A.G. de Kok and S.C. Graves, Eds., *Handbooks in OR & MS, Vol. 11*
© 2003 Elsevier B.V. All rights reserved.

Chapter 3

Supply Chain Design: Safety Stock Placement and Supply Chain Configuration

Stephen C. Graves

Leaders for Manufacturing Program and A. P. Sloan School of Management, Massachusetts Institute of Technology, Cambridge MA 02139-4307, USA
E-mail: sgraves@mit.edu

Sean P. Willems

Boston University School of Management, 595 Commonwealth Avenue, Boston MA 02215, USA; E-mail: willems@bu.edu

1 Introduction

The focus of this chapter is on safety stock placement in the design of a supply chain, as well as on the optimal configuration of the supply chain to minimize total supply chain cost. As our intent is not to cover all of supply chain design, we first need to position this chapter relative to the other work in this handbook on supply chain design. We also need to position our treatment of the safety stock placement problem relative to other chapters that address multi-echelon inventory systems.

There is a great range of decisions associated with the design of a supply chain. One might group the design decisions into three broad categories.

First, there are the traditional decisions of network design as applied to the design of a supply chain. The choice of nodes corresponds to questions about the number, location and sizing of facilities. The choice of arcs corresponds to setting the general logistics strategy in terms of who serves whom and by what transportation or production mode. Muriel and Simchi-Levi cover these models in Chapter 2 of this book.

Second, we mention the decisions that are made in product design that determine the topology, as well as the key economics, of the supply chain. Ideally, one would like to concurrently design the product and its supply chain so as to meet the market objectives for the product with the best performance of the supply chain. Lee and Swaminathan look at the impact of product design decisions on the supply chain, with a particular focus on understanding the tactic of postponement as a way to achieve product proliferation with a well performing supply chain.

Third, we note the design decisions that allow the supply chain to be responsive to uncertainty and variability. In Chapter 4, Bertrand addresses the general question of how to accomplish flexibility in a supply chain, for instance by means of having flexible facilities and/or capacity buffers, as well as through contracting mechanisms. In the current chapter, we examine a completely different tactic, the deployment of inventory as safety stock for addressing demand uncertainty. In particular we look at the strategic placement of safety stocks across a supply chain.

In this chapter, we also introduce a new design consideration for how to configure the supply chain. The configuration decision entails choosing how to source each step or stage in the supply chain, where there might be several options that vary in terms of lead-time and cost. For instance, the configuration decision includes decisions about choice of suppliers for raw materials, choice of transportation modes, and choice of processing options, which might vary in terms of technology and capacity.

As the majority of the chapter is on safety stock placement in a supply chain, there is a close connection between this chapter and the body of literature on multi-echelon inventory systems. In this handbook, there are three chapters that focus, to some degree, on multi-echelon inventory models, namely Axsäter (Chapter 10), Song and Zipkin (Chapter 11) and de Kok and Fransoo (Chapter 12). We see three distinctions between the focus of this chapter and the general literature on multi-echelon inventory systems, as treated in these other chapters.

First, the primary emphasis on the approaches studied in this chapter is in terms of providing decision support for supply chain design, rather than supply chain operation. By this, we mean that the intent is to determine the best overall strategy for deploying safety stock across the supply chain so as to buffer it against demand uncertainty. In particular, we are concerned with questions about where are the best places in the supply chain to position a safety stock, and how much is needed to protect the chain. In contrast with much of the multi-echelon inventory literature, the intent is not to find the inventory control policy for operating the supply chain.

Second, much of the multi-echelon literature focuses on specific network topologies such as serial, assembly or distribution systems. We know, however, that de Kok and Fransoo in Chapter 12 do explicitly describe a multi-echelon algorithm that applies to general network structures. The purpose of this chapter is to consider multi-echelon models that have been specifically designed for optimizing the placement of safety stocks in real-world supply chains. As such, we find that the network topologies of most supply chains are neither an assembly nor distribution system, and thus require different approaches. Admittedly, in order to make progress on these more complex systems, these approaches for safety stock placement require simplifications and at times, strong assumptions. As a consequence, these safety stock models lack some of the rigor found in the literature for multi-echelon systems. But, on the plus side, they have had substantial success in being applied in practice.

Third, we assume that the inventory policies throughout the supply chain just rely on local information and make local decisions in terms of their inventory management and replenishment. In contrast, the models in these other three chapters allow for a central decision maker to coordinate and control the actions at all stages in the supply chain.

The structure of the chapter is to consider first two approaches to safety stock placement, which we term the stochastic-service model and the guaranteed-service model. These two approaches provide an interesting contrast to how one models and analyzes a supply chain for the purposes of setting safety stocks. We then address the issue of how to optimally configure the supply chain. We introduce the notion of options for each stage in the supply chain, where the options differ in terms of lead-time and cost. We show how this work builds on the safety-stock placement models and we formulate an optimization model that finds the best choice of options and safety stock placement to minimize the total supply chain cost. We conclude the chapter with some reflections on this material and its applicability and value to practice, as well as reflections on opportunities for research.

2 Approaches to safety stock placement

In this chapter, we consider two approaches to optimizing safety stock levels in multi-echelon supply chains. Our intent is to compare and contrast these approaches in terms of their underlying assumptions, computational and modeling implications, and the nature of the results produced. Both approaches adopt a network representation of the supply chain, where nodes in the network correspond to stages in the supply chain and arcs denote the precedence relationship between stages. A stage represents a processing or transformation activity in the supply chain. Depending on the scope and granularity of the analysis being performed, the stage could represent anything from a single step in a manufacturing or distribution process to a collection of such steps to an entire assembly and test operation. Regardless of the level of detail chosen by the modeler, a stage corresponds to the material flow of a single item or a single family of items, and each stage is a candidate location for the placement of a safety stock of inventory. When it is necessary to distinguish the safety stocks for different items at the same location, then we need to replicate the stages. For example, if two products flow through a distribution center, we might model each product in the supply chain map by a stage that corresponds to that SKU at that distribution center.

The approaches also assume decentralized control throughout the supply chain. There is no central decision maker that coordinates and controls the actions at all of the stages in the supply chain. Instead, for the purposes of determining safety stocks, we assume that each stage in the supply chain manages its inventory with a simple control policy that takes inputs from adjacent upstream and downstream stages. Thus, in order to be implemented in practice, the final recommendations of the model with regard to safety stocks must be translated into the control policies that are in use throughout the supply chain. As a final note on this issue, saying the supply chain is subject to decentralized control is not equivalent to saying the supply chain is locally optimized. In an optimization context, the models attempt to find the safety stock levels, under the assumption of decentralized control, that minimize the total safety stock cost for the supply chain. For the approaches presented here, this requires global access to information to run the optimization and calculate the system's performance measures. In particular, demand information is passed from finished goods stages through the chain to raw materials and cost and lead-time data is passed in the opposite direction.

The two approaches differ in how they model the replenishment mechanism between stages in the supply chain. We refer to the two approaches as the *stochastic-service model* and *guaranteed-service model*. The stochastic-service model assumes the delivery or service time between stages can vary based on the material availability at the supplier stage. The guaranteed-service model assumes that each stage can quote a delivery or service time that it can always satisfy.

In the stochastic-service model, each stage in the supply chain maintains a safety stock sufficient to meet its service level target. In this setting, a stage that has one or more upstream-adjacent supply stages has to characterize its replenishment time taking into account the likelihood that these suppliers will meet a replenishment request from stock. Because the upstream suppliers will not always meet demand requests immediately from stock, each stage will occasionally experience a delay in obtaining its supplies from its upstream suppliers. Due to this stochastic delay, the replenishment time for the stage is also stochastic, even when the processing time at the stage is deterministic. The inventory level required at each stage to meet its service level target depends on its replenishment time. And the challenge in this work is in how to characterize these replenishment times given that a stage might have multiple upstream suppliers, and given that each of the upstream stages might also be dependent upon unreliable suppliers.

In the guaranteed-service model, each stage provides guaranteed service to its customer stages. In this setting, a supply stage sets a service time to its downstream customer and then must hold sufficient inventory so that it can always satisfy the service-time commitment. A key assumption in this model is to assume that demand is bounded for the purposes of making the service-time guarantee. As a consequence, the service-time guarantee can be accomplished with a finite stock of inventory. The guaranteed nature of these service times assures that the replenishment time for downstream stages is predictable and deterministic. This then allows the downstream stage to plan its inventory so that it can also make a service-time guarantee to its customers. In this work, the challenge is determining the best choice of service times within the supply chain that minimize the total supply-chain inventory and meet the service requirements for the supply-chain's customer.

The stochastic-service and guaranteed-service approaches both require strong assumptions in order to produce tractable models. The stochastic-service model assumes that the system behaves the same under all demand conditions. That is, each stage reacts in a predictable way whether there is ample inventory or there is a stock-out, and inventory is the only countermeasure available to deal with demand and supply uncertainty in the supply chain. If one were to make an analogy to a checkout clerk in a grocery store, the stochastic-service model assumes that the clerk behaves the same way if the line is one person or fifty people.

The guaranteed-service model makes an equally strong assumption. In order to provide guaranteed service, the guaranteed-service model assumes that the safety stock policy is only being designed to meet some portion of the demand, as specified by the demand bound. When demand exceeds the bound, the model does not attempt to address how the system will react. In effect, the guaranteed-service model assumes that inventory is held to handle some nominal level of uncertainty and that other responses or tactics are available to address demand or supply uncertainty beyond this nominal level. Continuing the grocery-checkout analogy, if the system were designed to process a maximum of 20 customers in a one-hour interval, then when 25 customers show up in an hour, the model does not say how exactly the additional customers would be served. In effect, the model just assumes outside measures are adopted to serve these customers in the specified time frame. (We note here that the control framework proposed in Chapter 12 assumes that other countermeasures are applied when planned lead times are threatened.)

The next two sections discuss in more detail the papers that have appeared in both streams of work.

2.1 *Stochastic-service model approach*

Lee and Billington (1993) develop a multi-echelon inventory model to reflect the decentralized supply chain structure they witnessed in Hewlett-Packard's DeskJet printer supply chain. Their goal was to produce a model that manufacturing and materials managers could use to evaluate different strategic decisions involved with the creation of a new-product supply chain. They model a supply chain as a collection of SKU-locations where each stage in the supply chain accepts as an exogenous input a service level target or a base stock policy. In the case where service level targets were inputs, the authors develop a single-stage base-stock calculation that, while approximate, is tractable. The single-stage base-stock level is a function of the replenishment lead-time at the stage, which includes the production lead-time, plus the effects from production downtime and random delays due to component shortages. Lee and Billington show how to propagate the single-stage model to multiple stages by developing expressions for the random delays induced on downstream stages from shortages from the base-stock policy of upstream stages.

