OPERATIONS RESEARCH
Vol. 56, No. 2, March–April 2008, pp. 437–452
ISSN 0030-364X | EISSN 1526-5463 | 08 | 5602 | 0437

DOI 10.1287/opre.1070.0461

Evaluation and Optimization of Installation Base-Stock Policies in Supply Chains with Compound Poisson Demand

Yao Zhao
Department of Management Science and Information Systems, Rutgers University,
Newark, New Jersey 07102, yaozhao@andromeda.rutgers.edu

In this paper, we establish an exact framework for a class of supply chains with at most one directed path between every two stages. External demands follow compound Poisson processes, the transit times are stochastic, sequential, and exogenous, and each stage controls its inventory by an installation base-stock policy under continuous review. Unsatisfied demand at each stage is fully backordered. This class of supply chains includes assembly, distribution, tree, and two-level general networks as special cases. We characterize the stockout delay for each unit of demand at each stage of the supply chain by developing an exact and unified approach that applies to various network topologies. We also present tractable approximations and decompositions that facilitate efficient evaluation and optimization (up to the approximations) of the base-stock policies in industry-size problems with a tree structure. We demonstrate the effectiveness of the solution by numerical studies.

Subject classifications: inventory/production: multi-item/echelon; operating characteristics; approximations; lead times; stochastic.
Area of review: Manufacturing, Service, and Supply Chain Operations.
History: Received January 2006; revisions received June 2006, January 2007, March 2007; accepted April 2007. Published online in *Articles in Advance* December 17, 2007.

1. Introduction

We consider a class of supply chains where there is at most one directed path between every two stages. A stage refers to a unique combination of a facility and a product. A path consists of arcs, each of which represents a direct demand and supply relationship between two stages. Each stage controls its inventory by an (installation) continuous-time base-stock policy. Unsatisfied demand is fully backordered. Our objective is twofold: (i) establishing an exact framework for these supply chains facing compound Poisson demand and stochastic sequential transit times, and (ii) developing numerically tractable approximations that allow efficient computation of the optimal or near-optimal base-stock levels.

This paper is an extension of Simchi-Levi and Zhao (2005) to supply chains with compound Poisson demand and a more general network structure. Below, we summarize works most related to this paper on modeling assumptions and solution approaches. We refer to Zipkin (2000), Graves and Willems (2003), and Simchi-Levi and Zhao (2005) for recent reviews of the related problems, motivations, and solution methods.

Exact analysis is provided for various serial and distribution systems facing compound Poisson demand; see Zipkin (1991), Forsberg (1995), Chen (1998), Axsater (2000, 2003), and references therein. Assembly systems are analytically more challenging than distribution systems because of the interaction among components. For assemble-to-order (ATO) systems, exact analyses and approximations are provided for models with either constant lead times (Song 1998, 2002; Hausman et al. 1998; Agrawal and Cohen 2001), i.i.d. lead times (Song and Yao 2002, Lu et al. 2003, Lu and Song 2005, Cheung and Hausman 1995), or stochastic and sequential lead times (Zhao and Simchi-Levi 2006). Song and Zipkin (2003) provides an excellent review.

Multilevel assembly networks with multiple endproducts and common components are more difficult to analyze than the assemble-to-order (ATO) systems because the lead time at each internal stage is endogenous, which depends on the supplying stages' fulfillment processes and their interactions. Literature in this area is limited.

De Kok and Visschers (1999) considers such a network with constant transit times and i.i.d. demand. They developed an algorithm based on Rosling (1989) to transform the network into a purely divergent system. Note that in this approach, the allocation of common components is determined before the components actually arrive in inventory. Lee and Billington (1993) considers the deskjet printer supply chain, which includes both assembly and distribution operations. Each stage controls its inventory by an installation periodic-review base-stock policy. The authors developed simple approximations for performance evaluation. For assembly systems, it is assumed that at most one supplier can be out of stock in any period. The same approximation is utilized by Ettl et al. (2000), which develops a model to evaluate and optimize supply chains with compound Poisson demand and i.i.d. transit times. Simchi-Levi and Zhao (2005) considers tree structure supply chains with stochastic sequential transit times. The authors derived sample path-based recursive equations and provided exact analyses for various network topologies facing Poisson demand. Approximations and decomposition techniques are developed for fast computation. However, the exact analysis on assembly systems is limited to the pure assembly systems.

Compound Poisson demand introduces additional challenges to the exact analysis of the multilevel assembly networks. For example, different units of a demand face statistically different stockout delays at each stage (Zipkin 1991, p. 405; Simchi-Levi and Zhao 2005). In addition, the intricate interaction and dependence among components are now driven not only by the common demand interarrival times, but also by the common demand sizes.

This paper provides an exact framework for a class of supply chains with compound Poisson demand and stochastic sequential transit times, where there is at most one directed path between each pair of stages. An important special case is the two-level general network with multiple end products and common components. We develop a unified approach to characterize the stockout delay for each unit of demand at each stage of the supply chain. In §2, we present an exact analysis for the two-level general network. In principle, the analysis extends to multilevel systems. In §3, we provide numerically tractable approximations (based on the exact analysis), which allow efficient optimization of tree structure supply chains subject to type 2 fill rate constraints within a certain committed service time. In §4, we demonstrate in various numerical examples the accuracy and efficiency of the approximation and the quality of the solution. We conclude the paper in §5 by identifying a few research directions.

Our approach is built on the flow unit method; see, e.g., Zipkin (1991), Axsater (1990), and Zhao and Simchi-Levi (2006). The key idea is that for each unit of an external demand, we identify, at each stage, not only the timing at which the corresponding order is placed, but also the index of the item in the corresponding order that satisfies this demand unit. Applying this idea to each stage leads to an exact and systematic way to characterize the system performance. However, the challenge lies in the complex dependencies among timings and indices in assembly networks due to the common demand size and interarrival time processes. Indeed, the joint probability distribution of the timings in the multiproduct ATO systems alone is an open question (Zhao and Simchi-Levi 2006). The approach here differs from that of Poisson demand (see, e.g., Simchi-Levi and Zhao 2005) because in the latter all demand units face statistically the same stockout delay, and one only needs to identify the timing (but not the index) of the corresponding order.

Although we provide a unified method for the supply chains with a quite general structure, there exist more efficient exact methods for two-level pure distribution systems, see, e.g., Forsberg (1995). A comparison between our approach and Forsberg (1995) is provided in §2.3.

2. Model and Analysis

We assume that each stage utilizes an installation continuous-time base-stock policy with a nonnegative base-stock level. We call the processing cycle time at each stage and the transportation lead time between every two stages the "transit times," and assume that they are stochastic, sequential, and exogenous (the "transit time" assumption by Svoronos and Zipkin 1991). External demands follow independent compound Poisson processes. The demand process faced by each internal stage is determined by the bill of materials, and is compound Poisson due to the continuous-time base-stock policy. Each stage fills its demand on a first-come, first-serve (FCFS) basis, and each individual demand unit can be satisfied separately. Unsatisfied demand is fully backordered. We assume that one unit of the final item at each stage requires one unit of each input item (see §5 for more discussion). Lastly, the service requirement at a stage can be specified by a type 2 fill rate (the fraction of demand satisfied) within a committed service time.

We map the supply chain into a graph $(\mathcal{N}, \mathcal{A})$ with the node set $\mathcal{N}$ and the arc set $\mathcal{A}$. A node represents a stage in the supply chain, and is denoted by $1, \dots, K$. An arc represents a pair of nodes $i, k \in \mathcal{N}$ that have a direct demand and supply relationship, and is denoted by $(i, k) \in \mathcal{A}$. It is convenient to assign an index n to each unit of a demand faced by a node, so that the smaller the n, the higher the priority of the demand unit. We define the following notation,

- $X_k(n)$: stockout delay for the nth unit of a demand at node k.
- $W_k(n)$: inventory holding time of the corresponding item that satisfies the nth unit of a demand at stage k.
- $\tilde{L}_k(n)$: total replenishment lead time for the nth unit of an order placed by node k.
- L_k: processing cycle time at node k.
- $L_{i,k}$: transportation lead time from node i to k, $(i, k) \in \mathcal{A}$.
- h_k: inventory holding cost per item per time unit at node k.
- s_k: base-stock level at node k.
- λ_k: customer arrival rate at node k.
- D_k: demand size at node k. D_k is an integer-valued positive random variable.

If node k faces external demand, we define τ_k and β_k to be the committed service time and the type 2 fill rate, respectively. Among these parameters, L_k, $L_{i,k}$, λ_k, D_k, τ_k, and β_k are inputs; s_k or $X_k(n)$ are decision variables. Let U_k be the maximum value of D_k. Following convention, we denote $a^+ = \max\{a, 0\}$, and let $E(\cdot)$, $V(\cdot)$, and $\sigma(\cdot)$ be the mean, variance, and standard deviation of a random variable, respectively.

2.1. Analysis of a Single Stage

Consider a node $k \in \mathcal{N}$ and suppose that a demand arrives at time t. We ask the following two questions: (1) When is the corresponding order placed at node k that satisfies the nth unit of this demand, where $n \geqslant 1$? (2) What is the index of the item in the corresponding order that satisfies the nth unit of this demand? Whereas one only needs to answer question (1) for Poisson demand, we have to answer both questions for compound Poisson demand.

We first define the following notations with respect to time t. Let $D_{k,1}, D_{k,2}, \ldots,$ be the sizes of demands that arrive prior to t, where $D_{k,1}$ is the size of the most recent demand prior to t, and so on. In a similar vein, let $\nu_{k,1}, \nu_{k,2}, \ldots,$ be the demand interarrival times prior to t, where $\nu_{k,1}$ is the time between the most recent demand and t, and so on. Figure 1 provides a visual aid.

Because of the continuous-review base-stock policy, each order is triggered by a demand. For any $n \geqslant 1$, we define $J_k(n)$ such that the corresponding order that satisfies the nth unit of the demand (at t) is placed at time $t - T_k(J_k(n))$ (see Figure 1), where

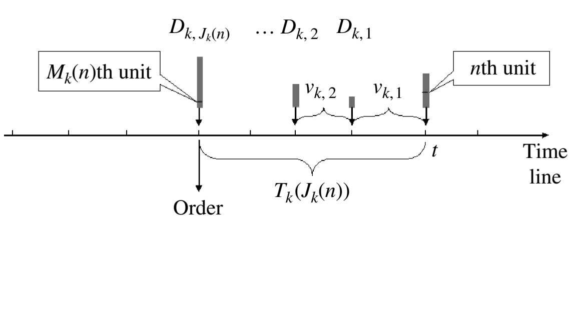
$$T_k(J_k(n)) = \sum_{j=1}^{J_k(n)} \nu_{k,j}. \tag{1}$$

We also define $M_k(n)$ to be the index of the item in the corresponding order that satisfies the nth unit of the demand (at t).

Clearly, if $n > s_k$, then the corresponding order for the nth unit of the demand (at t) must be placed at time t, i.e., $J_k(n) = 0$. This is true because the inventory position, s_k, is just enough to cover the s_kth unit of the demand. To answer the second question, we first note that there are s_k items on-hand or incoming right before t. By the transit time assumption, the nth unit of the demand is satisfied by the $M_k(n)$th item in the corresponding order, where $M_k(n) = n - s_k$.

If $n \leqslant s_k$, but $n + D_{k,1} > s_k$, then the corresponding order for the nth unit of the demand (at t) must be placed at time $t - \nu_{k,1}$, i.e., $J_k(n) = 1$. This is true because $n \leqslant s_k$ implies that the inventory position right before t is enough to cover the nth unit; thus, the corresponding order must be placed at or before $t - \nu_{k,1}$. On the other hand, $n + D_{k,1} > s_k$ implies that the inventory position right before $t - \nu_{k,1}$ is not sufficient to cover the nth unit; hence, the corresponding order must be placed at or after $t - \nu_{k,1}$.

Figure 1. The time line of a single-stage system.



To identify the item in the corresponding order that satisfies the nth unit, we combine $D_{k,1}$ with the demand at t. Then, the nth unit in the demand at t is the $(D_{k,1} + n)$th unit in the combined demand (due to FCFS). Because there are s_k items on-hand and incoming right before $t - \nu_{k,1}$, the nth unit in the demand (at t) must be satisfied by the $M_k(n)$th item in the corresponding order, where $M_k(n) = D_{k,1} + n - s_k$.

