusses the relationship between the subproblem and base-stock policies in the overall inventory system.

We finally note that besides providing a simple proof technique, the decomposition into single-unit single-customer subproblems leads to simple and efficient algorithms for calculating the base-stock levels. Even for several special cases of our model for which computational methods are already available, our algorithms are at least as efficient and provide an alternative method with potential advantages. These are listed at the end of §6, where the algorithms are presented.

The rest of the paper has six sections. Section 2 provides some background results on generic discrete-time decomposable dynamical systems. Section 3 provides a mathematical formulation of the problem and the necessary notation. Sections 4 and 5 contain the results for finite and infinite-horizon versions of the problem, respectively. Section 6 discusses the resulting algorithms for computing the optimal base-stock levels. Section 7 concludes the paper.

2. Preliminaries: Decomposable Systems

In this section, we introduce the problem of optimal control of a *decomposable* system and point out the decoupled nature of the resulting optimal policies. The result we provide is rather obvious, but we find it useful to state it explicitly, both for ease of exposition and also because it is a key building block for our subsequent analysis.

Following the notation in Bertsekas (1995), we consider a generic stationary discrete-time dynamical system of the form

$$x_{t+1} = f(x_t, u_t, w_t), \quad t = 0, 1, \dots, T - 1. \quad (1)$$

Here, x_t is the state of the system at time t , u_t a control to be applied at time t , w_t a stochastic disturbance, and T is the time horizon. We assume that x_t , u_t , and w_t are elements of given sets X , U , and W , respectively, and that $f: X \times U \times W \mapsto X$ is a given mapping. Finally, we assume that, given x_t and u_t , the random variable w_t is conditionally independent from the past and has a given conditional distribution.

We define a *policy* π as a sequence of functions, $\pi = (\mu_0, \mu_1, \dots, \mu_{T-1})$, where each $\mu_i: X \mapsto U$ maps the state x into a control $u = \mu_i(x)$. Let Π be the set of all policies.

Given an initial state x_0 and a policy π , the sequence x_t becomes a Markov chain with a well-defined probability distribution. For any time $t < T$ and any state $x \in X$, we define the cost-to-go $J_{i,T}^\pi(x)$ (from time t until the end of the horizon) by

$$J_{i,T}^\pi(x) = E \left\{ \sum_{\tau=t}^{T-1} \alpha^{\tau-t} \cdot g(x_\tau, \mu_\tau(x_\tau)) \mid x_t = x \right\},$$

where $g: X \times U \mapsto [0, \infty]$ is a given cost-per-stage function and $\alpha \in [0, 1]$ is a discount factor. The optimal cost-to-go function $J_{i,T}^*$ is defined by

$$J_{i,T}^*(x) = \inf_{\pi \in \Pi} J_{i,T}^\pi(x).$$

(Note that $J_{i,T}^\pi(x)$ and $J_{i,T}^*(x)$ can be infinite at certain states.) A policy π^* is said to be *optimal* if

$$J_{i,T}^{\pi^*}(x) = J_{i,T}^*(x) \quad \forall t, \forall x \in X.$$

When $t = 0$, we will use the simpler notation $J_T^\pi(x)$ and $J_T^*(x)$ instead of $J_{0,T}^\pi(x)$ and $J_{0,T}^*(x)$, respectively.

We now introduce the notion of a decomposable system. Loosely speaking, this is a system consisting of multiple (countably infinite) noninteracting subsystems that are driven by a common source of uncertainty, which evolves independently of the subsystems and is modulated by a Markov process s_t .

DEFINITION 2.1. A discrete-time dynamic programming problem of the form described above is said to be *decomposable* if it admits a representation with the following properties:

A1. The state space is a Cartesian product of the form $X = S \times \hat{X} \times \hat{X} \cdots$, so that any $x \in X$ can be represented as $x = (s, x^1, x^2, \dots)$ with $s \in S$ and $x^i \in \hat{X}$, for every $i \geq 1$.

A2. There is a set \hat{U} so that the control space U is the Cartesian product of countably many copies of \hat{U} , that is, any $u \in U$ can be represented as $u = (u^1, u^2, \dots)$ with $u^i \in \hat{U}$, for all $i \geq 1$.

A3. For each t , the conditional distribution of w_t given x_t and u_t , depends only on s_t .

A4. The evolution equation (1) for x_t is of the form

$$s_{t+1} = f^s(s_t, w_t), \quad x_{t+1}^i = \hat{f}(x_t^i, u_t^i, w_t), \quad \forall i \geq 1, \forall t,$$

for some functions $f^s: S \times W \mapsto S$ and $\hat{f}: \hat{X} \times \hat{U} \times W \mapsto \hat{X}$.

A5. The cost function g is additive, of the form

$$g(x_t, u_t) = \sum_{i=1}^{\infty} \hat{g}(s_t, x_t^i, u_t^i),$$

for some function $\hat{g}: S \times \hat{X} \times \hat{U} \mapsto [0, \infty)$.

A6. The sets \hat{X} and W are countable. The sets S and \hat{U} are finite.

In a decomposable system, any policy π can be represented in terms of a sequence of component mappings $\mu_i^i: X \mapsto \hat{U}$, so that

$$u_t^i = \mu_t^i(x_t) \quad \forall i, t.$$

We are especially interested in those policies under which the control u_t^i that affects the i th subsystem is chosen locally, without considering the state of the other subsystems, and using a mapping μ_t^i that is the same for all i .

DEFINITION 2.2. A policy π for a decomposable system is said to be *decoupled* if it can be represented in terms of mappings $\hat{\mu}_i: S \times \hat{X} \mapsto \hat{U}$, so that

$$u_t^i = \hat{\mu}_t^i(s_t, x_t^i) \quad \forall i, t.$$

For a decomposable system, the various state components x_t^1, x_t^2, \dots do not interact, the only coupling arising through the exogenous processes s_t and w_t . Because the costs are also additive, it should be clear that each subsystem can be controlled separately (that is, using a decoupled policy) without any loss of optimality. Furthermore, because all subsystems are identical, the same mapping $\hat{\mu}_t^i$ can be used in each subsystem. The required notation and a formal statement is provided below.

Each subsystem i defines a *subproblem*, with dynamics

$$s_{t+1} = f^s(s_t, w_t), \quad x_{t+1}^i = \hat{f}(x_t^i, u_t^i, w_t),$$

and costs per stage $\hat{g}(s_t, x_t^i, u_t^i)$. A policy $\hat{\pi}$ for a subproblem is of the form $\hat{\pi} = (\hat{\mu}_0^i, \hat{\mu}_1^i, \dots, \hat{\mu}_{T-1}^i)$, where each $\hat{\mu}_t^i$ is a mapping from $S \times \hat{X}$ into \hat{U} :

$$u_t^i = \hat{\mu}_t^i(s, x^i).$$

Let

$$\hat{J}_{i,T}^*(s, x^i)$$

be the optimal cost-to-go function for a subsystem that starts at time t from state (s, x^i) and evolves until the end of the horizon T . Note that this function is the same for all i , because we have assumed the subsystems to have identical dynamics and cost functions. Furthermore, because the control set \hat{U} is finite, an optimal policy is guaranteed to exist.

LEMMA 2.1. Consider a decomposable system.

1. For any $x = (s, x^1, x^2, \dots) \in X$ and any $t \leq T$, we have

$$J_{i,T}^*(x) = \sum_{i=1}^{\infty} \hat{J}_{i,T}^*(s, x^i).$$

2. There exists a decoupled policy π^* that is optimal, that is,

$$J_{i,T}^{\pi^*}(x) = J_{i,T}^*(x) \quad \forall t, \forall x \in X.$$

3. For any s, x^i , and any remaining time k , let $\widehat{U}_k^*(s, x^i) \subset \widehat{U}$ be the set of all decisions that are optimal for a subproblem, if the state of the subproblem at time $T - k$ is (s, x^i) . A policy $\pi = \{\mu_t^i\}$ is optimal if and only if for every i, t , and any $x = (s, x^1, x^2, \dots) \in X$ for which $J_{i,T}^*(x) < \infty$, we have

$$\mu_t^i(x) \in \widehat{U}_{T-t}^*(s, x^i).$$

The proof of the above result is straightforward and is omitted. Suffice it to say that we can pick an optimal policy for the subproblem and replicate it for all subsystems to obtain a decoupled and optimal policy. The last part of the lemma simply states that for any given x and t , a decision vector $u_t = (u_t^1, u_t^2, \dots)$ is optimal if and only if each component u_t^i of the decision is optimal for the i th subsystem viewed in isolation (except of course if the cost-to-go $J_{i,T}^*(x)$ is infinite, in which case all possible decisions are optimal). Let us also remark that the sets $\widehat{U}_k^*(s, x^i)$ of optimal decisions only depend on the remaining time k , but do not depend on the value of T . This is an immediate consequence of the stationarity of the problem.

3. Problem Formulation

We consider a single-product serial inventory system consisting of M stages, indexed by $1, \dots, M$. Customer demand can only be satisfied by units at stage 1. Any demand that is not immediately satisfied is backlogged. The inventory at stage m ($m = 1, \dots, M - 1$) is replenished by placing an order for units stored at stage $m + 1$. Stage M receives replenishments from an outside supplier with unlimited stock. For notational simplicity, we label the outside supplier as stage $M + 1$. We assume that the system is periodically reviewed and, therefore, a discrete-time model can be employed.

