No-Holdback Allocation Rules for Continuous-Time Assemble-to-Order Systems

Yingdong Lu   Jing-Sheng Song   Yao Zhao

IBM Research, Yorktown Heights, NY 10598, yingdong@us.ibm.com
Fuqua School of Business, Duke University, Durham, NC 27708, jssong@duke.edu
Department of Supply Chain Management & Marketing Sciences, Rutgers University
Newark, NJ 07102, yaozhao@andromeda.rutgers.edu

Abstract

This paper analyzes a class of common-component allocation rules, termed *no-holdback* (NHB) rules, in continuous-review assemble-to-order (ATO) systems with positive lead times. The inventory of each component is replenished following an independent base-stock policy. In contrast to the usually assumed first-come-first-served (FCFS) component allocation rule in the literature, a NHB rule allocates a component to a product demand only if it will yield immediate fulfillment of that demand. We identify metrics as well as cost and product structures under which NHB rules outperform all other component allocation rules. For systems with certain product structures, we obtain key performance expressions and compare them to those under FCFS. For general product structures, we present performance bounds and approximations. Finally, we discuss the applicability of these results to more general ATO systems.

1 Introduction

This paper considers a multiproduct assemble-to-order (ATO) system, in which components are kept in stock, but final products are assembled only after customer orders are realized. Each product is characterized by a fixed bill-of-materials (BOM), and different products may share common components. The assembly times of the products are negligible relative to the component replenishment lead times. This kind of system allows manufacturers to build products upon customer orders and hence to eliminate unwanted finished-product inventory. As such, these systems have become increasingly popular in industry. Incidentally, the same multi-component, multi-product system structure also applies to mail-order and online retailers as well as to distributors of multiple finished products. In such businesses, customers may order different, but possibly overlapping, subsets of items in stock and customer satisfaction is based on the fulfillment of the entire order.

Despite the widespread use of ATO systems, our understanding of the inventory-service tradeoff in

these systems remains limited. For instance, because the stockout of one component in an order delays the fulfillment of the entire order, the component replenishment policy that minimizes the total cost should be correlated across components. Also, because a common component can be used to fulfill several product orders, how one allocates that component among the products determines customer waiting times. Thus, a reasonable allocation rule should also be state dependent. However, we have little knowledge at present about the structures of the these replenishment and allocation policies; see §2 for more detail. As a result, only simple, heuristic control policies are implemented in practice, and the majority of the academic literature focuses on performance evaluation and optimization of specific types of control policies; see Song and Zipkin (2003) for a review. In particular, most prior works assume independent base-stock component replenishment policies, as we do in the current paper. For convenience, we refer to such systems as *base-stock (ATO) systems*. For continuous-review base-stock systems, in which an allocation decision must be made at the time of a customer order and at the time of a component replenishment delivery, most of the literature assumes the first-come-first-served (FCFS) allocation rule and studies the performance evaluation and optimization of the base-stock policies.

In the current paper, we study a continuous-review base-stock ATO system with several features common in the literature – Poisson demand processes, positive lead times, and complete backlogging. Unlike in previous works, we analyze the performance of a general class of non-FCFS allocation rules, termed *no-holdback* (NHB) rules in the base-stock system. For any given base-stock policy, we compare the effectiveness of NHB rules with that of FCFS and other allocation rules whenever possible. We show that, in some situations, NHB rules outperform any other allocation rules in the base-stock system.

Under the FCFS allocation rule, demand for each component is fulfilled in exactly the same sequence as it occurs. This is a *component-based* allocation rule, because the allocation of one component's inventory does not depend on the inventory status of other components. When a product demand arrives, if some of its components are available while others are not, the available components are put aside as committed stock. When a component is replenished, it is committed to the oldest backordered product that requires it. The drawback to this rule is that it commits available components to backlogged demand, creating extra waiting time for both the available components and for other demand. The virtue of this rule is the simplicity of its information requirements and implementation. Moreover, it is perceived to be fair to the customers and it guarantees order fulfillment within a certain time window.

By contrast, under the NHB rules, a demand is backordered if and only if at least one of its components is out of stock. An example is the *first-ready-first-served* (FRFS) rule, under which product demand is filled following the FCFS rule as long as all components demanded are available. However, when a product demand arrives and some of its components are not available, we assign the demand to

the last position in the backorder list and do not allocate or commit those available components to the order. That is, a unit of component inventory will not be allocated to an incoming order, unless such an allocation results in the completion of the order fulfillment. When a replenishment shipment arrives, we review the backorder list and satisfy the oldest backorder for which all required components are available. Under other NHB rules, one can satisfy the backorders for which all required components are available in different sequences. An NHB rule is a *product-based* allocation rule; the allocation of one component's inventory to a demand depends on the inventory status of other components. Consequently, it requires more information and more coordination than FCFS. It also may make some orders wait longer than under FCFS.

Both FCFS and FRFS are commonly seen in practice. Kapuscinski et al. (2004) describe an example of FRFS at Dell and Xu et al. (2006) describe an example of FCFS at Amazon.com.

Using a sample path argument, we show that, for base-stock systems with certain material-flow topologies, NHB rules minimize both the total backorder and the total on-hand inventory. The minimum is achieved for each sample path at any time, similar to the concept of *global optimality* studied in fluid networks; see, for example, Chen and Yao (1993). In addition, for these systems, assuming certain symmetric cost structures, NHB rules are optimal among all allocation rules in the sense that they minimize system-wide average inventory and backorder costs. The special product structures include the system-W in Figure 1 and its generalizations, where system-W comprises three components and two products, and each product comprises one product-specific component and one common component shared by both products. These results are established by constructing a modified demand process which reduces the original ATO systems to a single-item inventory system. To our knowledge, this is the first set of optimality results for allocation rules in multi-product ATO systems with different lead times. Unfortunately, this result is not true for general systems, we provide counterexamples to show that, under other product structures, such as system-N and system-M in Figure 1, FRFS (or any NHB rule) neither minimizes the total backorder nor minimizes the total on-hand inventory on all sample paths.

While the performance evaluation of the FCFS rule in a base-stock ATO system is certainly very challenging and, until the past decade or so, was considered impossible, the performance analysis of a state-dependent allocation rule, such as FRFS, is arguably even less tractable. For instance, due to committed inventories under FCFS, it is possible to analyze the component inventory and product backorders. Based on this, previous research developed tools to evaluate the product-based performance measures through the analysis of multivariate Poisson processes. This is no longer true for FRFS or NHB in general. A different approach must be used. Our basic approach is a sample path analysis. By examining flow conservation equations, we identify certain product structures for which we can
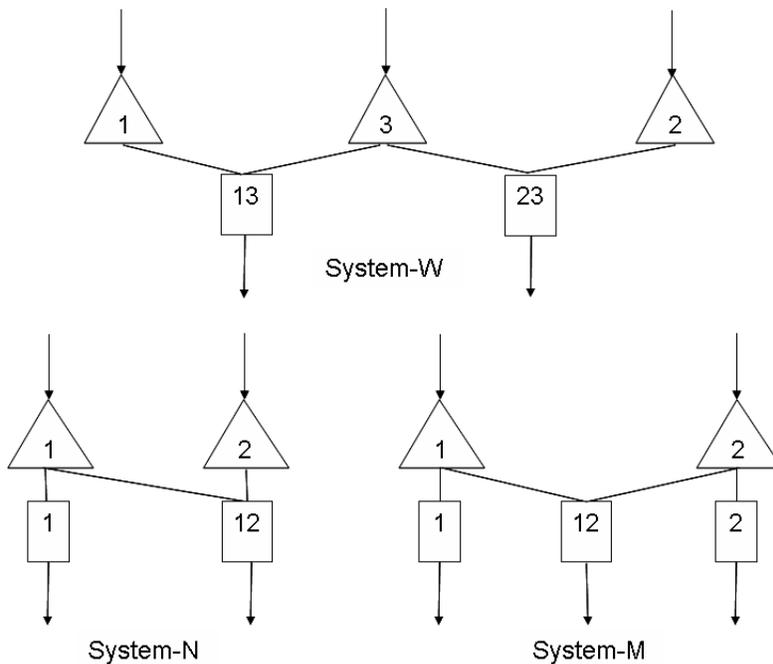
Figure 1: System-W, N, M.

obtain closed-form expressions for the total product backorders and order fill rates on any sample path. For the general system, we develop bounds and approximations for average component inventories and average backorders and order fill rates of individual products. Some of the bounds involve solving linear programs. The approximation of the average individual product backorders works very well in our numerical experiments.

Intuitively, FRFS may outperform FCFS in terms of the total backorders and order fill rates, and FCFS may outperform FRFS in terms of fill rates with a time window. It is of interest to confirm this intuition analytically and our analytical approaches provide a basis for doing so. In fact, these approaches allow us to provide a more accurate and complete comparison of the allocation rules. In particular, for any given base-stock policy, we prove that all NHB rules outperform FCFS on immediate fill-rates, but that FCFS may outperform an NHB rule on fill rates within a time window. We also prove that FRFS outperforms FCFS on the total backorders for system-W, N and M, but that FCFS may outperform FRFS on the backorders of some individual products.

The rest of the paper is organized as follow. §2 reviews the related literature. §3 introduces the basic notation, definition of NHB rules and performance measures. §4 presents the optimality results and discusses related issues. §5-§7 develops tools for evaluating system performance measures under any given base-stock policy and NHB allocation rule and compares the effectiveness of NHB and FCFS

whenever possible. §8 discusses extensions of the results to general replenishment policies, lead time models and demand processes, as well as general BOM matrices. Finally, §9 concludes the paper.

2   Literature Review

As mentioned above, inventory control of an ATO system consists of two decisions: inventory replenishment and inventory allocation. In general, the optimal decisions for any reasonable objective function are state-dependent; see Song and Zipkin (2003). For example, Benjaafar and ElHafsi (2006) considered a special case: an assembly system with a *single* end-product and multiple demand classes. Each component supplier has a finite production capacity. They show that under Markovian assumptions on demand and production, the optimal replenishment policy is a state-dependent base-stock policy, and the optimal allocation rule is a multi-level rationing policy that depends on the inventory levels of all other components. Dogru et al. (2007) considered a special case of the W system in which all components have identical constant lead times. They show that when a symmetric cost condition or a "balanced capacity" condition is satisfied, the optimal replenishment policy is a state independent base-stock policy, and the optimal allocation policy is an NHB policy (which they call *myopic policies*). Clearly, the optimal policy for the general *multi-product* ATO system having different component lead times, such as the one we consider here, is likely more complex. In fact, neither the optimal replenishment policy nor the optimal allocation rule is known for the general ATO system. In practice, only simple but possibly suboptimal replenishment and allocation policies are implemented. Correspondingly, the academic literature on ATO systems primarily focuses on these simple policies. In particular, most papers assume a certain allocation rule, and then evaluate and/or optimize the replenishment policy. Our paper also focuses on specific types of control policies. For the given type of replenishment policy, we evaluate the performance of NHB allocation rules and show that under certain conditions NHB rules are optimal in the allocation policy space.

To our knowledge, most papers that address continuous-review models (the focus of the present article) assume the FCFS allocation rule, and develop tools to evaluate and/or optimize the independent base-stock replenishment policies. See, for example, Song (1998, 2002), Song et al. (1999), Song and Yao (2002), Lu et al. (2003, 2005), and Zhao and Simchi-Levi (2006). For periodic-review models, because the demand during each period is batched and filled at the end of the period, different allocations rules have been considered. These include a fixed priority rule (Zhang, 1997), the FCFS rule (Hausman et al. 1998), and a fair-share rule (Agrawal and Cohen, 2001). Akcay and Xu (2004) studied a product-based allocation rule that makes optimal or near-optimal allocation decisions within each period. However, all these prior studies apply FCFS to demand between periods. In other words, these alternative allocation

5

rules, if implemented in a continuous-review environment, reduce to FCFS. Similarly, Bernstein et al. (2007) considered the W system in a single selling season that can be divided into multiple time buckets. At the end of each bucket, stock allocation decisions can be made. Their objective is to study the impact of demand aggregation on profits, sales and initial stock decisions. The model has neither replenishment nor lead time, so both the setting and the focus are quite different from the present paper. The policy choices in the asymptotic studies by Plambeck and Ward (2006a,b) on high-volume ATO systems are exceptions, but their setting also is quite different from ours. In their setting, the decision variables include the component production capacities, assembled product production sequence, and product prices. Once a component's production capacity is chosen, the production facility produces the component at full capacity, so there is no inventory decision.

Our paper differs from the studies cited above in several ways: (i) We analyze the base-stock ATO system under a different type of allocation rule – a NHB rule – that is non-FCFS. (ii) For any given base-stock policy, we compare the performances of NHB and FCFS whenever possible. (iii) We establish certain optimality properties of NHB rules over all allocation rules in the base-stock systems.

We are aware of only two other works that analyze non-FCFS rules in continuous-review ATO systems. First, Song and Zhao (2007) considered FRFS and the general NHB rules, but focused on a model with only one common component (including system-W), which is a special case of the model studied here. Also, these authors only studied performance evaluation of the fill rates, and did not consider backorders and optimality conditions. Second, Dogru et al. (2007) incidentally noticed the general class of the NHB rules. In particular, they considered a priority-based backlog clearing policy as well as a reservation policy. However, as mentioned above, their focus is the W system with equal constant lead times. The system considered here is much more general, which can include nonidentical constant and stochastic lead times. For such a system, we establish the optimality of the NHB rules assuming certain replenishment policies. Dogru et al. (2007) proved results on the joint optimal replenishment and allocation policy for their model, but these results do not carry over to the general system considered here. Interestingly, both studies identify similar cost structures to guarantee optimality, even though the approaches taken are very different. The analysis of Dogru et al. (2007) is based on a stochastic programming (SP) formulation and a lower bound approach, whereas we use a sample path approach, supplemented with a system reduction technique that reduces the original ATO systems to a single-item inventory system.