More generally, if $n + D_{k,1} + \cdots + D_{k,j-1} \leqslant s_k$ but $n + D_{k,1} + \cdots + D_{k,j} > s_k$ for $j = 2, 3, \ldots, s_k$, then (1) the corresponding order for the nth unit of the demand (at t) must be placed at time $t - \nu_{k,1} - \cdots - \nu_{k,j}$, i.e., $J_k(n) = j$; and (2) the nth unit of the demand (at t) is satisfied by the $M_k(n)$th item in the corresponding order, where

$$M_k(n) = D_{k,1} + \cdots + D_{k,j} + n - s_k. \tag{2}$$

$J_k(n)$ is related to the *renewal process* $\{N_k(\cdot)\}$, generated by the demand-size process $\{D_{k,j},\ j \geqslant 1\}$, as follows:

$$J_k(n) = \begin{cases} 0 & \text{if } n > s_k, \\ N_k(s_k - n) + 1 & \text{otherwise,} \end{cases} \tag{3}$$

$J_k(n) \in \{0, 1, 2, \ldots, s_k - n + 1\}$. $N_k(s)$ is the number of customer arrivals before the cumulative demand exceeds s.

$M_k(n)$ is related to the *remaining life process* $\{O_k(\cdot)\}$ associated with $\{N_k(\cdot)\}$ (see Kulkarni 1995, p. 433, for a definition). By Equation (2),

$$M_k(n) = \begin{cases} n - s_k & \text{if } n > s_k, \\ O_k(s_k - n) & \text{otherwise,} \end{cases} \tag{4}$$

$M_k(n) \in \{1, 2, \ldots, U_k\}$. $O_k(s)$ is the surplus of the cumulative demand over s when the cumulative demand first exceeds s.

$J_k(n)$ and $M_k(n)$ are generally dependent random variables. Given n, the joint probability mass function of $J_k(n)$ and $M_k(n)$ is given by

$$\Pr\{J_k(n) = 0, M_k(n) = m\} = 1_{\{n > s_k,\, m = n - s_k\}}, \tag{5}$$

$$\begin{aligned} &\Pr\{J_k(n) = j, M_k(n) = m\} \\ &\quad = \Pr\{N_k(s_k - n) = j - 1, O_k(s_k - n) = m\} \\ &\quad = \Pr\left\{\sum_{j=1}^{j} D_{k,j} = s_k - n + m, D_{k,j} \geqslant m\right\} \quad \text{for } j \geqslant 1. \end{aligned} \tag{6}$$

If $n \leqslant s_k$, $\Pr\{\sum_{j=1}^{j} D_{k,j} = s_k - n + m, D_{k,j} \geqslant m\} = \Pr\{D_{k,1} = s_k - n + m\}$ for $j = 1$ and $m = 1, \ldots, (U_k - s_k + n)^+$; $\Pr\{\sum_{j=1}^{j} D_{k,j} = s_k - n + m, D_{k,j} \geqslant m\} = \sum_{l=j-1}^{s_k-n} \Pr\{D_{k,1} + \cdots + D_{k,j-1} = l\} \Pr\{D_{k,j} = s_k - n - l + m\}$ for $j = 2, \ldots, s_k - n + 1$ and $m = 1, \ldots, \min\{jU_k - s_k + n, U_k\}^+$. In the special case of Poisson demand, Equations (3) and (4) imply that $\Pr\{J_k(n) = s_k\} = 1$ and $\Pr\{M_k(n) = 1\} = 1$, which is consistent with Simchi-Levi and Zhao (2005).

The stockout delay for the nth unit of a demand and the inventory holding time for the corresponding item that satisfies this unit are given by

$$X_k(n) = [\widetilde{L}_k(M_k(n)) - T_k(J_k(n))]^+, \tag{7}$$

$$W_k(n) = [T_k(J_k(n)) - \widetilde{L}_k(M_k(n))]^+. \tag{8}$$

Clearly, $T_k(j)$ is determined by the demand arrival process; $J_k(n)$ and $M_k(n)$ are determined by the demand size process. Unlike supply chains with Poisson demand (see Simchi-Levi and Zhao 2005, Proposition 3.9), $\widetilde{L}_k(M_k(n))$ now depends on $T_k(J_k(n))$ because $M_k(n)$ depends on $J_k(n)$.

The total replenishment lead time, $\widetilde{L}_k(\cdot)$, depends on the stockout delay(s) at all immediate supplier(s) of node k. Clearly, $\widetilde{L}_k(\cdot)$ depends on the network topology, which is analyzed in the following section.

2.2. Two-Level General Networks

Consider node sets $\mathcal{J}$ and $\mathcal{K}$, where each node in $\mathcal{J}$ supplies one or more nodes in $\mathcal{K}$, and each node in $\mathcal{K}$ is supplied by one or more nodes in $\mathcal{J}$. Nodes in $\mathcal{K}$ face independent compound Poisson demand. Each node in $\mathcal{J}$ faces an exogenous lead time, which may vary (statistically) across different units in an order. This assumption facilitates the extension of the analysis here to multilevel assembly networks. Within each set, there is no supply-demand relationship between every two nodes. Such a network is called a *two-level general network* (Lesnaia 2004). Clearly, there is at most one directed path between every two nodes in this network. Figure 2 gives two examples where system (a) has a tree structure but system (b) does not. In what follows, we first analyze these examples in detail, and then provide expressions in compact form for the general system.

Figure 2(a). We only analyze node 2 because node 1 represents a simpler case. For the nth unit of a demand that arrives at time t, $X_2(n)$ and $W_2(n)$ are given by Equations (7) and (8), where the joint distribution of $J_2(n)$ and $M_2(n)$ is characterized by Equations (5) and (6). Conditioning on $J_2(n) = j$ and $M_2(n) = m$, $T_2(j)$ follows Erlang distribution and

$$\widetilde{L}_2(m) = \max\{X_3(m) + L_{3,2}, X_4(m) + L_{4,2}\} + L_2,$$

where $L_{3,2}$, $L_{4,2}$, and L_2 are exogenous, $X_i(m)$ is the stockout delay for the mth unit of the demand received by node $i = 3, 4$ at time $t - T_2(J_2(n))$ (see Figure 3). $X_i(m) = [\widetilde{L}_i(M_i(m)) - T_i(J_i(m))]^+$ for $i = 3, 4$.

Figure 2. Examples of the two-level general network.

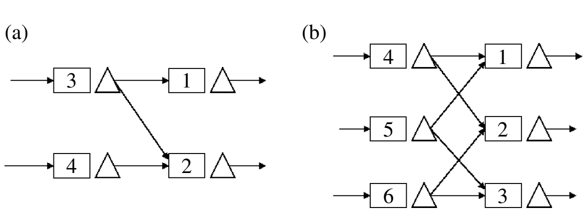
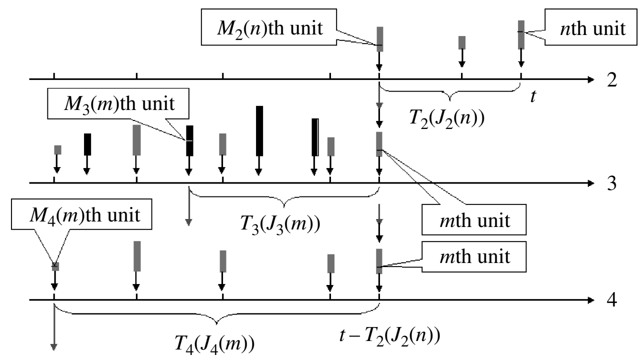


Figure 3. The time line of Figure 2(a).



Note that $T_2(\cdot)$ does not overlap with $T_i(\cdot)$ for $i = 3, 4$. It follows from the assumptions of compound Poisson demand and "transit time" that one can exactly evaluate system performance by first computing $\widetilde{L}_2(m)$ for all m, and then computing $X_2(n)$ and $W_2(n)$.

To characterize $\widetilde{L}_2(m)$, we consider the maximum of $X_i(m) + L_{i,2}$ for $i = 3, 4$, where $X_i(m)$, $i = 3, 4$, are dependent random variables because both of them depend on the demand process from node 2 (see Figure 3). To characterize the joint probability distribution of $X_3(m)$ and $X_4(m)$, we need to determine the joint probability distribution of $(J_i(m), M_i(m), T_i; i = 3, 4)$. Define the shorthand notation $p(j_i, m_i, t_i; i = 3, 4) = \Pr\{J_i(m) = j_i,\ M_i(m) = m_i,\ T_i(j_i) \geqslant t_i; i = 3, 4\}$. Because Equation (5) is an indicator function for $j_i = 0$, we consider $j_i > 0$ for $i = 3, 4$. By Equation (6),

$$p(j_i, m_i, t_i; i = 3, 4)$$
$$= \Pr\left\{\sum_{j=1}^{j_3} D_{3,j} = s_3 - m + m_3,\ D_{3,j_3} \geqslant m_3,\ T_3(j_3) \geqslant t_3;\right.$$
$$\left.\sum_{j=1}^{j_4} D_{4,j} = s_4 - m + m_4,\ D_{4,j_4} \geqslant m_4,\ T_4(j_4) \geqslant t_4\right\}, \tag{9}$$

where $D_{3,j}$ and $D_{4,j}$ are defined in the same way as $D_{k,j}$ in §2.1, but now with respect to time $t - T_2(J_2(n))$. $(J_i(m), M_i(m); i = 3, 4)$ are not independent of $(T_i, i = 3, 4)$ because all of them depend on the demand arrival processes.

Note that node 3 faces demand from both nodes 1 and 2, and node 4 faces demand only from node 2. Define $D_{1,j}$ and $D_{2,j}$ in the same way as $D_{k,j}$ in §2.1, but with respect to time $t - T_2(J_2(n))$, we can rewrite Equation (9) as follows:

$$p(j_i, m_i, t_i; i = 3, 4)$$
$$= \Pr\left\{\sum_{j=1}^{J_3^1} D_{1,j} + \sum_{j=1}^{J_3^2} D_{2,j} = s_3 - m + m_3,\ D_{\delta_3, J_3^{\delta_3}} \geqslant m_3,\right.$$
$$T_3(j_3) \geqslant t_3; \sum_{j=1}^{j_4} D_{2,j} = s_4 - m + m_4,$$
$$\left. D_{2,j_4} \geqslant m_4,\ T_4(j_4) \geqslant t_4\right\}, \tag{10}$$

where J_3^1 (J_3^2) is the number of node 1 (node 2) demands in the last j_3 demands at node 3, and δ_3 represents the node from which the last j_3th demand comes. Clearly, J_3^1 and J_3^2 are nonnegative integer random variables satisfying

$$J_3^1 + J_3^2 = j_3,$$

and δ_3 chooses a value of either one or two.

Enumerating all possible outcomes of (J_2^1, J_3^2, δ_3) and denoting them by (j_3^1, j_3^2, l), we arrive at

$$\begin{aligned} &p(j_i, m_i, t_i; i=3,4) \\ &= \sum_{\forall (j_3^1, j_3^2, l)} \Pr\left\{ \sum_{j=1}^{j_3^1} D_{1,j} + \sum_{j=1}^{j_3^2} D_{2,j} = s_3 - m + m_3, D_{l, j_3^l} \geqslant m_3; \right. \\ &\qquad \left. \sum_{j=1}^{j_4} D_{2,j} = s_4 - m + m_4, D_{2, j_4} \geqslant m_4 \right\} \\ &\quad \times \Pr\{T_3(j_3) \geqslant t_3, J_3^1 = j_3^1, J_3^2 = j_3^2, \delta_3 = l; T_4(j_4) \geqslant t_4\}. \end{aligned} \tag{11}$$

The equality holds here because conditioning on $(J_2^1, J_3^2, \delta_3) = (j_3^1, j_3^2, l)$, the arguments on demand sizes are independent of those on demand interarrival times.