To describe the evolution of the system, we need to specify the sources of uncertainty, the statistics of the demand, and the statistics of the lead times for the various orders.

(a) *Markovian exogenous uncertainty*: We assume that the customer demands and the order lead times are influenced by an exogenous finite-state Markov chain s_t , assumed to be time homogeneous and ergodic (irreducible and aperiodic).

(b) *Demand model*: The (nonnegative integer) demand d_t during period t is assumed to be Markov modulated. In particular, the probability distribution of d_t depends on the state s_t of the exogenous Markov chain and, conditioned on that state, is independent of the past history of the process. We also assume that $E[d_t | s_t = s] < \infty$ for every $s \in S$, and, to avoid trivial situations, that $E[d_t | s_t = s] > 0$ for at least one value of s .

(c) *Lead-time model*: We assume that the lead time between stage $m + 1$ and stage m is upper bounded by some integer l_m . We assume that the probability that an outstanding order arrives during the current period depends only on the amount of time since the order was placed, the exogenous state s_t , and the destination stage m and, given these, it is conditionally independent of the past history of the process. Finally, we assume that orders cannot overtake each other: An order cannot arrive at its destination before an earlier order does.

The lead-time model introduced above includes the obvious special case of deterministic lead times. It also includes a stochastic model of the type we describe next, and which extends the model of Kaplan (1970). At each time period t , there is a random variable ρ_t^m that determines which outstanding orders will arrive at stage m at time $t + 1$. More precisely, an outstanding order will be delivered at stage m if and only if it was placed ρ_t^m or more time units ago. Note that such a mechanism ensures that orders cannot overtake each other. Let $\rho_t = (\rho_t^1, \rho_t^2, \dots, \rho_t^M)$ be the vector of lead-time random variables associated with the various stages. We assume that the statistics of ρ_t are given in terms of a conditional probability distribution, given s_t . Notice that such a model allows for dependencies between the lead-time random variables corresponding to the same period but different stages. Furthermore, it can also capture intertemporal dependencies through the dynamics of s_t .

The cost structure that we use is fairly standard and consists of linear holding, ordering, and backorder costs. In more detail, we assume:

(a) For each stage m , there is an inventory holding cost rate h_m that gets charged at each time period to each unit at that stage. We assume that the holding cost rate h_{M+1} at the external supplier is zero. For concreteness, we also assume that after a unit is ordered and during its lead time, the holding cost rate charged for this unit is the rate corresponding to the destination echelon.

(b) For each stage m , there is an ordering cost c_m for initiating the shipment of a unit from stage $m + 1$ to stage m .

(c) There is a backorder cost rate b , which is charged at each time step for each unit of backlogged demand.

We assume that the holding cost (at stages other than the external supplier) and backorder cost parameters are positive, and that the ordering cost is nonnegative.

The detail-oriented reader may have noticed that the model has not been specified in full detail: We would still need to describe the relative timing of observing the demand, fulfilling the demand, placing orders, receiving orders, and charging the costs. Different choices with respect to these details result, in general, to slightly different optimal costs and policies. Whatever specific choices are made, the arguments used for our subsequent results remain unaffected. For specificity, however, we make the following assumption about delivery of units to customers: If a customer arrives during period t , a unit can be given to that customer only at time $t + 1$ or later.

3.1. State and Control Variables

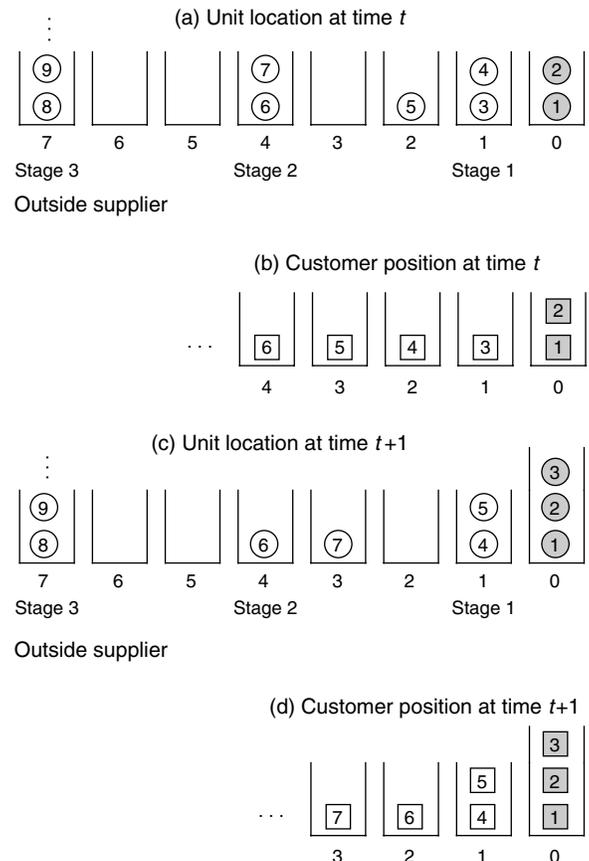
In §2, we described a generic discrete-time dynamic system. In this subsection, we define our choices for the state, control, and disturbance variables for the inventory control system being studied.

The traditional approach would be the following. The state would consist of a vector whose components are the number of units at each stage, the number of units that have been released by stage m and have been in transit for k time units (one component for each pair (m, k)), the size of the backlogged demand, and the state of the modulating Markov chain. The control would be the number of units to be released from each stage to the next. The demand in a given period, the various random variables associated with the random lead times (e.g., the random vector ρ_t in our earlier example), and the transition of the modulating chain s_t , would constitute the random disturbance. Obviously, such a choice is sufficient for optimization purposes because one does not need to distinguish between units that are at the same stage or between units that have been in transit for the same amount of time. However, we approach the problem differently. We treat each individual unit and each individual customer as distinguishable objects and then show that this results in a decomposable problem, with each unit-customer pair viewed as a separate subsystem. Towards this goal, we start by associating a unique label with each unit and customer.

At any given time, there will be a number of units at each stage or on order between two given stages. In addition, conceptually, we have a countably infinite number of units at the outside supplier, which we call stage $M + 1$. We will now introduce a set of conceptual unit locations in the system that can be used to describe where a unit is found and, if it is part of an outstanding order, how long ago it was ordered.

DEFINITION 3.1. *The location of a unit:* First, each of the actual stages in the system will constitute a location. Next, we insert $l_m - 1$ artificial locations between the locations corresponding to stages m and $m + 1$, for $m = 1, \dots, M$, to model the units in transit between these two stages. If a unit is part of an order between stages $m + 1$ and m that has been outstanding for k periods, $1 \leq k \leq l_m - 1$, then it will be in the k th location between stages $m + 1$ and m . Finally, for any unit that has been given to a customer, we define its location to be zero. Thus, the set of possible locations is $\{0, 1, \dots, N + 1\}$, where $N = \sum_{m=1}^M l_m$. We index the locations starting from location 0. Location 1 corresponds to stage 1. Location 2 corresponds to units that have been released $l_1 - 1$ times ago from stage 2. Location l_1 corresponds to units that have been released from stage 2 one time step earlier. Location $l_1 + 1$ corresponds to stage 2, etc. Location $N + 1$ corresponds to the outside supplier (stage $M + 1$). For example, in Figure 1(a), unit 5 is in location 2 at time t , which means that this unit has been released from stage 2 (location 4) at time $t - 2$.¹

Figure 1. Illustration of the system dynamics and the various definitions: Consider a system with $M = 2$ stages and deterministic lead times $l_1 = l_2 = 3$; this results in a total of $N + 2 = 8$ locations.



Notes. (a) Part (a) shows a configuration at some time t , in which there are seven units in the system and an infinity of additional units at the outside supplier. Units 1 and 2 have already been given to customers, in previous periods.

(b) Part (b) shows a customer configuration. Customers 1 and 2 arrived in previous periods and have already been given a unit each. Customer 3 has also arrived but has not yet been given a unit. Customers 4, 5, etc., have not yet arrived.

Suppose that the following decisions are made: $u_i^t = 1$ and $u_i^t = 0$ for $i \neq 7$. Suppose also that the new demand d_t turns out to be 2.

(c) Because there is only one customer in position 1, unit 3 (the one with lower index) is given to customer 3. Hence, this unit moves to location 0. Also, the released unit 7 moves one location.

(d) Customer 3 obtains unit 3 and moves to position 0. The two newly arrived customers (4 and 5) move to position 1. Every customer i , $i > 5$, moves by two positions.

DEFINITION 3.2. Let A be the set of unit locations corresponding to the actual stages of the system, including the outside supplier (location $N + 1$). Let A' be the same as A , except that location 1 is excluded. Also, let

$$v(z) = \max_{m \in A \text{ and } m \leq z} m \quad \forall z > 0,$$

so that starting from location z and going in the direction of decreasing z , $v(z)$ is the first location corresponding

to an actual stage. For example, in Figure 1(a), we have $A = \{1, 4, 7\}$, $A' = \{4, 7\}$, and $v(3) = 1$.