Ettl, Feigin, Lin and Yao (2000) also consider a supply chain context that is quite similar in spirit to the work of Lee and Billington (1993). The single-stage base-stock model in Ettl et al. (2000) makes a distinction between the *nominal* lead-time a stage quotes and the *actual* lead-time the stage experiences. The actual lead-time will exceed the nominal lead-time when there is a stock-out at a supplier. The authors develop an approximate characterization of the random variable for the actual lead-time. This approximation is based on assuming that at most one supplier is out of stock at any time instant, and then determining the stock-out probability for each supplier, given their service targets and this assumption. The authors use an $M/M/\infty$ model of the supplier's replenishment process to develop a bound on the expected delay induced by a stock-out. Weighting these delays by the stock-out probabilities for each supplier, and combining with the nominal lead-time provides the characterization of the actual lead–time for a single stage. Given this lead-time, a base-stock level is determined to assure a given service target for the stage. As in the case of Lee and Billington (1993), the single-stage model extends immediately to a multiple-stage supply chain. Indeed, given service level targets for every stage in the supply chain, it is possible to decompose the performance analysis of a multiple-stage system into the analysis of a series of single-stage base-stock systems.

In addition to performance analysis, Ettl et al. (2000) go on to place their supply chain model into an optimization context. The authors' objective function is to minimize the total inventory investment in the supply chain, defined as work-in-process inventory plus safety stock inventory. The decision variables are the safety factor (or service level) at each stage. The authors then develop expressions for the partial derivative of the objective function with respect to the safety factors. This formulation allows the authors to solve the resulting nonlinear programming problem using conjugate gradient methods.

Glasserman and Tayur (1995) consider a context very similar to that of Lee and Billington (1993) and Ettl et al. (2000) but go on to introduce capacity limits into their multi-echelon model. The introduction of production capacity requires each stage to operate a modified base-stock policy where at each period the stage will order the minimum of its capacity and the amount to bring its inventory position back to the base-stock level. The problem formulation of Glasserman and Tayur (1995) follows the framework of Clark and Scarf (1960) with the addition of capacity. The authors first develop recursions for stage inventories, production levels, and pipeline inventories. The authors develop estimates of the derivatives of the inventory requirements with respect to the base-stock levels, based on an infinitesimal perturbation analysis. They use these estimates to generate the gradient of the cost function, with which they can conduct a gradient-based search to find the optimal base-stock policy.

2.2 *Guaranteed-service model approach*

The guaranteed service-time approach traces its lineage back to the 1955 manuscript, which was later reprinted in 1988 (Kimball, 1988). In that paper, Kimball describes the mechanics of a single stage that operates a base-stock policy in the face of random but bounded demand. In particular, beyond the deterministic production time assumed at the stage, there is an incoming service time that represents the delivery time quoted from the stage's supplier and an outgoing service time representing the delivery time the stage quotes to its customer. Kimball further assumes that demand over any interval of time is bounded. Given this characterization, the base-stock level at the stage is set equal to the maximum demand over the net replenishment time, which is defined as the incoming service time plus the production time minus the outgoing service time.

Simpson (1958) develops a model to determine the optimal safety stocks in a serial supply chain. Simpson uses Kimball's work as the building block, coupling adjacent stages together through the use of service time. In particular, the incoming (or inbound) service time of a downstream stage is equal to the outgoing (or outbound) service time of the upstream stage, namely its supplier. The optimal stocking locations in the supply chain can then be found by determining the optimal service times in the supply chain. Simpson proves that for a serial supply chain the optimal service times satisfy an extreme point property where the outgoing service time at a stage will equal either zero or its incoming service time plus its production time. In terms of inventory, the optimal policy is an 'all or nothing' policy, in which a stage either has no safety stock or carries a decoupling safety stock, namely enough stock to decouple the downstream stages from the upstream stages. Simpson suggests an enumeration procedure to find the optimal service times.

Simpson also provides a rich interpretation for the bounded demand process. Rather than saying that bounded demand reflects the maximum demand the stage will see, the bound can instead reflect the maximum amount of demand the company wants to satisfy from safety stock. Under this interpretation, when demand exceeds the bound, the stage will have to resort to extraordinary measures, like expediting and overtime, to meet the demand requirement by means other than using safety stock.

Graves (1988) observes that the serial-line problem, as formulated by Simpson (1958), can be solved as a dynamic program. Inderfurth (1991), Inderfurth and Minner (1998), Graves and Willems (1996, 2000) extend Simpson's work to supply chains modeled as assembly networks, distribution networks, and spanning trees. In each case, the optimization problem is still to determine the service times that minimize the total cost for safety stock in the supply chain. The challenge is to determine an efficient approach to traverse the state space of the dynamic program. The definition of service time must also be expanded to include the cases where a stage can see more than one upstream or downstream stage. In the case of multiple upstream stages feeding a downstream stage, the papers assume that the downstream stage has to wait until the item with the longest service time arrives. In the case of an upstream stage supplying multiple downstream stages, the papers assume the upstream stage quotes the same service time to all of its adjacent downstream stages.

3 Model formulation

The single-stage base-stock policy is the common building block for all of the papers in this chapter. The differences in the approaches deal with how adjacent stages interact with one another and the assumptions about the operating behavior of the individual stages. This section develops both the underlying base-stock equations and the resulting multi-echelon problem formulations for both the guaranteed and stochastic-service models.

The goal of this section is to distill the two models into their simplest elements so that the reader can clearly see the similarities and differences between the two approaches. The goal is not to replicate the contents of the papers surveyed in the literature review nor is it to provide a careful development of the features of each model. Rather, the intent is to give a self-contained development that highlights the essence of each approach. We refer the reader to the specific papers for the critical details and refinements that are necessary for the successful implementation of each model.

To get started, we note the key similarities of the two approaches. In each model, each stage in the supply chain operates according to a base-stock policy. In each period, the stage observes demand and places a replenishment order on its suppliers equal to the observed demand. There are no capacity constraints. There is a common underlying review period for all stages in the supply chain. Demand is stationary and independent across nonoverlapping intervals, with mean demand per period of μ and a standard deviation of σ. Associated with each stage is a deterministic processing time (or lead-time) that includes all of the time required to transform the item at the stage. Once all of the stage's required inputs are available, the processing time includes any waiting time, manufacturing time and transportation time at the stage.

3.1 Stochastic-service model

In the stochastic-service model, each stage sets its base stock to meet a service level target, i.e., an upper bound on the probability that a stage is out of stock in any period and thus cannot meet customer demand directly from stock on hand. The service level target for external customers is an exogenous input to the model, usually dictated by market conditions. The service level target for internal customers is an input when the model is used for performance evaluation; alternatively, the service level target for internal customers is a decision variable when the model is placed within an optimization. We say that a stage provides stochastic service because a demand order will receive immediate service when stock is on hand, but will be subject to a random delay when the stage is out of stock.

The replenishment time at a stage equals the processing time at the stage plus any delay from upstream stages. If we denote the replenishment time at stage j as a random variable τ_j, the processing time as a constant L_j, and the delay for supplier i as a random variable Δ_i, then the replenishment time at stage j equals:

$$\tau_j = L_j + \max_{i:(i,j)\in A} \{\Delta_i\} \tag{3.1}$$

where A is the set of directed arcs in the network representation of the supply chain.

In the worst case, this delay might equal the entire replenishment time from its slowest supplier, e.g.,

$$\tau_j = L_j + \max_{i:(i,j)\in A} \{\tau_i\} \tag{3.2}$$

The development of an exact characterization of τ_j is extremely challenging. To illustrate, with N suppliers, there are 2^N–1 combinations of suppliers that might be out of stock in any period. For each stage that is out of stock, determining its delay requires considering where its first unallocated unit is in its replenishment process. Finally, there are the multi-echelon ramifications when a supplier's supplier is out of stock. As a consequence, one must make some simplifications to make the analysis of this model more tractable. We describe an approach here, which is loosely based on the development in Ettl et al. (2000).

For purposes of illustration, we assume that at most one supplier will stock out per period and the delay will equal the supplier's processing time. (This assembly assumption is also discussed in Chapter 12.) This allows us to express the expected replenishment time at stage j as:

$$E[\tau_j] = L_j + \sum_{i:(i,j)\in A} \pi_{ij} L_i \tag{3.3}$$

where π_{ij} is the probability that in a period stage i is causing a stock-out at stage j. Ettl et al. (2000) use this form of equation for the expected replenishment time, but use a bound on the expected delay, rather than the supplier's processing time as we have done in (3.3). They derive the bound on the expected delay by means of applying an M/M/∞ model to the supplier's replenishment process.

We assume the demand over the replenishment time is normally distributed with mean $\mu_j E[\tau_j]$ and with standard deviation $\sigma_j \sqrt{E[\tau_j]}$. We assume that the base stock is given by $B_j = \mu_j E[\tau_j] + k_j \sigma_j \sqrt{E[\tau_j]}$, where k_j is the safety factor necessary to achieve the service level target for the stage.

To determine the expected replenishment time, given by (3.3), we need to find an expression for π_{ij}. Ettl et al. (2000) propose the following calculation for π_{ij}:

$$\pi_{ij} = \frac{1-\Phi(k_i)}{\Phi(k_i)} \left(1 + \sum_{h:(h,j) \in A} \frac{1-\Phi(k_h)}{\Phi(k_h)} \right)^{-1} \tag{3.4}$$

where k_i denotes the safety factor at stage i and $\Phi(k_i)$ represents the cumulative distribution function for a standard normal random variable. In (3.4), the term in parentheses acts to normalize the probability that a stock-out does or does not occur, and the first part of the expression calculates the fraction of occurrences that are attributable to stage i.

Given the assumptions of normally distributed demand over the replenishment lead-time, we follow the development in Ettl et al. (2000) to get the following expression for the expected on-hand inventory for stage j:

$$E[I_j] = k_j \sigma_j \sqrt{E[\tau_j]} + \sigma_j \sqrt{E[\tau_j]} \int_{z=k_j}^{\infty} (z - k_j) \phi(z) \, dz \tag{3.5}$$

where τ_j is defined by (3.3) and (3.4), and $\phi()$ is the probability density function for a standard normal. The first term is the expected inventory level at stage j, equal to the base stock level net the expected demand. Since the on-hand inventory level cannot be negative, we need to augment the first term with the second term, which corresponds to the expected number of shortages or backorders.