The probability on demand sizes in Equation (11) can be evaluated exactly by identifying the common demand sizes shared by nodes 3 and 4; see the appendix for details. To compute the probability on demand interarrival times, we first note that the event $\{J_3^1 = j_3^1, J_3^2 = j_3^2, \delta_3 = 1\}$ means that among the last j_3 demand arrivals at node 3, j_3^1 of them are from node 1, and the j_3th demand is also from node 1. Hence, the event $\{T_3(j_3) \geqslant t_3, J_3^1 = j_3^1, J_3^2 = j_3^2, \delta_3 = 1\}$ for $j_3^1 > 0$ and $j_3^2 \geqslant 0$ is equivalent to the event $\{\sum_{j=1}^{j_3^1} \nu_{1,j} \geqslant t_3, \sum_{j=1}^{j_3^2} \nu_{2,j} < \sum_{j=1}^{j_3^1} \nu_{1,j} < \sum_{j=1}^{j_3^2+1} \nu_{2,j}\}$, where $\nu_{k,j}$ are the demand interarrival times at node k ($k = 1, 2$). Hence,

$$\Pr\{T_3(j_3) \geqslant t_3, J_3^1 = j_3^1, J_3^2 = j_3^2, \delta_3 = l; T_4(j_4) \geqslant t_4\}$$

$$= \begin{cases} \Pr\left\{ \sum_{j=1}^{j_3^1} \nu_{1,j} \geqslant t_3, \sum_{j=1}^{j_3^2} \nu_{2,j} < \sum_{j=1}^{j_3^1} \nu_{1,j} < \sum_{j=1}^{j_3^2+1} \nu_{2,j}; \sum_{j=1}^{j_4} \nu_{2,j} \geqslant t_4 \right\} & \text{if } l = 1, \\ \Pr\left\{ \sum_{j=1}^{j_3^2} \nu_{2,j} \geqslant t_3, \sum_{j=1}^{j_3^1} \nu_{1,j} < \sum_{j=1}^{j_3^2} \nu_{2,j} < \sum_{j=1}^{j_3^1+1} \nu_{1,j}; \sum_{j=1}^{j_4} \nu_{2,j} \geqslant t_4 \right\} & \text{if } l = 2, \end{cases} \tag{12}$$

where $j_3^1 > 0, j_3^2 \geqslant 0$ for $l = 1$, and $j_3^1 \geqslant 0$, $j_3^2 > 0$ for $l = 2$. These probabilities can be evaluated exactly by conditioning on the common Erlang random variables; see the appendix.

Given $p(j_i, m_i, t_i; i = 3, 4)$, we can express $\Pr\{X_2(n) \leqslant x\}$ as follows:

$$\Pr\{X_2(n) \leqslant x\} = \sum_{m,j} \Pr\{M_2(n) = m, J_2(n) = j\} \cdot \Pr\{\tilde{L}_2(m) - T_2(j) \leqslant x\}. \tag{13}$$

Conditioning on $T_2(j) = t_2$ and letting $x_3 = x + t_2 - L_2 - L_{3,2}$ and $x_4 = x + t_2 - L_2 - L_{4,2}$, then for $x_3 \geqslant 0$ and $x_4 \geqslant 0$,

$$\begin{aligned} &\Pr\{\tilde{L}_2(m) \leqslant x + t_2\} \\ &\quad = \Pr\{X_3(m) \leqslant x_3, X_4(m) \leqslant x_4\} \\ &\quad = \Pr\{T_3(J_3(m)) \geqslant \tilde{L}_3(M_3(m)) - x_3, \\ &\qquad\quad T_4(J_4(m)) \geqslant \tilde{L}_4(M_4(m)) - x_4\} \\ &\quad = \sum_{j_i, m_i; i=3,4} p(j_i, m_i, \tilde{L}_i(m_i) - x_i; i = 3, 4). \end{aligned} \tag{14}$$

Figure 2(b). This system differs from that in Figure 2(a) because each of the nodes 1, 2, and 3 is supplied by two nodes, and each of these two supplying nodes serves a different additional node. Below, we analyze node 1. The analysis of other nodes is analogous.

Consider the nth unit of a demand at time t. Conditioning on $M_1(n) = m$, then $\tilde{L}_1(m) = \max\{X_4(m) + L_{4,1}, X_5(m) + L_{5,1}\} + L_1$, where $X_4(m)$ and $X_5(m)$ are dependent due to the common demand process from node 1. To characterize the joint probability distribution of $X_4(m)$ and $X_5(m)$, we determine the joint probability distribution of $(J_i(m), M_i(m), T_i; i = 4, 5)$. Consider $j_i > 0$, $i = 4, 5$, and let $p(j_i, m_i, t_i; i = 4, 5) = \Pr\{J_i(m) = j_i, M_i(m) = m_i, T_i(j_i) \geqslant t_i; i = 4, 5\}$. Note that apart from the demand of node 1, each of nodes 4 and 5 serves an independent demand from nodes 2 and 3, respectively. By a similar analysis of Equation (10), we arrive at

$$\begin{aligned} &p(j_i, m_i, t_i; i = 4, 5) \\ &\quad = \Pr\left\{ \sum_{j=1}^{J_4^1} D_{1,j} + \sum_{j=1}^{J_4^2} D_{2,j} = s_4 - m + m_4, D_{\delta_4, J_4^{\delta_4}} \geqslant m_4, \right. \\ &\qquad T_4(j_4) \geqslant t_4; \sum_{j=1}^{J_5^1} D_{1,j} + \sum_{j=1}^{J_5^3} D_{3,j} = s_5 - m + m_5, \\ &\qquad \left. D_{\delta_5, J_5^{\delta_5}} \geqslant m_5, T_5(j_5) \geqslant t_5 \right\}, \end{aligned} \tag{15}$$

where (J_4^1, J_4^2) and (J_5^1, J_5^3) are nonnegative integer random variables satisfying

$$J_4^1 + J_4^2 = j_4, \qquad J_5^1 + J_5^3 = j_5.$$

δ_4 chooses a value of either one or two, and δ_5 chooses a value of either one or three.

Enumerating all possible outcomes of $(J_4^1, J_4^2, \delta_4; J_5^1, J_5^3, \delta_5) = (j_4^1, j_4^2, l_4; j_5^1, j_5^3, l_5)$, Equation (15) becomes

$$\begin{aligned} &p(j_i, m_i, t_i; i = 4, 5) \\ &= \sum_{\forall (j_4^1, j_4^2, l_4; j_5^1, j_5^3, l_5)} \Pr\left\{ \sum_{j=1}^{j_4^1} D_{1,j} + \sum_{j=1}^{j_4^2} D_{2,j} = s_4 - m + m_4, \right. \\ &\qquad \left. D_{l_4, j_4^{l_4}} \geqslant m_4; \sum_{j=1}^{j_5^1} D_{1,j} + \sum_{j=1}^{j_5^3} D_{3,j} = s_5 - m + m_5, D_{l_5, j_5^{l_5}} \geqslant m_5 \right\} \\ &\quad \times \Pr\{T_4(j_4) \geqslant t_4, J_4^1 = j_4^1, J_4^2 = j_4^2, \delta_4 = l_4; \\ &\qquad T_5(j_5) \geqslant t_5, J_5^1 = j_5^1, J_5^3 = j_5^3, \delta_5 = l_5\}. \end{aligned} \tag{16}$$

As for Figure 2(a), we can compute the probability on demand sizes by identifying the common demand sizes shared by nodes 4 and 5. Details are omitted.

The probability on demand interarrival times can be characterized in a way similar to Equation (12). For $(\delta_4, \delta_5) = (1, 1)$,

$$\Pr\{T_4(j_4) \geqslant t_4, J_4^1 = j_4^1, J_4^2 = j_4^2, \delta_4 = 1; T_5(j_5) \geqslant t_5, J_5^1 = j_5^1, J_5^3 = j_5^3, \delta_5 = 1\}$$
$$= \Pr\Biggl\{\sum_{j=1}^{j_4^1} \nu_{1,j} \geqslant t_4, \sum_{j=1}^{j_4^2} \nu_{2,j} < \sum_{j=1}^{j_4^1} \nu_{1,j} < \sum_{j=1}^{j_4^2+1} \nu_{2,j}; \sum_{j=1}^{j_5^1} \nu_{1,j} \geqslant t_5, \sum_{j=1}^{j_5^3} \nu_{3,j} < \sum_{j=1}^{j_5^1} \nu_{1,j} < \sum_{j=1}^{j_5^3+1} \nu_{3,j}\Biggr\}. \tag{17}$$

This probability can be evaluated exactly as follows: Conditioning on the common Erlang random variables $\sum_{j=1}^{j_4^1} \nu_{1,j} = \tau_4$ and $\sum_{j=1}^{j_5^1} \nu_{1,j} = \tau_5$, the conditional probability reduces to the product of two Poisson probabilities $\Pr\{\sum_{j=1}^{j_4^2} \nu_{2,j} < \tau_4 < \sum_{j=1}^{j_4^2+1} \nu_{2,j}\} \Pr\{\sum_{j=1}^{j_5^3} \nu_{3,j} < \tau_5 < \sum_{j=1}^{j_5^3+1} \nu_{3,j}\}$. A similar logic applies to the cases of $(\delta_4, \delta_5) = (1,3), (2,1),$ and $(2,3)$.

Given $p(j_i, m_i, t_i; i = 4, 5)$, one can express $\Pr\{X_1(n) \leqslant x\}$ in the same way as that of Figure 2(a). For brevity, we shall not repeat.

The General System. Essentially, the same analysis extends to any two-level general network. Let $[a_{i,k}]$ be the bill of material (BOM) matrix, i.e., assembling one unit at node k requires $a_{i,k}$ units from node i. By the unit flow assumption, $a_{i,k} = 0$ or 1. Let $\mathcal{I}_k = \{i \in \mathcal{I} \mid a_{i,k} = 1\}$ be the supplier set of node k, and $\mathcal{K}_i = \{k \in \mathcal{K} \mid a_{i,k} = 1\}$ be the customer set of node i.

For any node $k \in \mathcal{K}$, consider the nth unit of a demand that arrives at time t. $X_k(n)$ and $W_k(n)$ are given by Equations (7) and (8), where

$$\tilde{L}_k(m) = \max_{i \in \mathcal{I}_k} \{X_i(m) + L_{i,k}\} + L_k \quad \text{for any } m, \tag{18}$$

$$X_i(m) = [\tilde{L}_i(M_i(m)) - T_i(J_i(m))]^+, \quad i \in \mathcal{I}_k. \tag{19}$$

Clearly, the $X_i(m)$s are dependent due to the common demand processes shared by nodes $i \in \mathcal{I}_k$. We now derive compact form for the joint probability of $(J_i, M_i, T_i; i \in \mathcal{I}_k)$. Let $j_i > 0$, $i \in \mathcal{I}_k$,

$$p(j_i, m_i, t_i; i \in \mathcal{I}_k)$$
$$= \Pr\{J_i(m) = j_i, M_i(m) = m_i, T_i(j_i) \geqslant t_i; i \in \mathcal{I}_k\}$$
$$= \Pr\Biggl\{\sum_{j=1}^{j_i} D_{i,j} = s_i - m + m_i, D_{i,j_i} \geqslant m_i, T_i(j_i) \geqslant t_i; i \in \mathcal{I}_k\Biggr\}$$
$$= \Pr\Biggl\{\sum_{l \in \mathcal{K}_i} \sum_{j=1}^{J_i^l} D_{l,j} = s_i - m + m_i, D_{\delta_i, J_i^{\delta_i}} \geqslant m_i, T_i(j_i) \geqslant t_i; i \in \mathcal{I}_k\Biggr\}, \tag{20}$$

where J_i^l represents the number of demands from node $l \in \mathcal{K}_i$ in the last j_i demands realized at node i. Clearly,

$$\sum_{l \in \mathcal{K}_i} J_i^l = j_i \quad \forall\, i \in \mathcal{I}_k.$$

δ_i represents the node from which the last j_ith demand at node i comes, $\delta_i \in \mathcal{K}_i$.

Enumerating all possible outcomes of $(J_i^l, l \in \mathcal{K}_i, \delta_i; i \in \mathcal{I}_k) = (j_i^l, l \in \mathcal{K}_i, l_i; i \in \mathcal{I}_k)$ yields

$$p(j_i, m_i, t_i; i \in \mathcal{I}_k)$$
$$= \sum_{\forall (j_i^l, l \in \mathcal{K}_i, l_i; i \in \mathcal{I}_k)} \Pr\Biggl\{\sum_{l \in \mathcal{K}_i} \sum_{j=1}^{j_i^l} D_{l,j} = s_i - m + m_i, D_{l_i, j_i^{l_i}} \geqslant m_i; i \in \mathcal{I}_k\Biggr\}$$
$$\times \Pr\{T_i(j_i) \geqslant t_i, J_i^l = j_i^l, l \in \mathcal{K}_i, \delta_i = l_i; i \in \mathcal{I}_k\}. \tag{21}$$

Similar to Figures 2(a) and 2(b), $\Pr\{\sum_{l \in \mathcal{K}_i} \sum_{j=1}^{j_i^l} D_{l,j} = s_i - m + m_i, D_{l_i, j_i^{l_i}} \geqslant m_i\}$ can be computed by identifying and conditioning on the common demand sizes shared by nodes $i \in \mathcal{I}_k$. By a similar logic as that for Equation (17), we must have

$$\Pr\{T_i(j_i) \geqslant t_i, J_i^l = j_i^l, l \in \mathcal{K}_i, \delta_i = l_i; i \in \mathcal{I}_k\}$$
$$= \Pr\Biggl\{\sum_{j=1}^{j_i^{l_i}} \nu_{l_i,j} \geqslant t_i, \sum_{j=1}^{j_i^l} \nu_{l,j} < \sum_{j=1}^{j_i^{l_i}} \nu_{l_i,j} < \sum_{j=1}^{j_i^l+1} \nu_{l,j}, \; l \in \mathcal{K}_i \text{ and } l \neq l_i; i \in \mathcal{I}_k\Biggr\}. \tag{22}$$

To compute this probability, one can aggregate interarrival times to obtain the minimum number of independent Erlang random variables that are shared across different arguments. Conditioning on common Erlang random variables, the probability breaks into products of marginal probabilities (either Poisson or Erlang), each associated with a demand process.