Finally, there is a large body of literature on single-item inventory systems with multiple demand classes; see, for example, Deshpande et al. (2003) and references therein. These systems can be viewed as special cases of an ATO system containing only a single common component but no product-specific

components. Thus, there is no assembly issue. In particular, the fulfillment of a demand does not need to consider the availability of the product-specific components. Most papers in this literature study variants of a rationing (or reservation) policy, which is a hold-back rule, whereas we consider NHB rules.

3 Model Description and Preliminaries

In this section, we introduce the basic model setting. In §8 we will discuss several generalizations.

3.1 Demand and Supply Processes

We consider a continuous-review assemble-to-order system of m different components indexed by $i \in \mathcal{I} = \{1, 2, \cdots, m\}$. We use $K \subset \mathcal{I}$ to index a product if the product consists of one unit of each item in K and 0 units in $\mathcal{I} \setminus K$ (general BOM matrices are discussed in §8). The demand process $\{D^K(t), t \geq 0\}$ for product K forms a Poisson process with rate λ^K, which is independent of the demand process of any other product and of all other events. Let $\mathcal{K}$ be the set of all products. For each component i, let $\mathcal{K}_i$ denote the set of products that require component i. Clearly, the demand process for component i, $\{D_i(t) = \sum_{K \in \mathcal{K}_i} D^K(t), t \geq 0\}$ is a Poisson process with rate $\lambda_i = \sum_{K \in \mathcal{K}_i} \lambda^K$.

Demands are filled based on an allocation rule, such as, FCFS, FRFS, or other NHB rules, as described in §1. A demand is considered backlogged until it is satisfied completely; partial fulfillment is not allowed.

For each component i, an independent base-stock policy with base-stock level s_i (a nonnegative integer) is followed to control inventory replenishment. That is, whenever the inventory position (= net inventory + inventory on-order) is below s_i, we place an order to bring the inventory position back to s_i. Otherwise, we do not order. Assuming that the initial on-hand inventory of component i equals s_i, this is a one-for-one ordering policy: the arrival of each demand requiring component i triggers a replenishment order of one unit for this component. Note that the component replenishment policy $\mathbf{s} = (s_1, ..., s_m)$ is independent of the component allocation rule.

Let L_i be the replenishment lead time for component i. For simplicity, for most of this paper, we assume that L_i is constant but not necessarily identical for each component.

For any real numbers u and v, denote $u \vee v = \max\{u, v\}$, $u \wedge v = \min\{u, v\}$, and $u^+ = \max\{u, 0\}$.

Take any $t \geq 0$, denote $O_i(t)$ to be the number of outstanding orders of component i at time t. Because $O_i(t) = D_i(t) - D_i(t - L_i)$, it has a Poisson distribution with mean $\lambda_i L_i$. Let

$$B_i(t, s_i) \quad = \quad \text{total shortage of component } i \text{ at } t = [O_i(t) - s_i]^+.$$

Note that $B_i(t, s_i)$ depends only on the base-stock policy, not on the allocation policy. It is the cumulative

supply less the cumulative demand of component i by time t. Thus, depending on the type of allocation policy, a positive $B_i(t, s_i)$ may not imply zero on-hand inventory of component i at t. (It is actually the number of component i backorders at t under the FCFS rule.) Let O_i and $B_i(s_i)$ be the steady state limits of $O_i(t)$ and $B_i(t, s_i)$, respectively. Then O_i also has a Poisson distribution with mean $\lambda_i L_i$ and

$$B_i(s_i) = [O_i - s_i]^+.$$

The joint distribution of $(O_1(t), ..., O_m(t))$, as well as the existence and exact form of its stationary limit, have been established by Lu et al. (2003).

3.2 Allocation Rules and Performance Measures

For any given base-stock policy $\mathbf{s}$, we are interested in comparing the performance under different allocation rules. Denote Π to be the set of all possible allocation rules. Under any allocation rule $\pi \in \Pi$, let

$$
\begin{aligned}
I_i(t, \mathbf{s}, \pi) &= \text{on-hand inventory of component } i \text{ at time } t \\
B^K(t, \mathbf{s}, \pi) &= \text{number of product-}K \text{ backorders at } t \\
I_i(\mathbf{s}, \pi) &= \text{on-hand inventory of component } i \text{ in steady state} \\
B^K(\mathbf{s}, \pi) &= \text{number of product-}K \text{ backorders in steady state} \\
W^K(\mathbf{s}, \pi) &= \text{waiting time of a product-}K \text{ backorder in steady state} \\
f^{K,w}(\mathbf{s}, \pi) &= \text{product-}K \text{ fill rate with time window } w \\
&= \text{probability of satisfying a product-}K \text{ order within a time window } w \\
&= \mathsf{P}[W^K(\mathbf{s}, \pi) \leq w] \\
f^K(\mathbf{s}, \pi) &= \text{immediate fill rate of product-}K = f^{K,0}(\mathbf{s}, \pi).
\end{aligned}
$$

For a single-component product $K = \{j\}$ or a two-component product $K = \{i, j\}$, we simply denote B^K by B^j or B^{ij}, respectively. The same notation applies to other performance measures.

We also define $I(t, \mathbf{s}, \pi) = \sum_{i=1}^m I_i(t, \mathbf{s}, \pi)$ to be the total on-hand component inventory at time t. This is an appropriate measure when the component costs are the same. Let $I(\mathbf{s}, \pi) = \sum_{i=1}^m I_i(\mathbf{s}, \pi)$. Similarly, $B(t, \mathbf{s}, \pi) = \sum_{K \in \mathcal{K}} B^K(t, \mathbf{s}, \pi)$ is the total product backorder at t, and $B(\mathbf{s}, \pi) = \sum_{K \in \mathcal{K}} B^K(\mathbf{s}, \pi)$.

Let h_i be the unit inventory holding cost rate of component i and b^K be the unit backorder penalty cost rate for product K. Then, for any given base-stock policy $\mathbf{s}$ and any $\pi \in \Pi$, the total cost incurred at time t is

$$C(t, \mathbf{s}, \pi) = \sum_{i=1}^m h_i I_i(t, \mathbf{s}, \pi) + \sum_{K \in \mathcal{K}} b^K B^K(t, \mathbf{s}, \pi) \tag{1}$$

and the steady-state average system cost is

$$C(\mathbf{s}, \pi) = \sum_{i=1}^{m} h_i \mathsf{E}[I_i(\mathbf{s}, \pi)] + \sum_{K \in \mathcal{K}} b^K \mathsf{E}[B^K(\mathbf{s}, \pi)]. \tag{2}$$

In this paper, we sometimes pay particular attention to two allocation rules – FRFS and FCFS. For this reason, we shall use $\widehat{M}$ and $\widetilde{M}$ to denote the performance measure M under FRFS and FCFS, respectively. Examples include $\widehat{f}^{K,w}$ and $\widetilde{f}^{K,w}$, $\widehat{C}(\mathbf{s})$ and $\widetilde{C}(\mathbf{s})$.

Definition 1 *An allocation rule $\pi \in \Pi$ is a no-holdback (NHB) rule if the following condition is satisfied*

$$B^K(t, \mathbf{s}, \pi) * \min\{I_i(t, \mathbf{s}, \pi), i \in K\} = 0 \quad \text{for all } K \in \mathcal{K}, t \geq 0. \tag{3}$$

That is, a demand for product K is backordered if and only if at least one of its components runs out of on-hand inventory. Any rule which does not satisfy this condition is called a holdback rule. We denote $\mathcal{N}$ to be the set of all NHB rules.

Clearly, FRFS is an NHB rule but FCFS is a holdback rule. Other examples of a holdback rule include rationing policies and rules that dedicate stock to each product.

3.3 Flow Conservation Equations

We now describe the flow conservation equations that are essential to our analysis. Consider any component $i \in \mathcal{I}$. Suppose at time $t = 0$, the on-hand stock level is s_i and there is no backorder. For any $t > 0$, because a demand always triggers a replenishment order instantaneously, the total number of replenishment orders placed for component i by t is $D_i(t)$. Note that, except for the outstanding orders $O_i(t)$, all other orders have arrived by t, and the total inventory replenished by t is $D_i(t) - O_i(t)$. Under any allocation policy π, because each backorder $B^K(t, \mathbf{s}, \pi), K \in \mathcal{K}_i$ does not consume a stock of component i, the on-hand inventory of component i at time t is

$$I_i(t, \mathbf{s}, \pi) \;\; = \;\; s_i + D_i(t) - O_i(t) - [D_i(t) - \sum_{K \in \mathcal{K}_i} B^K(t, \mathbf{s}, \pi)] = s_i - O_i(t) + \sum_{K \in \mathcal{K}_i} B^K(t, \mathbf{s}, \pi). \tag{4}$$

Because a shortage of item i must result in a backordered demand, $I_i(t, \mathbf{s}, \pi) \geq 0$, we obtain

$$\sum_{K \in \mathcal{K}_i} B^K(t, \mathbf{s}, \pi) \geq B_i(t, s_i) = (O_i(t) - s_i)^+, \qquad i = 1, 2, \ldots, m. \tag{5}$$

Note that Eqs. (4)-(5) hold for any allocation rule π.

Now consider any NHB rule $\pi \in \mathcal{N}$. Note that by condition (3), $I_i(t, \mathbf{s}, \pi) > 0, \forall i \in K$ cannot hold simultaneously with $B^K(t, \mathbf{s}, \pi) > 0$. So either $B^K(t, \mathbf{s}, \pi) = 0$ or $B^K(t, \mathbf{s}, \pi) > 0$ and there exists $i \in K$

such that $I_i(t, \mathbf{s}, \pi) = 0$. By Eq. (4), $I_i(t, \mathbf{s}, \pi) = 0$ implies $B^K(t, \mathbf{s}, \pi) \leq (O_i(t) - s_i)^+$ for any $K \in \mathcal{K}_i$. Therefore, we must have

$$B^K(t, \mathbf{s}, \pi) \leq \max_{i \in K}\{(O_i(t) - s_i)^+\} = \max_{i \in K}\{B_i(t, s_i)\}, \qquad K \in \mathcal{K}, \quad \pi \in \mathcal{N}. \tag{6}$$

4 Optimality Results

In this section, we identify system structures and performance metrics under which FRFS (and any NHB rule) is optimal among all allocation rules. These results are inspired by and build on some properties of the single-item system presented in §4.1. We then provide counterexamples for the cases that the optimality conditions fail.

4.1 Single-Component Systems

In the special case of one-component systems ($m = 1$), we drop the index of the component and product, so condition (3) for the NHB rules reduces to

$$B(t, s, \pi) * I(t, s, \pi) = 0, \qquad \pi \in \mathcal{N}, t \geq 0, \tag{7}$$

where s is the base-stock level. Clearly, the traditional FCFS and LCFS (last-come-first-served) all satisfy this condition. Since there is only one component, FRFS is identical to FCFS for this system. A rationing policy is not an NHB rule.

We observe the following properties (see Appendix for a proof):

Lemma 1 *In a single-component system ($m = 1$), under any given inventory replenishment policy that is independent of allocation rules:*

(a) *All NHB rules have identical on-hand inventory and identical number of backorders on any sample path. In particular,*

$$I(t, \pi) = \widehat{I}(t) = \widetilde{I}(t), \qquad B(t, \pi) = \widehat{B}(t) = \widetilde{B}(t), \qquad t \geq 0, \quad \pi \in \mathcal{N}. \tag{8}$$

(Note that, for ease of exposition, we have omitted the sample path notation "ω" from the above expressions. This omission is followed throughout the paper.)

(b) *The FCFS rule has no more backorders and no more on-hand inventory than any hold-back rule on any sample path. That is,*

$$I(t, \pi) \geq \widehat{I}(t) = \widetilde{I}(t), \qquad B(t, \pi) \geq \widehat{B}(t) = \widetilde{B}(t), \qquad t \geq 0, \quad \pi \in \Pi. \tag{9}$$

(c) *Among all possible rules, an NHB rule minimizes both the total number of backorders and the total on-hand inventory on any sample path. Thus, $\widehat{C}(s) = \widetilde{C}(s) = \min\{C(s, \pi), \ \pi \in \Pi\}$.*

10

4.2 System-W and Its Generalizations

In this subsection, we focus on a special case: system-W and its generalizations. System-W is a three-component, two-product ATO system, where product-$\{i, 3\}$ requires components i and 3, $i = 1, 2$. Thus components 1 and 2 are product-specific while component 3 is a common component (see Figure 1). Our goal is to decompose this system into three single-item systems, one for each component, and show that the allocation rule only affects the system containing component 3 with an appropriately constructed demand process. In this way, we reduce the analysis of the original system to that of a single-item system.

The following observation for general ATO systems is useful for the system reduction. Let the phrase "first-available-first-use" (FAFU) denote filling demand with items in the sequence as the items become available (i.e., as they arrive at the inventory location).

Observation 1 *Consider an ATO system with a replenishment policy, supply (lead times) and demand processes independent of allocation rules. Given any allocation rule, one can always modify the rule in the following two ways without changing the backorder of each product and on-hand inventory of each item: (a) the demand for each product is filled on a FCFS basis, which implies that the backorders of each product are also filled on a FCFS basis, and (b) the inventory of each component is withdrawn on a FAFU basis.*

Proof. The proof involves a straightforward swapping of demands or inventory items. For simplicity, we omit the details. □

Note that Observation 1 (a) does not imply demand for all products can be satisfied on a FCFS basis, because different products have different BOMs. That is, a demand for product 1 can be filled before a demand for product 2, even though the latter may have arrived before the former. Observation 1 (b) implies that the product-specific components can be committed (or assigned) to demand on a FCFS basis. Therefore, we only need to consider how to allocate common-component inventories.