We index the countably infinite pool of units by the non-negative integers. We assume that *the indexing is chosen at time 0 in increasing order of their location*, breaking ties arbitrarily.

Let us now turn to the customer side of the model, which we describe using a countably infinite pool of past and potential future customers, with each such customer treated as a distinguishable object. At any given time, there is a finite number of customers that have arrived and whose demand is either satisfied or backlogged. In addition, we envisage a countably infinite number of potential customers that may arrive to the system at a future period. Consider the system at time 0. Let k be the number of customers that have arrived whose demand is already satisfied. We index them as customers $1, \dots, k$ in any arbitrary order. Let l be the number of customer that have arrived whose demand is backlogged. We index them as customers $k+1, k+2, \dots, k+l$ in any order. The remaining (countably infinite) customers are assigned indices $k+l+1, k+l+2, \dots$ in order of their arrival times to the system, breaking ties arbitrarily, starting with the earliest arrival time. Of course, we do not know the exact arrival times of future customers, but we can still talk about a “next customer,” a “second to next customer,” etc. This way, we index the past and potential future customers at time 0. We now define a quantity that we call “the position of a customer.”

DEFINITION 3.3. *The position of a customer:* Suppose that at time t a customer i has already arrived and its demand is satisfied. We define the position of such a customer to be zero. Suppose that the customer has arrived but its demand is backlogged. Then, we define the position of the customer to be one. If, on the other hand, customer i has not yet arrived but customers $1, 2, \dots, m$ have arrived, then the position of customer i at time t is defined to be $i - m + 1$. In particular, a customer whose position at time t is k will have arrived by the end of the current period if and only if $d_t \geq k - 1$. For example, in Figure 1(b), customer 5 has a position of 3 at time t . Customer 5 will have arrived at time $t + 1$ if and only if two or more customers (customers 4 and 5) arrive during period t .

Now that we have labeled every unit and every customer, we can treat them as distinguishable objects and, furthermore, we can think of unit i and customer i as forming a pair. This pairing is established at time 0, when indices are assigned, taking into account the initial unit locations and customer positions, and is to be maintained throughout the planning horizon.

We are now ready to specify the state and control variables for the problem of interest. For each unit-customer pair i , $i \in \mathbb{N}$, we have a vector (z_t^i, y_t^i) , with $z_t^i \in Z = \{0, 1, \dots, N+1\}$ and $y_t^i \in Y = \mathbb{N}_0$, where z_t^i is the location of unit i at time t , and y_t^i is the position of customer i

at time t . The state of the system consists of a countably infinite number of such vectors, one for each unit-customer pair, and finally a variable $s_t \in S$ for the state of the modulating Markov chain, i.e.,

$$x_t = \{s_t, (z_t^1, y_t^1), (z_t^2, y_t^2), \dots\}.$$

The control vector is an infinite binary sequence $u_t = (u_t^1, u_t^2, \dots)$, where the i th component u_t^i corresponds to a “release” or “hold” decision for the i th unit. An action u_t^i has an effect only if the corresponding unit i is at a nonartificial location other than location 1, (i.e., $z_t^i \in A'$). If the unit is in such a location, $u_t^i = 0$ corresponds to holding it at its current location, and $u_t^i = 1$ corresponds to releasing it, and the unit will arrive at the next nonartificial location $v(z_t^i - 1)$ after a stochastic lead time (that satisfies the conditions explained at the beginning of §3). The movement of units that are at artificial locations (in between stages, i.e. locations outside A) is solely governed by the randomness in the lead times.

If a customer i is at position $y_t^i \geq 2$, in the next period, it moves to position $(y_t^i - d_t + 1)^+ + 1$. Finally, units that are in location 1 and customers that are in position 1 move to unit location 0 and customer position 0, respectively, in the following way: Out of the available units in location 1 and arrived customers waiting for a unit (customers in position 1), k of them with the lowest indices automatically move to location 0 and position 0, respectively, where k is the minimum of the number of units at location 1 and customers in position 1. Once a unit moves to location 0 or a customer moves to position 0, they stay there.

The random disturbance at time t consists of the demand d_t , random variables that model the uncertainty in the lead times (e.g., the vector ρ_t of lead-time random variables in our earlier example), and whatever additional exogenous randomness is needed to drive the Markov chain s_t .

We will refer to the above-described model of the serial inventory system as the “main model.” Clearly, the main model is a sufficient description of the overall system, albeit not the most compact one. Let $J_{t,T}^\pi(x)$ and $J_{t,T}^*(x)$ be the cost of policy π and the optimal cost, respectively, under the main model, starting from state x at time t until the end of the horizon T . We use the shorthand versions $J_T^\pi(x)$ and $J_T^*(x)$, if $t = 0$.

3.2. Policy Classification

We now define various classes of policies for the main model, state-dependent echelon base-stock policies being one particular class. In the next section, we will show that the search for an optimal policy can be restricted to any one of these policy classes, without sacrificing performance.

As a first step, we define a class of states that we call *monotonic states*.

DEFINITION 3.4. A state $x_t = \{s_t, (z_t^1, y_t^1), (z_t^2, y_t^2), \dots\}$ is called *monotonic* if and only if the unit locations are monotonic functions of the unit labels, that is,

$$i < j \Rightarrow z_t^i \leq z_t^j.$$

Note that x_0 is always a monotonic state, by construction. The state x_t shown in Figure 1 is monotonic, but the state x_{t+1} is not.

DEFINITION 3.5. Here, we define three classes of policies. Let $x_t = \{s_t, (z_t^1, y_t^1), (z_t^2, y_t^2), \dots\}$ and $u_t = (u_t^1, u_t^2, \dots)$ be the state and control variables as defined in §3.1. Consider a policy $\pi = \{\mu_t^i\}$, where each $\mu_t^i: X \mapsto \widehat{U} = \{0, 1\}$ prescribes the control for the i th unit at time t , according to $u_t^i = \mu_t^i(x_t)$.

Monotonic policies: The policy π is monotonic if it guarantees that a monotonic state x_t always results in a next state x_{t+1} that is monotonic. Intuitively, a policy is monotonic if and only if units can never overtake each other.

Decoupled policies: This is essentially the same as Definition 2.2 in the preceding section. We call a policy *decoupled* if it can be represented in terms of mappings $\hat{\mu}_t: S \times Z \times Y \mapsto \widehat{U}$, so that

$$\mu_t^i(x) = \hat{\mu}_t(s_t, z_t^i, y_t^i) \quad \forall i, t.$$

In other words, a *decoupled policy* is a policy where the decision of whether or not to release a unit from its current location can be written as a function of the state of the modulating Markov chain, the location of the unit, and the position of the corresponding customer. Moreover, the function is the same for every unit.

State-dependent echelon base-stock policies: A policy is a state-dependent echelon base-stock policy if for every t , every state x , every location $z \in A'$, and every $s \in S$, there exists a value² $S_t^{v(z-1)}(s)$ such that

$$= \min \left\{ \begin{array}{l} \overbrace{|\{i \mid z^i = z, \mu_t^i(x) = 1\}|}^{\text{\# units released from } z} \\ \left(\begin{array}{l} \text{basestock level of } v(z-1) \\ S_t^{v(z-1)}(s) \\ \text{echelon inventory position at } v(z-1) \\ - \left(\overbrace{|\{i \mid 1 \leq z^i \leq z-1\}|}^{\text{Backlog}} - |\{i \mid y^i = 1\}| \right) \end{array} \right)^+ \\ \underbrace{|\{i \mid z^i = z\}|}_{\text{\# units at } z} \end{array} \right\},$$

We are using here the notation $|B|$ to denote the cardinality of a set B , and the notation $(a)^+$ to denote $\max\{0, a\}$. In other words, such a policy operates as follows: For every unit location $z > 1$ corresponding to an actual stage ($z \in A'$), the policy calculates the echelon inventory position at the next actual stage (unit location $v(z-1) \in A$) and releases enough units (to the extent that they are available) to raise this number to a target value $S_t^{v(z-1)}(s)$. The echelon inventory position at the actual stage that corresponds to unit location $v(z-1)$ is the total number of units at locations $1, \dots, v(z-1)$, plus the units in transit towards location $v(z-1)$ (i.e., units at locations $v(z-1)+1, \dots, z-1$), minus the backlogged demand.

Note that if the initial state is a monotonic state (which we always assume to be the case) and one uses a monotonic policy, the state of the system at any time in the planning horizon will be monotonic.

We say that a set of policies Π_0 is *optimal* (respectively, *M-optimal*) for the main model if

$$\inf_{\pi \in \Pi_0} J_{i,T}^\pi(x) = J_{i,T}^*(x)$$

for all t and all states x (respectively, all monotonic states x).

In the next section, we show that the intersection of the sets of monotonic and decoupled policies is M-optimal for the main model. Proposition 3.1 below will then imply that state-dependent echelon base-stock policies are also M-optimal. We then show in the next section (Theorem 4.1) that state-dependent echelon base-stock policies are optimal at all states, not just monotonic ones. This is the main result of this paper.