If demand follows independent Poisson processes, then $J_i(m) \equiv s_i$ and $M_i(m) \equiv 1$. Zhao and Simchi-Levi (2006) characterizes the first two moments of $T_i(s_i)$, but leaves their joint probability distribution as an open question. Following the analysis here, the joint probability distribution of $T_i(s_i)$ is given as follows:

$$\Pr\{T_i(s_i) \geqslant t_i; i \in \mathcal{I}_k\}$$
$$= \sum_{\forall (j_i^l, l \in \mathcal{K}_i, l_i; i \in \mathcal{I}_k)} \Pr\{T_i(j_i) \geqslant t_i, J_i^l = j_i^l, l \in \mathcal{K}_i, \delta_i = l_i; i \in \mathcal{I}_k\}, \tag{23}$$

where $\sum_{l \in \mathcal{K}_i} J_i^l = s_i$ and the right-hand side is given by Equation (22). In other words, by conditioning on the number of arrivals from each demand type and the demand type of the last s_ith arrival, one can characterize the joint distribution of $T_i(s_i)$, $i \in \mathcal{I}_k$.

In principle, the analysis of the two-level general networks extends to multilevel assembly networks with at most one directed path between each pair of nodes. However, the exact analysis is more involved because in addition to the dependence among nodes on the same level, we also need to characterize the dependence among nodes across different levels. In other words, we need to characterize the joint probability of (J_k, M_k, T_k) for nodes on all levels.

Finally, we characterize the impact of the common demand size and interarrival time processes on system performance. See the appendix for a proof.

Proposition 2.1. *In a two-level general network, consider an assembly node $k \in \mathcal{K}$ and its suppliers $i \in \mathcal{I}$, where $(i,k) \in \mathcal{A}$. For any $m \geqslant 1$ and $\tau \geqslant 0$,*

$$\begin{aligned}
&\Pr\{X_i(m) + L_{i,k} \leqslant \tau, (i,k) \in \mathcal{A}\}\\
&= \Pr\{[\tilde{L}_i(M_i(m)) - T_i(J_i(m))]^+ + L_{i,k} \leqslant \tau, (i,k) \in \mathcal{A}\}\\
&\geqslant \Pr\{[\tilde{L}_i(M_i'(m)) - T_i(J_i'(m))]^+ + L_{i,k} \leqslant \tau, (i,k) \in \mathcal{A}\} \quad (24)\\
&\geqslant \Pi_{(i,k)\in\mathcal{A}} \Pr\{X_i(m) + L_{i,k} \leqslant \tau\}, \quad (25)
\end{aligned}$$

where $M_i'(m)$ and $J_i'(m)$ are independent copies of $M_i(m)$ and $J_i(m)$.

Because $\tilde{L}_k(m) = \max_{(i,k)\in\mathcal{A}}\{X_i(m) + L_{i,k}\} + L_k$, Proposition 2.1 has two implications: First, the total lead time at node k is stochastically larger than or equal to that of an analogous system with independent demand-size processes (inequality (24)). Second, the total lead time at node k with independent demand-size processes but a dependent demand interarrival process are stochastically larger than or equal to that of an analogous system with both independent demand-size processes and independent interarrival time processes (inequality (25)).

2.3. Networks of Special Structure

For networks with special structure, such as the pure assembly or the pure distribution systems, the analysis can be significantly simplified.

Pure Assembly Systems. Consider a pure assembly system where nodes $i = 1, 2, \ldots, I$ supply node 0, and node 0 is the only customer of each node i. For any $m \geqslant 1$, $\tilde{L}_0(m) = \max_{i=1,2,\ldots,I}\{X_i(m) + L_{i,0}\} + L_0$ and $X_i(m) = [\tilde{L}_i(M_i(m)) - T_i(J_i(m))]^+$ for all i.

The joint probability of $(J_i, M_i, T_i; i = 1, 2, \ldots, I)$ can be characterized as follows:

$$\begin{aligned}
&\Pr\{J_i(m) = j_i, M_i(m) = m_i, T_i(j_i) \geqslant t_i; i = 1, 2, \ldots, I\}\\
&\quad = \Pr\{J_i(m) = j_i, M_i(m) = m_i; i = 1, 2, \ldots, I\}\\
&\qquad \cdot \Pr\{T_i(j_i) \geqslant t_i; i = 1, 2, \ldots, I\}. \quad (26)
\end{aligned}$$

The equation holds because $(J_i, M_i; i = 1, 2, \ldots, I)$ now depends only on the demand-size process.

We index the component nodes so that $s_1 \leqslant s_2 \leqslant \cdots \leqslant s_I$. If $j_1 < j_2 < \cdots < j_I$, then the probability mass function of $(J_i, M_i; i = 1, 2, \ldots, I)$ can be broken down into products of marginal probabilities as follows:

$$\begin{aligned}
&\Pr\{J_i(m) = j_i, M_i(m) = m_i; i = 1, 2, \ldots, I\}\\
&\quad = \Pr\left\{\sum_{j=1}^{j_i} D_{0,j} = s_i - m + m_i, D_{0,j_i} \geqslant m_i; i = 1, 2, \ldots, I\right\}\\
&\quad = \Pr\left\{\sum_{j=1}^{j_1} D_{0,j} = s_1 - m + m_1, D_{0,j_1} \geqslant m_1\right\}\\
&\qquad \cdot \Pr\left\{\sum_{j=j_1+1}^{j_2} D_{0,j} = s_2 + m_2 - s_1 - m_1, D_{0,j_2} \geqslant m_2\right\}\\
&\qquad \times \cdots \times \Pr\left\{\sum_{j=j_{I-1}+1}^{j_I} D_{0,j} = s_I + m_I - s_{I-1} - m_{I-1},\right.\\
&\qquad\qquad \left. D_{0,j_I} \geqslant m_I\right\}. \quad (27)
\end{aligned}$$

The second equality holds because all components share the identical demand-size process $\{D_{0,j}, j \geqslant 1\}$; see Figure 4 for a graphic illustration. If $j_i = j_{i+1}$ for some i, then we replace the term $\Pr\{\sum_{j=j_i+1}^{j_{i+1}} D_{0,j} = s_{i+1} + m_{i+1} - s_i - m_i, D_{0,j_{i+1}} \geqslant m_{i+1}\}$ in Equation (27) by $1_{\{m_{i+1}+s_{i+1}=m_i+s_i\}}$.

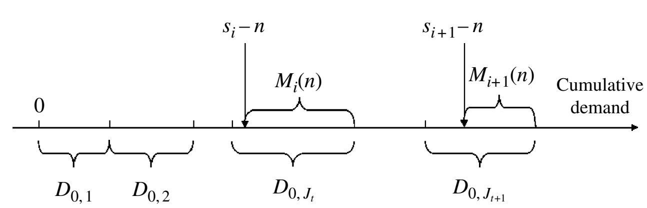
Because all components face the identical demand arrival process, Equation (22) reduces to

$$\begin{aligned}
&\Pr\{T_i(j_i) \geqslant t_i, i = 1, 2, \ldots, I\}\\
&\quad = \Pr\left\{\sum_{j=1}^{j_i} \nu_{0,j} \geqslant t_i; i \in \mathcal{I}_k\right\}\\
&\quad = \Pr\left\{\sum_{j=1}^{j_1} \nu_{0,j} \geqslant t_1, \sum_{j=1}^{j_1} \nu_{0,j} + \sum_{j=j_1+1}^{j_2} \nu_{0,j} \geqslant t_2, \ldots,\right.\\
&\qquad \left. \sum_{j=1}^{j_1} \nu_{0,j} + \sum_{j=j_1+1}^{j_2} \nu_{0,j} + \cdots + \sum_{j=j_{I-1}+1}^{j_I} \nu_{0,j} \geqslant t_I\right\}, \quad (28)
\end{aligned}$$

where $\nu_{0,j}$ are demand interarrival times at node 0. This equation confirms Equation (8) of Zhao and Simchi-Levi (2006).

An assembly system with compound Poisson demand is analytically more challenging than an analogous system with Poisson demand because of the dependence due to common demand-size process (Equation 27).

Figure 4. J_i and M_i in a pure assembly system.



Pure Distribution Systems. Consider a distribution system where node 0 supplies nodes $k = 1, 2, \ldots, K$. For a node k, consider the nth unit of a demand that arrives at time t. $X_k(n)$ and $W_k(n)$ are given by Equations (7) and (8), where $\tilde{L}_k(m) = X_0(m) + L_{0,k} + L_k$ and $X_0(m) = [\tilde{L}_0(M_0(m)) - T_0(J_0(m))]^+$.

Pure distribution systems are more tractable than pure assembly systems because $(J_k(n), M_k(n), T_k(j))$ are independent across $k = 0, 1, \ldots, K$. Thus, we can decompose the distribution system into $K+1$ single-stage systems as follows: We first characterize $X_0(m)$ and $\tilde{L}_k(m)$ for all m, then determine $X_k(n)$ and $W_k(n)$ for each k and n.

Forsberg (1995) provides a more efficient exact method for two-level distribution systems. For comparison, let $p_{k,n}$ be the probability that a randomly picked demand unit is the nth unit of a demand at node k, or equivalently, the long-run proportion of demand units that are the nth unit of a demand. By Sigman (2001),

$$p_{k,n} = \Pr\{D_k \geqslant n\}/E(D_k). \tag{29}$$

Then, the expected inventory holding time at node k is given by

$$\begin{aligned}\sum_n p_{k,n} E(W_k(n)) &= \sum_n p_{k,n} \sum_{j_k, m} \Pr\{J_k(n) = j_k, M_k(n) = m\} \sum_{j_0} \Pr\{J_0(m) = j_0\} \\ &\quad \times E((T_k(j_k) - [L_0 - T_0(j_0)]^+ - L_{0,k} - L_k)^+). \end{aligned} \tag{30}$$

The method of Forsberg (1995) is computationally more efficient than Equation (30) because for each unit of a retailer k's order, it identifies the corresponding order of node 0 that satisfies this unit, as well as the corresponding demand at the retailer that will be satisfied by this unit; see Equation (3) of Forsberg (1995). Therefore, effectively it ignores n in Equation (30) and directly sums over all possible m. The method of Forsberg (1995) can also be extended to handle stochastic sequential transit times.

For two-level distribution systems, we also point out that Axsater (2000) provides a more general exact method to handle both compound Poisson demand and batch-ordering policies. The advantage of our approach is that it applies to supply chains with a much more general network structure (than distribution systems) in a uniform way. In §3, we shall derive approximations based on our approach to facilitate fast computation.

2.4. Performance Measures

To determine the cost measures and service levels at node k, we consider a randomly picked demand unit. Let X_k be the stockout delay of a randomly picked demand unit at node k, and W_k be the inventory holding time of the corresponding item at node k that satisfies a randomly picked demand unit. Then,

$$\Pr\{X_k \leqslant x\} = \sum_{n \geqslant 1} p_{k,n} \Pr\{X_k(n) \leqslant x\}, \tag{31}$$

$$\Pr\{W_k \leqslant w\} = \sum_{n \geqslant 1} p_{k,n} \Pr\{W_k(n) \leqslant w\}, \tag{32}$$

where $p_{k,n}$ is the probability that a randomly picked demand unit is the nth unit of a demand at node k; see Equation (29).