Now, consider a system-W under any given base-stock policy $\mathbf{s}$ and allocation policy π. Because a demand can be satisfied only when both the common component and its product-specific component are available, at any given t, the backorders and on-hand inventory of the system can be characterized by the following four elements:

1. On-hand inventory of the product-specific components not yet assigned to a demand. (The target demand has not yet arrived.) Denote these $Y_i(t, \mathbf{s}, \pi)$, $i = 1, 2$.
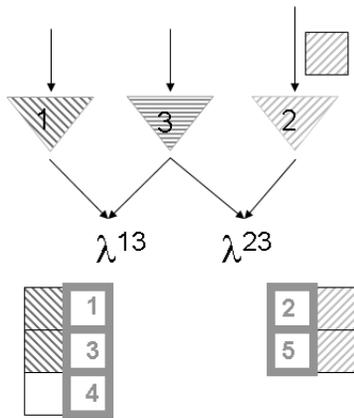
Figure 2: Four categories and modified demand process.

2. Backorders due to missing product-specific components $B_i^{i3}(t, \mathbf{s}, \pi)$, (which include backorders missing both components i and 3), $i = 1, 2$.

3. Backorders either because on-hand inventory of component 3 is missing or because the allocation rule commits the available component 3 inventory to other demand and their assigned product-specific component inventory items $(B_3^{i3}(t, \mathbf{s}, \pi), X_i(t, \mathbf{s}, \pi))$, where

$$X_i(t, \mathbf{s}, \pi) = B_3^{i3}(t, \mathbf{s}, \pi), \quad I_i(t, \mathbf{s}, \pi) = X_i(t, \mathbf{s}, \pi) + Y_i(t, \mathbf{s}, \pi), \quad i = 1, 2. \tag{10}$$

4. On-hand inventory of the common component $I_3(t, \mathbf{s}, \pi)$.

Note that only elements 3 and 4 depend on the common component allocation rule. Consequently, to study the impact of the allocation rules for this system, we only need to consider a single-item inventory system (termed *the reduced system-W*) consisting of the common component 3 and a *modified demand process for component 3* defined as follows: The modified arrival time of a demand is the maximum of this demand's original arrival time and the time when its product-specific component becomes available. A modified demand for component 3 thus refers to the combination of a demand and its assigned product-specific component. In fact, the modified demand process differs from the original demand process only when an arrival increases the element 2 (i.e., when a demand arrives and there is a unit of product-specific component ready but no component 3) and when there is a transition from element 2 to element 3 (i.e., when a backlogged order resulting from the absence of both components receives an allocation of the product-specific component).

Figure 2 provides a graphic illustration under FRFS. The demands are represented by the pairs of squares in front of their product-specific components; the squares with thick frames are the demands for the common component 3. The original demand arrivals are numbered from 1 to 5 chronologically, at

12

times $t_1 < ... < t_5$. The blank squares indicate missing component inventory, while the shaded squares indicate assigned inventory. This snapshot of the system at the arrival of demand 5, t_5, shows that ever since the arrival of demand 1, component-3 inventory has been missing. While component 2 has been available the entire time, component 1 has been one unit short since the arrival of demand 4. The square on top of the inventory location of component 2 indicates one unit of uncommitted component-2 inventory. By definition, $Y_1(t_5, \mathbf{s}, \pi) = 0$ and $Y_2(t_5, \mathbf{s}, \pi) = 1$; $B_1^{13}(t_5, \mathbf{s}, \pi) = 1$ but $B_2^{23}(t_5, \mathbf{s}, \pi) = 0$. Furthermore, $B_3^{13}(t_5, \mathbf{s}, \pi) = B_3^{23}(t, \mathbf{s}, \pi) = 2$. Finally, $I_3(t, \mathbf{s}, \pi) = 0$. Suppose the next event is the arrival of component 1 at t_6; then the modified demand process during $[t_1, t_6]$ is the arrival of demands 1,2,3,5,4 at times t_1, t_2, t_3, t_5, t_6, respectively.

We are now ready to show:

Theorem 1 (**System-W**) *Consider a system-W under any given base-stock policy* $\mathbf{s}$:

(a) (**Equivalence**) *All NHB rules have identical total on-hand inventory and identical total backorder on any sample path. In particular,* $I(t, \mathbf{s}, \pi) = \widehat{I}(t, \mathbf{s})$, $B(t, \mathbf{s}, \pi) = \widehat{B}(t, \mathbf{s})$, *for all* $t \geq 0$ *and* $\pi \in \mathcal{N}$.

(b) (**Total Inventory/Backorder**) *Among all possible rules, any NHB rule minimizes both the total on-hand inventory and the total backorder on any sample path. That is,* $I(t, \mathbf{s}, \pi) \geq \widehat{I}_i(t, \mathbf{s})$ *and* $B(t, \mathbf{s}, \pi) \geq \widehat{B}(t, \mathbf{s})$, *for all* $t \geq 0$ *and* $\pi \in \Pi$. *In particular,* $\widetilde{I}(t, \mathbf{s}) \geq \widehat{I}(t, \mathbf{s})$ *and* $\widetilde{B}(t, \mathbf{s}) \geq \widehat{B}(t, \mathbf{s})$, $t \geq 0$.

(c) (**Optimality Condition**) *If* $b^{13} + h_1 = b^{23} + h_2$, *then any NHB rule is the optimal allocation rule that minimizes the system-wide inventory-holding and backorder-penalty cost among all possible rules on any sample path. That is,* $\widehat{C}(t, \mathbf{s}) = \min_{\pi \in \Pi} C(t, \mathbf{s}, \pi)$, $t \geq 0$, *and* $\widehat{C}(\mathbf{s}) = \min_{\pi \in \Pi} C(\mathbf{s}, \pi)$.

Proof. (a) In the reduced system-W, consider any NHB rule that allocates the common component to the modified demand process. Then, the same rule is also an NHB rule for the common component in the original system-W. Conversely, any NHB rule for the common component in system-W is also an NHB rule for the reduced system-W.

In the reduced system-W, it follows from Lemma 1 (a) that all NHB rules have identical on-hand inventory of the common component and an identical number of backorders of the modified demand along any sample path. Because elements 1 and 2 are independent of the common component allocation rule, the result follows from the fact that different products require the same number of product-specific components.

(b) For the reduced system-W with the common component and the modified demand process, consider a NHB rule that satisfies the modified demand process on a FCFS basis. Clearly, this is an

NHB rule for the common component in the original system-W. By part (a), this rule yields an identical total number of backorders and on-hand inventory as does any other NHB rule (including FRFS) in system-W. For convenience, we call this rule θ.

For any given holdback rule for the common component in system-W, consider its sample path.

- **Case 1**: If demands are satisfied in the same sequence as under θ, due to the holdback feature, this policy must hold back the common component inventory for some time before following θ. Therefore it must result in no less total inventory and no fewer total backorders than θ.

- **Case 2**: If demands are not satisfied in the sequence as under θ, then we can modify the rule so that demands are satisfied in the sequence as under θ without adversely affecting performance. To see this, consider two demands in the modified demand process where demand 1 arrives at $t_1 > 0$ while demand 2 arrives at $t_2 > t_1$. Suppose demand 2 is satisfied at $t_3 > t_2$ and demand 1 is satisfied at $t_4 > t_3$. Because different products require the same number of product-specific components, we can use the common component that satisfies demand 2 (or 1) to satisfy demand 1 (or 2) at t_3 (or t_4) without changing either the total component inventory or the total number of backorders. After the swapping, we return to **Case 1**.

Finally, we note that elements 1 and 2 are independent of allocation rules. The proof is thus completed. (Note that θ is not necessarily equivalent to FRFS because the backorders may be filled in different sequences under these rules.)

(c) Because elements 1 and 2 are independent of allocation rules, we can subtract their associated costs from the total cost and consider only the holding cost for the common component and the cost for backordering a demand and holding its product-specific component. That is, for any given t, the relevant cost is

$$
\begin{aligned}
R(t, \mathbf{s}, \pi) &= h_3 I_3(t, \mathbf{s}, \pi) + \sum_{i=1}^{2} [h_i X_i(t, \mathbf{s}, \pi) + b^{i3} B_3^{i3}(t, \mathbf{s}, \pi)] \\
&= h_3 I_3(t, \mathbf{s}, \pi) + \sum_{i=1}^{2} (h_i + b^{i3}) B_3^{i3}(t, \mathbf{s}, \pi).
\end{aligned}
$$

Applying Lemma 1 (c) to the reduced system-W, we know that FRFS (and any NHB rule) minimizes both the on-hand inventory of the common component and the backorders of the modified demand on each sample path. So

$$
I_3(t, \mathbf{s}, \pi) \geq \widehat{I}_3(t, \mathbf{s}) \quad \text{and} \quad \sum_{i=1}^{2} B_3^{i3}(t, \mathbf{s}, \pi) \geq \sum_{i=1}^{2} \widehat{B}_3^{i3}(t, \mathbf{s}).
$$

14

Because $b^{13} + h_1 = b^{23} + h_2$, we have

$$R(t, \mathbf{s}, \pi) \;\geq\; h_3 \widehat{I}_3(t, \mathbf{s}) + (h_1 + b^{13}) \sum_{i=1}^{2} \widehat{B}_3^{i3}(t, \mathbf{s}) = \widehat{R}(t, \mathbf{s}), \quad t \geq 0, \quad \pi \in \Pi,$$

proving the result. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

The optimality condition in part (c) of Theorem 1 implies that backordering a demand and holding its product-specific component(s) costs the same for all products. It is a generalization of the common assumption of homogenous products from single-item inventory systems to assembly systems.

Now we consider the general system-W in which the BOM set of each product-K satisfies: $K = A \cup S^K$, where A is a subset of components shared by all products, and S^K is the subset of components that are specific to product-K. Following a logic similar to that of system-W, we divide the backorders and on-hand inventory of a general system-W into five elements:

1. On-hand inventory of the product-specific components not yet assigned to a demand. Denote these $Y_i(t, \mathbf{s}, \pi)$, $i \in \bigcup_{K \in \mathcal{K}} S^K$.

2. Backorders due to missing at least one product-specific component $B_{S^K}^K(t, \mathbf{s}, \pi)$, $K \in \mathcal{K}$.

3. On-hand inventory of the product-specific components assigned to backorders in element 2, $Z_i(t, \mathbf{s}, \pi)$, $i \in \bigcup_{K \in \mathcal{K}} S^K$.

4. Backorders due to missing at least one common component or holdback allocation rules and the complete set of their assigned product-specific component inventory items $(B_A^K(t, \mathbf{s}, \pi), \{X_i(t, \mathbf{s}, \pi), i \in S^K\})$, where

$$X_i(t, \mathbf{s}, \pi) = B_A^K(t, \mathbf{s}, \pi), \quad I_i(t, \mathbf{s}, \pi) = X_i(t, \mathbf{s}, \pi) + Y_i(t, \mathbf{s}, \pi) + Z_i(t, \mathbf{s}, \pi), \quad i \in S^K, K \in \mathcal{K}. \quad (11)$$

5. On-hand inventory of the common component $I_i(t, \mathbf{s}, \pi)$, $i \in A$.

Note that only elements 4 and 5 depend on the allocation rule for the common components. As with system-W, we can define the *modified demand process* for a general system-W as follows: The modified arrival time of a demand is the maximum of the demand's arrival time and the time when all its product-specific components become available. A modified demand for components in A thus refers to a demand and the complete set of its assigned product-specific components. Theorem 1 can be generalized as follows (see Appendix for a proof):

Theorem 2 *(**General System-W**) Consider the general system-W under any base-stock policy* $\mathbf{s}$*.

15

(a) If $b^K + \sum_{i \in S^K} h_i$ is identical across all K, then an NHB rule is optimal among all allocation rules in minimizing the system-wide costs.

(b) If $|S^K|$ (the size of S^K) is identical for all K, then among all possible rules, any NHB rule achieves the lowest total on-hand inventory and the fewest total backorders on any sample path.

(c) If $|S^K|$ is not identical, then among all possible rules, any NHB rule achieves the minimum total backorders on any sample path. In addition, the NHB rule that prioritizes demand for products with greater $|S^K|$ achieves the lowest total on-hand inventory among all NHB rules.

The simplest case of the general system-W is system-N. Applying Theorem 2, we know that a NHB rule is optimal if $b^1 = b^{12} + h_2$; a NHB rule achieves the minimum total backorders among all possible rules; and a NHB rule with priority given to demand for product-12 achieves the lowest total on-hand inventory among all NHB rules.

Note that one cannot generalize Theorem 2 (c) to all possible allocation rules because, even for system-W, a NHB rule does not always outperform a holdback rule with respect to the total on-hand inventory on all sample paths, §4.3 provides a counterexample.

If the general system-W is asymmetric, i.e., $b^K + \sum_{i \in S^K} h_i$ is product-dependent, Theorem 2 (a) does not hold. Different NHB rules may have different total costs. Clearly, a NHB rule (for the common components) that gives priority to backorders with larger $b^K + \sum_{i \in S^K} h_i$ results in lower total cost than FRFS.

4.3 Counterexamples

For an ATO system with product structure other than system-W, Theorem 1 may not hold, although some partial results hold. In this subsection, we present two counterexamples in which the W product structure is destroyed. In all the counterexamples, we assume the following initial condition: $I_i(0, \mathbf{s}, \pi) = s_i$ for all i and any allocation rule π. We denote t_l^K as the lth demand epoch of $\{D^K(t), t \geq 0\}$, and r_{il} the lth replenishment arrival epoch of component i. Let $r_1 = \min_i \{r_{i1}\}$.

The first example shows that FRFS (or any other NHB rule) does not necessarily minimize the total on-hand inventory along all sample paths for general systems.