LEMMA 3.1. *If a policy is monotonic and decoupled, then for every unit location $z > 1$ that corresponds to an actual stage ($z \in A'$), the underlying subproblem policy $\hat{\mu}_t(s, z, y)$ has to be nonincreasing in y over the set $\{y \mid y \geq 1\}$.*

PROOF. If the underlying subproblem policy $\hat{\mu}_t(s, z, y)$ does not have the claimed property, then when the decoupled policy is applied to the main model, it will be possible for units to overtake each other, contradicting monotonicity. \square

PROPOSITION 3.1. *Suppose that π is a monotonic and decoupled policy. Then, there exists a state-dependent echelon base-stock policy that agrees with π at every monotonic state.*

PROOF. A monotonic and decoupled policy releases units whose corresponding customers are in positions less than or equal to a certain threshold. Let $y_t^*(s, z)$ be the threshold for releasing units at location z , when the Markov chain is in state s , at time t . At a monotonic state, this is equivalent to trying to set the echelon inventory position at the next downstream stage $v(z-1)$ to $y_t^*(s, z) - 1$, which is a state-dependent echelon base-stock policy with $S_t^{v(z-1)}(s) = y_t^*(s, z) - 1$. \square

4. Finite-Horizon Analysis

In this section, we consider the finite-horizon model and establish the optimality of state-dependent echelon base-stock policies for the main model. We start with the observation that the set of monotonic policies is optimal.

PROPOSITION 4.1. *The set of monotonic policies is optimal for the main model.*

PROOF. Because units are identical, releasing units with smaller indices no later than units with higher indices can be done without loss of optimality. \square

Note that the main model is not decomposable because of the coupling that arises when available units are delivered to available customers. Indeed, if there are two units at location 1, and if only one customer has arrived, only one unit will move to location 0, the one that has the lower index. Thus, the dynamics of the unit that was not delivered is affected by the presence of the unit that was delivered, violating condition A4 for decomposable systems in Definition 2.1.

We now introduce a modified model, to be referred to as the “surrogate model,” that is identical to the main model except for the way units and customers move from unit location 1 and customer position 1, respectively. In particular, in the surrogate model, we have

$$\text{if } (z_t^i, y_t^i) = (1, 1), \quad \text{then } (z_{t+1}^i, y_{t+1}^i) = (0, 0),$$

$$\text{if } z_t^i = 1 \text{ and } y_t^i > 1, \quad \text{then } z_{t+1}^i = 1,$$

$$\text{if } y_t^i = 1 \text{ and } z_t^i > 1, \quad \text{then } y_{t+1}^i = 1.$$

In other words, a unit i moves to location 0 and the corresponding customer i moves to position 0 if and only if unit i is available in location 1 and customer i has arrived and is in position 1. This modification removes the coupling that was present in the main model, and makes the surrogate model decomposable. Note that the state and control spaces of the main and surrogate models are identical. Hence, any policy π can be applied to either the main or the surrogate model. Let $V_{t,T}^\pi(x)$ and $V_{t,T}^*(x)$ be the cost of policy π and the optimal cost, respectively, under the surrogate model, starting from state x at time t until the end of the horizon T . We use the shorthand versions $V_T^\pi(x)$ and $V_T^*(x)$, if $t = 0$. As before, we say that a set of policies Π_0 is optimal (respectively, M-optimal) for the surrogate model if $\inf_{\pi \in \Pi_0} V_{t,T}^\pi(x) = V_{t,T}^*(x)$ for every state x (respectively, for every monotonic state x). The next step is to relate the main model to the surrogate model. This is done in the following proposition.

PROPOSITION 4.2. (a) $J_{t,T}^\pi(x) \leq V_{t,T}^\pi(x)$ for all π, t, T , and x .

(b) Under a monotonic policy π , $J_{t,T}^\pi(x) = V_{t,T}^\pi(x)$ for every monotonic state x .

(c) $J_{t,T}^*(x) = V_{t,T}^*(x)$ for every monotonic state x .

(d) The set of monotonic policies is M-optimal for the surrogate model.

PROOF. (a) The only difference between the main model and the surrogate model arises when there are available units in location 1 and available customers in position 1, but some of them do not move to location 0 and position 0 under the surrogate model because their indices do not match. Therefore, under the same policy, the surrogate model can never incur a lower cost, but can incur a higher cost if the above-described situation happens.

(b) Starting from a monotonic state, and under a monotonic policy, the dynamics and costs of the main model and

the surrogate model are identical because the units arrive at location 1 in increasing sequence, implying that the situation described in part (a) cannot happen.

(c)–(d) These follow from Proposition 4.1 and parts (a)–(b) of this proposition. Indeed, if Π_m is the class of monotonic policies, we have

$$\begin{aligned} V_{t,T}^*(x) &\leq \inf_{\pi \in \Pi_m} V_{t,T}^\pi(x) = \inf_{\pi \in \Pi_m} J_{t,T}^\pi(x) \\ &= \inf_{\pi \in \Pi} J_{t,T}^\pi(x) = J_{t,T}^*(x) \leq V_{t,T}^*(x) \end{aligned}$$

for every monotonic state x . \square

Because the surrogate model is decomposable, it consists of a series of subproblems as in §2, with corresponding optimal cost-to-go functions $\hat{J}_{t,T}^*(s, z^i, y^i)$. From now on, whenever we refer to a “subproblem,” we will mean this single-unit, single-customer subproblem.

PROPOSITION 4.3. The set of decoupled policies is optimal for the surrogate model. Furthermore,

$$J_{t,T}^*(x) = V_{t,T}^*(x) = \sum_{i=1}^{\infty} \hat{J}_{t,T}^*(s, z^i, y^i)$$

for every t and every monotonic state x .

PROOF. The system under the surrogate model is decomposable, and Lemma 2.1 implies the optimality of decoupled policies. \square

The rest of the proof will proceed as follows. We will first show that there exists a decoupled policy for the surrogate model that is optimal and also monotonic. Using parts (b) and (c) of Proposition 4.2, this will imply that this same policy is M-optimal for the main model. We will make use of the following definition.

DEFINITION 4.1. For any $k, s \in S, z \in Z, y \in Y$, let $\hat{U}_k^*(s, z, y) \subset \{0, 1\}$ be the set of all decisions that are optimal for a subproblem if it is found at state (s, z, y) at time $t = T - k$, that is, k time steps before the end of the horizon.

The next lemma establishes that if an optimal subproblem policy releases a certain unit when the position of the corresponding customer is y , then it is also optimal (in the subproblem) to release the unit when the position of the corresponding customer is smaller than y . This is intuitive because as the customer comes closer, there is more urgency to move the unit towards stage 1.

LEMMA 4.1. For every (k, s, z, y) , with $z \in A'$, if $\hat{U}_k^*(s, z, y) = \{1\}$, then $1 \in \hat{U}_k^*(s, z, y')$ for every $y' < y$.

PROOF. Let $z \in A'$. Suppose that there exist some (k, s, z, y) and (k, s, z, y') , with $y' < y$, such that $\hat{U}_k^*(s, z, y) = \{1\}$ and $\hat{U}_k^*(s, z, y') = \{0\}$. Let $t = T - k$. Consider a monotonic state x_t for the surrogate model such that $(z_t^i, y_t^i) = (z, y)$, $(z_t^j, y_t^j) = (z, y')$, and $J_{t,T}^*(x_t) < \infty$. Note that because $y' < y$, we must have $j < i$. Then, according to

Lemma 2.1(3), the decision u_t under any optimal policy for the surrogate model must satisfy $u_t^i = 1$ and $u_t^j = 0$. This means that the higher-indexed unit i will move ahead of unit j , and the new state will not be monotonic. Therefore, a monotonic policy cannot be M-optimal for the surrogate model, which contradicts Proposition 4.2(d). \square

PROPOSITION 4.4. *The set of monotonic and decoupled policies is optimal for the surrogate model and M-optimal for the main model.*

PROOF. Let us fix t, s , and z . Let $k = T - t$ be the number of remaining time steps. If $z \notin A'$, we let $\hat{\mu}_t(s, z, y) = 1$ for all y . (These are clearly optimal decisions.) Now suppose that $z \in A'$. We consider three cases.

(a) If there are infinitely many y for which $\hat{U}_k^*(s, z, y) = \{1\}$, then by Lemma 4.1, $1 \in \hat{U}_k^*(s, z, y)$ for every y , and we let $\hat{\mu}_t(s, z, y) = 1$ for all y .

(b) If there is no y for which $\hat{U}_k^*(s, z, y) = \{1\}$, we let $\hat{\mu}_t(s, z, y) = 0$ for all y .

(c) If there is a largest y for which $\hat{U}_k^*(s, z, y) = \{1\}$, call it y^* , then we have $0 \in \hat{U}_k^*(s, z, y)$ for every $y > y^*$, and by Lemma 4.1, we have $1 \in \hat{U}_k^*(s, z, y)$ for every $y < y^*$. We then let $\hat{\mu}_t(s, z, y) = 1$ if and only if $y \leq y^*$.