By Little's Law, the expected backorders and on-hand inventory at node k are given by

$$E(B_k) = \lambda_k E(D_k) E(X_k) = \lambda_k E(D_k) \sum_n p_{k,n} E(X_k(n)), \tag{33}$$

$$E(I_k) = \lambda_k E(D_k) E(W_k) = \lambda_k E(D_k) \sum_n p_{k,n} E(W_k(n)). \tag{34}$$

The type 2 fill rate, β_k, within a committed service time τ_k, is given by

$$\beta_k = \sum_n p_{k,n} \Pr\{X_k(n) \leqslant \tau_k\}. \tag{35}$$

If node k is an assembly node, then in addition to the on-hand inventory of the finished good, I_k, we also need to consider the component inventories, $I_k^i\ \forall (i,k) \in \mathcal{A}$. These inventories are held at node k without being processed because the corresponding units of other required components are not yet replenished. By Little's Law and Equation (18),

$$E(I_k^i) = \lambda_k E(D_k) \sum_n p_{k,n} E(W_k^i(n)), \tag{36}$$

$$W_k^i(n) = \max_{\{l \mid (l,k) \in \mathcal{A}\}} \{X_l(M_k(n)) + L_{l,k}\} - X_i(M_k(n)) - L_{i,k}. \tag{37}$$

3. Approximations and Optimization

Section 2 provides a basis for exact evaluation of small-size problems (see §4.1 for an example). For larger problems, exact evaluation is time demanding due to the dependent stockout delays in assembly networks. In this section, we develop numerically tractable approximations on the stockout delay, cost, and fill rate. We also formulate the optimization problem.

Approximations on Stockout Delay and Cost. The class of supply chains considered in this paper introduces four challenges for exact evaluation: (1) the dependence among parallel stages in assembly networks, (2) the dependence between M_k and J_k at each node k, (3) the need to compute $X_k(n)$ for each n, and finally, (4) the need to compute probability distribution for $X_k(n)$.

To enable fast evaluation and optimization of system performance, we first ignore the dependence in assembly networks by focusing only on the marginal probability distributions of the stockout delays. Hence, the supply chain can be decomposed into multiple single-stage systems where each stage can be characterized separately. Second, we ignore the dependence between M_k and J_k at each node k.

To resolve the third challenge, we replace $X_k(n) = [\tilde{L}_k(M_k(n)) - T_k(J_k(n))]^+\ \forall n$ by an approximation

$$X_k \approx [\tilde{L}_k - T_k]^+, \tag{38}$$

where X_k is the stockout delay for a randomly picked demand unit at node k, $\tilde{L}_k = \max_{(i,k)\in\mathcal{A}}\{X_i + L_{i,k}\} + L_k$, and $T_k = T_k(J_k(N))$, where N is the index of a randomly picked demand unit at node k.

This approximation can be justified as follows. Given s_k, λ_k, and D_k, as well as the distribution of $\tilde{L}_k(m)$ for all m, we can compute the probability distribution of X_k as follows (Equation (7)):

$$\begin{aligned}
&\Pr\{X_k \leqslant x\}\\
&= \sum_{n>0} p_{k,n} \sum_{m\geqslant 1,\, j\geqslant 0} \Pr\{J_k(n)=j, M_k(n)=m\}\\
&\quad \cdot \Pr\{\tilde{L}_k(m) - T_k(j) \leqslant x\}\\
&\approx \sum_{n>0} p_{k,n} \sum_{m\geqslant 1} \Pr\{M_k(n)=m\} \sum_{j\geqslant 0} \Pr\{J_k(n)=j\}\\
&\quad \cdot \Pr\{\tilde{L}_k(m) - T_k(j) \leqslant x\}\\
&\to \sum_{m\geqslant 1} p_{k,m} \sum_{n>0} p_{k,n} \sum_{j\geqslant 0} \Pr\{J_k(n)=j\} \Pr\{\tilde{L}_k(m) - T_k(j) \leqslant x\}\\
&\qquad \text{as } s_k - n \to \infty. \qquad (39)
\end{aligned}$$

The "≈" comes from our approximation of ignoring the dependence between M_k and J_k. The asymptotic result holds because by a property of renewal processes, $\Pr\{M_k(n) = m\} \to p_{k,m}$ as $s_k - n \to \infty$ (Kulkarni 1995, p. 435), where $p_{k,m}$ is defined in Equation (29) and is independent of n.

Equation (39) implies that $X_k \approx (\tilde{L}'_k - T_k)^+$, where $\tilde{L}'_k$ is the lead time for a randomly picked order unit at node k. To justify Equation (39), we note that $\tilde{L}'_k$ is identical to $\tilde{L}_k$ if node k and all its immediate supplying nodes face the identical demand-size process. This is true in serial and pure assembly systems. However, in distribution systems, $\tilde{L}_k$ is only an approximation of $\tilde{L}'_k$.

To resolve the fourth challenge, we utilize a two-moment approximation proposed by Zipkin (1991) to compute $E(I_k)$ and the first two moments of X_k; see the appendix. If node k is an assembly node, then by Equations (36) and (37), the expected component inventory level at node k can be approximated by

$$E(I_k^i) \approx \lambda_k E(D_k) E\Big[\max_{\{l \mid (l,k)\in\mathcal{A}\}}\{X_l + L_{l,k}\} - X_i - L_{i,k}\Big], \quad (40)$$

where the mean and variance of the maximum of independent random variables can be computed by Clark's two-moment approximation (Clark 1961).

Approximations on Fill Rates. We now develop approximations for service levels based on compound distribution (Zipkin 2000, §C.2.3.8) and renewal theory. If node k faces external demand, then it follows from Equations (35) and (38) that the type 2 fill rate within τ_k at node k can be approximated by

$$\sum_{n\geqslant 1} p_{k,n} \Pr\{\tilde{L}_k - T_k(J_k(n)) \leqslant \tau_k\} \geqslant \beta_k. \quad (41)$$

Define $\Delta_{k,n} = \tilde{L}_k - T_k(J_k(n)) - \tau_k$. Given $\tilde{L}_k$, s_k, τ_k, and demand parameters λ_k and D_k, we approximate $\Delta_{k,n}$ by a normal random variable with the mean and variance determined by (see, e.g., Zipkin 2000, §C.2.3.8)

$$E(\Delta_{k,n}) = E(\tilde{L}_k) - E(J_k(n))/\lambda_k - \tau_k, \quad (42)$$

$$V(\Delta_{k,n}) = V(\tilde{L}_k) + E(J_k(n))/\lambda_k^2 + V(J_k(n))/\lambda_k^2. \quad (43)$$

We now determine $E(J_k(n))$ and $V(J_k(n))$. By Equation (3), if $s_k < n$, then $J_k(n) = 0$, and therefore $E(J_k(n)) = V(J_k(n)) = 0$. If $s_k \geqslant n$, then $J_k(n) = 1 + N_k(s_k - n)$, where $E(N_k(s_k - n))$ and $V(N_k(s_k - n))$ are approximated by their asymptotic value $(s_k - n)/E(D_k)$ and $(s_k - n)V(D_k)/E^3(D_k)$ (as $s_k - n \to \infty$), respectively (Kulkarni 1995, Theorem 8.7).

Optimization. For convenience, let $\bar{X}_k$ be a vector representing the stockout delays at the immediate suppliers of node k, i.e., $\bar{X}_k = (X_i \mid (i,k) \in \mathcal{A})$. Given the mean and variance of $\bar{X}_k$ ($E(\bar{X}_k)$ and $V(\bar{X}_k)$), $E(X_k)$, and the demand parameters λ_k and D_k, we can determine s_k, $V(X_k)$, and the safety-stock carrying costs H_k at node k, where

$$H_k(E(\bar{X}_k), V(\bar{X}_k), E(X_k)) = h_k E(I_k) + \sum_{(i,k)\in\mathcal{A}} h_i E(I_k^i). \quad (44)$$

We can formulate a mathematical program for compound Poisson demand in the same way as that for Poisson demand; see Simchi-Levi and Zhao (2005) for details. The program here differs from that of Poisson demand in two ways: (1) s_k, $V(X_k)$, and H_k are computed in a different way by $\bar{X}_k$ and $E(X_k)$ due to the compound demand processes; (2) the service level is computed in a different way. Simchi-Levi and Zhao (2005) develops an algorithm based on dynamic programming (DP) to compute the optimal or near-optimal base-stock levels for tree-structure supply chains with Poisson demand and stochastic sequential transit times. The same algorithm applies here. We refer the reader to Simchi-Levi and Zhao (2005) for an extended discussion on the DP algorithm and its complexity.

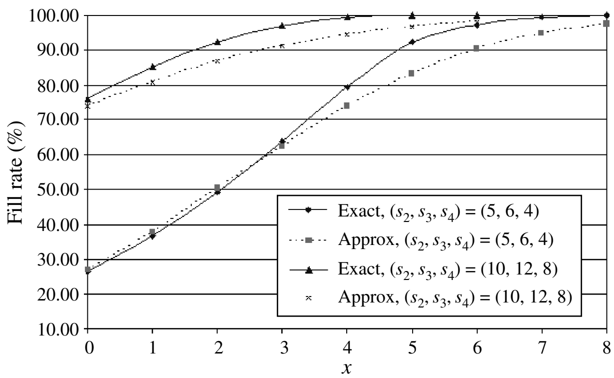
4. Numerical Studies

The objective of this section is threefold: (1) developing insights into the conditions under which the approximation may or may not be sufficiently accurate, (2) illustrating the efficiency of the approximation, and (3) demonstrating the quality of the solution found by the DP algorithm.

4.1. Figure 2(a)

We present an exact evaluation of the system in Figure 2(a) based on the closed-form expressions in Equations (11), (12), and (14). Our objectives are to compare the exact evaluation and the approximation in their numerical efficiency and to test the accuracy of the approximation. The accuracy of the approximation in larger systems is tested in §§4.2–4.3.

Figure 5. A comparison between the exact evaluation and the approximation on $\Pr\{X_2 \leqslant x\}$.

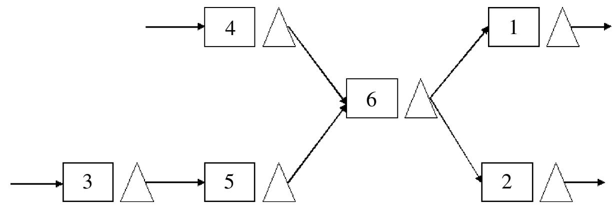


We assume zero lead times from external suppliers and zero transportation lead times between every two nodes, but nonzero processing times at all nodes. We consider stochastic processing times at nodes 3 and 4. In particular, we assume that L_3 (L_4) follows uniform distribution between zero and four (zero and three). Other types of distribution are considered in §§4.2–4.3. For simplicity, we let $L_2 = 5$. Without loss of generality, let $\lambda_1 = 0.4$ and $\lambda_2 = 0.6$. The demand sizes at nodes 1 and 2 have probability mass function $(\Pr\{D_1 = 1\}, \ldots, \Pr\{D_1 = 5\}) = (0.2, 0.3, 0.2, 0.2, 0.1)$ and $(\Pr\{D_1 = 1\}, \ldots, \Pr\{D_1 = 4\}) = (0.4, 0.3, 0.2, 0.1)$, respectively.

Figure 5 presents the exact evaluation and the approximation of $\Pr\{X_2 \leqslant x\}$ for two sets of base-stock levels. The numerical result indicates that the approximation is reasonably accurate overall, although sizable errors can occur. For example, at $(s_2, s_3, s_4) = (5, 6, 4)$, the approximation is very accurate when $x < 4$. However, the error grows to around 8.5% at $x = 5$. This can be explained as follows: Because $L_2 = 5$, the fill rate becomes heavily dependent on the stockout delays from nodes 3 and 4 as x increases to 5. Because the Clark's (1961) two-moment approximation of the maximum of random variables is based on normal distribution, it can generate sizable errors when the stockout delays at nodes 3 and 4 are far from being normally distributed. We also observe that the approximation considerably underestimates the fill rate for $x \geqslant 4$ (at $(s_2, s_3, s_4) = (5, 6, 4)$) and for all x (at $(s_2, s_3, s_4) = (10, 12, 8)$). This observation confirms Proposition 2.1—because the approximation ignores the dependences in assembly systems, it can overestimate $\tilde{X}_2$.