We can now develop an expression for the total safety stock cost across the supply chain. We let C_j^S denote the per unit holding cost of safety stock at stage j. C_j^S is typically determined by multiplying the cumulative cost of the product at stage j by a holding cost rate. Given this cost characterization, we let $\mathbf{C}^{\text{ssm}}$ denote the total safety stock cost of the stochastic-service model. Then,

$$\mathbf{C}^{\text{ssm}} = \sum_{j=1}^{N} \mathrm{C}_j^{\mathrm{S}} \sigma_j \sqrt{E[\tau_j]} \left(k_j + \int_{z=k_j}^{\infty} (z - k_j) \phi(z) \, dz \right) \tag{3.6}$$

We can use (3.6) for performance evaluation in a supply chain, namely to find the inventory requirements and costs for a given set of service level targets or safety factors. We can also place (3.6) in an optimization context, where the objective is to minimize safety stock cost and the decision variables are the service level targets, or equivalently the safety factors, at each stage in the supply chain.

3.2 Guaranteed-service model

In the guaranteed-service model, each stage sets its base stock so as to guarantee that it can meet its service-time commitment to its customers. That is, each stage will quote a guaranteed service or delivery time to its downstream customers, who know that this commitment will be met with certainty. The service time for external customers is an exogenous input, just as with the service level target for the stochastic-service model. The service time for internal customers can be either an input or a decision variable, depending upon whether the model is being used for performance evaluation or for optimization.

As we have noted earlier, the guarantee applies to a bounded demand process. We specify for each stage j a function $D_j(t)$ that represents the maximum demand over t consecutive periods for which we will guarantee the service commitment. For each stage j, the model finds the base stock that satisfies the stage's service time commitment, provided that the demand time series is always within the demand bound given by $D_j(t)$. In a typical application, similar to the stochastic-service model, one might assume that actual demand at stage j is normally distributed with mean demand per period of μ_j and a standard deviation of σ_j. Then a common way to set the demand bound is as follows:

$$D_j(t) = t\mu_j + k_j\sigma_j\sqrt{t},$$

where k_j is a given safety factor. When demand exceeds the demand bound, then the safety stock in the system will not be adequate to assure the service times. We assume that in this case of extraordinary demand, some correspondingly extraordinary measures are taken to augment the safety stock so that the demand can be served. Alternatively, one might view demand in excess of this bound as being lost or somehow being served from another source.

For the guaranteed-service model, the replenishment time at a stage does not drive base stock requirements, but, rather, it is the net replenishment time that is of importance. In order to understand net replenishment time, we first define the concept of service time. Service time is the amount of time that elapses between when a downstream stage places an order on an upstream stage and when the order is delivered by the upstream stage to the downstream stage and is available to begin processing at that stage. Each stage in the supply chain quotes a service time to its downstream (customer) stages and it is quoted service times from its upstream (supplier) stages. We describe the service time that stage j quotes its customers as the outbound service time, denoted by s_j^{out}. The inbound service time at stage j is denoted by s_j^{in}. Since stage j cannot start its processing activities until it receives all of the required inputs, we can state the inbound service time at stage j in terms of the outbound service times for its suppliers:

$$s_j^{in} = \max_{i:(i,j)\in A} \left\{s_i^{out}\right\}. \tag{3.7}$$

Now, for the guaranteed-service model the replenishment time at stage j is:

$$\tau_j = s_j^{in} + L_j. \tag{3.8}$$

Since both the service time and processing time (by assumption) are deterministic constants, we have that the replenishment time for this model is also deterministic. The net replenishment time for stage j is the replenishment time minus the stage's outbound service time, i.e., $s_j^{in} + L_j - s_j^{out}$. In this model, we set the base stock for each stage to cover the maximum demand over its net replenishment time, as will be shown next.

We assume each stage j starts at time 0 with initial inventory $I_j(0) = B_j$. Given the assumptions of guaranteed service and the definition of the service times, the inventory at time t, $I_j(t)$, equals

$$I_j(t) = B_j - \sum_{v=0}^{t-s_j^{out}} d_j(v) + \sum_{w=0}^{t-L_j-s_j^{in}} d_j(w), \tag{3.9}$$

where $d_j(t)$ denotes the demand in period t. In period t, stage j completes into its inventory the replenishment order that was placed in period $t - L_j - s_j^{in}$. Correspondingly, in period t, stage j must serve the replenishment orders placed by its customers in period $t - s_j^{out}$. We can simplify (3.9) as,

$$I_j(t) = B_j - \sum_{v=t-L_j-s_j^{in}+1}^{t-s_j^{out}} d_j(v). \tag{3.10}$$

In order to satisfy the service-time guarantee, we need to set the base stock B_j so that the inventory on hand $I_j(t)$ is always non-negative. That is, we will want to set

$$B_j \geq \sum_{v=t-L_j-s_j^{in}+1}^{t-s_j^{out}} d_j(v). \tag{3.11}$$

In words, we need for the base stock to equal (or exceed) the maximum possible demand over the net replenishment time. But given the assumption of a bounded demand process, then we can set $B_j = D_j(s_j^{in} + L_j - s_j^{out})$ and be assured that Eq. (3.11) holds.

For illustration, assume we set the demand bound as given earlier. Then we choose $B_j = (s_j^{in} + L_j - s_j^{out})\mu_j + k_j\sigma_j\sqrt{s_j^{in} + L_j - s_j^{out}}$. We can immediately find that the expected inventory on-hand at stage j equals:

$$E[I_j] = k_j\sigma_j\sqrt{s_j^{in} + L_j - s_j^{out}}. \tag{3.12}$$

We can use (3.12) to determine the inventory requirements for a given setting of the service times in a supply chain. We can also incorporate (3.12) into an optimization to find the best choice of service times. The objective function for the optimization could be the total holding cost for safety stocks, which we denote by $\mathbf{C}^{\mathrm{gsm}}$ and state as:

$$\mathbf{C}^{\mathrm{gsm}} = \sum_{j=1}^{N} C_j^S k_j\sigma_j\sqrt{s_j^{in} + L_j - s_j^{out}}. \tag{3.13}$$

In the optimization model, one minimizes the objective (3.13) with the decision variables being the service times at the stages in the supply chain and subject to constraints (3.7) to relate the inbound to the outbound service times, and non-negativity constraints on the net replenishment times.

4 Heavy industry and consumer packaged goods example

In this section we apply the two approaches for safety stock placement to two examples. Our intent is two-fold. First, we wish to show the applicability of these approaches to different industries. The examples presented in Lee and Billington (1993); Ettl et al. (2000); Graves and Willems (2000) are all drawn from the high-technology industry. Here we will present examples from two other industries, heavy industry and consumer packaged goods, to demonstrate the characteristics of their supply chains and the differences in the structure of their optimal solutions. Our second purpose is to illustrate how the results of the two approaches can differ, and then to discuss the implications for implementing these models.

4.1 Bulldozer assembly and manufacturing

In this section we present the assembly and manufacturing process for a bulldozer. Figure 1 presents the bulldozer's supply chain map.

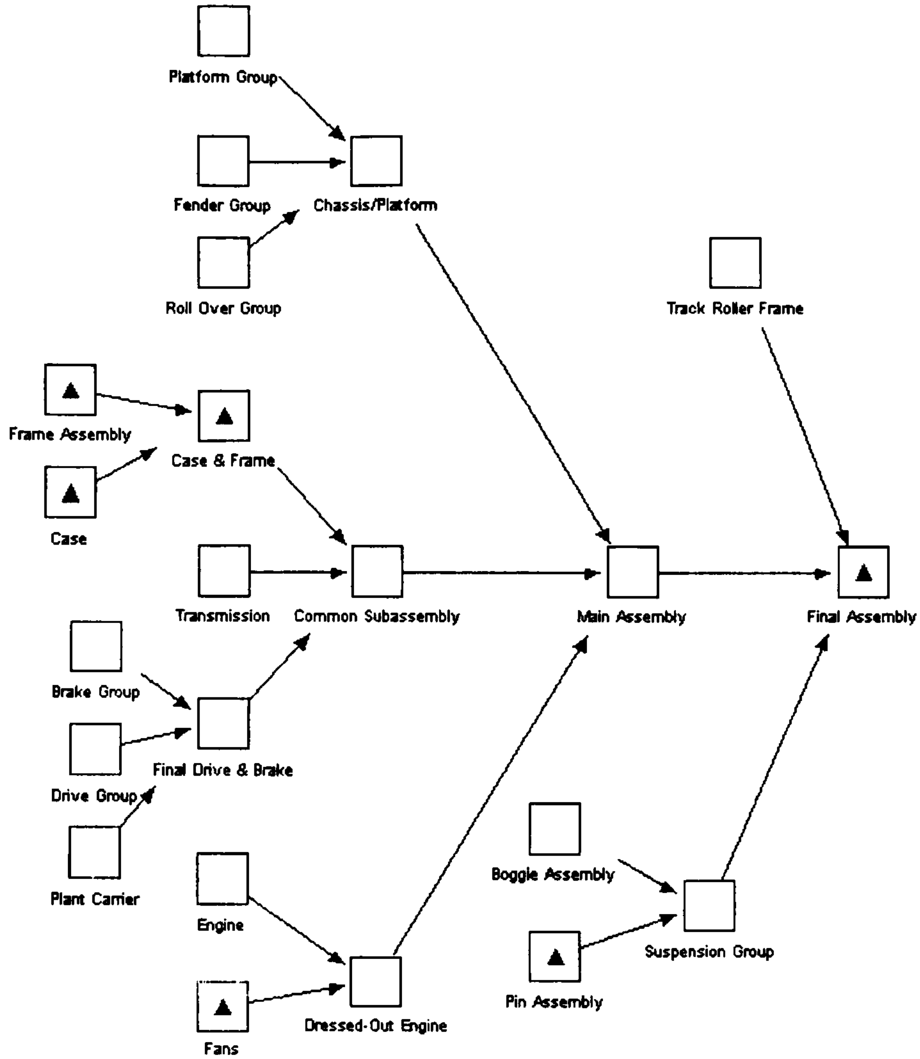

Fig. 1. Bulldozer Supply Chain Map.

At a high level, a bulldozer is put together in three operations. In the common subassembly step, the transmission, drivetrain, and brake system are attached to the case and frame. The main assembly process then installs the chassis and the engine to the common subassembly. In final assembly the track and suspension are installed. A real-world map of this supply chain exceeds 1000 stages with many of the stages shown in Fig. 1 spawning their own large supply chains. Whereas we have combined many stages in order to present the supply chain, the structure of the supply chain accurately represents the manufacturing process and flow for bulldozers. In addition, we have

Table 1
Parameters for Bulldozer Supply Chain

Stage name	Nominal time	Stage cost ($)
Boggie assembly	11	575
Brake group	8	3850
Case	15	2200
Case & frame	16	1500
Chassis/platform	7	4320
Common subassembly	5	8000
Dressed-out engine	10	4100
Drive group	9	1550
Engine	7	4500
Fans	12	650
Fender group	9	900
Final assembly	4	8000
Final drive & brake	6	3680
Frame assembly	19	605
Main assembly	8	12,000
Pin assembly	35	90
Plant carrier	9	155
Platform group	6	725
Roll over group	8	1150
Suspension group	7	3600
Track roller frame	10	3000
Transmission	15	7450

made two major simplifications to the supply chain. First, we ignore the customization process in this analysis; in effect, we are modeling the supply chain for a stock bulldozer that will be modified at the dealer. Second, we have not modeled all of the different variations the bulldozer can come in. Besides adding complexity to the example, after adding the different variations, the network is no longer a spanning tree, and thus requires a solution technique presented in Humair and Willems (2003). Table 1 provides the cost and nominal lead-time information for the supply chain.