The above-described procedure is repeated for every t, s , and z . This results in functions $\hat{\mu}_t$ that satisfy $\hat{\mu}_t(s, z, y) \in \hat{U}_k^*(s, z, y)$ for all (t, s, z, y) . According to Lemma 2.1(3), choosing the decision according to $\hat{\mu}_t$ for each unit at each time step constitutes an optimal (and also decoupled) policy for the surrogate model. Furthermore, by our construction, $\hat{\mu}_t(s, z, y)$ is a monotonically nonincreasing function of y . It follows that this decoupled policy that we constructed is also a monotonic policy, thus establishing the existence of a monotonic and decoupled policy that is optimal for the surrogate model. By parts (b) and (c) of Proposition 4.2, this policy is M-optimal for the main model. \square

The fact that an optimal policy $\{\hat{\mu}_t\}$ for the subproblem can be chosen so that it is nonincreasing in y can also be established using a traditional inductive argument, based on the dynamic programming recursion for the subproblem. For example, Achy-Brou (2001) studies the recursion for the infinite-horizon single-unit single-customer problem with deterministic lead times and i.i.d. demands, and provides a (somewhat lengthy) algebraic derivation. In contrast, the proof given here relies only on qualitative arguments.

THEOREM 4.1. *The set of state-dependent echelon base-stock policies is optimal for the main model.*

PROOF. Propositions 3.1 and 4.4 imply that state-dependent echelon base-stock policies are M-optimal. If such policies are optimal starting from a monotonic state, they are also optimal starting from a nonmonotonic state because all units are identical and the number of units released under a state-dependent echelon base-stock policy does not depend on the labels of particular units, but on the number of units in different locations. \square

Note that the base-stock levels $S_t^{v(z-1)}(s) = y_t^*(s, z) - 1$ are readily determined once an optimal subproblem policy and the corresponding sets $\hat{U}_k^*(s, z, y)$ for the single-unit, single-customer subproblem are available.

5. Infinite-Horizon Analysis

This section provides the main results for the case where the planning horizon is infinite. The proofs and some supporting results can be found in the online appendix that is available as part of the online version at <http://orjournal.informs.org>. We study both the expected total discounted cost criterion and the average cost per unit time criterion. We start with the part of the analysis that is common to both criteria.

In the infinite-horizon setting, we focus on stationary policies. A stationary policy is one of the form (μ, μ, \dots) , with $\mu: X \mapsto U$, so that the decision at each time is a function of the current state but not of the current time. We refer to a stationary policy of this type as policy μ and let Ω denote the set of all stationary policies.

Similarly, for the subproblems, we refer to a stationary policy of the form $(\hat{\mu}, \hat{\mu}, \dots)$, with $\hat{\mu}: S \times Z \times Y \mapsto \hat{U}$, as policy $\hat{\mu}$. Given a fixed discount factor $\alpha \in [0, 1]$, let $\hat{J}_\infty^{\hat{\mu}}(s, z, y)$ and $\hat{J}_\infty^*(s, z, y)$ be the infinite-horizon expected total discounted cost of policy $\hat{\mu}$ and the corresponding optimal cost, respectively. Let $\hat{J}_T^{\hat{\mu}}(s, z, y)$ be the expected total discounted cost of using the stationary policy $\hat{\mu}$ in a subproblem over a finite horizon of length T . We will still use the definitions introduced in §§2 and 3.2, which have obvious extensions to the infinite-horizon case.

DEFINITION 5.1. For any $s \in S, z \in Z, y \in Y$, let $\hat{U}_\infty^*(s, z, y) \subset \{0, 1\}$ be the set of all decisions that are optimal if an (infinite-horizon) subproblem is found at state (s, z, y) .

The next lemma relates the finite- and infinite-horizon versions of the single-unit, single-customer subproblem.

LEMMA 5.1. *For any fixed $\alpha \in [0, 1]$, and any s, z, y , we have*

$$\lim_{T \rightarrow \infty} \hat{J}_T^*(s, z, y) = \hat{J}_\infty^*(s, z, y).$$

PROPOSITION 5.1. *There exists an optimal policy $\hat{\mu}^*$ for the infinite-horizon single-unit, single-customer subproblem such that $\hat{\mu}^*(s, z, y)$ is a monotonically nonincreasing function of y .*

We now establish that under any optimal subproblem policy, including the optimal subproblem policy $\hat{\mu}^*$ from Proposition 5.1, a unit is never released when the position of the corresponding customer is above a certain threshold. This result is used in the analysis of the average cost case, but is also instrumental in establishing bounds on the running time of the algorithms to be presented in §6.

Without loss of generality, we assume that at the actual stages, the holding cost rates are strictly increasing in the

direction of physical flow. (It is well known that optimal policies for systems where the holding costs are not increasing can be found by studying a related system where they are.) This assumption will remain in effect for the remainder of the paper.

LEMMA 5.2. *There exists a positive integer Y_{\max} with the following property: If $y > Y_{\max}$ and $z \in A'$, then $\hat{U}_t^*(s, z, y) = \hat{U}_\infty^*(s, z, y) = \{0\}$ for all t and s .*

PROPOSITION 5.2. *Let μ^* be the stationary, decoupled policy for the main model that uses the optimal subproblem policy $\hat{\mu}^*$ of Proposition 5.1 for each unit-customer pair. Then there exists a stationary state-dependent echelon base-stock policy that agrees with μ^* at every monotonic state.*

We have so far constructed a stationary monotonic and decoupled policy μ^* . This policy is constructed as a limit of optimal policies for the corresponding finite-horizon problems. It should then be no surprise that μ^* is optimal for the infinite-horizon problem. However, some careful limiting arguments are needed to make this rigorous. This is the subject of the rest of this section.

5.1. Discounted Cost Criterion

In this subsection, we focus on the infinite-horizon expected total discounted cost. In particular, the cost of a stationary policy μ , starting from an initial state $x = \{s, (z^1, y^1), (z^2, y^2), \dots\}$, is defined as

$$J_\infty^\mu(x) = \lim_{T \rightarrow \infty} E \left\{ \sum_{t=0}^{T-1} \alpha^t \cdot g(x_t, \mu(x_t)) \mid x_0 = x \right\},$$

where $\alpha \in [0, 1]$. The infinite-horizon optimal cost is defined by

$$J_\infty^*(x) = \inf_{\mu \in \Omega} J_\infty^\mu(x).$$

We say that a set of policies Ω_0 is optimal (respectively, M-optimal) for the main model if

$$\inf_{\mu \in \Omega_0} J_\infty^\mu(x) = J_\infty^*(x)$$

for all states x (respectively, for all monotonic states x).

A stationary policy can be used over any time horizon, finite or infinite. Let $J_T^\mu(x)$ be the expected total discounted cost of using the stationary policy μ during a finite planning horizon of length T , starting with the initial state $x_0 = x$. We then have

$$J_\infty^\mu(x) = \lim_{T \rightarrow \infty} J_T^\mu(x).$$

Recall that $J_T^*(x)$ is defined as the optimal expected cost with a planning horizon from time 0 until time T , given that $x_0 = x$. Therefore, $J_T^*(x) \leq J_T^\mu(x)$ for any stationary policy μ .

By Proposition 4.3, we have

$$J_T^*(x) = \sum_{i=1}^{\infty} \hat{J}_T^*(s, z^i, y^i)$$

for any monotonic state x . Hence, for every monotonic state x , we have

$$\begin{aligned} J_\infty^*(x) &= \inf_{\mu \in \Omega} \lim_{T \rightarrow \infty} J_T^\mu(x) \\ &\geq \lim_{T \rightarrow \infty} \inf_{\mu \in \Omega} J_T^\mu(x) \\ &\geq \lim_{T \rightarrow \infty} J_T^*(x) \\ &= \lim_{T \rightarrow \infty} \sum_{i=1}^{\infty} \hat{J}_T^*(s, z^i, y^i) \\ &= \sum_{i=1}^{\infty} \lim_{T \rightarrow \infty} \hat{J}_T^*(s, z^i, y^i) \\ &= \sum_{i=1}^{\infty} \hat{J}_\infty^*(s, z^i, y^i), \end{aligned}$$

where the interchange of the limit and the summation is warranted by the monotone convergence theorem because the functions \hat{J}_T^* are nonnegative and monotonically nondecreasing in T .

The above inequality provides a lower bound for the optimal cost. Now consider the decoupled policy μ^* from Proposition 5.2, which uses an optimal subproblem policy $\hat{\mu}^*$ for each unit-customer pair. The cost of μ^* is

$$J_\infty^{\mu^*}(x) = \sum_{i=1}^{\infty} \hat{J}_\infty^{\hat{\mu}^*}(s, z^i, y^i) = \sum_{i=1}^{\infty} \hat{J}_\infty^*(s, z^i, y^i).$$

For a monotonic state x , this is equal to the lower bound, which establishes the M-optimality of μ^* for the main model. This fact leads to the following main result of this section.

THEOREM 5.1. *The set of state-dependent echelon base-stock policies is optimal for the main model under the infinite-horizon discounted cost criterion.*

5.2. Average Cost Criterion

In this subsection, we study the average cost per unit time criterion. The average cost of a policy μ , starting from an initial state $x = \{s, (z^1, y^1), (z^2, y^2), \dots\}$, is defined as

$$\begin{aligned} \lambda^\mu(x) &= \limsup_{T \rightarrow \infty} \frac{1}{T} E \left\{ \sum_{t=0}^{T-1} g(x_t, \mu(x_t)) \mid x_0 = x \right\} \\ &= \limsup_{T \rightarrow \infty} \frac{1}{T} J_T^\mu(x). \end{aligned}$$

The optimal average cost is defined as

$$\lambda^*(x) = \inf_{\mu \in \Omega} \lambda^\mu(x).$$

As before, a set of policies Ω_0 is said to be optimal (respectively, M-optimal) for the main model if

$$\inf_{\mu \in \Omega_0} \lambda^\mu(x) = \lambda^*(x)$$

for every state x (respectively, for all monotonic states x).