Although sizable errors can occur, the approximation is attractive because of its numerical efficiency. Let $s_{\max} = \max\{s_2, s_3, s_4\}$. The complexity of the approximation is $O(s_{\max})$ and $O(U_2)$, whereas the complexity of the exact evaluation is $O(s_{\max}^{11})$ and $O(U_2^4)$. By storing some intermediary quantities such as incomplete convolutions of Erlang random variables, we can reduce the computational complexity of the latter to $O(s_{\max}^8)$ and $O(U_2^4)$. On a Pentium 1.73 GHz laptop, the exact evaluation for cases with $(s_2, s_3, s_4) = (10, 12, 8)$ takes around two to three seconds, whereas the approximation takes less than $1/10^5$ a second. Because the exact evaluation is time demanding for larger systems (in the number of nodes and arcs, as well as in stock levels), we use Monte Carlo simulation (a sampling method based on the exact analysis in §2) to test the accuracy of the approximation in §§4.2–4.3.

Figure 6. The six-stage example.



4.2. A Six-Stage System

To further test the accuracy of the approximation and to demonstrate the quality of the solution, we consider a six-stage production-distribution system (see Figure 6). We ignore all transportation lead times, but assume that the processing cycle times follow Erlang distributions with parameters $E(L)$ and n (see, e.g., Zipkin 2000, p. 457). According to convention, we assume that the inventory holding cost increases as one moves downstream in the supply chain, i.e., $(h_1, h_2, h_3, h_4, h_5, h_6) = (4, 4, 1, 1.5, 2, 3)$. Without loss of generality, we let $\lambda_1 = 0.7$ and $\lambda_2 = 0.3$. The demand sizes at nodes 1 and 2 follow discrete normal distributions with probability mass functions $(\Pr\{D_1 = 1\}, \Pr\{D_1 = 2\}, \ldots, \Pr\{D_1 = 7\}) = (0.00621,\ 0.0606,\ 0.2417,\ 0.383, 0.2417, 0.0606, 0.00621)$ and $(\Pr\{D_1 = 3\}, \Pr\{D_1 = 4\}, \ldots, \Pr\{D_1 = 9\}) = (0.00621,\ 0.0606,\ 0.2417,\ 0.383, 0.2417, 0.0606,\ 0.00621)$, respectively.

Compound Poisson Demand. We first consider compound Poisson demand. To test the accuracy of the approximation, we use the DP algorithm to find a solution and then use Monte Carlo simulation to evaluate the solution. We run 10^4 independent replications for each parameter set and calculate the 95% confidence interval (CI) for the performance measure. The CI for the fill rate is less than 1%, and the CI for the cost is less than 3% of the simulated cost.

We conduct two numerical studies. In the first study, we fix the parameters for the processing times $\{n^1, n^2, n^3, n^4, n^5, n^6\} = \{5, 7, 6, 7, 8, 5\}$ and $\{E(L_1), E(L_2), E(L_3), E(L_4), E(L_5), E(L_6)\} = \{6, 6, 1, 2, 5, 6\}$, but vary τ_1, τ_2 and β_1, β_2, where τ_k, $k = 1, 2$ chooses value from $\{0, 2, 4\}$, and β_k, $k = 1, 2$ chooses value from $\{0.85, 0.9, 0.95, 0.99\}$. Thus, this study has a total of 144 instances. In the second study, we fix $\tau_1 = 1$ and $\tau_2 = 2$. For each $\beta_1 = \beta_2 = 0.85, 0.90, 0.95, 0.99$, we study 100 instances with randomly generated processing time parameters, where $E(L_k) \sim \text{Uniform}(1, 2, \ldots, 10)$ and $n^k \sim \text{Uniform}(1, 2, \ldots, 10)$ for $k = 1, 2, \ldots, 6$.

Table 1. The accuracy of the approximation in the first study.

Avg. abs. percentage error in cost	Max. abs. percentage error in cost	Avg. abs. error in β_1	Max. abs. error in β_1	Avg. abs. error in β_2	Max. abs. error in β_2
0.52%	1.7%	0.46%	1.66%	1.24%	3.9%

Table 1 demonstrates the average and maximum for the absolute percentage error in cost and the absolute error in β_1 and β_2 in the first study. The absolute percentage error in cost is the absolute difference between the simulated cost and the cost generated by the DP algorithm divided by the simulated cost.

Table 2 summarizes the results of the second study. Tables 1 and 2 demonstrate that overall the approximation is reasonably accurate for the parameter sets examined here. The largest error is a 3.9% error on β_2 (Table 1); the corresponding instance has $\beta_2 = 0.85$, which indicates that the errors may increase as β decreases. This observation is partially confirmed by Table 2, which shows that the average and maximum errors of the cost and β_2 are increasing as β_1 and β_2 decrease. However, the trend is not clear for β_1.

Given its accuracy and the numerical tractability, the approximation is attractive for a wide range of parameters of interest. However, we should point out that if $\tau_k \geqslant L_k + L_6$, $k = 1, 2$ (although not likely to happen in practice), the approximation can be much less accurate, e.g., errors around 10% are recorded in the fill rates. This observation is confirmed in §4.1.

To demonstrate the quality of the DP solution, we compare it to a solution found by a simulation-based search algorithm. Because the simulated fill rates of the DP solution may not exactly match the targets, we adjust the input fill rates and run the DP algorithm repetitively until the target fill rates fall into the 95% confidence interval of the simulated fill rates. In the search algorithm, we first identify an upper bound s'_k for the base-stock level at each node. Then, for a base-stock level vector $(s_3, s_4, s_5, s_6) \in \bigotimes_{k=3,4,5,6}\{0, \lfloor s'_k/10 \rfloor, \lfloor 2s'_k/10 \rfloor, \ldots, s'_k\}$, we use simulation to evaluate system cost and choose s_1 and s_2 so that the simulated fill rates closely match the targets. Table 3 summarizes the results where $\tau_1 = 1$, $\tau_2 = 2$, $(E(L_1), E(L_2), E(L_3), E(L_4), E(L_5), E(L_6)) = (3, 3, 4, 4, 3, 1)$, and $(n^1, n^2, n^3, n^4, n^5, n^6) = (6, 5, 2, 8, 6, 3)$. All costs are given by simulation.

The percentage difference in costs in Table 3 is defined as the difference between the cost of the DP solution and the cost of the search-based solution divided by the cost of the search-based solution. On a Pentium 1.73 GHz laptop, the search algorithm takes hours to solve for one instance, whereas the DP algorithm takes about one to two seconds. We first note that in all cases, the cost of the DP solution is reasonably close to that of the search-based solution. The DP solutions tend to perform better as the target fill rates increase. This observation is consistent with our results on the accuracy of the approximation (Tables 1 and 2). Because the approximation tends to be more accurate for higher target fill rates, the DP algorithm (based on the approximation) finds better solutions as the fill rates increase.

Renewal Demand Processes. To further demonstrate the potential of the approximation, we consider renewal batch demand where the interarrival time, ν, follows an Erlang distribution with parameters λ and n_ν. The demand sizes are independent of the interarrival times. Because the approximation requires $E(Y_k) < V(Y_k)$ for all nodes k, we assume that the demand size at node 1 (2) follows a Poisson distribution with mean 4 (6). All other parameters remain unchanged.

Under renewal batch demand, the demand process faced by the distribution node (node 6) is nonrenewal. To resolve this issue, we approximate the demand process at node 6 by a renewal batch process, following Whitt (1982). We also modify the two-moment approximation to accommodate renewal batch demand. See the appendix for details. To test the accuracy of the approximation, we conduct the two similar numerical studies as in the case of compound Poisson demand. In the first study, we vary $n_\nu = 1, 2, 4, 16$, whereas in the second study, we set $n_\nu = 4$.

Tables 4 and 5 summarize the results. First, the average and maximum errors in cost are sufficiently small in all cases. Second, although the average errors in fill rates are relatively small, e.g., $\leqslant$4.6% (Table 4), the maximum errors increase considerably as n_ν increases (i.e., as the demand processes deviate from compound Poisson); see Table 4. Examination of the instances corresponding to the largest

Table 2. The accuracy of the approximation in the second study.

$\beta_1 = \beta_2 =$	Avg. abs. percentage error in cost (%)	Max. abs. percentage error in cost (%)	Avg. abs. error in β_1 (%)	Max. abs. error in β_1 (%)	Avg. abs. error in β_2 (%)	Max. abs. error in β_2 (%)
0.85	0.67	3.51	0.93	2.48	2.02	3.71
0.90	0.64	2.54	0.93	3.09	1	2.53
0.95	0.46	1.79	1.10	3.28	0.64	2.42
0.99	0.32	1.36	0.79	3.04	0.45	1.10

Table 3. Comparison between the solutions found by the DP and the search algorithm.

(β_1, β_2)	$(s_1, s_2, s_3, s_4, s_5, s_6)$ The search solution	$(s_1, s_2, s_3, s_4, s_5, s_6)$ The DP solution	Cost (search)	Cost (DP)	Percent diff. in costs (%)
(0.90, 0.90)	(24, 17, 23, 7, 6, 21)	(27, 20, 6, 0, 15, 21)	133.66	140.77	5.32
(0.90, 0.96)	(23, 22, 27, 0, 0, 28)	(22, 23, 12, 11, 18, 21)	153.68	161.43	5.04
(0.95, 0.90)	(30, 17, 13, 15, 19, 14)	(33, 20, 6, 0, 15, 21)	156.33	163.60	4.65
(0.96, 0.96)	(29, 22, 27, 15, 12, 14)	(28, 22, 6, 0, 15, 31)	175.90	182.90	3.98
(0.98, 0.95)	(37, 21, 18, 7, 12, 21)	(36, 21, 6, 0, 9, 34)	203.40	208.90	2.70
(0.95, 0.99)	(29, 34, 23, 0, 0, 28)	(28, 33, 6, 0, 9, 34)	220.60	222.80	1
(0.98, 0.99)	(36, 33, 27, 3, 0, 28)	(36, 33, 6, 0, 9, 34)	248.03	255.47	3

errors in fill rates (in Table 4) reveals that the instances also have relatively small target fill rates (e.g., 85%). This observation is confirmed by Table 5, which clearly demonstrates that as the target fill rates increase from 85% to 99%, both the average and the maximum errors in the fill rates and cost tend to decrease.

In summary, the approximation works well on cost for renewal batch demand. However, it may result in sizable errors on fill rates if the target fill rates are relatively low and the distribution of the interarrival times is far from exponential.

4.3. Larger-Size Examples

In this section, we demonstrate the accuracy and efficiency of the approximation using examples motivated by real-world problems.

The first example is inspired by the bulldozer supply chain (see Graves and Willems 2003), which is a multistage assembly network with 22 nodes and 21 arcs. We keep all the data on the cost and the expected processing time unchanged, but consider stochastic processing and transportation lead times (see Appendix III of Simchi-Levi and Zhao 2005 for details). We assume that the external demand follows compound Poisson process with $\lambda = 1$ and $(\Pr\{D_1 = 1\}, \Pr\{D_1 = 2\}, \ldots, \Pr\{D_1 = 7\}) = (0.00621, 0.0606, 0.2417, 0.383, 0.2417, 0.0606, 0.00621)$. The target customer service at the final assembly is set by τ and β, where $\tau = 0$ and $\beta \in \{0.85, 0.9, 0.95, 0.99\}$.

To study the impact of the lead-time uncertainty on the accuracy of the approximation, we assume that all processing times and transportation lead times follow Erlang distributions with identical coefficient of variation, i.e., identical $1/\sqrt{n}$, where $n \in \{4, 9, 16\}$. This example takes around one minute for the DP algorithm to generate a solution on a Pentium 1.73 GHz laptop. We compare the approximation here to a simpler one that only considers the mean stockout delays, e.g., METRIC (Sherbrooke 1968). For simplicity, we call the latter "one-moment" approximation and the former "two-moment" approximation.

For each parameter set, we first use the DP algorithm (based on the two-moment approximation) to determine the base-stock level at each node. Then, we use the one-moment approximation as well as simulation to compute the total cost and fill rate. In Table 6, the numbers without (with) parentheses show the absolute percentage difference in costs between the simulation and the two-moment (one-moment) approximation. Similarly, in Table 7, the numbers without (with) parentheses represent the absolute difference in fill rates between the simulation and the two-moment (one-moment) approximation.

Table 6 shows that the two-moment approximation is sufficiently accurate for cost in all cases, and it tends to be more accurate as the target fill rate increases. Table 7 illustrates that the two-moment approximation for the fill rate is reasonably accurate when the transit time coefficients of variation (c.v.s) are relatively small or when the target fill rate is high. When the transit time c.v.s are relatively large, e.g., $\geqslant 0.33$, the two-moment approximation may perform poorly on the fill rates.