Example 1 *Consider system-N (see Figure 1), which consists of two components and two products. Product 1 only requires component 1, product-$\{1, 2\}$ requires both components 1 and 2. Let $s_1 = 2$ and $s_2 = 1$. Consider a sample path where $0 < t_1^1 < t_2^1 < t_1^{12} < r_1$ and there are no other events in $[0, r_1]$.*

Then under FRFS,

$$\widehat{I}_1(t, \mathbf{s}) = 2, \quad \widehat{I}_2(t, \mathbf{s}) = 1, \quad 0 \le t < t_1^1$$

$$\widehat{I}_1(t, \mathbf{s}) = 1, \quad \widehat{I}_2(t, \mathbf{s}) = 1, \quad t_1^1 \le t < t_2^1$$

$$\widehat{I}_1(t, \mathbf{s}) = 0, \quad \widehat{I}_2(t, \mathbf{s}) = 1, \quad t_2^1 \le t < t_1^{12}$$

$$\widehat{I}_1(t, \mathbf{s}) = 0, \quad \widehat{I}_2(t, \mathbf{s}) = 1, \quad t_1^{12} \le t < r_1.$$

Consider a rationing policy δ that reserves one unit of the component 1 inventory for product-$\{1,2\}$, and apply FCFS otherwise. Then,

$$I_1(t, \mathbf{s}, \delta) = 2, \quad I_2(t, \mathbf{s}, \delta) = 1, \quad 0 \le t < t_1^1$$

$$I_1(t, \mathbf{s}, \delta) = 1, \quad I_2(t, \mathbf{s}, \delta) = 1, \quad t_1^1 \le t < t_2^1$$

$$I_1(t, \mathbf{s}, \delta) = 1, \quad I_2(t, \mathbf{s}, \delta) = 1, \quad t_2^1 \le t < t_1^{12}$$

$$I_1(t, \mathbf{s}, \delta) = 0, \quad I_2(t, \mathbf{s}, \delta) = 0, \quad t_1^{12} \le t < r_1.$$

Thus, on this sample path, FRFS can have a higher total on-hand inventory than δ. Note that δ is a holdback rule.

Next, we show that FRFS (or any other NHB rule) does not necessarily minimize the total number of backorders along all sample paths for general systems.

Example 2 *Consider system-M (see Figure 1): a two-component and three-product system, where product i requires component i only, $i = 1, 2$, and product=$\{1,2\}$ requires both components 1 and 2. $s_1 = 2$ and $s_2 = 2$. Consider a sample path where $0 < t_1^2 < t_1^1 < t_1^{12} < t_2^1 < t_2^2 < r_1$ and there are no other events in $[0, r_1]$. Then under FRFS,*

$$\widehat{B}^1(t, \mathbf{s}) = \widehat{B}^2(t, \mathbf{s}) = \widehat{B}^{12}(t, \mathbf{s}) = 0, \quad 0 \le t < t_1^2$$

$$\widehat{B}^1(t, \mathbf{s}) = \widehat{B}^2(t, \mathbf{s}) = \widehat{B}^{12}(t, \mathbf{s}) = 0, \quad t_1^2 \le t < t_1^1$$

$$\widehat{B}^1(t, \mathbf{s}) = \widehat{B}^2(t, \mathbf{s}) = \widehat{B}^{12}(t, \mathbf{s}) = 0, \quad t_1^1 \le t < t_1^{12}$$

$$\widehat{B}^1(t, \mathbf{s}) = \widehat{B}^2(t, \mathbf{s}) = \widehat{B}^{12}(t, \mathbf{s}) = 0, \quad t_1^{12} \le t < t_2^1$$

$$\widehat{B}^1(t, \mathbf{s}) = 1, \quad \widehat{B}^2(t, \mathbf{s}) = \widehat{B}^{12}(t, \mathbf{s}) = 0, \quad t_2^1 \le t < t_2^2$$

$$\widehat{B}^1(t, \mathbf{s}) = 1, \quad \widehat{B}^2(t, \mathbf{s}) = 1, \quad \widehat{B}^{12}(t, \mathbf{s}) = 0, \quad t_2^2 \le t < r_1.$$

Consider a rationing policy δ which reserves one unit of component i to product-i, and follows FCFS

otherwise. Then,

$$B^1(t, \mathbf{s}, \delta) = B^2(t, \mathbf{s}, \delta) = B^{12}(t, \mathbf{s}, \delta) = 0, \quad 0 \le t < t_1^2$$

$$B^1(t, \mathbf{s}, \delta) = B^2(t, \mathbf{s}, \delta) = B^{12}(t, \mathbf{s}, \delta) = 0, \quad t_1^2 \le t < t_1^1$$

$$B^1(t, \mathbf{s}, \delta) = B^2(t, \mathbf{s}, \delta) = B^{12}(t, \mathbf{s}, \delta) = 0, \quad t_1^1 \le t < t_1^{12}$$

$$B^1(t, \mathbf{s}, \delta) = B^2(t, \mathbf{s}, \delta) = 0, \quad B^{12}(t, \mathbf{s}, \delta) = 1, \quad t_1^{12} \le t < t_2^1$$

$$B^1(t, \mathbf{s}, \delta) = B^2(t, \mathbf{s}, \delta) = 0, \quad B^{12}(t, \mathbf{s}, \delta) = 1, \quad t_2^1 \le t < t_2^2$$

$$B^1(t, \mathbf{s}, \delta) = B^2(t, \mathbf{s}, \delta) = 0, \quad B^{12}(t, \mathbf{s}, \delta) = 1, \quad t_2^2 \le t < r_1.$$

Thus, on this sample path, we can have a total number of backorders under FRFS greater than that under δ. Note that δ is a holdback rule.

These examples show that when products differ in the number of components required or have different common components (a departure from the W structure), FRFS (or any NHB rule) may not outperform other rules in terms of minimizing total inventory or total number of backorders. For any given base-stock policy, if the objective is to reduce the total number of backorders, then a rationing policy that reserves inventory for products requiring the smallest amount of components may be better. On the other hand, if the objective is to reduce total on-hand inventory, then one may choose a policy that reserves inventory for products requiring the largest amount of components.

5 Performance Evaluation: Backorders

In this section, we focus on the evaluation of product backorders under FRFS, assuming a given base-stock policy $\mathbf{s}$. It is straightforward to extend the results to any other NHB rule. We first analyze a few systems with special product structures, where we present closed-form exact expressions for the total backorder (§5.1). We then provide bounds and an approximation for product backorders in systems with general product structures (§5.2). Finally, we compare FCFS and FRFS on the total number of backorders for system-M (§5.3).

5.1 Exact Results for Total Number of Backorders

This subsection presents closed-form exact expressions for the total number of backorders for some special systems. In particular, we obtain these closed-form expressions for system-W and systems with nested structure. Again, we would like to emphasize that these results hold for all sample paths. The proofs are included in Appendix.

5.1.1 System-W and Its Generalizations

Theorem 3 *(a) For system-W under any base-stock policy $\mathbf{s}$ and FRFS, we have*

$$\widehat{B}(t, \mathbf{s}) = \max\{B_3(t, s_3), B_1(t, s_1) + B_2(t, s_2)\},$$
$$\widehat{B}(\mathbf{s}) = \max\{B_3(s_3), B_1(s_1) + B_2(s_2)\}.$$

(b) For the general system-W under any base-stock policy $\mathbf{s}$ and FRFS, we have

$$\widehat{B}(t, \mathbf{s}) = \max\left\{\max_{i \in A} B_i(t, s_i), \sum_K \max_{i \in S^K}\{B_i(t, s_i)\}\right\},$$
$$\widehat{B}(\mathbf{s}) = \max\left\{\max_{i \in A} B_i(s_i), \sum_K \max_{i \in S^K}\{B_i(s_i)\}\right\}.$$

Lu et al. (2005) showed that the backorders for system-W under FCFS are given by $\widetilde{B}^{13}(t, \mathbf{s}) = \max\{B_3^{13}(t, \mathbf{s}), B_1(t, s_1)\}$ and $\widetilde{B}^{23}(t, \mathbf{s}) = \max\{B_3^{23}(t, \mathbf{s}), B_2(t, s_2)\}$ where $B_3^{13}(t, \mathbf{s}) + B_3^{23}(t, \mathbf{s}) = B_3(t, s_3)$. Thus, the total number of backorders under FCFS is $\max\{B_3^{13}(t, \mathbf{s}), B_1(t, s_1)\} + \max\{B_3^{23}(t, \mathbf{s}), B_2(t, s_2)\}$, which is greater than or equal to the total number of backorders under FRFS, $\max\{B_3(t, s_3), B_1(t, s_1) + B_2(t, s_2)\}$, on any sample path.

5.1.2 Nested Structure

Consider a system with n products. The BOM sets of products satisfy

$$\emptyset = K_0 \subset K_1 \subseteq K_2 \subseteq \cdots \subseteq K_n = \{1, 2, \cdots, m\}.$$

We call such a BOM structure a nested structure. The simplest system with a nested structure is system-N.

Theorem 4 *(a) For system-N under any given base-stock policy $\mathbf{s}$ and FRFS, we have:*

$$\widehat{B}(t, \mathbf{s}) = \max\{B_1(t, s_1), B_2(t, s_2)\}, \quad t \geq 0$$
$$\widehat{B}(\mathbf{s}) = \max\{B_1(s_1), B_2(s_2)\}.$$

(b) For a nested system under any given base-stock policy $\mathbf{s}$ and FRFS, we have:

$$\widehat{B}(t, \mathbf{s}) = \max_{1 \leq \ell \leq n}\{\max_{i \in K_\ell \backslash K_{\ell-1}} B_i(t, s_i)\} = \max_{1 \leq i \leq m}\{B_i(t, s_i)\}, \quad t \geq 0, \tag{12}$$
$$\widehat{B}(\mathbf{s}) = \max_{1 \leq i \leq m}\{B_i(s_i)\}. \tag{13}$$

5.2 Bounds and Approximations

For general ATO systems under base-stock policy $\mathbf{s}$ and FRFS, the exact expressions as in Theorems 3-4 cannot be expected. However, we can develop some bounds and an approximation for the expected number of backorders of individual products.

Let the notation $1_{\{i,j\in K\}} = 0$ mean that components i and j are not required by product K simultaneously, i.e., $K \notin \mathcal{K}_i \cap \mathcal{K}_j$. Consider any subset $S \subset \mathcal{I}$ such that for any pair $i, j \in S$, we have $1_{\{i,j\in K\}} = 0$ for all $K \in \mathcal{K}$, i.e., $\mathcal{K}_i \cap \mathcal{K}_j = \emptyset$. By Eq. (5),

$$\sum_{K\in\mathcal{K}_i} \widehat{B}^K(t,\mathbf{s}) \geq B_i(t,s_i), \quad i = 1, ..., m,$$

$$\sum_{K\in\bigcup_{i\in S}\mathcal{K}_i} \widehat{B}^K(t,\mathbf{s}) \geq \sum_{i\in S} B_i(t,s_i), \quad t \geq 0,$$

hold for every sample path and for all S. Taking the average over all sample paths, and denoting $x^K = \mathsf{E}[\widehat{B}^K(\mathbf{s})]$, $K \in \mathcal{K}$, and $\beta_i = \mathsf{E}[B_i(s_i)]$, $i = 1, 2, \ldots, m$. The following linear program (LP) gives a lower bound on $\mathsf{E}[\widehat{B}^{K_0}(\mathbf{s})]$ for each $K_0 \in \mathcal{K}$,

$$\min \quad x^{K_0} \tag{14}$$

$$s.t. \quad \sum_{K\in\bigcup_{i\in S}\mathcal{K}_i} x^K \geq \sum_{i\in S} \beta_i, \qquad \forall S,$$

$$\sum_{K\in\mathcal{K}_i} x^K \geq \beta_i, \qquad \forall i = 1, 2, \cdots, m,$$

$$x^K \geq 0, \qquad \forall K \in \mathcal{K}.$$

For system-W, we have $S = \{1, 2\}$. We can solve this LP easily and obtain a lower bound on $\mathsf{E}[\widehat{B}^{13}(\mathbf{s})]$, which is β_1. Similarly, a lower bound on $\mathsf{E}[\widehat{B}^{23}(\mathbf{s})]$ obtained from the LP is β_2. In other words,

$$\mathsf{E}[\widehat{B}^{13}(\mathbf{s})] \geq \mathsf{E}[B_1(s_1)], \quad \mathsf{E}[\widehat{B}^{23}(\mathbf{s})] \geq \mathsf{E}[B_2(s_2)]. \tag{15}$$

We now develop an upper bound for x^K. By Part (b) of Theorem 3, we know that for a general system-W with two products K_1 and K_2, the total number of backorders, $\widehat{B}^{K_1}(t,\mathbf{s}) + \widehat{B}^{K_2}(t,\mathbf{s})$, equals

$$\max\{\max_{i\in K_1\cap K_2} B_i(t,s_i), \max_{i\in K_1\backslash K_2} B_i(t,s_i) + \max_{j\in K_2\backslash K_1} B_j(t,s_j)\}.$$

But in a general ATO system, the right-hand side will only be larger with contributions from other products. Therefore, we have

$$\widehat{B}^{K_1}(t,\mathbf{s}) + \widehat{B}^{K_2}(t,\mathbf{s}) \leq \max\{\max_{i\in K_1\cap K_2} B_i(t,s_i), \max_{i\in K_1\backslash K_2} B_i(t,s_i) + \max_{j\in K_2\backslash K_1} B_j(t,s_j)\}.$$

Incorporating this inequality for any pair of two products as well as Eq. (6), the following LP gives an upper bound to $\mathsf{E}[\widehat{B}^{K_0}(\mathbf{s})]$, for each $K_0 \in \mathcal{K}$

$$\max \quad x^{K_0} \tag{16}$$

$$\text{s.t.} \quad \sum_{K \in \bigcup_{i \in S} \mathcal{K}_i} x^K \geq \sum_{i \in S} \beta_i, \qquad \forall S,$$

$$\sum_{K \in \mathcal{K}_i} x^K \geq \beta_i, \qquad i = 1, 2, \cdots, m,$$

$$x^{K_1} + x^{K_2} \leq \mathsf{E}\left[\max\{\max_{i \in K_1 \cap K_2} B_i(s_i), \max_{i \in K_1 \setminus K_2} B_i(s_i) + \max_{j \in K_2 \setminus K_1} B_j(s_j)\}\right], \qquad \forall K_1, K_2 \in \mathcal{K},$$

$$0 \leq x^K \leq \mathsf{E}\left[\max\{B_i(s_i), i \in K\}\right], \qquad \forall K \in \mathcal{K}.$$

For system-W, we can explicitly solve the above LP to obtain upper bounds on each of the product backorders, yielding

$$\mathsf{E}[\widehat{B}^{13}(\mathbf{s})] \quad \leq \quad \mathsf{E}[\max\{B_1(s_1), B_3(s_3) - B_2(s_2)\}], \tag{17}$$

$$\mathsf{E}[\widehat{B}^{23}(\mathbf{s})] \quad \leq \quad \mathsf{E}[\max\{B_2(s_2), B_3(s_3) - B_1(s_1)\}]. \tag{18}$$

These upper bounds are tighter than those by Eq. (6).