For any monotonic state x , we have

$$\begin{aligned} \lambda^*(x) &= \inf_{\mu \in \Omega} \limsup_{T \rightarrow \infty} \frac{1}{T} J_T^\mu(x) \\ &\geq \limsup_{T \rightarrow \infty} \frac{1}{T} \inf_{\mu \in \Omega} J_T^\mu(x) \\ &\geq \limsup_{T \rightarrow \infty} \frac{1}{T} J_T^*(x) \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=1}^{\infty} \hat{J}_T^*(s, z^i, y^i). \end{aligned} \quad (2)$$

The right-hand side of the above inequality is a lower bound on the optimal infinite-horizon average cost. We show (in the online appendix) that the monotonic and decoupled policy μ^* from Proposition 5.2 achieves this lower bound, and is therefore M-optimal. This leads to our main result of this section.

THEOREM 5.2. *The set of state-dependent echelon base-stock policies is optimal for the main model under the infinite-horizon average cost criterion.*

We close this section by providing a characterization of the infinite-horizon average cost. Let \bar{d} be the expected demand per unit time, in steady state. In particular,

$$\bar{d} = \sum_{s \in S} \lim_{t \rightarrow \infty} P(s_t = s) E[d_t | s_t = s].$$

Note that the limits defining the steady-state probabilities $\lim_{t \rightarrow \infty} P(s_t = s)$ exist and are independent of s_0 because we have assumed that s_t is irreducible and aperiodic.

Except for finitely many customers, the expected cost incurred by successive customers is of the form $\hat{J}_\infty^*(s, N + 1, y)$ for ever-increasing values of y . Over a time interval of length T , about $\bar{d} \cdot T$ customers are expected to arrive, suggesting that the average cost per unit time is of the form $\bar{d} \cdot \lim_{y \rightarrow \infty} \hat{J}_\infty^*(s, N + 1, y)$. The proposition that follows shows that the above limit exists and that the above intuition is correct. We will need, however, a minor assumption on the nature of the demand process. We say that the demand process $\{d_t\}$ is of the *lattice* type if there exists an integer $\ell > 1$ such that for every s , the conditional distribution of d_t , given $s_t = s$, is concentrated on the integer multiples of ℓ . Otherwise, we say that $\{d_t\}$ is of the *nonlattice* type.

PROPOSITION 5.3. *Suppose that the demand process is of the nonlattice type. Then,*

(a) *There exists a constant C , such that*

$$\lim_{y \rightarrow \infty} \hat{J}_\infty^*(s, N + 1, y) = C \quad \forall s \in S.$$

(b) *For every state x such that the number of units in locations other than $N + 1$ is finite, we have*

$$\lambda^*(x) = C\bar{d}.$$

In particular, the optimal average cost is the same for all such initial states.

6. Algorithmic Issues

In the preceding sections, we have shown that state-dependent echelon base-stock policies are optimal. In this section, we develop algorithms for determining optimal echelon base-stock levels. The proofs of the results in this section can be found in the online appendix.

While proving the optimality of state-dependent echelon base-stock policies, we established the existence of a monotonic and decoupled policy, which is optimal for the surrogate model and M-optimal for the main model. This policy releases units from an actual stage $z \in A'$ if and only if the position of the corresponding customer is less than or equal to a threshold $y_t^*(s, z)$ and agrees with a state-dependent echelon base-stock policy with base-stock levels $S^{v(z-1)}(s) = y_t^*(s, z) - 1$ at every monotonic state. This state-dependent echelon base-stock policy is optimal for the main model. This means that by solving the single-unit, single-customer subproblem that underlies the monotonic and decoupled policy, we can find the threshold levels $y_t^*(s, z)$, and thereby the optimal base-stock levels. Hence, besides providing a simple proof technique, the decomposition of the problem into single-unit, single-customer subproblems gives rise to efficient algorithms as well. Instead of applying a dynamic programming algorithm on the larger problem involving all units and customers, we can calculate optimal base-stock levels by simply computing an optimal policy for a subproblem involving a single unit-customer pair.

6.1. Subproblem Formulation

The subproblem to be solved is as follows. Given a single unit and a single customer, the goal is to move the unit through the serial system in a way that minimizes the expected total holding costs and backorder costs. There are uncertainties associated with both the time it takes for a released unit to move from a stage to another (lead-time) and with the way the position of the corresponding customer changes (based on the sequence of demand realizations).

We assume that lead times for the overall system follow the extension of Kaplan's stochastic lead-time model given in §3. Recall that in that model, at each time period t , there is a random variable ρ_t^m that determines which outstanding orders will arrive at stage m . More precisely, a unit will arrive at its destination at time $t + 1$ if and only if it was released ρ_t^m or more time units ago. Let $\rho_t = (\rho_t^1, \rho_t^2, \dots, \rho_t^M)$ be the vector of lead-time variables associated with the various stages. The statistics of ρ_t are given in terms of a conditional probability distribution, given the state s_t of the modulating Markov chain:

$$P(\rho_t^m = q | s_t = s), \quad q = 1, \dots, l_m, s \in S, m = 1, \dots, M.$$

An alternative way of describing the statistics of the lead-time model is to give the distribution of the time it takes for

a unit to reach its destination, conditional on the state of the modulating Markov chain when released. Given the distribution of the lead-time random variables ρ_t , it is possible to calculate this actual lead-time distribution, and vice versa.

As stated in §3, we assume that if the unit is in transit, the holding cost rate that is charged is the rate of the destination echelon. In addition, we assume that the sequence of events within a period is as follows. First, the unit may arrive at its destination stage if it was released previously (depending on the previous period's lead-time random variable). The resulting new state is determined and observed. Then, the decision of whether or not to release the unit in this period is made. Finally, the demand and other random variables are realized, and holding and/or backorder costs are charged.

Let (s_t, z_t, y_t) be the state of the subproblem at time t , where s_t is the state of the modulating Markov chain at time t , z_t is the location of the unit at the beginning of period t (after the move of the previous period is completed), and y_t is the position of the customer at the beginning of period t .

Let $u_t \in \{0, 1\}$ be the control variable at time t , where

$$u_t = \begin{cases} 1, & \text{if the unit is released from its current location,} \\ 0, & \text{if the unit is kept at its current location.} \end{cases}$$

Of course, this decision can have an effect only if the location corresponds to an actual stage. Otherwise, we still allow a choice of zero or one for u_t ; however, this choice has no bearing on the evolution of the system or the costs charged. The location of a unit was defined in §3.1 and indicates whether the unit is at an actual stage, or in transit between two actual stages for a specific number of periods, or has been given to a customer (location 0). Recall that A was defined to be the set of locations corresponding to actual stages in the original system, including the outside supplier (location $N + 1$), and A' is the same set with location 1 omitted.

The evolution of the system is affected by a vector $w_t = (d_t, r_t, \bar{s}_t)$ of random variables, whose components are as follows:

- d_t : demand at time t . It has a distribution conditional on the state of the modulating Markov chain. Let D be the largest possible demand in a period. (Let $D = \infty$ if the support of the demand distribution is infinite.)
- r_t : affects the evolution only when the unit is in transit. This random variable takes on the value one if the unit will reach its destination (i.e., moves to $v(z_t - 1)$), or zero if it will stay in transit³ (i.e., moves to $z_t - 1$). Its distribution depends on z_t and s_t . Given the probability distribution (conditioned on s_t) of the lead-time random vector $\rho_t = (\rho_t^1, \rho_t^2, \dots, \rho_t^M)$, the conditional probability distribution of r_t given (s_t, z_t) can easily be calculated.
- \bar{s}_t : a random variable determining the state of the modulating Markov chain in the next period, i.e., $s_{t+1} = \bar{s}_t$.