Tables 6 and 7 also show that the two-moment approximation is much more accurate for both the cost and the fill rate than the one-moment approximation in almost all cases, especially when the transit time c.v.s are high. Indeed, the one-moment approximation always and substantially overestimates the fill rate except when the target fill rate is close to one.

The second example is inspired by the battery supply chain (see Graves and Willems 2003), which is a production-distribution network with 20 nodes and 19 arcs. We simplify the original network by combining the nodes

Table 4. The accuracy of the approximation in the first study, renewal demand.

n_ν	Avg. abs. percentage error in cost (%)	Max. abs. percentage error in cost (%)	Avg. abs. error in β_1 (%)	Max. abs. error in β_1 (%)	Avg. abs. error in β_2 (%)	Max. abs. error in β_2 (%)
1	0.52	1.75	0.25	1.18	1	3.57
2	0.46	2.54	0.97	2.67	1.87	5.10
4	0.47	1.89	1.40	3.15	3.04	6.74
16	0.61	2.48	2	3.53	4.60	9.05

Table 5. The accuracy of the approximation in the second study, renewal demand.

Target $\beta_1 = \beta_2 =$	Avg. abs. percentage error in cost (%)	Max. abs. percentage error in cost (%)	Avg. abs. error in β_1 (%)	Max. abs. error in β_1 (%)	Avg. abs. error in β_2 (%)	Max. abs. error in β_2 (%)
0.85	1.08	3.78	1.44	5.77	4.11	9.55
0.90	0.82	3.65	1.83	3.77	3.72	7.32
0.95	0.55	2.60	2.23	4.73	3.02	5.42
0.99	0.36	1.23	1.76	4.77	1.46	3.86

EMD, Spun Zinc, and Separater into a single node (denoted by ESS) because their lead times are identical.

We keep all parameters the same as in Graves and Willems (2003), except that we consider stochastic processing times that follow Erlang distributions (with identical parameter n). We define demand in units of 5,000 to allow for fast simulation. External demands are modeled by compound Poisson processes where the demand sizes follow negative binomial distributions. We set $\lambda = 1$ for each external demand, and choose $E(D)$ and $V(D)$ to match the mean and standard deviation of the demand in one period. $\tau = 0$ for all nodes facing external demand.

For each combination of $n \in \{4, 9, 16\}$ and $\beta \in \{0.85, 0.9, 0.95\}$, we use the DP algorithm to generate a solution (which takes around 20 seconds on a Pentium 1.73 GHz laptop), then we use the one-moment approximation and simulation to compute the total cost and fill rates. The numerical study shows that both the two-moment and the one-moment approximations are sufficiently accurate on cost. The maximum absolute percentage error (relative to the simulated cost) recorded is 4.58% (for the one-moment approximation) and 3.8% (for the two-moment approximation). In most cases, the percentage error is below 2%. Table 8 summarizes the result on fill rates, where the numbers without (with) parentheses represent the average and the maximum of the absolute differences in fill rates between the simulation and the two-moment (one-moment) approximation over all nodes facing external demand. Table 8 shows that the two-moment approximation always outperforms the one-moment approximation. The improvement of the former over the latter is quite significant in almost all cases.

Comparing Tables 7 and 8, we find that the transit time c.v. has less impact on the accuracy of the two-moment approximation in the bulldozer supply chain than in the battery supply chain. This is true because the former is an extensive assembly network, whereas the latter is a mixture of assembly and distribution networks. In assembly systems, Clark's (1961) approximation is utilized to compute the maximum of random variables. It can accumulate sizable errors as the number of assembly operations increases and the lead-time distributions deviate from normal.

5. Concluding Remarks

In this paper, we provide an exact framework to analyze a class of supply chains with at most one directed path between every two stages. The external demands follow independent compound Poisson processes, the transit times are stochastic, sequential, and exogenous, and each stage controls its inventory by an (installation) continuous-review base-stock policy. Based on the exact analysis, we present tractable approximations and demonstrate their effectiveness by numerical studies. The exact framework and approximations can be modified to handle supply chains under periodic-review base-stock policies; see Zhao (2006) for an extended discussion.

To conclude the paper, we identify a number of future research directions.

• Real-world supply chains may have multiple directed paths between two stages, i.e., acyclic supply chains. In addition, the bill of material (BOM) structure can be nonunit. An exact analysis of these supply chains requires the joint probability distribution of the transit times. To see this, consider a simple acyclic supply chain where a product is assembled from one unit of two subassemblies, each of which requires one unit of a common component. Clearly, two units of the common component are assembled into one unit of the product along different paths. To determine the stockout delay for the product, one has to characterize the joint replenishment processes of consecutive units for the common component, which requires information on the joint probability distribution of the transit times. Nonunit BOM structure poses a similar challenge. Consider a single-stage system with two types of demand. Each type requires a different number of units, and a demand

Table 6. The accuracy of the approximations in cost in the bulldozer supply chain.

	$\beta = 0.85$ 2 (1)-moment (%)	0.90 2 (1)-moment (%)	0.95 2 (1)-moment (%)	0.99 2 (1)-moment (%)
$n = 4$ ($c.v. = 0.5$)	2.08 (12.2)	1.5 (12.5)	0.70 (11.9)	0.16 (10.5)
9 (0.33)	1.40 (10.1)	0.9 (10.9)	0.40 (9.9)	0.16 (8.7)
16 (0.25)	2.55 (7.07)	2.3 (7.8)	1.68 (7.5)	0.10 (6.2)
($c.v. = 0$)	1.90 (1)	1.9 (2.5)	1.57 (3.17)	1.05 (2.9)

Table 7. The accuracy of the approximations in fill rate in the bulldozer supply chain.

	$\beta=0.85$ 2 (1)-moment (%)	0.90 2 (1)-moment (%)	0.95 2 (1)-moment (%)	0.99 2 (1)-moment (%)
$n=4$ ($c.v.=0.5$)	6.50 (21.3)	7 (16.8)	6.50 (11.4)	4 (5)
9 (0.33)	4.25 (18.6)	4.10 (13.9)	3.96 (8.8)	2 (3)
16 (0.25)	2.17 (16)	2.40 (11.8)	1.88 (6.5)	0.88 (1.86)
($c.v.=0$)	0.40 (10.7)	0.26 (7.8)	0.01 (4.26)	0.29 (1.28)

is satisfied only when all the required units are available. Because different units may have to wait for each other, additional inventory holding costs may occur. Exact evaluation of these holding costs requires information on the joint distribution of the transit times.

- In this paper, we ignore the batch-size constraints. We note that Axsater (2000) provides an exact method to evaluate two-level distribution systems with compound Poisson demand and batch-ordering policies. Song (2000) and Zhao and Simchi-Levi (2006) present exact analyzes for various ATO systems with batch-ordering policies. However, it is not clear how to exactly evaluate multilevel supply chains of more general structures (than distribution) with batch-ordering policies.

Appendix

We first compute the probability on demand sizes in Equation (11). Let $j_3>0$ and $j_4>0$. For simplicity, let $\Omega_D=\Pr\{\sum_{j=1}^{j_3^1}D_{1,j}+\sum_{j=1}^{j_3^2}D_{2,j}=s_3-m+m_3, D_{l,j_3^l}\geqslant m_3; \sum_{j=1}^{j_4}D_{2,j}=s_4-m+m_4, D_{2,j_4}\geqslant m_4\}$. We discuss two cases: $l=1$ and $l=2$. For $l=1$, $j_3^1>0$ and $j_3^2\geqslant 0$. We consider three subcases:

(1) If $j_3^2=j_4$, $\Omega_D=\Pr\{\sum_{j=1}^{j_3^1}D_{1,j}=s_3+m_3-s_4-m_4, D_{1,j_3^1}\geqslant m_3\}\Pr\{\sum_{j=1}^{j_4}D_{2,j}=s_4-m+m_4, D_{2,j_4}\geqslant m_4\}$, where each of the right-hand-side probabilities can be computed in a way similar to Equation (6).

(2) If $j_3^2>j_4$, $\Omega_D=\Pr\{\sum_{j=1}^{j_3^1}D_{1,j}+\sum_{j=j_4+1}^{j_3^2}D_{2,j}=s_3+m_3-s_4-m_4, D_{1,j_3^1}\geqslant m_3\}\Pr\{\sum_{j=1}^{j_4}D_{2,j}=s_4-m+m_4, D_{2,j_4}\geqslant m_4\}$.

(3) If $j_3^2<j_4$, $\Omega_D=\Pr\{\sum_{j=1}^{j_3^1}D_{1,j}+\sum_{j=1}^{j_3^2}D_{2,j}=s_3-m+m_3, D_{1,j_3^1}\geqslant m_3; \sum_{j=1}^{j_3^2}D_{2,j}+\sum_{j=j_3^2+1}^{j_4}D_{2,j}=s_4-m+m_4, D_{2,j_4}\geqslant m_4\}$. One can compute this probability by conditioning on $\sum_{j=1}^{j_3^2}D_{2,j}$.

For $l=2$, $j_3^1\geqslant 0$ and $j_3^2>0$. We consider three subcases:

(1) If $j_3^2=j_4$, $\Omega_D=\Pr\{\sum_{j=1}^{j_3^1}D_{1,j}=s_3+m_3-s_4-m_4\}\Pr\{\sum_{j=1}^{j_4}D_{2,j}=s_4-m+m_4, D_{2,j_3^2}\geqslant \max\{m_3,m_4\}\}$.

(2) If $j_3^2>j_4$, $\Omega_D=\Pr\{\sum_{j=1}^{j_3^1}D_{1,j}+\sum_{j=j_4+1}^{j_3^2}D_{2,j}=s_3+m_3-s_4-m_4, D_{2,j_3^2}\geqslant m_3\}\Pr\{\sum_{j=1}^{j_4}D_{2,j}=s_4-m+m_4, D_{2,j_4}\geqslant m_4\}$.

(3) If $j_3^2<j_4$, $\Omega_D=\Pr\{\sum_{j=1}^{j_3^1}D_{1,j}+\sum_{j=1}^{j_3^2}D_{2,j}=s_3-m+m_3, D_{2,j_3^2}\geqslant m_3; \sum_{j=1}^{j_3^2}D_{2,j}+\sum_{j=j_3^2+1}^{j_4}D_{2,j}=s_4-m+m_4, D_{2,j_4}\geqslant m_4\}$. One can compute this probability by conditioning on $\sum_{j=1}^{j_3^1}D_{1,j}$.

We then compute the probability on demand interarrival times in Equation (11). For simplicity, let $\Omega_T=\Pr\{T_3(j_3)\geqslant t_3, J_3^1=j_3^1, J_3^2=j_3^2, \delta_3=l; T_4(j_4)\geqslant t_4\}$. We discuss two cases: $l=1$ and $l=2$. For $l=1$, we consider three subcases:

(1) If $j_4<j_3^2$, then by Equation (12),

$$\Omega_T=\int_{\max\{t_3,t_4\}}^{\infty}\int_{t_4}^{\tau_3}\Pr\left\{\sum_{j=j_4+1}^{j_3^2}\nu_{2,j}<\tau_3-\tau_4<\sum_{j=j_4+1}^{j_3^2+1}\nu_{2,j}\right\} \times P\left\{\sum_{j=1}^{j_4}\nu_{2,j}=\tau_4\right\}d\tau_4 P\left\{\sum_{j=1}^{j_3^1}\nu_{1,j}=\tau_3\right\}d\tau_3,$$

where $P\{\cdot\}$ denotes density functions for Erlang random variables, and the probability inside the integral is given by Poisson mass function.

(2) If $j_4=j_3^2$, then

$$\Omega_T=\int_{\max\{t_3,t_4\}}^{\infty}\int_{t_4}^{\tau_3}\Pr\{\nu_{2,j_3^2+1}>\tau_3-\tau_4\}P\left\{\sum_{j=1}^{j_4}\nu_{2,j}=\tau_4\right\}d\tau_4 \cdot P\left\{\sum_{j=1}^{j_3^1}\nu_{1,j}=\tau_3\right\}d\tau_3.$$

Table 8. The accuracy of the approximations in fill rates in the battery supply chain.