Finally, we can develop an approximation for the expected number of individual backorders by taking the average of the lower and upper bounds. For example, for system-W, using the above lower and upper bounds (15)- (18), we obtain

$$\mathsf{E}[\widehat{B}^{13}(\mathbf{s})] \quad \approx \quad \frac{1}{2}\left\{\mathsf{E}[B_1(s_1)] + \mathsf{E}[\max\{B_1(s_1), B_3(s_3) - B_2(s_2)\}]\right\},$$

$$\mathsf{E}[\widehat{B}^{23}(\mathbf{s})] \quad \approx \quad \frac{1}{2}\left\{\mathsf{E}[B_2(s_2)] + \mathsf{E}[\max\{B_2(s_2), B_3(s_3) - B_1(s_1)\}]\right\}.$$

Interestingly, adding up the above two approximations yields

$$\mathsf{E}[\widehat{B}(\mathbf{s})] \quad \approx \quad \mathsf{E}[\max\{B_3(s_3), B_1(s_1) + B_2(s_2)\}] = \mathsf{E}[\widehat{B}(\mathbf{s})].$$

Thus the approximation becomes exact for the average total number of backorders.

We now test the accuracy of these two bounds and the approximation in a numerical example. We consider a system-W with $\lambda^{13} = 7$ and $\lambda^{23} = 4$, and 100 randomly generated sets of base-stock levels following a uniform integer distribution between 0 and 10. The average number of backorders of product-$\{1, 3\}$ and product-$\{2, 3\}$ are obtained through simulation with percentage error ≤ 0.005. Figure 3 compares the simulation result, the bounds and the approximation for $\mathsf{E}[\widehat{B}^{13}]$ and $\mathsf{E}[\widehat{B}^{23}]$ over the 100 instances.

Interestingly, the figure shows that although the bounds can deviate significantly from the simulated values, they demonstrate a certain symmetric behavior – The bounds coincide with and deviate from
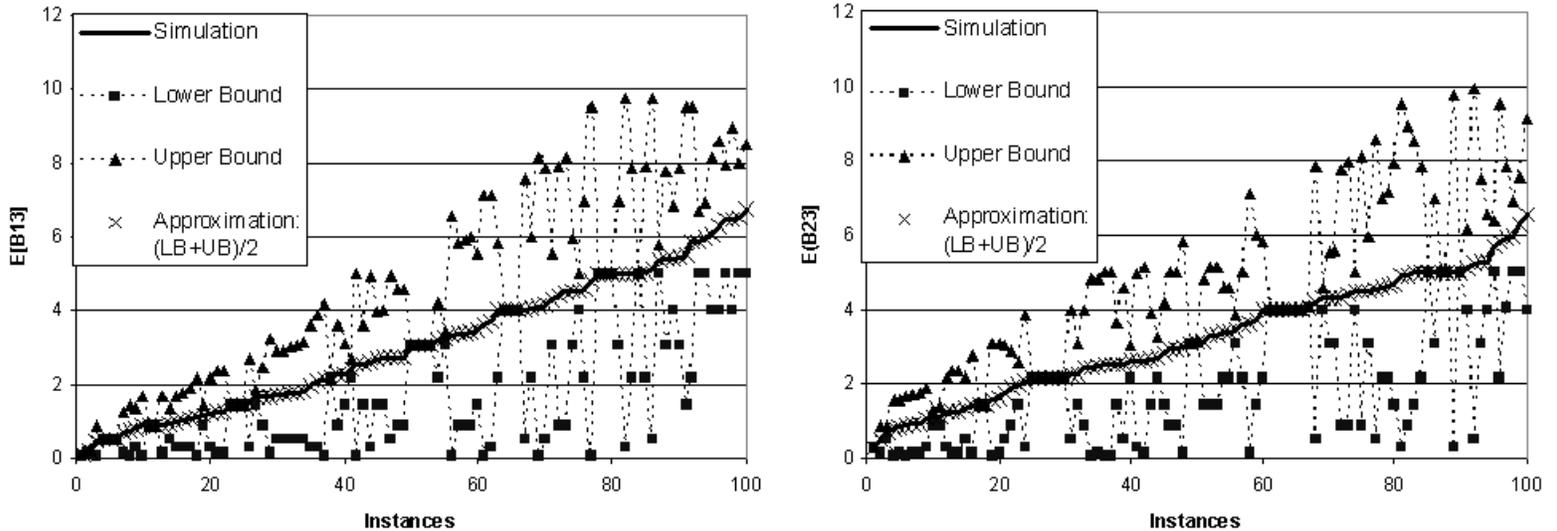
Figure 3: Bounds, approximations and simulation results

the simulation results at exactly the same instances. When deviating, they deviate roughly the same amount but in opposite directions. As a result, the approximation, which is the simple average of the bounds, is quite accurate. Indeed, one can hardly tell the difference between the approximate and the simulated values in the figure.

5.3   FCFS vs. FRFS on Total Backorders

By Theorems 1-2, we know that for the general system-W, under any given base-stock policy, FRFS has fewer total backorders than FCFS on any sample path. We now show that this is also true for system-M.

Let $\leq_{st}$ and $\geq_{st}$ denote *stochastically smaller than* and *stochastically greater than* in the sense of the usual stochastic order. We can show (see Appendix for a proof):

Theorem 5 *For system-M under any base-stock policy* $\mathbf{s}$, *we have*

(a) $\widehat{B}^{12}(t, \mathbf{s}) \geq \widetilde{B}^{12}(t, \mathbf{s})$ *on all sample paths. Consequently,* $\widehat{B}^{12}(\mathbf{s}) \geq_{st} \widetilde{B}^{12}(\mathbf{s})$.

(b) $\widehat{B}(t, \mathbf{s}) \leq \widetilde{B}(t, \mathbf{s})$ *on all sample paths. Consequently,* $\widehat{B}(\mathbf{s}) \leq_{st} \widetilde{B}(\mathbf{s})$.

Thus, while FCFS results in fewer product-12 backorders than FRFS due to committed stock, it underperforms FRFS on the total number of backorders. We point out that Theorem 5 (a) does not hold for all NHB rules, e.g., a NHB rule that gives priority to product-12 backorders.

22

6 Performance Evaluation: Average Inventory and Average Cost

We now discuss the evaluation of the average inventory and average cost. By Eq. (4), under any given base-stock policy $\mathbf{s}$ and NHB rule π, we have

$$I_i(\mathbf{s}, \pi) = s_i - O_i + \sum_{K \in \mathcal{K}_i} B^K(\mathbf{s}, \pi), \quad i \in \mathcal{I}.$$

Thus, as long as we can compute the average number of backorders of each product, we can evaluate the average inventory for each component:

$$\mathsf{E}[I_i(\mathbf{s}, \pi)] = s_i - \lambda_i L_i + \sum_{K \in \mathcal{K}_i} \mathsf{E}[B^K(\mathbf{s}, \pi)], \quad i \in \mathcal{I}.$$

Thus, one can readily compute the long-run average system cost $C(\mathbf{s}, \pi)$, expressed in Eq. (2), if $\mathsf{E}[B^K(\mathbf{s}, \pi)]$ can be obtained easily.

Unfortunately, as discussed in §5, so far, we can only obtain approximations of $\mathsf{E}[B^K(\mathbf{s}, \pi)]$, for each K. Hence, we can only obtain approximations of $\mathsf{E}[I_i(\mathbf{s}, \pi)]$ and $C(\mathbf{s}, \pi)$. Because the approximation of a general system is not in closed-form, it is difficult to discuss the optimization of $\mathbf{s}$ for any given π. We shall leave this to future research.

On the other hand, under certain special product structures, such as system-W, we can obtain closed-form approximations of the expected number of backorders for each product, as illustrated in §5.2. Thus, one can evaluate the approximate system costs for any base-stock policy relatively easily. In addition, for these systems, when the cost parameters satisfy certain symmetric conditions, it is possible to express the average cost in terms of total number of backorders, for which we have closed-form exact expressions. In all these cases, it is possible to develop procedures to optimize over the base-stock levels.

7 Performance Evaluation: Fill Rates

In this section, we study the evaluation of order-based fill rates. We first present closed-form bounds for immediate fill rates in §7.1, and show that these bounds are exact in a few special cases. We then provide a bound for the fill rates under FRFS within a time window in §7.2. Finally, we provide a comparison between FRFS (or any NHB rule) and FCFS in §7.3.

7.1 Order Fill Rates

We assume FRFS; the results can be easily extended to any NHB rule. We first present the following lemma (see Appendix for a proof).

Lemma 2 *For any ATO system under base-stock policy* **s** *and FRFS, the product-K fill rate is*

$$\widehat{f}^K(\mathbf{s}) = \mathsf{P}\{s_i - O_i + \sum_{J \in \mathcal{K}_i, J \neq K} \widehat{B}^J(\mathbf{s}) > 0, \ i \in K\}. \tag{19}$$

Eq. (19) is not a closed-form expression and therefore is not computable because the distribution of $\widehat{B}^J(\mathbf{s})$ is not known. The following results provide closed-form expressions for systems with certain special structures or bounds for general systems.

Theorem 6 *For any ATO system under base-stock policy* **s** *and FRFS, consider Eq. (19).*

(a) *If $J \subset K$, then we can set $\widehat{B}^J(\mathbf{s}) = 0$ in Eq. (19).*

(b) *If $J = K_J \cup S_J$ where $K_J \subseteq K$ and S_J is completely product-specific, then we can set $\widehat{B}^J(\mathbf{s}) = \max\{B_i(s_i), i \in S_J\}$ in Eq. (19).*

(c) *If $J = K_J \cup S_J \cup A$ where $K_J \subseteq K$, $(S_J \cup A) \cap K = \emptyset$, S_J is completely product-specific, but each component in A is shared by other products, then*

$$\max\{B_i(s_i), i \in S_J\} \ \leq \ \widehat{B}^J(\mathbf{s}) \ \leq \ \max\{B_i(s_i), i \in S_J \cup A\}. \tag{20}$$

Moreover, replacing $\widehat{B}^J(\mathbf{s})$ in Eq. (19) by $\max\{B_i(s_i), i \in S_J\}$ and $\max\{B_i(s_i), i \in S_J \cup A\}$ results in a lower and an upper bound on $\widehat{f}^K(\mathbf{s})$, respectively.

Proof. We prove the sample-path version of the results. Suppose a demand of product-type K arrives at time t. By Eq. (19), the probability that this demand will be fulfilled is

$$\widehat{f}^K(t, \mathbf{s}) = \mathsf{P}\{s_i - O_i(t) + \sum_{J \in \mathcal{K}_i, J \neq K} \widehat{B}^J(t, \mathbf{s}) > 0, i \in K\}. \tag{21}$$

To prove part (a), we consider any sample path ω such that

$$s_i - O_i(t) + \sum_{J \in \mathcal{K}_i, J \neq K} \widehat{B}^J(t, \mathbf{s}) > 0, \quad \forall i \in K. \tag{22}$$

Because $\widehat{I}_i(t, \mathbf{s}) > 0, \forall i \in K$ and $\widehat{B}^K(t, \mathbf{s}) > 0$ cannot hold simultaneously, we must have $\widehat{B}^K(t, \mathbf{s}) = 0$ and $\widehat{B}^J(t, \mathbf{s}) = 0$ for any $J \subset K$. This implies that the probability of Eq. (21) must be no larger than itself but with these $\widehat{B}^J(t, \mathbf{s}) = 0$, i.e.,

$$\mathsf{P}\{s_i - O_i(t) + \sum_{J \in \mathcal{K}_i, J \not\subseteq K} \widehat{B}^J(t, \mathbf{s}) > 0, i \in K\}. \tag{23}$$

On the other hand, because $\widehat{B}^J(t, \mathbf{s}) \geq 0$ for all J, the probability of Eq. (21) must be no smaller than the probability of Eq. (23). The proof of part (a) is thus completed.

For part (b), we consider any sample path ω such that inequalities (22) hold. Let the set of products that satisfy the condition of Part (b) be $\mathcal{X}_K$. Consider $J \in \mathcal{X}_K$. If $\widehat{B}^J(t, \mathbf{s}) > \max\{(O_i(t) - s_i)^+, i \in S_J\}$, then $\widehat{I}_i(t, \mathbf{s}) > 0, \forall i \in J$, which contradicts the fact that $\widehat{B}^J(t, \mathbf{s}) > 0$. By inequality (5), $\widehat{B}^J(t, \mathbf{s}) = \max\{(O_i(t) - s_i)^+, i \in S_J\}$ for the sample path ω. This implies that the probability in Eq. (21) must be no larger than itself but with these $\widehat{B}^J(t, \mathbf{s}) = \max\{(O_i(t) - s_i)^+, i \in S_J\}$, that is,

$$\mathsf{P}\{s_i - O_i(t) + \sum_{J \in \mathcal{K}_i \backslash \mathcal{X}_K, J \neq K} \widehat{B}^J(t, \mathbf{s}) + \sum_{J \in \mathcal{X}_K} \max\{(O_j(t) - s_j)^+, j \in S_J\} > 0, i \in K\}. \quad (24)$$

On the other hand, inequality (5) also implies that the probability in Eq. (21) must be no smaller than the probability in Eq. (24).