The cumulative cost of the product is already $28,990 when the common subassembly stage is complete. The chassis/platform and engine subcomponents each contribute $7095 and $9250, respectively, to the cost of the product. The total cost for the bulldozer, upon completion of final assembly, is $72,600. Because of the lean manufacturing initiative at the company, lead-times are quite low given the complexity of the product. The average daily demand is 5 and the daily standard deviation is 3. Assuming 260 days per year, the cost of goods sold is $94,380,000. The company applies an annual holding cost rate of 30% when calculating inventory costs.

For the guaranteed-service model, we set the demand bound to correspond to the 95th percentile of demand, and thus, set the safety factor as $k = 1.645$. Figure 1 graphically presents the optimal solution to the guaranteed-service model; a triangle within a stage designates that the stage

Table 2
Optimal Service Times and Safety Stock Costs under Guaranteed-Service Model

Stage name	Service time	Stage safety stock cost ($)
Boggie assembly	11	0
Brake group	8	0
Case	0	12,614
Case & frame	15	6373
Chassis/platform	16	0
Common subassembly	20	0
Dressed-out engine	20	0
Drive group	9	0
Engine	7	0
Fans	10	1361
Fender group	9	0
Final assembly	0	607,969
Final drive & brake	15	0
Frame assembly	0	3904
Main assembly	28	0
Pin assembly	21	499
Plant carrier	9	0
Platform group	6	0
Roll over group	8	0
Suspension group	28	0
Track roller frame	10	0
Transmission	15	0

holds a safety stock. Since the underlying network is a spanning tree with 22 nodes and 21 arcs, we can optimize the network with the algorithm from Graves and Willems (2000). The resulting optimal service times and holding costs for the safety stock are displayed in Table 2.

The optimal inventory policy does not demonstrate the clear decoupling policy that one often sees in the guaranteed-service model. There is a large safety stock at final assembly, which is necessary to provide immediate service to the distribution department of the company. The safety stock at final assembly is sized to cover the demand variability over the net replenishment time for the stage of 28 days. The remaining stages, for the most part, carry no safety stock; the exceptions are a few long lead-time stages where safety stock is held so as to keep the net replenishment time for final assembly to 28 days. In total, the annual holding cost for the safety stock in the supply chain is $633,000.

To understand the solution better, we repeated the optimization but with a constraint that forced the common subassembly to have a service time of zero and thus to hold a safety stock. A priori one might suspect that having a safety stock at the common subassembly stage would lead to a good if not optimal solution, as this would seem to be a logical point to decouple the chain. From the resulting optimization, we found that the chassis/platform and dressed-out engine stages also quote service times of zero and are thus decoupling points. However, the annual holding cost for safety stock increases by nearly 10%, from $633,000 to $693,000. In contrast to our experience with supply chains for high-tech products, we observe that the incredibly expensive nature of the components and the relatively short lead-times make it uneconomical to develop local decoupling points.

We begin the analysis of the stochastic-service model by determining the range of allowable service targets per stage. To maintain consistency with the presentation of the guaranteed-service model, we again assumed a 95% service level at the final assembly stage. For the other stages, the service level is a parameter that is to be set or serve as a decision variable in an optimization. We need for the resulting service levels to be consistent with the assumptions that were made in the development of (3.3), our calculation of the expected replenishment time. In particular, we assumed that for each stage, at most one of its supplier stages stocks out in any period. In order to make this assumption operational, we set an upper bound on the probability that two or more of a stage's suppliers stock out in a period, namely 0.10. Thus, we restrict the choice of service levels so that the probability that a stage has two or more suppliers out of stock is no more than 0.10.

In the bulldozer supply chain, stages have between one and three suppliers. For stages with three suppliers, we impose a lower bound of 0.80 on the service level for each supplier. For stages with two suppliers, we impose a lower bound of 0.68 on the service level for each supplier. As justification for these lower bounds, we observe that with an assumption that the stock-out events of the suppliers are independent, then setting the service levels to these lower bounds results in the stage having a probability of 0.90 that at most one supplier is out of stock. There is no claim that this is the best way to implement the stochastic-service model; rather, we argue that this seems a reasonable way to proceed with the model based on the assumptions that underlie its development, and given the purposes of this chapter. Finally, we will also use 68% as the lower bound on the service level for the case of a sole supplier.

In Table 3 we report the results for the stochastic-service model when the service level for each stage is set to its lower bound. The table displays the expected lead-time and annual holding cost for the safety stock for each stage in the bulldozer supply chain. We cannot guarantee that this is the best solution for the stochastic-service model for this example. However, we did conduct an extensive grid search over the service levels and found the lower bounds on the service levels always to be binding.

As expected, every stage carries a safety stock sufficient to cover the expected lead-time. On a percentage basis, two types of stages have expected lead-times that differ significantly from their nominal times. First, there are those stages that have short nominal lead-times; final assembly is an

Table 3
Nominal and Expected Lead-times for the Stochastic-Service Model

Stage name	Nominal lead-time	Service level (%)	Expected lead-time	Stage safety stock cost ($)
Boggie assembly	11	68	11.00	1160
Brake group	8	80	8.00	9342
Case	15	68	15.00	5181
Case & frame	16	80	24.24	18,184
Chassis/platform	7	80	10.29	19,521
Common subassembly	5	80	10.29	79,764
Dressed-out engine	10	80	14.61	30,328
Drive group	9	80	9.00	3989
Engine	7	68	7.00	7240
Fans	12	68	12.00	1369
Fender group	9	80	9.00	2316
Final assembly	4	95	7.57	299,472
Final drive & brake	6	80	9.71	24,693
Frame assembly	19	68	19.00	1604
Main assembly	8	80	11.14	164,194
Pin assembly	35	68	35.00	324
Plant carrier	9	80	9.00	399
Platform group	6	80	6.00	1524
Roll over group	8	80	8.00	2791
Suspension group	7	80	18.15	15,589
Track roller frame	10	80	10.00	8139
Transmission	15	80	15.00	24,754

example of this kind of stage. Second, there are stages with suppliers that have long nominal lead-times; the suspension group is an example of this kind of stage.

Whereas the safety stock in final assembly is now much less than in the case of the guaranteed-service model, overall we find that there is about 12% more inventory cost with the stochastic-service model. In the stochastic-service model, safety stock is the single countermeasure to address the demand variability in the supply chain. In the guaranteed-service model, safety stock is used to protect against demand variability up to the demand bound. The model assumes that other countermeasures, including expediting and overtime, are utilized when demand exceeds the demand bound. While the guaranteed-service model does not quantify the cost of these other countermeasures, understanding the gap between the two models presented gives some indication of the relative benefit of using only safety stock as a countermeasure versus other operational tactics. Figure 2 displays the total annual holding cost as a function of the service level for the external customer for two different types of policies for each model.

The top two lines represent what one could reasonably consider being the base case for each of the models. For the stochastic-service model, each stage maintains a service level equal to final assembly's service level. For the

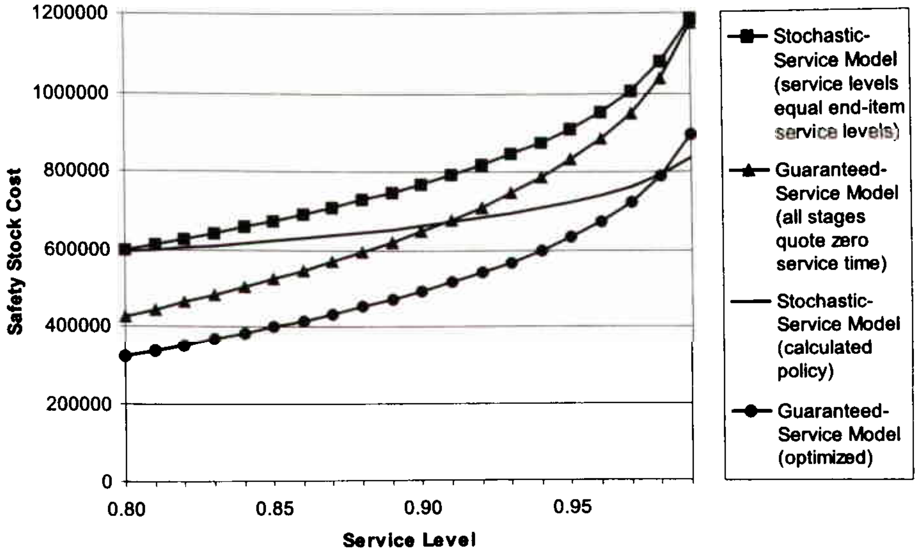

Fig. 2. Safety Stock Cost as a Function of Service Level in Bulldozer Supply Chain.

guaranteed-service model, each stage quotes a zero service time. For each model, the base case has each stage holding significant safety stock so as to decouple it from the other stages in the supply chain. For extremely high service levels, the two models are virtually identical, which is not a surprise given the assumptions in the models. As service levels decrease, there is a difference in cost between the models, due to the fact that the replenishment times for the stochastic-service model increase due to delays from the suppliers. In this example, an increase in the replenishment time at many stages has a large impact on the stage's inventory requirement, due to the expensive nature of the product and the relatively short nominal lead-times.

The lower two lines represent the best inventory policy identified for the stochastic-service model and the optimal inventory policy for the guaranteed-service model. The difference between the two policies allows the manager to quantify the cost of using countermeasures other than inventory. For example, at the 95% service level the cost difference is $80,000 and at 85% service the difference is $224,000; Table 4 provides the numerical values for each of the four policies.

Table 4 also helps determine the appropriate demand bound. These calculations allow us to trade off the cost of safety stock against the cost of other tactics, like expediting and subcontracting, which can also be employed to satisfy demand. If these other tactics are cheaper than holding the higher levels of safety stock, then it makes financial sense to adopt a safety stock policy that only meets 95%, or less, of the possible demand realizations.