The dynamics of the system are stationary. Costs are incurred until the state becomes $(s, 0, 0)$ for some s , or until the end of the horizon is reached. States of the form $(s, z, 0)$ with $z > 0$, or $(s, 0, y)$ with $y > 0$ are impossible (“degenerate”) because if the unit is given to a customer, the customer should have received a unit and vice versa. We assume that the initial state is not degenerate, and we will define the system dynamics so that no degenerate state will ever be reached. The state of the system in the next period is given by a mapping $\hat{f}(s_t, z_t, y_t, u_t, w_t) = (s_{t+1}, z_{t+1}, y_{t+1})$ as follows:

1. The dynamics of the modulating Markov chain: $s_{t+1} = \bar{s}_t$ for every t .
2. The evolution of the location of the unit: For every t ,

$$z_{t+1} = \begin{cases} 0, & \text{if } z_t = 0, \\ 1 - \mathbb{1}_{y_t=1}, & \text{if } z_t = 1, \\ (z_t - 1) \cdot (1 - r_t) + v(z_t - 1) \cdot r_t, & \text{if } z_t \geq 2 \text{ and } z_t \notin A, \\ (z_t - u_t) \cdot (1 - r_t) + v(z_t - u_t) \cdot r_t, & \text{if } z_t \geq 2 \text{ and } z_t \in A, \end{cases}$$

where we use the notation $\mathbb{1}_B$ to denote the indicator function of an event B , that is,

$$\mathbb{1}_B = \begin{cases} 1, & \text{if } B \text{ occurs,} \\ 0, & \text{otherwise.} \end{cases}$$

3. The evolution of the position of the customer: For every t ,

$$y_{t+1} = \begin{cases} y_t - d_t, & \text{if } y_t - d_t \geq 1 \text{ and } y_t > 1, \\ 1, & \text{if } y_t - d_t < 1 \text{ and } y_t > 1, \\ 1 - \mathbb{1}_{z_t=1}, & \text{if } y_t = 1, \\ 0, & \text{if } y_t = 0. \end{cases}$$

The one-period costs are stationary and are defined by

$$\hat{g}(s_t, z_t, y_t, u_t) = \hat{h}_{z_t} (1 - \mathbb{1}_{z_t=1} \cdot \mathbb{1}_{y_t=1}) + b \cdot \mathbb{1}_{y_t=1} \cdot (1 - \mathbb{1}_{z_t=1}) + \hat{c}_{z_t} \cdot u_t. \quad (3)$$

We define $\hat{h}_{N+1} = 0$, $\hat{h}_0 = 0$, and $\hat{c}_0 = 0$, and set \hat{h}_i and \hat{c}_i to be the appropriate holding cost rate and order cost rate, respectively, for location i . The holding cost rate for units in artificial locations is assumed to be the rate associated with the downstream actual stage. Order costs for locations other than the actual stages are set to zero. Then, the first term of (3) gives the holding cost, the second term gives the backorder cost, and the third term gives the ordering cost.

6.2. Finite-Horizon Algorithm

We use dynamic programming to compute optimal policies for the subproblem. In particular, the algorithm described below carries out the standard Bellman recursion until all the optimal base-stock levels are determined. It finds threshold levels $y_t^*(s, z)$ for every location corresponding to an actual stage of the system (except for location 1), such that it is optimal to release a unit from that particular location if and only if the position of the corresponding customer is less than or equal to the threshold. The base-stock level of a location $v(z - 1)$ (the location corresponding to the next actual stage after z) is then determined from this threshold.

Input and Output of the Finite-Horizon Algorithm (FHA):

Input:

1. One-period cost function $\hat{g}(s, z, y, u)$ for every (s, z, y, u) .
2. System dynamics function $\hat{f}(s, z, y, u, w)$ for every (s, z, y, u, w) .
3. Conditional probability distribution of $w_t = (d_t, r_t, \bar{s}_t)$ given (s_t, z_t) . (The same distribution for all t .)
4. List of locations corresponding to actual stages, A ; time horizon T .

Output: Threshold values $y_t^*(s, z)$ for every s, t , and $z \in A'$.

Initialization of the Finite-Horizon Algorithm (FHA):

$$\begin{aligned} \hat{J}_T^*(s, z, y) &= \hat{J}_T^*(s, 0, 0) = 0 \quad \forall s, z, y, t, \\ y_t^*(s, z) &= -\infty \text{ and } K_t(s, z) = 0 \quad \forall s, z \in A', t, \\ y &= 0. \end{aligned}$$

The algorithm goes through progressively increasing values of y until all the base-stock levels are determined. In particular, the largest value of y considered by the algorithm is the largest optimal base-stock level over all stages, exogenous states, and time periods. To keep track of which base-stock levels are determined, we use the indicator variables $K_t(s, z)$. At the beginning of the algorithm, $K_t(s, z) = 0$ for all t, s , and $z \in A'$, indicating that none of the base-stock levels is determined yet. During the course of the algorithm, we let $K_t(s, z) = 1$, after the threshold level for location z (and therefore the base-stock level for the actual stage corresponding to location $v(z - 1)$) and exogenous state s at time t is determined.

Let

$$V_t(s, z, y, u) = \hat{g}(s, z, y, u) + E[\hat{J}_{t+1}^*(\hat{f}(s, z, y, u, w))].$$

Recursion of the Finite-Horizon Algorithm (FHA):

while ($K_t(s, z) = 0$ for some $s, z \in A'$, and t) **do**
 $y = y + 1$
for $t = T - 1, \dots, 0$,
for $z = 1, \dots, N + 1$,

for $s = 1, \dots, |S|$,
if $z \in A'$ and $K_t(s, z) = 0$ **then**
 $\hat{J}_t^*(s, z, y) = \min_{u \in \{0, 1\}} V_t(s, z, y, u)$
 $\hat{U}_{T-t}^*(s, z, y) = \{u \mid u \in \arg \min_{u \in \{0, 1\}} V_t(s, z, y, u)\}$
if $\hat{U}_{T-t}^*(s, z, y) = \{1\}$ **then** $y_t^*(s, z) = y$
if $\hat{U}_{T-t}^*(s, z, y) \neq \{1\}$ **then** $K_t(s, z) = 1$
else $J_t^*(s, z, y) = V_t(s, z, y, 0)$
next s
next z
next t
end while

PROPOSITION 6.1. (a) *FHA terminates after at most Y_{\max} iterations of the outer (while) loop.*

(b) *Let $y_t^*(s, z)$ be determined through FHA. A state-dependent echelon base-stock policy with base-stock level $S_t^{v(z-1)}(s) = y_t^*(s, z) - 1$ for each $s, z \in A'$, and t , is optimal for the main model.*

(c) *The complexity of FHA is $O(N \cdot Y_{\max} \cdot \min\{Y_{\max}, D\} \cdot |S|^2 \cdot T)$.*

Note that FHA will terminate when the outer loop reaches the largest base-stock level. In particular, the largest base-stock level or any bound on its size (such as Y_{\max}) does not need to be known a priori. The term Y_{\max} is used only to give a complexity estimate. Moreover, in the complexity estimate, it can be replaced by any a priori known upper bound on the base-stock levels.

Among the inputs to the algorithm, the costs, system dynamics, distribution of demand, and the dynamics of the exogenous Markov chain are quite standard items. The only input that is a little different is the probability distribution of r_t , the random variable that determines whether or not a unit in transit will arrive to its final destination in the current period. In practice, one is more likely to start with an actual lead-time distribution. Assuming that the lead-time distribution is compatible with Kaplan's model (and is not Markov modulated), one can easily recover a compatible distribution for the random variables ρ_t^m , and from these a compatible distribution for the random variables r_t (conditional on z_t). For example, suppose that location z corresponds to actual stage $m + 1 > 1$, and that L_t is the lead time of a unit released from location z at time t . The probability $P(\rho_t^m = 1)$ is immediately determined because it is equal to $P(L_t = 1)$. We then have

$$\begin{aligned} P(L_t = 2) &= P(\rho_t^m \neq 1) \cdot (P(\rho_{t+1}^m = 1) + P(\rho_{t+1}^m = 2)) \\ &= P(\rho_t^m \neq 1) \cdot (P(\rho_t^m = 1) + P(\rho_t^m = 2)), \end{aligned}$$

from which $P(\rho_t^m = 2)$ is determined, and we can continue similarly. The case of Markov-modulated lead times is, however, more complicated.

6.3. Infinite-Horizon Algorithm

We now describe an algorithm for the infinite-horizon problem. The same algorithm is used for both the discounted cost and the average cost criteria.

Let

$$V_\infty(s, z, y, u) = \hat{g}(s, z, y, u) + \alpha \cdot E[\hat{J}_\infty^*(\hat{f}(s, z, y, u, w))],$$

where α is the discount factor (set $\alpha = 1$ if the objective is to minimize the average cost). The input to the algorithm is the same as in the finite-horizon case (except for T). The output is a set of threshold values $y_\infty^*(s, z)$ for every s and every $z \in A'$.

Initialization of the Infinite-Horizon Algorithm (IHA):

$$\begin{aligned} \hat{J}_\infty^*(s, 0, 0) &= 0 \quad \forall s, \\ y_\infty^*(s, z) &= -\infty \text{ and } K_\infty(s, z) = 0 \quad \forall s, z \in A', \\ y &= 0. \end{aligned}$$

Recursion of the Infinite-Horizon Algorithm (IHA):

```

while ( $K_\infty(s, z) = 0$  for some  $(s, z \in A')$ ) do
   $y = y + 1$ 
  for  $z = 1, \dots, N + 1$ ,
    (*) Compute  $\hat{J}_\infty^*(s, z, y)$  for all  $s$ 
    if  $z \in A'$  then
      for  $s = 1, \dots, |S|$ 
         $\hat{U}_\infty^*(s, z, y) = \{u \mid u \in \arg \min_{u \in \{0, 1\}} V_\infty(s, z, y, u)\}$ 
        if  $\hat{U}_\infty^*(s, z, y) = \{1\}$  then  $y_\infty^*(s, z) = y$ 
        if  $\hat{U}_\infty^*(s, z, y) \neq \{1\}$  then  $K_\infty(s, z) = 1$ 
      next  $s$ 
    next  $z$ 
end while

```

PROPOSITION 6.2. (a) IHA terminates after at most Y_{\max} iterations of the outer (while) loop.