For part (c), we also consider any sample path ω such that inequalities (22) hold. By inequality (5), the first inequality of (20) holds. If $\widehat{B}^J(t, \mathbf{s}) > \max\{(O_i(t) - s_i)^+, i \in S_J \cup A\}$, then $\widehat{I}_i(t, \mathbf{s}) > 0, \forall i \in J$ which contradicts the fact that $\widehat{B}^J(t, \mathbf{s}) > 0$. Thus, the second inequality of (20) holds.

To prove the upper bound, we note that for each such ω, inequalities (22) still hold if one replaces $\widehat{B}^J(t, \mathbf{s})$ by $\max\{(O_i(t) - s_i)^+, i \in S_J \cup A\}$. Hence, the resulting probability is an upper bound on $\widehat{f}^K(t, \mathbf{s})$. The lower bound can be proved in a similar way. $\qquad \square$

Theorem 6 indicates that for ATO systems with certain network topologies, an exact and closed-form expression is available for the order-based fill rates. For general ATO systems, we can develop lower and upper bounds. Below are a few important examples.

- **System-W**: Here components 1 and 2 are product-specific, while component 3 is common. By Eq. (19),

$$\widehat{f}^{13}(\mathbf{s}) = \mathsf{P}\{s_1 - O_1 > 0; s_3 - O_3 + \widehat{B}^{23}(\mathbf{s}) > 0\}$$
$$\widehat{f}^{23}(\mathbf{s}) = \mathsf{P}\{s_2 - O_2 > 0; s_3 - O_3 + \widehat{B}^{13}(\mathbf{s}) > 0\}.$$

By Theorem 6 Part (b), the above can be simplified to

$$\widehat{f}^{13}(\mathbf{s}) = \mathsf{P}\{s_1 - O_1 > 0; s_3 - O_3 + B_2(s_2) > 0\}$$
$$= \mathsf{P}\{s_1 - O_1 > 0; s_3 - O_3 + (O_2 - s_2)^+ > 0\}$$
$$\widehat{f}^{23}(\mathbf{s}) = \mathsf{P}\{s_2 - O_2 > 0; s_3 - O_3 + B_1(s_1) > 0\}$$
$$= \mathsf{P}\{s_2 - O_2 > 0; s_3 - O_3 + (O_1 - s_1)^+ > 0\}$$

These expressions are also given in Song and Zhao (2007) using a different approach.

- **General System-W**: Here, $K = S^K \cup A$, $K \in \mathcal{K}$; see §5.1.1. By Theorem 6 Part (b),

$$
\begin{aligned}
\widehat{f}^K(\mathbf{s}) &= \; \mathsf{P}\{s_i - O_i > 0, i \in S^K; s_i - O_i + \sum_{J \neq K} \widehat{B}^J(\mathbf{s}) > 0, i \in A\} \\
&= \; \mathsf{P}\{s_i - O_i > 0, i \in S^K; s_i - O_i + \sum_{J \neq K} \max\{B_j(s_j), j \in S^J\} > 0, i \in A\}.
\end{aligned}
$$

- **Nested System**: For the general system with nested structure (§5.1.2), we have the following theorem (see Appendix for a proof).

  Theorem 7 *For a nested system, the fill rate of a product K_j, $1 \leq j \leq n$, is given by*

$$
\begin{aligned}
\widehat{f}^{K_j}(\mathbf{s}) &= \; \mathsf{P}\{s_i - O_i + \widehat{B}^{K_{j+1}}(\mathbf{s}) + \ldots + \widehat{B}^{K_n}(\mathbf{s}) > 0, i \in K_j\} \\
&= \; \mathsf{P}\{s_i - O_i + \max\{B_l(s_l), l \in K_n \setminus K_j\} > 0, i \in K_j\}.
\end{aligned}
$$

  Theorem 7 clearly shows that $\widehat{f}^{K_1}(\mathbf{s}) \geq \widehat{f}^{K_2}(\mathbf{s}) \geq \cdots \geq \widehat{f}^{K_n}(\mathbf{s})$.

- **System-M**: We have two components and three products – product-$\{1\}$, $\{2\}$ and $\{1, 2\}$. By Theorem 6,

$$
\begin{aligned}
\widehat{f}^{12}(\mathbf{s}) &= \; \mathsf{P}\{s_1 - O_1 > 0, s_2 - O_2 > 0\}, \\
\mathsf{P}\{s_1 - O_1 > 0\} &\leq \widehat{f}^1(\mathbf{s}) \leq \mathsf{P}\{s_1 - O_1 + B_2(s_2) > 0\}, \\
\mathsf{P}\{s_2 - O_2 > 0\} &\leq \widehat{f}^2(\mathbf{s}) \leq \mathsf{P}\{s_2 - O_2 + B_1(s_1) > 0\}.
\end{aligned}
$$

7.2   Time Window $w > 0$

The order fill rate within a time window $w > 0$ (see definition in §3.2) is much more difficult to characterize because the dynamics of demand fulfillment processes during a time period of w is very complex. Because different NHB rules may have different fill rates within a time window, we focus on FRFS and develop the following lower bound (see Appendix for a proof).

Lemma 3 *For any ATO system under the FRFS rule, $\widehat{f}^{K,w}(t, \mathbf{s})$, the probability that a product-K order arriving at t will be fulfilled by $t + w$, satisfies:*

$$
\widehat{f}^{K,w}(t, \mathbf{s}) \geq \mathsf{P}\{s_i - O_i(t+w) + \sum_{J \in \mathcal{K}_i, K \subseteq J} D^J[t, t+w] + \sum_{J \in \mathcal{K}_i, K \setminus J \neq \emptyset} \widehat{B}^J(t+w, \mathbf{s}) > 0, i \in K\}. \quad (25)
$$

Note that strict inequality can hold in Eq. (25) because the demand can be satisfied prior to $t + w$ even if the arguments on the right-hand-side of Eq. (25) do not hold. By inequality (5), we must have the following lower bound for $\widehat{f}^{K,w}(t, \mathbf{s})$,

Corollary 1 *Consider any ATO system, a product K and one of its components i. For any product $J \in \mathcal{K}_i, K \setminus J \neq \emptyset$, if its product-specific components $S_J \neq \emptyset$ ($S_J = \emptyset$), then replacing $\widehat{B}^J(t+w, \mathbf{s})$ in Eq. (25) by $\max\{(O_i(t+w) - s_i)^+, i \in S_J\}$ (by 0) results in a closed-form lower bound for $\widehat{f}^{K,w}(t, \mathbf{s})$.*

7.3 FCFS vs. FRFS on Fill Rates

We now compare the effectiveness of FCFS and FRFS on order fill rates (see Appendix for a proof).

Theorem 8 *Given any base-stock policy, an ATO system under the FRFS allocation rule always out-performs an analogous ATO system under the FCFS rule in the order fill rates.*

Intuitively, under the FRFS rule, each backordered demand triggers an order for its components, but no stock is committed to it. Effectively, the available inventory of components becomes higher than those under the FCFS rule. Because Eq. (21) holds for any NHB rule, Theorem 8 applies to any NHB rule.

Although FRFS always outperforms FCFS on immediate order fill rates, it may be inferior to FCFS on order fill rates within a time window $w > 0$.

To see this, suppose a demand for product-type K arrives at time t. Under the FCFS rule, the available stock at time $t + w$ is $(s_i + D_i[0, t + w) - O_i(t + w) - D_i[0, t))^+$. This is true because all demand for component i arriving after t must be satisfied after the demand at time t. Since $O_i(t + w) = D_i[t + w - L_i, t + w)$, the fill rate for product K within time window w is

$$
\begin{aligned}
\widetilde{f}^{K,w}(t, \mathbf{s}) &= \mathsf{P}\{s_i - O_i(t + w) + D_i[t, t + w) > 0, i \in K\} \\
&= \mathsf{P}\{s_i - D_i[t - L_i + w, t) > 0, i \in K\}.
\end{aligned}
\tag{26}
$$

Clearly, if $w > \max\{L_i, i \in K\}$, then $D_i[t - L_i + \tau, t) = 0$ and therefore $f^{K,w}(t, \mathbf{s}) = 1$. Thus, under the FCFS rule, a demand for product K is guaranteed to be filled within a time window of $\max\{L_i, i \in K\}$.

Under the FRFS rule, however, a demand for product K can wait longer than $\max\{L_i, i \in K\}$. Moreover, there may not exist a time window within which the demand is guaranteed to be filled, as the following examples demonstrate.

Example 3 *Consider system-W under FRFS. One can easily create a sample path in which product-1 waits longer than $\max\{L_1, L_3\}$ because it is possible that just before component 1 becomes available, component 3 becomes unavailable.*

Example 4 *Consider system-M under FRFS. One can easily create a sample path in which a demand for product-$\{1, 2\}$ waits for an arbitrarily long time because at any time t, either component 1 or 2 is not available.*

It is straightforward to show that these examples are valid for any NHB rule. These examples show that although the FRFS rule yields higher order-based fill rates at $w = 0$ relative to the FCFS rule, the fill rates can approach 1 (as $w \to \infty$) at a slower rate than fill rates under the FCFS rule.

8 Extensions

Up to now, we have assumed base-stock policies, constant lead times and Poisson demand processes in the above analysis, however, most of our results hold under more general inventory ordering policies, lead time models and demand processes. This is largely because these results were obtained through a sample path analysis.

First, the optimality results in §4 (Theorems 1-2) hold for any inventory ordering policy that is independent of the allocation rules, such as those depending only on inventory positions (e.g., the (r, nQ) policies). These results also hold for stochastic replenishment lead times, which may be either i.i.d. lead times or exogenous and sequential stochastic lead times (see Zipkin 2000, Chapter 7 for a definition). In addition, these results hold for any demand process under the assumption that demand can be split, i.e., each unit of demand can be satisfied separately.

Second, the performance evaluation results in §§5-7, such as the order fill rates (Eq. 19), Theorems 3-8, and the bounds and the approximation (in §5.2), hold for both i.i.d. lead times and for exogenous and sequential lead times, as well as for any point demand process under the assumption that demand can be split. For ordering policies (other than the base-stock policy) that depend only on inventory positions, the flow-conservation equation must be modified accordingly.

We now discuss ATO systems with a general BOM matrix, assuming everything else stays the same as in the basic model. Let a_i^K be the amount of component i required by product-K. Eq. (3) becomes

$$B^K(t, \mathbf{s}, \pi) * \min\{[I_i(t, \mathbf{s}, \pi) - a_i^K + 1]^+, i \in K\} = 0 \quad \text{for all } K \in \mathcal{K}, t \geq 0. \tag{27}$$

Eq. (4) becomes

$$I_i(t, \mathbf{s}, \pi) = s_i - O_i(t) + \sum_{K \in \mathcal{K}_i} a_i^K B^K(t, \mathbf{s}, \pi). \tag{28}$$

The optimality results in Theorem 1 of §4, hold as long as $a_3^{13} = a_3^{23}$, i.e., all products require the same amount of the common component. This is true because we can redefine the flow unit for the common component, which reduces the system to one with a unit BOM matrix. Note that the same assumption is also made by Dogru et al. (2007). Similarly, Theorem 2 holds as long as all products require the same amount of each common component, i.e., $a_i^K = a_i^{K'}$ for any $K, K' \in \mathcal{K}$ and any $i \in A$.

Theorems 3-5 on the total number of backorders in §5 hold if for any component $i \in \mathcal{I}$, a_i^K is a constant for all $K \in \mathcal{K}$. Otherwise, these results may not hold.

As for the performance evaluation results on fill rates in §7, note that Eq. (19) becomes

$$\widehat{f}^K(\mathbf{s}) = \mathsf{P}\{s_i - O_i + \sum_{J \in \mathcal{K}_i, J \neq K} a_i^J \widehat{B}^J(\mathbf{s}) \geq a_i^K, \ i \in K\}. \tag{29}$$

Theorem 6 holds only if more conditions are specified. Without loss of generality, we assume $a_i^K = 1$ for each $K \in \mathcal{K}$ and its product-specific component i. Specifically, for a general BOM matrix, Theorem 6 becomes

Corollary 2 *For any ATO system, consider Eq. (29).*

(a) *If $J \subset K$ and $a_i^J \leq a_i^K$ for all $i \in J$, then we can set $\widehat{B}^J(\mathbf{s}) = 0$ in Eq. (29).*

(b) *If $J = K_J \cup S_J$ where $K_J \subseteq K$ and $a_i^J \leq a_i^K$ for all $i \in K_J$, and S_J is completely product-specific, then we can set $\widehat{B}^J(\mathbf{s}) = \max\{B_i(s_i), i \in S_J\}$ in Eq. (29).*

(c) *If $J = K_J \cup S_J \cup A$ where $K_J \subseteq K$ and $a_i^J \leq a_i^K$ for all $i \in K_J$, $(S_J \cup A) \cap K = \emptyset$, S_J is completely product-specific, but each component in A is shared by other products, then*

$$\max\{B_i(s_i), i \in S_J\} \ \leq \ \widehat{B}^J(\mathbf{s}) \ < \ \max\{B_i(s_i)/a_i^J + 1, i \in S_J \cup A\}. \tag{30}$$

Moreover, replacing $\widehat{B}^J(\mathbf{s})$ in Eq. (29) by $\max\{B_i(s_i), i \in S_J\}$ and $\max\{B_i(s_i)/a_i^J + 1, i \in S_J \cup A\}$ results in a lower and an upper bound on $\widehat{f}^K(\mathbf{s})$, respectively.

The proof of this corollary follows that of Theorem 6; we omit the details. Finally, Theorem 8 holds for any BOM matrix.

9 Concluding Remarks

In this paper, we have analyzed a class of non-FCFS component allocation rules, called NHB rules, for continuous-time base-stock ATO systems. We have identified conditions under which the NHB rules are optimal among all allocation rules for minimizing total backorder and inventory costs. We have also developed tools for evaluating the performance of a NHB rule. Using these results, we have compared the effectiveness of two widely used allocation rules, FCFS and FRFS, and, hence, have elucidated the merits and limitations of a class of state-dependent allocation rules. These results, as well as the analytical approach, constitute one of the first attempts to study non-FCFS allocation rules for these systems. We hope this effort will facilitate future research analyzing other types of state-dependent allocation rules, such as the rationing policies (a type of hold-back rules), leading to better understanding of effective control policies for ATO systems.