Table 4
Safety Stock Cost as a Function of Service Level in Bulldozer Supply Chain

Service level	Stochastic-service model (service levels equal end-item service levels)	Guaranteed-service model (all stages quote zero service time)	Stochastic-service model (calculated policy)	Guaranteed-service model (optimized)
0.80	596,618	425,062	593,788	323,743
0.81	611,063	443,382	599,044	337,697
0.82	625,931	462,306	604,532	352,110
0.83	641,281	481,902	610,279	367,035
0.84	657,215	502,252	616,330	382,534
0.85	673,780	523,452	622,711	398,680
0.86	691,037	545,616	629,459	415,562
0.87	709,091	568,885	636,623	433,284
0.88	728,073	593,428	644,268	451,977
0.89	748,150	619,459	652,469	471,803
0.90	769,527	647,249	661,322	492,969
0.91	792,481	677,150	670,952	515,742
0.92	817,409	709,633	681,531	540,483
0.93	844,880	745,350	693,297	567,686
0.94	875,651	785,240	706,570	598,068
0.95	911,043	830,735	721,877	632,719
0.96	953,156	884,186	740,048	673,429
0.97	1,006,022	949,896	762,629	723,477
0.98	1,078,523	1,037,248	792,955	790,007
0.99	1,185,142	1,174,924	836,583	894,866

4.2 Battery manufacturing and distribution

Figure 3 presents a supply chain map for a single battery product line. This supply chain depicts the manufacturing and packaging process for one size of battery that is sold in three regions in three types of packaging. In this setting, the battery manufacturing stage represents the manufacturing of a single size like AAA, AA, C, or D. The battery size is produced in a single bulk manufacturing facility. Finished batteries are then sent to three pack locations that produce specialty battery packages. For example, the package that comprises an end-item SKU is distinguished by the number of batteries included, the artwork on the package, and the inclusion or exclusion of items like RFID tags, hangers, bar codes, and price labels. Each SKU is sent to the company's three distribution centers (DCs) in the United States for distribution to regional markets. The nominal lead-time and direct cost added for each stage are displayed in Table 5. Table 5 demonstrates the commodity nature of the business. Materials have relatively short lead-times and the cost per item is extremely low. Whereas the cumulative cost of a bulldozer is $72,600, the total unit cost of a battery is less than one dollar. Indeed, the material and process cost to package the battery is on the same order of magnitude as the cost of the battery.

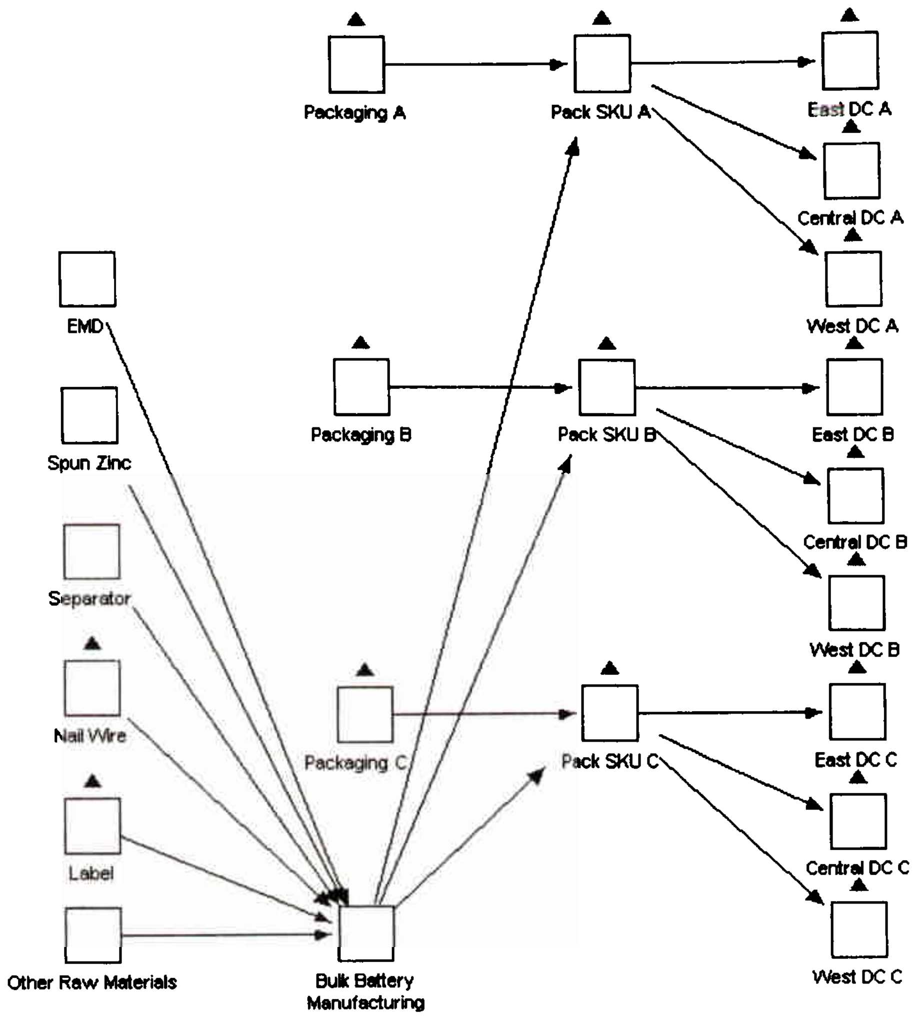

Fig. 3. Graphical Representation of the Battery Supply Chain.

The company's holding cost rate is 25%. The daily demand parameters are given in Table 6. Assuming 260 days per year, the expected cost of goods sold per year is $53,569,000. On a daily basis the demand is highly variable with the coefficient of variation being more than two in several instances. Clearly, the daily demand of a SKU at a DC is not well modeled as being from a normal distribution. Nevertheless, we note that for each model we are effectively assuming that the demand over a stage's replenishment time is normally distributed, which seems more plausible given the range of nominal lead-times. Furthermore, for all of the upstream stages, the

Table 5
Parameters for Battery Supply Chain

Stage name	Nominal time	Stage cost ($)
Bulk battery manufacturing	5	0.07
Central DC A	6	0.02
Central DC B	6	0.01
Central DC C	4	0.01
East DC A	4	0.00
East DC B	4	0.01
East DC C	4	0.01
EMD	2	0.13
Label	28	0.06
Nail wire	24	0.02
Other raw materials	1	0.24
Pack SKU A	11	0.07
Pack SKU B	11	0.12
Pack SKU C	9	0.24
Packaging A	28	0.16
Packaging B	28	0.24
Packaging C	28	0.36
Separator	2	0.02
Spun zinc	2	0.05
West DC A	5	0.01
West DC B	8	0.03
West DC C	6	0.06

demand is pooled over several regions and, except for packaging, over all SKUs. Hence, the assumption of normality seems defensible for the purposes at hand.

For the guaranteed-service model, we set the demand bound again to correspond to the 95th percentile of the demand process. Figure 3 graphically displays the optimal safety stock locations when the guaranteed-service model is optimized; a triangle within a stage denotes safety stock being held at the stage. The resulting optimal service times are displayed in Table 7. The interesting result here is that the bulk manufacturing facility does not hold any safety stock. Instead the three packing locations are the decoupling points in the supply chain. The intuition is that the packing locations are able to pool the demand variability for the three DCs and also pool the variability over the lead-time from the bulk manufacturing plant. This is more cost effective than holding inventory at the bulk plant but then having the regional DCs holding a safety stock that covers not only their lead-time but the pack lead-times as well. The optimal annual holding cost for safety stock is $853,000.

For the stochastic service method, we have a 95% service level target for each SKU at each regional DC. As we did with the bulldozer example, we assess a lower bound on the service level provided by each supplier to a stage, where the lower bound depends on the number of suppliers. In the battery

Table 6
Demand Information for Battery Supply Chain

Stage name	Mean demand	Standard deviation of demand
Central DC A	43,422	67,236
Central DC B	16,350	39,552
Central DC C	5536	11,213
East DC A	67,226	109,308
East DC B	15,765	34,079
East DC C	6416	14,125
West DC A	65,638	119,901
West DC B	10,597	23,277
West DC C	3519	6576

Table 7
Optimal Service Times using Guaranteed-Service Model

Stage name	Service time	Stage safety stock cost ($)
Bulk battery manufacturing	7	0
Central DC A	0	56,889
Central DC B	0	38,245
Central DC C	0	11,066
East DC A	0	73,716
East DC B	0	26,907
East DC C	0	13,940
EMD	2	0
Label	2	23,361
Nail wire	2	7163
Other raw materials	1	0
Pack SKU A	0	251,253
Pack SKU B	0	94,741
Pack SKU C	0	37,573
Packaging A	7	52,953
Packaging B	7	25,852
Packaging C	7	13,022
Separator	2	0
Spun zinc	2	0
West DC A	0	91,507
West DC B	0	26,531
West DC C	0	8279

supply chain, six suppliers supply the bulk battery manufacturing plant. To maintain a probability of 0.90 that at most one stage will be out of stock in a period, each of the six supplier stages must have a service level of at least 0.91. The bulk battery is combined with packaging at each pack location. We set the lower bound on the service level for these two inputs for the pack location to be 0.68. Since the three pack locations are themselves

Table 8
Expected Lead-times and Safety Stock Costs for the Stochastic-Service Model

Stage name	Nominal lead-time	Service level (%)	Expected lead-time	Stage safety stock cost ($)
Bulk battery manufacturing	5	68	8.66	54,467
Central DC A	6	95	6.55	60,191
Central DC B	6	95	6.55	40,465
Central DC C	4	95	4.32	11,649
East DC A	4	95	4.55	79,616
East DC B	4	95	4.55	29,060
East DC C	4	95	4.32	14,674
EMD	2	91	2.00	11,799
Label	28	91	28.00	20,375
Nail wire	24	91	24.00	6288
Other raw materials	1	91	1.00	15,402
Pack SKU A	11	95	19.00	261,404
Pack SKU B	11	95	19.00	98,568
Pack SKU C	9	95	17.00	39,220
Packaging A	28	68	28.00	25,117
Packaging B	28	68	28.00	12,263
Packaging C	28	68	28.00	6177
Separator	2	91	2.00	1815
Spun zinc	2	91	2.00	4538
West DC A	5	95	5.55	97,628
West DC B	8	95	8.55	27,775
West DC C	6	95	6.32	8606

finished goods locations, we assume that both the pack locations and the nine DC-SKU pairs all set a service level equal to the customer service level target, namely 0.95. Again, we find that the best solution for the stochastic-service model seems to set the service level to the lower bound at each stage. Table 8 displays the nominal and expected lead-times for the battery supply chain.

On a percentage basis, the bulk manufacturing stage and the three packaging stages have the greatest increase between nominal and expected lead-time. The difference is attributed to the fact that these stages each have one or more suppliers with significant nominal lead-times. The total safety stock cost under the stochastic-service model is $927,000.