(b) Let $y_\infty^*(s, z)$ be determined through IHA. A state-dependent echelon base-stock policy with base-stock level $S_\infty^{v(z-1)}(s) = y_\infty^*(s, z) - 1$ for each s and $z \in A'$ is optimal for the main model.

(c) The complexity of IHA is $O(M \cdot Y_{\max} \cdot |S|^3 + N \cdot Y_{\max} \cdot \min\{Y_{\max}, D\} \cdot |S|^2)$.

We note that the complexity estimate relies on an efficient algorithm for computing $\hat{J}_\infty^*(s, z, y)$ at step (*). This, in turn, relies on the solution, for every (s, z, y) , of an optimal stopping problem with state space cardinality $|S|$, to which an efficient, $O(|S|^3)$, version of policy iteration applies. See the online appendix for the details.

The algorithms reported in this section are fairly efficient in terms of complexity. There is no other work that presents optimal algorithms for a multiechelon inventory control problem with Markov-modulated stochastic lead times and demands, so we cannot provide a direct comparison with existing methods. However, the decomposition of the problem into single unit-customer pairs can be applied to special cases that have been studied before. For example, the serial system with deterministic lead times and i.i.d. demands, i.e., the model of Clark and Scarf (1960), is such a special case, as is the serial system with deterministic lead times and Markovian demands studied in Chen and Song

(2001). Gallego and Zipkin (1999) note that the algorithm of Chen and Zheng (1994), even though originally developed for systems with constant lead times, can be used to find the best base-stock policy⁴ for systems with stochastic lead times similar to the ones studied in this paper (but not Markov modulated). For all of these problems, we are not aware of any methods that are more efficient than the single-unit, single-customer approach. Moreover, our approach possesses some potential advantages. First and foremost, when the support of the demand distribution is infinite, our algorithms do not need to resort to approximations by truncating the distribution at a certain point, whereas existing methods do. (Note that our algorithms do not need to know the probability that the demand is greater than the largest base-stock level, or an a priori upper bound on the optimal base-stock levels.) Second, we provide finite and infinite-horizon algorithms, and our infinite-horizon algorithm can be used to compute optimal policies for both discounted cost and average cost criteria, by simply changing the value of the discount factor. On the contrary, some of the more efficient algorithms for multiechelon systems (such as the algorithm of Chen and Zheng 1994) are only applicable to infinite-horizon average cost problems. Finally, our method is very simple to understand and to implement, because it only involves a problem with a single unit and a single customer.

7. Conclusions

In this paper, we have analyzed an uncapacitated serial inventory system with Markov-modulated demands and Markov-modulated stochastic lead times, in the absence of order crossing. We proved that state-dependent echelon base-stock policies are optimal for both finite and infinite-horizon formulations. We also provided an efficient algorithm to calculate the base-stock levels. All this was done using a different approach than the standard one in the multiechelon inventory control literature, namely, the decomposition into single unit-customer subproblems, as opposed to a stage-by-stage decomposition.

Our approach readily extends to several variations of the problem. For example, the holding, shortage, and cost coefficients could be exogenous, Markov-modulated stochastic processes—in particular, stochastic costs for ordering from the external supplier—can be used to model situations involving fluctuating market prices. Furthermore, for the finite-horizon case, the dynamics do not need to be stationary. Another problem variation involves a backorder cost that accumulates nonlinearly with time or, equivalently, a backorder cost rate that increases with the amount of time a customer has been waiting for a unit. This variation is easily handled as follows: Instead of having a single position (position 1) for backlogged customers, we can have a range of positions (no more than N such positions are needed) that encode the length of time that a customer has been backlogged.⁵

Our approach bears more fruit than was provided in this paper. For example, Muharremoglu (2002) shows that it can be applied to problems with expediting options, problems that incorporate fixed lot sizes between echelons, and to assembly systems. In addition, we believe that this approach can form the basis for obtaining approximate solutions to harder problems (Achy-Brou 2001 discusses some possibilities), in which the structure of the optimal policy is potentially more complicated.

8. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at <http://or.journal.informs.org>.

Endnotes

1. Song and Zipkin (1996b) deal with a single-stage system with Markov-modulated lead times and use a concept similar to the location of a unit in this paper.
2. We allow $S_i^{v(z-1)}(s)$ to take values from the set $\mathbb{N}_0 \cup \{-\infty\} \cup \{\infty\}$, so that the base-stock level can either be a nonnegative integer (the usual case), equal to ∞ (no inventory is kept at z), or equal to $-\infty$ (no units are shipped from z when the Markov chain is in state s).
3. Recall that starting from location z and going forward, $v(z)$ is the first location corresponding to an actual stage (Definition 3.2).
4. Gallego and Zipkin (1999) did not claim that base-stock policies are optimal for such serial systems with stochastic lead times, just that this particular algorithm would find the optimal policy within the class of echelon base-stock policies.
5. We thank a referee for suggesting this variation.

Acknowledgments

This research was supported by the National Science Foundation under grant ACI-9873339. The authors thank the area editor, the associate editor, and the anonymous referees for providing numerous suggestions that improved the exposition of the paper.

A previous version of this paper was titled “Echelon Base Stock Policies in Uncapacitated Serial Inventory Systems.”

References

Achy-Brou, A. 2001. A new approach to multistage serial inventory systems. Master's thesis, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA.

Axsäter, S. 1990. Simple solution procedures for a class of two-echelon inventory problems. *Oper. Res.* **38** 64–69.

Axsäter, S. 1993a. Exact and approximate evaluation of batch-ordering policies for two-level inventory systems. *Oper. Res.* **41** 777–785.

Axsäter, S. 1993b. Optimization of order-up-to- s policies in two echelon inventory systems with periodic review. *Naval Res. Logist.* **40** 245–253.

Bertsekas, D. 1995. *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont, MA.

Beyer, D., S. Sethi. 1997. Average cost optimality in inventory models with Markovian demands. *J. Optim. Theory Appl.* **92** 487–526.

Chen, F., J. S. Song. 2001. Optimal policies for multi-echelon inventory problems with Markov-modulated demand. *Oper. Res.* **49**(2) 226–234.

Chen, F., Y. Zheng. 1994. Lower bounds for multi-echelon stochastic inventory systems. *Management Sci.* **40** 1426–1443.

Cheng, F., S. Sethi. 1999. Optimality of state dependent (s, S) policies in inventory models with Markov-modulated demand and lost sales. *Production Oper. Management* **8** 183–192.

Clark, A., S. Scarf. 1960. Optimal policies for a multi-echelon inventory problem. *Management Sci.* **6** 475–490.

Ehrhardt, R. 1984. (s, s) policies for a dynamic inventory model with stochastic lead times. *Oper. Res.* **32** 121–132.

Federgruen, A., P. Zipkin. 1984. Computational issues in an infinite horizon, multi-echelon inventory model. *Oper. Res.* **32** 818–836.

Gallego, G., P. Zipkin. 1999. Stock positioning and performance estimation in serial production-transportation systems. *Manufacturing Service Oper. Management* **1** 77–88.

Graves, S. 1996. A multi-echelon inventory model with fixed replenishment intervals. *Management Sci.* **42** 1–18.

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

Kaplan, R. S. 1970. A dynamic inventory model with stochastic lead times. *Management Sci.* **16** 491–507.

Muharremoglu, A. 2002. A new perspective on multi-echelon inventory systems. Ph.D. thesis, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA.

Nahmias, S. 1979. Simple approximations or a variety of dynamic lead-time lost-sales inventory models. *Oper. Res.* **27** 904–924.

Parker, R., R. Kapuscinski. 2004. Optimal policies for a capacitated two-echelon inventory system. *Oper. Res.* **52** 739–755.

Rosling, K. 1989. Optimal inventory policies for assembly systems under random demands. *Oper. Res.* **37** 565–579.

Sethi, S., F. Cheng. 1997. Optimality of (s, S) policies in inventory models with Markovian demand. *Oper. Res.* **45** 931–939.

Song, J. S., P. Zipkin. 1992. Evaluation of base-stock policies in multiechelon inventory systems with state-dependent demands. Part I. State-independent policies. *Naval Res. Logist.* **39** 715–728.

Song, J. S., P. Zipkin. 1993. Inventory control in a fluctuating demand environment. *Oper. Res.* **41** 351–370.

Song, J. S., P. Zipkin. 1996a. Evaluation of base-stock policies in multiechelon inventory systems with state-dependent demands. Part II. State-dependent policies. *Naval Res. Logist.* **43** 381–396.

Song, J. S., P. Zipkin. 1996b. Inventory control with information about supply conditions. *Management Sci.* **42** 1409–1419.

Svoronos, A., P. Zipkin. 1991. Evaluation of one-for-one replenishment policies for multi-echelon inventory systems. *Management Sci.* **37** 68–83.

Veinott, A. 1966. The status of mathematical inventory theory. *Management Sci.* **12** 745–777.

Zipkin, P. 1986. Stochastic leadtimes in continuous-time inventory models. *Naval Res. Logist. Quart.* **33** 763–774.