Acknowledgment

References

[1] Agrawal, N. and M. Cohen (2001). Optimal material control and performance evaluation in an assembly environment with component commonality. *Naval Research Logistics*, **48**, 409-429.

[2] Akcay, Y. and S. H. Xu (2004). Joint inventory replenishment and component allocation optimization in an assemble-to-order system. *Management Science*, **50**, 99-116.

[3] Benjaafar, S. and M. ElHafsi (2006). Production and inventory control of a single product Assemble-to-Order system with multiple customer classes. *Management Science*, **52**, 1896-1912.

[4] Bernstein, F., G. DeCroix and Y. Wang (2007). Allocation policies based on demand aggregation in an assemble-to-order system. Working Paper. Duke University, Durham, NC.

[5] Chen, H. and D. D. Yao (1993). Dynamic Scheduling Control of a Multi-Class Fluid Network. *Operations Research*, **41**, 1104-1115.

[6] Deshpande, V., M. A. Cohen and K. Donohue (2003). A threshold inventory rationing policy for service-differentiated demand classes. *Management Science*, **49**, 683-703.

[7] Dogru, M. K., M. I. Reiman and Q. Wang (2007). A stochastic programming based inventory policy for assemble-to-order systems with application to the W model. Working Paper, Bell Laboratories Ireland. Dublin, Ireland.

[8] Hausman, W. H., H. L. Lee, and A. X. Zhang (1998). Joint demand fulfillment probability in a multi-item inventory system with independent order-up-to policies. *European Journal of Operational Research*, **109**, 646-659.

[9] Kapuscinski, R., R. Zhang, P. Carbonneau, R. Moore and B. Reeves (2004). Inventory decisions in Dell's supply chain. *Interfaces*, **34**, 191-205.

[10] Lu, Y., Song, J. S. and D. D. Yao (2003). Order fill rate, lead-time variability and advance demand information in an Assemble-to-order system. *Operations Research*, **51**, 292-308.

[11] Lu, Y., Song, J. S. and D. D. Yao (2005). Backorder minimization in multiproduct assemble-to-order systems. *IIE Transactions*, **37**, 763-774.

[12] Plambeck, E. and A. Ward (2006a). Note: A separation principle for a class of assemble-to-order systems with expediting. *Operations Research* **55**, 603-609.

[13] Plambeck, E. and A. Ward (2006b). Optimal control of a high-volume assemble-to-order system. *Mathematics of Operations Research* **31**, 453-471.

[14] Song, J. S. (1998). On the order fill rate in a multi-item, base-stock inventory system. *Operations Research*, **46**, 831-845.

[15] Song, J. S. (2002). Order-Based backorders and their implications in multi-item inventory systems. *Management Science*, **48**, 499-516.

[16] Song, J.-S., S. Xu and B. Liu (1999). Order-fulfillment performance measures in an assemble-to-order system with stochastic lead times. *Operations Research* **47**, 131-149.

[17] Song, J. S. and D. D. Yao (2002). Performance analysis and optimization in Assemble-to-order systems with random lead-times. *Operations Research*, **50**, 889-903.

[18] Song, J. S. and Y. Zhao (2007). The value of component commonality in a dynamic inventory system with lead times. To appear in *Manufacturing and Service Operations Management*.

[19] Song, J. S. and P. Zipkin (2003). Supply chain operations: Assemble-to-order systems. Chapter 11 in *Handbooks in Operations Research and Management Science*, Vol. 11: Supply Chain Management.

[20] Xu, P. J., R. Allgor and S. Graves (2006). The benefits of re-evaluating real-time order fulfillment decisions. Working Paper. Sloan School of Management, MIT, Cambridge, MA.

[21] Zhang, A. X. (1997). Demand fulfillment rates in an assemble-to-order system with multiple products and dependent demands. *Production and Operations Management*, **6**, 309-323.

[22] Zhao, Y. and D. Simchi-Levi (2006). Performance analysis and evaluation of Assemble-to-Order systems with stochastic sequential lead times. *Operations Research*, **54**, 706-724.

[23] Zipkin, P (2000). *Foundations of Inventory Management.* McGraw Hill, Boston.

Appendix

Proof for Lemma 1. (a) We consider all NHB rules in the same probability space, i.e., under identical initial conditions and demand processes. Note that the replenishment process is identical for all $\pi \in \mathcal{N}$ because the replenishment policy is independent of allocation rules. Under any $\pi \in \mathcal{N}$, there are four types of events on a sample path:

1. When a demand arrives, there is excessive inventory on-hand. Therefore the demand is satisfied immediately by Eq. (7).

2. When a demand arrives, there is no on-hand inventory. Therefore the demand is backordered.

3. When a replenishment item arrives, there are no backorders. Therefore the item is kept in inventory.

4. When a replenishment item arrives, there are some backorders. The item is allocated to satisfy one of the backorders.

Note that, in each of these events, all $\pi \in \mathcal{N}$ yield identical on-hand inventory and an identical number of backorders. The proof is now completed.

(b) If $I(t) > 0$ at some time t, then all the demand up to time t must have been satisfied, so $I(t)$ is the smallest inventory that can be achieved at time t. A similar argument can be applied to obtain the inequality for the backorders.

Part (c) follows directly from Part (a) and Part (b). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

Proof for Theorem 2. (a) Note that since elements 1-3 are independent of allocation rules, we can subtract their cost from the total cost function and only consider the costs associated with elements 4 and 5. That is, for any time t, the relevant cost is

$$
\begin{aligned}
R(t, \mathbf{s}, \pi) &= \sum_{i \in A} h_i I_i(t, \mathbf{s}, \pi) + \sum_{K \in \mathcal{K}} \left[\sum_{i \in S^K} h_i X_i(t, \mathbf{s}, \pi) + b^K B_A^K(t, \mathbf{s}, \pi) \right] \\
&= \sum_{i \in A} h_i I_i(t, \mathbf{s}, \pi) + \sum_{K \in \mathcal{K}} \left[\sum_{i \in S^K} h_i + b^K \right] B_A^K(t, \mathbf{s}, \pi).
\end{aligned}
$$

We now reduce the system with multiple common components to one with a single common component. Because all products require the same number of units for each common component in set A, all modified demand can be considered as the same product type, and each component in A can be considered specific to this product type. By Observation 1, one can assign each component in set A to the modified demand on a FCFS basis without changing performance. We now define the *modified common*

32

component as follows: A modified common component is a complete set of components in A that meet the same demand, and the replenishment time of a modified common component is the earliest time such that all components in the set are replenished. We further note that at time t, the inventory cost of the modified common component is $[\sum_{i \in A} h_i] \min_{i \in A}\{I_i(t, \mathbf{s}, \pi)\}$. Because other on-hand inventory of components in A, $I_i(t, \mathbf{s}) - \min_{i \in A}\{I_i(t, \mathbf{s}, \pi)\}$, depends only on the inventory replenishment process, applying the proof of Theorem 1 (c) to the modified common component yields the result.

(b) Since $|S^K|$ is identical for all products K, the result follows directly by Theorem 1(b).

(c) If $|S^K|$ is product-dependent, the proof of Theorem 1 (b) still applies to the total backorder in the general system-W. The proof for the second part involves a straightforward swapping of demands or inventory items of common components. For simplicity, we omit the details. $\qquad\square$

Proof of Theorem 3. (a) For any $t \geq 0$, define $\widehat{B}_3^{i3}(t, \mathbf{s})$ to be the number of backorders of product-$i3$ at t that are due to a shortage of only the common component 3. Because of the W-structure, we have

$$\widehat{B}^{i3}(t, \mathbf{s}) = B_i(t, s_i) + \widehat{B}_3^{i3}(t, \mathbf{s}), \quad i = 1, 2. \tag{31}$$

If

$$\widehat{B}_3^{13}(t, \mathbf{s}) \vee \widehat{B}_3^{23}(t, \mathbf{s}) > 0, \tag{32}$$

then according to FRFS, $\widehat{I}_3(t, \mathbf{s}) = 0$. By Eq. (4), $s_3 - O_3(t) + \widehat{B}^{13}(t, \mathbf{s}) + \widehat{B}^{23}(t, \mathbf{s}) = 0$, and thus

$$B_3(t, s_3) = \widehat{B}^{13}(t, \mathbf{s}) + \widehat{B}^{23}(t, \mathbf{s}).$$

By Eqs. (31)-(32), $B_3(t, s_3)$ is greater than $B_1(t, s_1) + B_2(t, s_2)$, so the result holds. On the other hand, if $\widehat{B}_3^{13}(t, \mathbf{s}) \vee \widehat{B}_3^{23}(t, \mathbf{s}) = 0$, then by FRFS, $\widehat{I}_3(t, \mathbf{s}) \geq 0$. It follows by Eqs. (4) and (31) that

$$B_3(t, s_3) \leq \widehat{B}^{13}(t, \mathbf{s}) + \widehat{B}^{23}(t, \mathbf{s}) = B_1(t, s_1) + B_2(t, s_2).$$

So, again, the result holds. Part (a) is thus proved.

(b) Define $\widehat{B}_A^K(t, \mathbf{s})$ to be the number of backorders of product-K at t that are due to a shortage of at least one common component. Thus,

$$\widehat{B}^K(t, \mathbf{s}) = \max_{i \in S^K}\{B_i(t, s_i)\} + \widehat{B}_A^K(t, \mathbf{s}), \tag{33}$$

If

$$\max_{K \in \mathcal{K}}\{\widehat{B}_A^K(t, \mathbf{s})\} > 0, \tag{34}$$

then by FRFS, there exists at least one component, say $l \in A$, such that $\widehat{I}_l(t, \mathbf{s}) = 0$. By Eq. (4),

$$B_l(t, s_l) = \sum_{K \in \mathcal{K}} \widehat{B}^K(t, \mathbf{s}).$$

33

Note that $\widehat{I}_i(t, \mathbf{s}) \geq 0$ for all $i \in A$, so $B_i(t, s_i) \leq \sum_{K \in \mathcal{K}} \widehat{B}^K(t, \mathbf{s})$. In summary, we must have

$$\sum_{K \in \mathcal{K}} \widehat{B}^K(t, \mathbf{s}) = B_l(t, s_l) = \max\{\max_{i \in A}\{B_i(t, s_i)\}, \sum_K \max_{i \in S^K}\{B_i(t, s_i)\}\},$$

so the result holds. On the other hand, if $\max_{K \in \mathcal{K}}\{\widehat{B}_A^K(t, \mathbf{s})\} = 0$, then by Eq. (33), $\sum_{K \in \mathcal{K}} \widehat{B}^K(t, \mathbf{s}) = \sum_{K \in \mathcal{K}} \max_{i \in S^K}\{B_i(t, s_i)\}$. Also by FRFS rule, $\widehat{I}_i(t, \mathbf{s}) \geq 0$, for all $i \in A$. It follows from Eq. (4) that $O_i(t) - s_i \leq \sum_{K \in \mathcal{K}_i} \widehat{B}^K(t, \mathbf{s})$ for all $i \in A$. Thus

$$B_i(t, s_i) \leq \sum_{K \in \mathcal{K}} \widehat{B}^K(t, \mathbf{s}) = \sum_{K \in \mathcal{K}} \max_{i \in S^K}\{B_i(t, s_i)\}, \quad i \in A.$$

The proof is completed. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Proof of Theorem 4. For Part (a), we only need to show the first result; the second result follows immediately from the first. We consider two cases: (i) if $B_1(t, s_1) \geq B_2(t, s_2)$, then component 1 is in short supply, and all the backorders are waiting for component 1, so the total number of backorders is $B_1(t, s_1)$. More specifically, $\widehat{B}^{12}(t, \mathbf{s}) \leq \max\{B_1(t, s_1), B_2(t, s_2)\} = B_1(t, s_1)$ by Eq. (6). By Eq. (5), $\widehat{B}^1(t, \mathbf{s}) + \widehat{B}^{12}(t, \mathbf{s}) \geq B_1(t, s_1)$. If $\widehat{B}^1(t, \mathbf{s}) + \widehat{B}^{12}(t, \mathbf{s}) > B_1(t, s_1)$, then by Eq. (4), $\widehat{I}_1(t, \mathbf{s}) > 0$ and $\widehat{B}^1(t, \mathbf{s}) = 0$. Thus $\widehat{B}^{12}(t, \mathbf{s}) > B_1(t, s_1)$, which contradicts $\widehat{B}^{12}(t, \mathbf{s}) \leq B_1(t, s_1)$. Hence, $\widehat{B}^1(t, \mathbf{s}) + \widehat{B}^{12}(t, \mathbf{s}) = B_1(t, s_1)$. (ii) If $B_1(t, s_1) < B_2(t, s_2)$, then we know that there is no product 1 backordered, hence the total number of backorders is $B_2(t, s_2)$. This is true because applying Eqs. (5)-(6) to product-12 yields $\widehat{B}^{12}(t, \mathbf{s}) = B_2(t, s_2)$. By Eq. (4), $\widehat{I}_1(t, \mathbf{s}) > 0$ and $\widehat{B}^1(t, \mathbf{s}) = 0$.