To gain more insight into the role that different countermeasures may play in this supply chain, Figure 4 displays the cost for each policy under different end-item service levels. Varying the service levels at all end-item stages does not change the structure of the optimal policy under the guaranteed-service model. For the stochastic-service model, the nine DCs and the three pack locations had their service levels changed to equal the current end-item service level while all the other stages maintained their existing levels from Table 8. The safety stock costs are shown graphically in Figure 4 and reported in Table 9. One can view the cost difference as being the additional

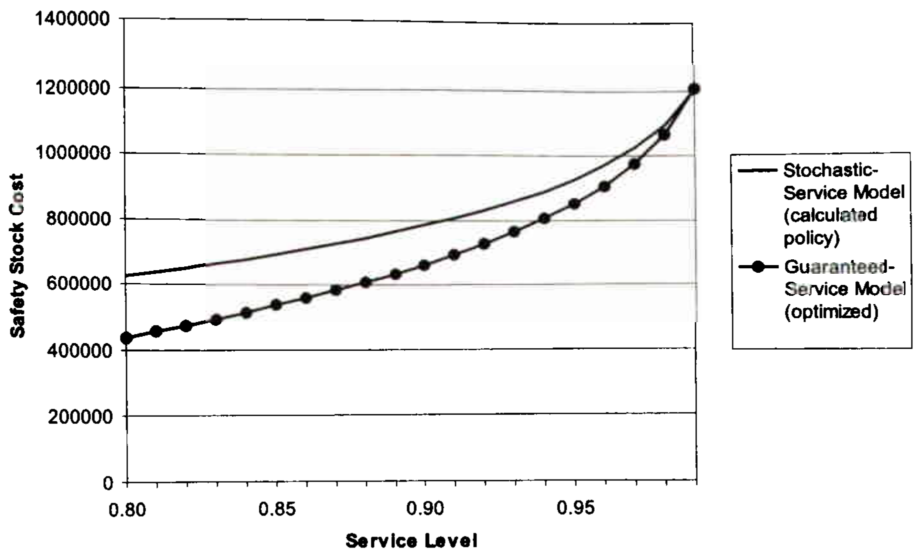

Fig. 4. Safety Stock Cost as a Function of Service Level in Battery Supply Chain.

Table 9
Safety Stock Cost as a Function of Service Level in Battery Supply Chain

Service level	Stochastic-service model (calculated policy)	Guaranteed-service model (optimized)
0.80	627,155	436,454
0.81	639,637	455,266
0.82	652,624	474,696
0.83	666,184	494,818
0.84	680,421	515,713
0.85	695,400	537,481
0.86	711,203	560,239
0.87	727,950	584,132
0.88	745,791	609,333
0.89	764,911	636,061
0.90	785,539	664,596
0.91	807,974	695,298
0.92	832,635	728,652
0.93	860,105	765,326
0.94	891,159	806,285
0.95	927,098	853,000
0.96	969,960	907,883
0.97	1,023,570	975,355
0.98	1,096,198	1,065,047
0.99	1,201,556	1,206,414

cost for the stochastic-service model to handle the full range of demand realizations by means of safety stock.

5 Supply-chain configuration

Lee and Billington (1993), Ettl et al. (2000), Graves and Willems (2000) all report on industrial applications in which their work was used to optimize safety stocks in a supply chain. These applications provide compelling evidence of the financial impact that optimizing inventory levels can have in practice, as reductions of 25–50% in the safety-stock holding cost are common. However, this work starts after most of the design decisions for the supply chain have been set, namely the topology of the network and the key cost and lead-time parameters. In effect, these models are being applied to existing supply chains where the only design options available are in terms of whether to locate safety stocks at a stage, and if so, how much to maintain.

In Chapter 2 of this handbook, Muriel and Simchi-Levi present and consider one category of supply-chain design problems, referred to as network design problems. The intent of the network design problem is typically to determine the optimal manufacturing and distribution network for a company's entire product line. The most common approach is to formulate a large-scale mixed-integer linear program that captures the relevant fixed and variable operating costs for each facility and each major product family. The fixed costs are usually associated with the investment and/or overhead costs for opening and operating a facility, or with placing a product family in a facility. The variable costs include not only the manufacturing, procurement and distributions costs, but also the tariffs and taxes that depend on the network design. Network design focuses on the design of two or three major echelons in the supply chain for multiple products. Due to the nature of the problem being solved, network design is typically solved every two to five years.

In this section, we consider another type of supply-chain design problem that arises after the network design for the supply chain has been set. These decisions determine the *total supply chain cost*, which we define to be the cost of goods sold (COGS), plus the inventory holding costs for the pipeline inventory and for the safety stock. For example, in the bulldozer supply chain presented earlier, the total supply chain cost consists of the annual COGS of $94,380,000, plus the annual holding cost for pipeline stock of $2,007,000, plus the annual holding cost for safety stock of $633,000. For the battery supply chain, there is a similar cost breakdown with annual COGS of $55,364,000, and annual holding costs for pipeline and safety stock of $942,000 and $853,000. In both examples COGS is an order of magnitude larger than the total inventory cost, while the pipeline stock exceeds the safety stock.

For every supply chain a company launches, there is a set of decisions that are made after the network design and that act to configure each stage in the network. In particular, these decisions result in determining the key operating parameters for each stage, including the lead-time and cost added at the stage. These decisions determine the total supply chain cost. For instance, the company must decide whether to source a part locally or globally. The company must decide whether to dedicate machinery to a manufacturing process or to conduct the manufacturing process on shared equipment. The company must decide the transportation mode to move product into a distribution channel. These decisions affect more than just safety stock cost. They also affect the cost of goods sold, pipeline stock cost, quality costs, and time-to-market costs.

To inform these decisions, we introduce a problem that we refer to as the supply-chain configuration problem. For this problem, we address how to configure the supply chain for a new product for which the product's design has already been decided and the topology for the supply-chain network has been set. The central question is to determine what suppliers, parts, processes, and transportation modes to select at each stage in the supply chain. For each stage, we have a set of options that are differentiated, at a minimum, by their lead-times and their direct costs added. Our supply-chain design framework considers the total supply chain cost, equal to the cost of goods sold, plus the inventory holding costs for both safety stock and pipeline stock. The supply chain configuration problem chooses a sourcing option for each stage of the supply chain so as to minimize the sum of these costs.

In the next section, we expand our discussion of options as the fundamental construct in the supply chain configuration problem. We then present the model formulation for the supply-chain configuration problem. The final section presents an example of the approach applied to the bulldozer example presented earlier.

5.1 Option definition

The supply-chain configuration problem is based on the same assumptions made for the safety stock optimization problem, with one significant difference. In the safety stock optimization problem, a stage represents a processing or transformation activity. That is, it is a defined task that will take a certain amount of time at a certain cost. In the supply-chain configuration problem, we still model the supply chain as a network of stages but now a stage represents functionality that must be provided. The essence of the configuration problem is to decide how best to satisfy this functionality in the context of the overall supply chain.

For each stage, we assume that we can specify one or more options that can satisfy the stage's required activity. For example, if a stage represents the procurement of a metal housing, then one option might be a locally based high-cost provider and another option could be a low-cost international supplier.

For each stage, we assume that we will select a single option. Thus, we do not permit the possibility of having dual or multiple sources for a single activity or stage; this might be a topic for further research.

We characterize an option at a stage by its direct cost added and its processing time or lead-time. When a stage reorders, the processing or lead-time is the time to process an item at the stage, provided all of the inputs are available. An option's direct cost represents the direct material and direct labor costs associated with the option. If the option were the procurement of a raw material from a vendor, then the direct costs would be the purchase price including transportation and the labor cost to unpack and inspect the product.

In practice, there might be other dimensions or attributes upon which different options are evaluated. For instance, different suppliers might differ in terms of the quality of the product they supply. Similarly, different options for a manufacturing activity might differ in terms of the amount of capacity that could be made available to the supply chain. We do not consider these other attributes in this presentation. In effect, we assume that the different options at a stage are the same on all attributes except for lead-time and cost. Admittedly, this is a simplification of reality. We leave it to future research to extend the work presented here to address this additional complexity.

We will present the configuration problem for the case of guaranteed service. One could also develop the supply-chain configuration model with the assumption that stages provide stochastic service, but we do not do this here. Rather, we will follow the development of the supply-chain configuration problem for the guaranteed-service model, as given by Graves and Willems (2002).

The model assumptions for the supply-chain configuration problem are the same as for the guaranteed-service safety-stock problem presented earlier in this chapter. We assume that each stage j promises a guaranteed service time s_j^{out} by which the stage will satisfy its demand, either from internal or external customers. Similarly, we define s_j^{in} to be the inbound service time for stage j, which equals the maximum of the service times quoted to stage j by its suppliers. We assume each stage operates according to a periodic review policy with a common review period. We assume the demand process for any finished good is stationary with mean demand per period of μ and a standard deviation of σ. For the purpose of determining the safety stock, we assume that we are given a bound on the demand process for each stage.

If we let n denote the nth option at stage j, then L_{jn} and C_{jn} represent the processing time and cost added associated with the nth option at stage j. The choice of option at a stage will have an impact on the cost of goods sold, on the amount of safety stock and pipeline stock at the stage, and the holding costs. For a given option n, the stage's contribution to the COGS is $C_{jn}\mu_j$

per period. As before, we can calculate the stage's safety stock to be $k_j\sigma_j\sqrt{s_j^{in} + L_{jn} - s_j^{out}}$. Since the choice of an option at a stage decides the lead-time at the stage, the work-in-process or pipeline stock now depends on the option chosen. In particular, the pipeline stock at a stage will equal $\mu_j L_{jn}$ when option n is selected. Finally, the option choice will also affect the holding cost rate because it depends on the cumulative cost at the stage; the option choices at the stage and at any upstream suppliers determine the stage's cumulative cost, and thus the holding cost rate.

5.2 *Model formulation*

We can formulate the supply chain configuration problem as a non-linear mixed-integer optimization problem where the decision variables are the binary variable for option selection and the services times.