		$\beta=0.85$, 2 (1)-moment (%)	0.90, 2 (1)-moment (%)	0.95, 2 (1)-moment (%)
$n=4$	Max:	2.80 (11.5)	2 (9.1)	1.50 (5.6)
	Avg:	1.80 (7.6)	1.07 (5.7)	0.73 (3.17)
9	Max:	3.50 (9.67)	2.50 (7.3)	1.40 (4.3)
	Avg:	1.90 (5.7)	1.40 (4)	0.85 (1.9)
16	Max:	3.50 (9)	3.16 (6)	1.70 (3.4)
	Avg:	2.26 (4.4)	1.76 (2.74)	1 (1.2)

(3) If $j_4 > j_3^2$, then

$$\Omega_T = \int_{t_3}^{\infty}\int_0^{\tau_3} \Pr\left\{\nu_{2,\,j_3^2+1} > \tau_3-\tau_4,\ \sum_{j=j_3^2+1}^{j_4}\nu_{2,\,j} \geqslant t_4-\tau_4\right\} \cdot P\left\{\sum_{j=1}^{j_3^2}\nu_{2,\,j}=\tau_4\right\} d\tau_4 P\left\{\sum_{j=1}^{j_3^1}\nu_{1,\,j}=\tau_3\right\} d\tau_3.$$

One can compute this probability by further conditioning on $\nu_{2,\,j_3^2+1}$.

For $l=2$, we consider three subcases:

(1) If $j_4 < j_3^2$, then by Equation (12),

$$\Omega_T = \int_{t_4}^{\infty}\int_{(t_3-\tau_4)^+}^{\infty} \Pr\left\{\sum_{j=1}^{j_3^1}\nu_{1,\,j} < \tau_3+\tau_4 < \sum_{j=1}^{j_3^1+1}\nu_{1,\,j}\right\} \times P\left\{\sum_{j=j_4+1}^{j_3^2}\nu_{2,\,j}=\tau_3\right\} d\tau_3 P\left\{\sum_{j=1}^{j_4}\nu_{2,\,j}=\tau_4\right\} d\tau_4.$$

(2) If $j_4 = j_3^2$, then

$$\Omega_T = \int_{\max\{t_3,t_4\}}^{\infty} \Pr\left\{\sum_{j=1}^{j_3^1}\nu_{1,\,j} < \tau < \sum_{j=1}^{j_3^1+1}\nu_{1,\,j}\right\} P\left\{\sum_{j=1}^{j_3^2}\nu_{2,\,j}=\tau\right\} d\tau.$$

(3) If $j_4 > j_3^2$, then

$$\Omega_T = \int_{t_3}^{\infty} \Pr\left\{\sum_{j=1}^{j_3^1}\nu_{1,\,j} < \tau < \sum_{j=1}^{j_3^1+1}\nu_{1,\,j}\right\} \cdot \Pr\left\{\sum_{j=j_3^2+1}^{j_4}\nu_{2,\,j} \geqslant t_4-\tau\right\} \times P\left\{\sum_{j=1}^{j_3^2}\nu_{2,\,j}=\tau\right\} d\tau.$$

Proof of Proposition 2.1. To prove inequality (24), we condition on the transit times and demand interarrival times. We first show that $X_i(m)$ is a nondecreasing function of the demand sizes. By Equations (3) and (4), $J_i(m)$ and $M_i(m)$ are independent of demand sizes if $m > s_i$. If $m \leqslant s_i$, then $J_i(m)=j$, where j is the smallest integer such that $\sum_{J=1}^{j} D_{i,\,J} > s_i - m$, and $M_i(m)=\sum_{J=1}^{j} D_{i,\,J}+m-s_i$. If we increase $D_{i,\,J}$ for $1 \leqslant J \leqslant j$, then we have two cases: (1) $J_i(m)$ does not change but $M_i(m)$ increases. Hence, by the transit time assumption, $\widetilde{L}_i(M_i(m))$ and $X_i(m)$ are nondecreasing in $D_{i,\,J}$. (2) $J_i(m)$ decreases, which means that the order is delayed. By the transit time assumption, $X_i(m)$ is again nondecreasing in $D_{i,\,l}$. Because all other demand sizes have no impact on $X_i(m)$, $X_i(m)$ is a nondecreasing function of the demand sizes.

Because the demand sizes are independent, they are associated random variables (Tong 1980, Theorem 5.2.2). Because $X_i(m)$, $(i,k) \in \mathcal{A}$ are nondecreasing functions of the demand sizes, they are also associated random variables (Tong 1980, Theorem 5.2.3). Unconditioning on the transit times and interarrival times, we obtain inequality (24) by Theorem 5.2.4 of Tong (1980).

To prove inequality (25), we condition on the transit times and demand sizes. Because $T_i(J_i(m))$ is nondecreasing in demand interarrival times, $X_i(m)$ is nonincreasing in demand interarrival times. By the same arguments as those for inequality (24), we obtain inequality (25). □

We utilize a two-moment approximation for X_k (see Zipkin 1991): Given $E(\widetilde{L}_k)$ and $V(\widetilde{L}_k)$, s_k and the demand process λ_k and D_k, we compute $E(X_k)$ and $V(X_k)$ as follows:

(1) Compute the mean and variance of the lead-time demand, Y_k, where

$$E(Y_k) = E(\widetilde{L}_k)\lambda_k E(D_k), \tag{45}$$

$$V(Y_k) = \lambda_k E(D_k^2)E(\widetilde{L}_k) + (\lambda_k E(D_k))^2 V(\widetilde{L}_k). \tag{46}$$

Then, fit the lead-time demand distribution by a negative binomial distribution, which matches the first two moments.

(2) Compute the expected backorders $E(B_k)$ and the variance of the backorders $V(B_k)$ by

$$B_k = (Y_k - s_k)^+. \tag{47}$$

(3) Finally, compute $E(X_k)$, $V(X_k)$, and $E(I_k)$ as follows:

$$E(X_k) = E(B_k)/\lambda_k/E(D_k), \tag{48}$$

$$V(X_k) = (V(B_k) - \lambda_k E(D_k^2)E(X_k))/(\lambda_k E(D_k))^2, \tag{49}$$

$$E(I_k) = s_k - E(Y_k) + E(B_k). \tag{50}$$

We refer the reader to Zipkin (2000) for more discussions.

For renewal batch demand, let ν_k be an arbitrary interarrival time for node k. It follows from Zipkin (2000, §C.2.3.8) and Kulkarni (1995, Theorem 8.7) that Equations (46) and (49) become

$$V(Y_k) = \lambda_k V(D_k)E(\widetilde{L}_k) + [\lambda_k^3 V(\nu_k)E(\widetilde{L}_k) + \lambda_k^2 V(\widetilde{L}_k)]E(D_k)^2, \tag{51}$$

$$V(X_k) = (V(B_k) - \lambda_k V(D_k)E(X_k))/(\lambda_k E(D_k))^2 - \lambda_k V(\nu_k)E(X_k). \tag{52}$$

Consider Example 1 in §4.2 where the interarrival time at node 1 (2) follows an Erlang distribution with λ_1 and $V(\nu_1)$ (λ_2 and $V(\nu_2)$). By the asymptotic method of Whitt (1982), we can approximate the demand process at node 6 by a renewal process with

$$\lambda_6 \approx \lambda_1 + \lambda_2, \tag{53}$$

$$V(\nu_6) \approx \frac{\lambda_1^3}{\lambda_6^3}V(\nu_1) + \frac{\lambda_2^3}{\lambda_6^3}V(\nu_2). \tag{54}$$

The demand-size distribution at node 6 is approximated by a mixture of the demand-size distributions at node 1 (with probability λ_1/λ_6) and node 2 (with probability λ_2/λ_6). Finally, for renewal process, Equation (43) becomes

$$V(\Delta_{k,\,n}) = V(\widetilde{L}_k) + E(J_k(n))V(\nu_k) + V(J_k(n))/\lambda_k^2. \tag{55}$$

Acknowledgments

The author thanks the associate editor and anonymous referees for constructive comments and suggestions that have significantly improved the paper. The author also thanks Paul Zipkin and other faculty members at Duke University for the helpful discussions and comments on an earlier version of this paper. This research was supported in part by a Faculty Research Grant from the Rutgers Business School, Newark and New Brunswick.

References

Agrawal, N., M. A. Cohen. 2001. Optimal material control in an assembly system with component commonality. *Naval Res. Logist.* **48** 409–429.

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

Axsater, S. 2000. Exact analysis of continuous review (R, Q) policies in two-echelon inventory systems with compound Poisson demand. *Oper. Res.* **48** 686–696.

Axsater, S. 2003. Supply chain operations: Serial and distribution inventory systems. A. G. de Kok, S. C. Graves, eds. *Handbooks in Operations Research and Management Science*, Chapter 10, Vol. 11. *Supply Chain Management*: *Design, Coordination and Operation*. North-Holland Publishing Company, Amsterdam, The Netherlands, 525–559.

Chen, F. 1998. On (R, NQ) policies in serial inventory systems. S. Tayur, R. Ganeshan, M. J. Magazine, eds. *Quantitative Model for Supply Chain Management*. Kluwer Academic Publishers, Boston, MA, 72–109.

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

Clark, C. E. 1961. The greatest of a finite set of random variables. *Oper. Res.* **9** 145–162.

de Kok, T. G., J. W. C. H. Visschers. 1999. Analysis of assembly systems with service level constraints. *Internat. J. Production Econom.* **59** 313–326.

Ettl, M., G. E. Feigin, G. Y. Lin, D. D. Yao. 2000. A supply network model with base-stock control and service requirements. *Oper. Res.* **48** 216–232.

Forsberg, R. 1995. Optimization of order-up-to-S policies for two-level inventory systems with compound Poisson demand. *Eur. J. Oper. Res.* **81** 143–153.

Graves, S. C., S. P. Willems. 2003. Supply chain design: Safety stock placement and supply chain configuration. A. G. de Kok, S. C. Graves, eds. *Handbooks in Operations Research and Management Science*, Chapter 3, Vol. 11. *Supply Chain Management*: *Design, Coordination and Operation*. North-Holland Publishing Company, Amsterdam, The Netherlands, 95–132.

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

Kulkarni, V. G. 1995. *Modeling and Analysis of Stochastic Systems*. Chapman & Hall, New York.

Lee, H. L., C. Billington. 1993. Material management in decentralized supply chains. *Oper. Res.* **41** 835–847.

Lesnaia, E. 2004. Optimizing safety stock placement in general network supply chains. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Lu, Y., J. S. Song. 2005. Order-based cost optimization in assemble-to-order systems. *Oper. Res.* **53** 151–169.

Lu, Y., J. S. Song, D. D. Yao. 2003. Order fill rate, lead-time variability and advance demand information in an assemble-to-order system. *Oper. Res.* **51** 292–308.

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

Sherbrooke, C. 1968. METRIC: A multi-echelon technique for recoverable item control. *Oper. Res.* **16** 122–141.

Sigman, K. 2001. Lecture notes on $GI/M/1$ queue. Teaching notes, Columbia University, New York.

Simchi-Levi, D., Y. Zhao. 2005. Safety stock positioning in supply chains with stochastic lead times. *Manufacturing Service Oper. Management* **7** 295–318.

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

Song, J. S. 2000. A note on assemble-to-order systems with batch ordering. *Management Sci.* **46** 739–943.

Song, J. S. 2002. Order-based backorders and their implications in multi-item inventory systems. *Management Sci.* **48** 499–516.

Song, J. S., D. D. Yao. 2002. Performance analysis and optimization in assemble-to-order systems with random lead times. *Oper. Res.* **50** 889–903.

Song, J. S., P. Zipkin. 2003. Supply chain operations: Assemble-to-order systems. A. G. de Kok, S. C. Graves, eds. *Handbooks in Operations Research and Management Science*, Chapter 11, Vol. 11. *Supply Chain Management*: *Design, Coordination and Operation*. North-Holland Publishing Company, Amsterdam, The Netherlands, 561–596.

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

Tong, Y. L. 1980. *Probability Inequalities in Multivariate Distributions*. Academic Press, New York.

Whitt, W. 1982. Approximating a point process by a renewal process: Two basic methods. *Oper. Res.* **30** 125–147.

Zhao, Y. 2006. Evaluation and optimization of installation base-stock policies in supply chains with compound Poisson demand. Working paper, Rutgers University, Newark, NJ.

Zhao, Y., D. Simchi-Levi. 2006. Performance analysis and evaluation of assemble-to-order systems with stochastic sequential lead times. *Oper. Res.* **54** 706–724.

Zipkin, P. 1991. Evaluation of base-stock policies in multiechelon inventory systems with compound-Poisson demands. *Naval Res. Logist.* **38** 397–412.

Zipkin, P. 2000. *Foundations of Inventory Management*. McGraw Hill, Boston, MA.