For Part (b), we prove by induction. When $n = 1$ it is trivial. Assume that the statement is true for n, we now examine the case $n + 1$. Consider any sample path at time t. If

$$\max_{i \in K_{n+1} \backslash K_n} B_i(t, s_i) \;\; < \;\; \max_{1 \leq \ell \leq n}\{\max_{i \in K_\ell \backslash K_{\ell-1}} B_i(t, s_i)\},$$

we then know that

$$\max_{i \in K_{n+1} \backslash K_n} B_i(t, s_i) \leq B^{n+1}(t, \mathbf{s}) \leq \max_{1 \leq \ell \leq n}\{\max_{i \in K_\ell \backslash K_{\ell-1}} B_i(t, s_i)\}.$$

According to the FRFS policy, there must exist at least one component having $B^{n+1}(t, \mathbf{s})$ units backlogged due to product $n + 1$. Now, let us consider the sub-system of product $1, \cdots, n$. The item backlogging for each $i \in K_n$ then becomes $[B_i(t, s_i) - B^{n+1}(t, \mathbf{s})]^+$. This is true because, in the sub-system, for each component, there are $B^{n+1}(t, \mathbf{s})$ units of demand that can be ignored. Thus, by the induction hypothesis, the total number of backorders in the sub-system is equal to

$$\max_{1 \leq \ell \leq n}\{\max_{i \in K_\ell \backslash K_{\ell-1}}[B_i(t, s_i) - B^{n+1}(t, \mathbf{s})]^+\} = \max_{1 \leq \ell \leq n}\{\max_{i \in K_\ell \backslash K_{\ell-1}} B_i(t, s_i)\} - B^{n+1}(t, \mathbf{s}).$$

34

Combining the above two facts, we conclude that the total number of backorders is

$$\max_{1 \le \ell \le n} \{ \max_{i \in K_\ell \setminus K_{\ell-1}} B_i(t, s_i) \}.$$

On the other hand, if

$$\max_{i \in K_{n+1} \setminus K_n} B_i(t, s_i) \ge \max_{1 \le \ell \le n} \{ \max_{i \in K_\ell \setminus K_{\ell-1}} B_i(t, s_i) \},$$

then we know that all the backordered components are due to K_{n+1}. Therefore, the total number of backorders equals $\max_{i \in K_{n+1} \setminus K_n} B_i(t, s_i)$. □

Proof for Theorem 5. To prove Part (a), we consider a sample path ω under FCFS and the nth demand for product-12. Suppose there are n^1 demands for product-1 and n^2 demands for product-2 that arrived prior to this demand. According to FCFS, exactly n^{th} demands for product-1 are assigned with component 1 inventory earlier than the demand for product-12, namely $\tilde{n}^1 = n^1$. Similarly, $\tilde{n}^2 = n^2$.

Consider the same ω and the same n^{th} demand for product-12 under FRFS. According to FRFS, any demand for product-1 that arrives prior to the nth demand for product-12 must be filled earlier than the latter (this can easily be shown by contradiction). So the number of demands for product-1 that are assigned component 1 inventory earlier than the nth demand for product-12 $\hat{n}^1 \ge n^1$. Similarly, $\hat{n}^2 \ge n^2$. Also note that by Observation 1, product-12 demands are satisfied on a FCFS basis, thus the nth demand for product-12 must be satisfied by inventories replenished later (or not earlier) than those under FCFS. Because the demand arrival time does not change (i.e., identical ω), the waiting time of the nth demand for product-12 must be as least as long as that under FCFS. This is true because under both FCFS and FRFS, a demand is satisfied at the earliest time when the demand occurs and all its assigned components become available. Because this statement applies to each product-12 demand, we must have $\hat{B}^{12} \ge \tilde{B}^{12}$ on every ω.

Now we prove Part (b). For simplicity, we suppress the dependence of backorders on t and base-stock levels. For system-M under FCFS, by Lu et al. (2005), $\tilde{B}^1 = B_1^1$, $\tilde{B}^2 = B_2^2$ and $\tilde{B}^{12} = \max\{B_1^{12}, B_2^{12}\}$, where $B_1 = B_1^1 + B_1^{12}$ and $B_2 = B_2^2 + B_2^{12}$. Thus, $\tilde{B} \ge B_1 + B_2 - \tilde{B}^{12}$.

For system-M under FRFS, by Eq. (5), $\hat{B}^1 + \hat{B}^{12} \ge B_1$ and $\hat{B}^2 + \hat{B}^{12} \ge B_2$. By Eq. (6) $\hat{B}^1 \le B_1$, $\hat{B}^2 \le B_2$ and $\hat{B}^{12} \le \max\{B_1, B_2\}$.

We now consider three cases for B_1 and B_2 on a sample path ω.

1. $B_1 = 0, B_2 = 0$. In this case, it is easy to show that $\hat{B} = 0 = \tilde{B}$.

2. $B_1 > 0, B_2 = 0$ or $B_1 = 0, B_2 > 0$. Consider the first case $B_1 > 0, B_2 = 0$. Under FCFS, $\tilde{B}^2 = 0$ and $\tilde{B}^1 + \tilde{B}^{12} = B_1$. Under FRFS, $\hat{B}^2 = 0$ and $\hat{B}^1 + \hat{B}^{12} = B_1$. The last equation holds because if

$\widehat{B}^1 + \widehat{B}^{12} > B_1$, then $\widehat{I}_1 > 0$ and $\widehat{B}^1 = 0$ by Eq. (3); further note that $\widehat{B}^{12} > 0 = B_2$ and so $\widehat{I}_2 > 0$. These facts violate Eq. (3) and creates a contradiction. Thus $\widetilde{B} = \widehat{B}$. The same logic applies to the second case.

3. $B_1 > 0, B_2 > 0$. Under FRFS, if $\widehat{B}^1 + \widehat{B}^{12} = B_1$ and $\widehat{B}^2 + \widehat{B}^{12} = B_2$, then $\widehat{B} = B_1 + B_2 - \widehat{B}^{12}$. However, under FCFS, we always have $\widetilde{B} \geq B_1 + B_2 - \widetilde{B}^{12}$. By Part (a) of this theorem, $\widetilde{B} \geq \widehat{B}$.

Under FRFS, if $\widehat{B}^1 + \widehat{B}^{12} > B_1$ and $\widehat{B}^2 + \widehat{B}^{12} = B_2$, then $\widehat{B}^1 = 0$ by Eq. (3), which implies $\widehat{B} = B_2$. However, under FCFS, we always have $\widetilde{B}^2 + \widetilde{B}^{12} \geq B_2$ and thus $\widetilde{B} \geq B_2$. Thus, $\widetilde{B} \geq \widehat{B}$. The case of $\widehat{B}^1 + \widehat{B}^{12} = B_1$ and $\widehat{B}^2 + \widehat{B}^{12} > B_2$ can be handled similarly.

The proof is now completed. $\qquad\square$

Proof of Lemma 2. By Eq. (4), in the event that a product-K demand arrives at time t and is satisfied upon its arrival, $\widehat{I}_i(t, \mathbf{s}) > 0$ for all $i \in K$, that is,

$$s_i - O_i(t) + \sum_{J \in \mathcal{K}_i} \widehat{B}^J(t, \mathbf{s}) > 0, \quad i \in K.$$

By condition (3), $\widehat{B}^K(t, \mathbf{s}) = 0$ in the same event, and so the fill rate at time t,

$$\widehat{f}^K(t, \mathbf{s}) = \mathsf{P}\{s_i - O_i(t) + \sum_{J \in \mathcal{K}_i} \widehat{B}^J(t, \mathbf{s}) > 0, \ i \in K\} \leq \mathsf{P}\{s_i - O_i(t) + \sum_{J \in \mathcal{K}_i, J \neq K} \widehat{B}^J(t, \mathbf{s}) > 0, \ i \in K\}.$$

On the other hand, $B^K(t, \mathbf{s}) \geq 0$ for all t, so

$$\mathsf{P}\{s_i - O_i(t) + \sum_{J \in \mathcal{K}_i} \widehat{B}^J(t, \mathbf{s}) > 0, \ i \in K\} \geq \mathsf{P}\{s_i - O_i(t) + \sum_{J \in \mathcal{K}_i, J \neq K} \widehat{B}^J(t, \mathbf{s}) > 0, \ i \in K\}.$$

In summary,

$$\widehat{f}^K(t, \mathbf{s}) = \mathsf{P}\{s_i - O_i(t) + \sum_{J \in \mathcal{K}_i, J \neq K} \widehat{B}^J(t, \mathbf{s}) > 0, \ i \in K\}.$$

By the PASTA (Poisson Arrivals See Time Averages) property, the proof is completed. $\qquad\square$

Proof for Theorem 7. By Theorem 6 Part (a), we have $\widehat{f}^{K_n}(\mathbf{s}) = \mathsf{P}\{s_i - O_i > 0, i \in K_n\}$; by Theorem 6 Part (b), we have $\widehat{f}^{K_{n-1}}(\mathbf{s}) = \mathsf{P}\{s_i - O_i + \max\{B_l(s_l), l \in K_n \setminus K_{n-1}\} > 0, i \in K_{n-1}\}$. For $j < n - 1$, Eq. (19) and Theorem 6 Part (a) give,

$$\widehat{f}^{K_j}(\mathbf{s}) = \mathsf{P}\{s_i - O_i + \widehat{B}^{K_{j+1}}(\mathbf{s}) + \ldots + \widehat{B}^{K_n}(\mathbf{s}) > 0, i \in K_j\}.$$

Consider any sample path ω where

$$s_i - O_i(t) + \widehat{B}^{K_{j+1}}(t, \mathbf{s}) + \ldots + \widehat{B}^{K_n}(t, \mathbf{s}) > 0, \quad i \in K_j.$$

On ω, we must have $\widehat{B}^{K_1}(t, \mathbf{s}) = 0, \ldots, \widehat{B}^{K_j}(t, \mathbf{s}) = 0$ by condition (3). By Theorem 4,

$$\sum_{1 \leq l \leq n} \widehat{B}^{K_l}(t, \mathbf{s}) = \sum_{j+1 \leq l \leq n} \widehat{B}^{K_l}(t, \mathbf{s}) = \max_{1 \leq i \leq m}\{B_i(t, s_i)\}, \quad \text{on} \quad \omega.$$

However, for $i \in K_j$, $B_i(t, s_i) = (O_i(t) - s_i)^+ < \widehat{B}^{K_{j+1}}(t, \mathbf{s}) + \ldots + \widehat{B}^{K_n}(t, \mathbf{s})$ on ω. Thus,

$$\sum_{j+1 \leq l \leq n} \widehat{B}^{K_l}(t, \mathbf{s}) = \max_{i \in K_n \backslash K_j}\{B_i(t, s_i)\}, \quad \text{on} \quad \omega.$$

This fact implies that

$$\widehat{f}^{K_j}(t, \mathbf{s}) \leq \mathsf{P}\{s_i - O_i(t) + \max\{B_l(t, s_l), l \in K_n \setminus K_j\} > 0, i \in K_j\}.$$

On the other hand, by Eq. (5), $\widehat{B}^{K_n}(t, \mathbf{s}) \geq \max_{l \in K_n \backslash K_{n-1}}\{B_l(t, s_l)\}$, $\widehat{B}^{K_{n-1}}(t, \mathbf{s}) + \widehat{B}^{K_n}(t, \mathbf{s}) \geq \max_{l \in K_{n-1} \backslash K_{n-2}}\{B_l(t, s_l)\}, \ldots, \widehat{B}^{K_{j+1}}(t, \mathbf{s}) + \ldots + \widehat{B}^{K_n}(t, \mathbf{s}) \geq \max_{l \in K_{j+1} \backslash K_j}\{B_l(t, s_l)\}$. Thus, $\widehat{B}^{K_{j+1}}(t, \mathbf{s}) + \ldots + \widehat{B}^{K_n}(t, \mathbf{s}) \geq \max_{l \in K_n \backslash K_j}\{B_l(t, s_l)\}$. This fact implies that

$$\widehat{f}^{K_j}(t, \mathbf{s}) \geq \mathsf{P}\{s_i - O_i(t) + \max\{B_l(t, s_l), l \in K_n \setminus K_j\} > 0, i \in K_j\}.$$

Applying PASTA, the proof is completed. $\qquad \square$

Proof of Lemma 3. Consider any sample path ω. If a demand at time t for product-K is still backordered at time $t + w$, then the following statements must hold: (i) there exists at least one $i \in K$ such that

$$s_i + D_i[0, t+w) - O_i(t+w) - \sum_{J \in \mathcal{K}_i}[D^J[0, t+w) - \widehat{B}^J(t+w, \mathbf{s})]$$
$$= s_i - O_i(t+w) + \sum_{J \in \mathcal{K}_i}\widehat{B}^J(t+w, \mathbf{s}) < 1.$$

(ii) $\widehat{B}^J(t+w, \mathbf{s}) \geq D^J[t, t+w)$ for any $J \in \mathcal{K}_i, K \subseteq J$.

(i) holds because if otherwise, a demand for product K realized at time $t + w$ could be satisfied immediately. This would contradict the assumption that the demand for product K realized at time t is still backordered. (ii) holds because the demand for product K realized at time t has priority over demand for any product J with $K \subseteq J$ realized after t.

By statements (i) and (ii), the probability of a demand for product K realized at time t waiting through $t + w$ is bounded from above by

$$\mathsf{P}\{s_i - O_i(t+w) + \sum_{J \in \mathcal{K}_i, K \subseteq J} D^J[t, t+w) + \sum_{J \in \mathcal{K}_i, K \backslash J \neq \emptyset} \widehat{B}^J(t+w, \mathbf{s}) < 1, \exists i \in K\}.$$

Therefore, Eq. (25) holds. The proof is now completed. $\qquad \square$

Proof of Theorem 8. We compare the FCFS and FRFS rules in every sample path ω (i.e., in the same probability space). Let us focus on component $i \in \mathcal{I}$. Suppose at time $t = 0$, the on-hand stock level is s_i. At time t, suppose a demand for product type-K arrives. Then under FCFS, the fill rate for this demand is (see Song 1998),

$$\widetilde{f}^K(t, \mathbf{s}) = \mathsf{P}\{s_i - O_i(t) > 0, i \in K\}. \tag{35}$$

By the non-negativity of $\widehat{B}^K(t, \mathbf{s})$, it follows from Eqs. (35) and (21) that

$$\widehat{f}^K(t, \mathbf{s}) \geq \widetilde{f}^K(t, \mathbf{s}) \quad \text{for any } K \text{ and } t. \tag{36}$$

The proof is now completed. $\qquad\square$