P

$$\min\sum_{i=1}^{N}\Big[\alpha c_i\big[D_i(s_i^{in}+t_i - s_i^{out}) - (s_i^{in}+t_i-s_i^{out})\mu_i\big]+\alpha\Big(c_i-\frac{x_i}{2}\Big)t_i\mu_i+\beta x_i\mu_i\Big]$$

s.t.

$$\sum_{n=1}^{O_i} L_{in}y_{in} - t_i = 0 \qquad \text{for } i = 1,2,\ldots,N \qquad (5.1)$$

$$\sum_{n=1}^{O_i} C_{in}y_{in} - x_i = 0 \qquad \text{for } i = 1,2,\ldots,N \qquad (5.2)$$

$$c_i - \sum_{j:(j,i)\in A} c_j - x_i = 0 \qquad \text{for } i = 1,2,\ldots,N \qquad (5.3)$$

$$s_i^{in} \geq s_j^{out} \qquad \text{for } i = 1,2,\ldots,N, j:(j,i)\in A \qquad (5.4)$$

$$s_i^{in} + t_i - s_i^{out} \geq 0 \qquad \text{for } i = 1,2,\ldots,N \qquad (5.5)$$

$$s_j^{out} \leq S_j \qquad \text{for all demand nodes } j \qquad (5.6)$$

$$s_i^{in}, s_i^{out} \geq 0 \text{ and integer} \qquad \text{for } i = 1,2,\ldots,N \qquad (5.7)$$

$$\sum_{n=1}^{O_i} y_{in} = 1 \qquad \text{for } i = 1,2,\ldots,N \qquad (5.8)$$

$$y_{in} \in \{0,1\} \qquad \text{for } i = 1,2,\ldots,N, 1 \leq n \leq O_i \qquad (5.9)$$

where O_i, number of options to choose from at stage i; C_{in}, direct cost added of the nth option at stage i; L_{in}, lead-time of the nth option at stage i; $D_i()$, maximum demand function for stage i; α, scalar representing the holding cost rate; β, scalar converting the model's underlying time unit into the company's time interval of interest; μ_i, mean demand rate at stage i; c_i, cumulative cost at stage i; t_i, selected option's lead-time at stage i; x_i, selected option's cost at stage i; y_{in}, indicator variable which equals 1 if stage i's nth option is selected and 0 otherwise; S_j, maximum service time permitted for demand node j.

The objective function has three terms, each corresponding to a component of the total supply chain cost. The first term represents stage i's safety stock cost, which is a function of the stage's net replenishment time and demand characterization. The holding cost at stage i equals the cumulative cost of the product at stage i times the holding cost rate. The second term expresses the pipeline stock cost as the product of the holding cost rate, the average cost of the product at the stage, and the expected amount of pipeline stock. The third term, cost of goods sold (COGS), represents the total cost of all the units that are delivered to customers during a company-defined interval of time. The incremental contribution to COGS is calculated at each stage by a product of the average demand at the stage, the option's cost, and a scalar β, which expresses COGS in the same units as pipeline stock cost and safety stock cost. (For instance, one might set α and β so that all terms are expressed as annual costs or as the total costs over the lifetime of the supply chain.)

The cost and time associated with the option chosen at each stage is given in (5.1) and (5.2). Constraint (5.3) calculates the cumulative cost at each stage. Constraints (5.4–5.7) assure that the service times are feasible. In particular, the incoming service time at every stage is at least as large as the largest service time quoted to the stage, the net replenishment time of each stage is non-negative, the maximum service times to the customer must be no greater than the user-defined maximums, and service times must be non-negative and integer. The last two constraints, (5.8) and (5.9), enforce the sole sourcing of options.

Graves and Willems (2002) describe how to solve **P** by dynamic programming when the underlying network is a spanning tree.

While **P** clearly uses safety stock optimization as a building block, it also exhibits behavior that is far more complex than just optimizing safety stock. For safety stock optimization, inventory stocking decisions at one stage in the supply chain affect adjacent downstream stages in the supply chain through the downstream stage's net replenishment time. In the supply chain configuration problem, inventory decisions again affect downstream adjacent stages, but the cost at the current stage has an impact on all stages that are downstream of the current stage, not just those that are adjacent. On an intuitive level, **P** is balancing the increase in COGS against the decrease in inventory-related costs. One can reduce inventory related costs by choosing more responsive options, but at the cost of an increase to the COGS. A key realization is that this tradeoff cannot be properly considered by solving **P** one stage at a time, in isolation; rather, one needs to consider the impact of configuration decisions on the entire supply chain in order to produce the globally optimal solution. The benefit to **P** comes from globally balancing the potential increase in COGS with the benefits one gets from being able to reduce inventory costs.

5.3 Example

To gain more insight into the supply chain configuration problem, we will revisit the bulldozer supply chain discussed earlier in the chapter. There are two types of stages in the bulldozer supply chain: procurement and assembly stages. Procurement stages are stages that do not have any incoming arcs; they represent the purchase of materials outside the supply chain. All of the other stages in the supply chain are assembly stages, at which one or more components are combined in the process.

For the example, there are two options per stage. The stage lead-times and costs from the original presentation correspond to the standard option at each stage. If the stage is a procurement stage, this is the existing procurement arrangement. If the stage is an internal assembly stage, this is the traditional manufacturing method at the stage. All procurement stages also have a consignment option where the supplier is responsible for providing immediate delivery to the bulldozer line. Each assembly stage has an expedited option that corresponds to the company investing in process improvement opportunities to decrease the stage's lead-time. These second options are not based on actual data at the company, but they are indicative of the kinds of option costs we have seen in similar supply chains.

We calculate the cost of the consignment option by the following formula: for each one-week reduction in the supplier lead-time, the supplier will increase the purchase price of the part by 0.75%. This is a similar structure to the kind of arrangements that we have encountered before in practice; see Graves and Willems (2002) and Willems (1999). Typically, the cost increase for a week's reduction ranges from 0.5% to 1% of the original purchase price. The increase in price represents the cost to the supplier for bearing the additional inventory holding cost.

For the expedited assembly option, we classify the required improvement activity at a stage as easy, medium or hard. An easy improvement activity might include the assignment of additional labor resources to the task or the dedication of some minor equipment. For the purposes of this analysis, the cost of an easy improvement is $97,500. By dividing this by the average annual demand, we convert this into a per unit approximation of the cost increase, namely $75 per unit. Most of the upstream assembly stages fall into the camp of assembly operations that could be easily improved. Medium improvement activities cost $150 per unit and hard improvements cost $300 per unit. As the cost of the improvement increases, significant redesign and additional human labor are often required. Final assembly is the only stage that is classified as hard to redesign.

The costs and associated lead-times for each option are presented in Table 10.

Figure 5 depicts the service times that correspond to optimizing the supply chain configuration problem.

Among the 22 stages, the optimal supply chain configuration selects the higher cost, shorter lead-time option for only six of the stages. The procurement stages with the higher cost, shorter lead-time option are the brake group, fender group, and plant carrier. The assembly stages with the higher cost, shorter lead-time option are the common subassembly, dressed-out engine, and main assembly. All of the other stages continue to use their original option.

The optimal inventory policy has also changed in reaction to the different options selected. In the original safety stock optimization, the common subassembly and dressed-out engine held no safety stock and both quoted a service time of twenty days to the main assembly. The chassis/platform also held no safety stock, but quoted its maximum possible service time of sixteen days. With the reduction in the processing time at the common subassembly, the brake group and the plant carrier, the common subassembly is now able to quote a service time of eight days to the main assembly. To achieve this eight-day service time requires that two of its suppliers, transmission and case & frame, must now hold a safety stock. Furthermore, stages in the other two sub-networks that supply main assembly are also holding safety stock so that the inbound service time to main assembly remains at eight days. This is a good example of how subtle changes in the configuration of some stages have a dramatic impact on the resulting safety stock policy

Table 11 summarizes the results from optimizing the supply chain configuration and compares the results to those for a solution that keeps the original option at each stage and optimizes the resulting guaranteed-service model.

We observe that when we optimize the safety stock, there are savings in annual holding cost of $198,000 relative to a base case in which each stage holds a safety stock and quotes a service time of zero. When we optimize the configuration, we find an additional savings in total supply chain costs of $371,000. We see, as expected, that when we optimize the supply-chain configuration we actually increase the COGS but get an overall savings due to lower inventory holding costs.

Based on Graves and Willems (2002) and the work presented here, we are able to formulate some initial hypotheses about the behavior of optimal supply chain configurations. First, the further upstream the supply chain, the less likely is it that we choose a higher cost option. Higher-cost options increase the cost at a stage, which not only increases the COGS but also the holding cost for all of the pipeline and safety stock at downstream stages. Furthermore, since the cumulative cost at an upstream stage is typically

Table 10
Option values for Bulldozer Supply Chain Configuration

Stage name	Option description	Option time	Option cost ($)
Boggie assembly	Standard procurement	11	575
	Consignment	0	584
Brake group	Standard procurement	8	3850
	Consignment	0	3896
Case	Standard procurement	15	2200
	Consignment	0	2250
Case & frame	Standard assembly	16	1500
	Expedited assembly	4	1575
Chassis/platform	Standard assembly	7	4320
	Expedited assembly	2	4395
Common subassembly	Standard assembly	5	8000
	Expedited assembly	2	8075
Dressed-out engine	Standard assembly	10	4100
	Expedited assembly	3	4175
Drive group	Standard procurement	9	1550
	Consignment	0	1571
Engine	Standard procurement	7	4500
	Consignment	0	4557
Fans	Standard procurement	12	650
	Consignment	0	662
Fender group	Standard procurement	9	900
	Consignment	0	912
Final assembly	Standard assembly	4	8000
	Expedited assembly	1	8300
Final drive & brake	Standard assembly	6	3680
	Expedited assembly	2	3755
Frame assembly	Standard procurement	19	605
	Consignment	0	622
Main assembly	Standard assembly	8	12,000
	Expedited assembly	2	12,150
Pin assembly	Standard procurement	35	90
	Consignment	0	95
Plant carrier	Standard procurement	9	155
	Cosnsignment	0	157
Platform group	Standard procurement	6	725
	Consignment	0	732
Roll over group	Standard procurement	8	1150
	Consignment	0	1164
Suspension group	Standard assembly	7	3600
	Expedited assembly	2	3675
Track roller frame	Standard procurement	10	3000
	Consignment	0	3045
Transmission	Standard procurement	15	7450
	Consignment	0	7618

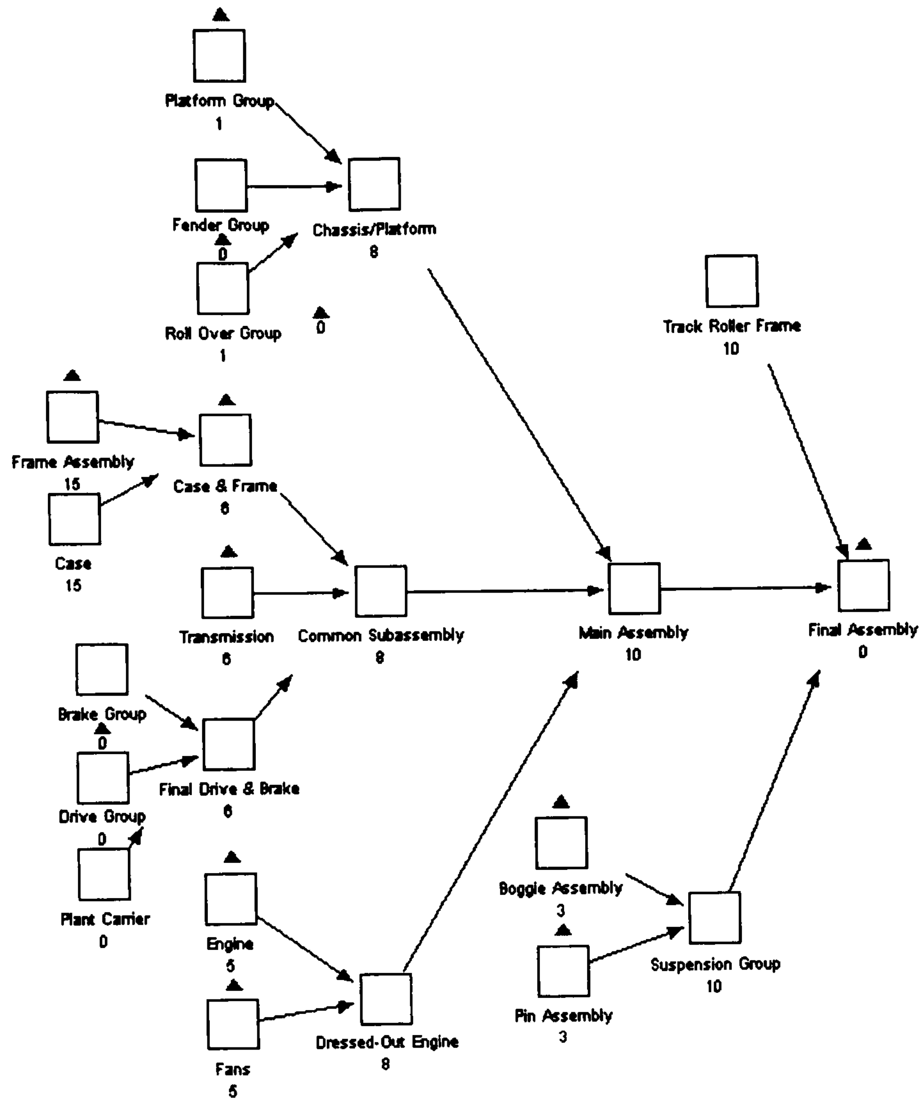

Fig. 5. Optimal Service Times for Bulldozer Supply Chain.

relatively small, it is not that costly to just hold a decoupling safety stock at the upstream stage, thereby making its effective lead-time to the rest of the supply chain zero. Therefore, when choosing a higher cost option at an upstream stage, the inventory savings will have to be truly dramatic to justify the higher cost to the supply chain.

Conversely, we note that a higher-cost, shorter lead-time option is more likely to be attractive at a downstream stage. For instance, we find for stages that represent the transportation of a finished good to an end customer, a

Table 11
Comparison of Optimal Safety Stock and Supply Chain Cofiguration

Cost category	Results from safety stock optimization ($)	Optimal supply chain configuration ($)	Numerical difference ($)	Percentage difference (%)
Cost of goods sold	94,380,000	94,848,000	468,000	0.50%
Total safety stock cost	632,719	499,786	(132,933)	−21.01%
Total pipeline stock cost	2,006,843	1,300,328	(706,514)	−35.21%
Total supply chain cost	97,019,561	96,648,114	(371,447)	−0.38%

faster but more expensive transportation mode is able to pay for itself in terms of inventory savings.

Second, the greatest potential for supply chain configuration occurs when the supply chain has a structure where the selection of options across the chain makes different sub-networks of the supply chain similarly responsive, that is, have the same service times. As an example, making one subassembly very responsive will not likely be cost-effective unless the other components can either be equally responsive, or it is cost-effective to decouple them with inventory so that effectively the subassemblies in the echelon are all similarly responsive. If one of the stages cannot be made more responsive, then the high-cost subassembly could be produced with cheaper options at no penalty to the supply chain's performance.

6 Conclusion

In this chapter we have presented two general approaches to safety stock placement, and have introduced the supply chain configuration problem. The safety stock placement work, as evidenced by the applications cited in the papers, has proven itself to be of value to practice. Both approaches quantify the impact that demand uncertainty has in supply chains. By taking a system-wide view of the problem, these models are able to mitigate the impact of this uncertainty in a cost-effective manner.

We find that there is a significant opportunity to improve the total supply chain cost by jointly optimizing sourcing and inventory decisions during the configuration of the supply chain. The earlier that these supply chain considerations can be incorporated into product and sourcing decisions, the more leverage we have – we see this in this chapter in terms of value of getting the configuration right vis-à-vis solely optimizing safety stocks.

Nevertheless, we wish to conclude this chapter with some thoughts about research opportunities.

Product life cycles are increasingly short, with products within a product family continually being introduced and terminated. Supply chains need to be designed to accommodate this. In particular, demand is never stationary and there is huge uncertainty and risk over a product life cycle. In particular, the risk applies not just to the inventory holding cost in the chain, but also to the inventory investment required to fill the chain since enough demand may not materialize to empty out the chain. There is a need for good models and approaches to determine how to evolve the supply chain to handle generations of products.

Many products see seasonal or cyclic demand. Furthermore, there are often different service targets, holding cost rates, and/or costs for stock outs in different periods. Characterizing both optimal and reasonable approaches for planning safety stocks across a supply chain and across a seasonal cycle is worthy of research.

Supply chains are most certainly not limited to just demand uncertainty. Other types of uncertainty, such as lead-time, capacity, and yield uncertainty, can be equally important. While this chapter has not covered these kinds of uncertainty, the models presented can adapt to these issues, albeit with additional assumptions and often in an ad hoc fashion. Nevertheless, there are opportunities for the development of more general and comprehensive methods for handling the full range of supply chain uncertainties.

Properly designing contracts is another opportunity. This can be thought of as a different form of the configuration decision. In this case, we are looking to establish contracts throughout the supply chain so as to get the best overall performance. In particular, one would expect to design contracts with, say, suppliers so that there is some consistency in how the supply chain is able to respond to upswings (or downswings) in demand. Furthermore, one would hope to understand how to design and coordinate contracts across a number of suppliers or channel partners so as to distribute the risks and rewards in the most economic way.

We have found empirically an interesting analog between supply chains and project management networks. As with a project management network, we find when applying the guaranteed-service model, there is a critical path that underlies the optimal safety stock policy. The difference is that instead of a lead-time-weighted critical path, there is a critical path that is driven by cumulative cost, maximum replenishment time and safety stock policy. For example, appropriately buffering a long lead-time part makes its effective lead-time to the system zero. Identifying and characterizing the components of the critical path in the supply chain is a potentially fertile area to begin the development of new solution approaches.

As a final opportunity, we would hope to see the continuing deployment of models in this chapter to practice. This should provide an opportunity to examine, test, and validate the underlying assumptions of these models. To wit, the stochastic service and guaranteed-service models offer two different perspectives on how the world works. Can we determine which is right? Can we say anything about which is more common? Is either of them right? Or is there a better perspective? We hope that some future research will be able to conduct a careful empirical study of how well these models match reality, as well as how good is the decision support that they provide.

Acknowledgement

This research has been supported in part by the MIT Leaders for Manufacturing Program, a partnership between MIT and major U.S. manufacturing firms; and by the Singapore-MIT Alliance, an engineering education and research collaboration among the National University of Singapore, Nanyang Technological University, and MIT. The second author was also supported by the Boston University School of Management Junior Faculty Research Program.

References

Axsäter, S. (2003). Supply Chain Operations: Serial and Distribution Inventory Systems, in: A. G. de Kok, S. C. Graves (eds.), *Handbooks in Oper. Res. And Management Sci, Supply Chain Management: Design, Coordination and Operation*, Vol. 11. Amsterdam, The Netherlands, North-Holland Publishing Company. Chapter 10.

Bertrand, W. (2003). Supply Chain Design: Flexibility Considerations, in: A. G. de Kok, S. C. Graves (eds.), *Handbooks in Oper. Res. And Management Sci. Supply Chain Management: Design, Coordination and Operation*, Vol. 11. Amsterdam, The Netherlands, North-Holland Publishing Company. Chapter 4.

Clark, A. J., H. Scarf (1960). Optimal Policies for a Multi-Echelon Inventory Problem. *Management Science* 6, 475–490.

de Kok, A. G., J. Fransoo (2003). Planning Supply Chain Operations: Definition and Comparison of Planning Concepts, in: A. G. de Kok, S. C. Graves (eds.), *Handbooks in Oper. Res. And Management Sci. Supply Chain Management: Design, Coordination and Operation*, Vol. 11. Amsterdam, The Netherlands, North-Holland Publishing Company. Chapter 12.

Ettl, M., G. E. Feigin, G. Y. Lin, D. D. Yao (2000). A Supply Network Model with Base-Stock Control and Service Requirements. *Operations Research*. 48, 216–232.

Glasserman, P., S. Tayur (1995). Sensitivity Analysis for Base-Stock Levels in Multiechelon Production-Inventory Systems. *Management Science* 41, 263–281.

Graves, S. C. (1988). Safety Stocks in Manufacturing Systems. *Journal of Manufacturing and Operations Management* 1, 67–101.

Graves, S. C., S. P. Willems (1996). Strategic Safety Stock Placement in Supply Chains, *Proceedings of the 1996 MSOM Conference*, Hanover, NH.

Graves, S. C., S. P. Willems (2000). Optimizing Strategic Safety Stock Placement in Supply Chains. *Manufacturing & Service Operations Management* 2, 68–83.

Graves, S. C., S. P. Willems (2002). Optimizing the Supply-Chain Configuration for New Products, MIT Working Paper, January, 32 pages.

Humair, S., S. P. Willems (2003). Optimal Inventory Placement in Networks with Clusters of Commonality, Working Paper, January, 19 pages.

Inderfurth, K. (1991). Safety Stock Optimization in Multi-Stage Inventory Systems. *International Journal of Production Economics* 24, 103–113.

Inderfurth, K., S. Minner (1998). Safety Stocks in Multi-Stage Inventory Systems under Different Service Levels. *European Journal of Operational Research* 106, 57–73.

Kimball, G. E. (1988). General Principles of Inventory Control. *Journal of Manufacturing and Operations Management* 1, 119–130.

Lee, H. L., C. Billington (1993). Material Management in Decentralized Supply Chains. *Operations Research* 41, 835–847.

Muriel, A., D. Simchi-Levi (2003). Supply Chain Design and Planning – Applications of Optimisation Techniques for Strategic and Tactical Models, in: A. G. de Kok, S. C. Graves (eds.), *Handbooks in Oper. Res. And Management Sci. Supply Chain Management: Design, Coordination and Operation*, Vol. 11. The Netherlands, North-Holland Publishing Company, Amsterdam. Chapter 2.

Simpson, K. F. (1958). In-Process Inventories. *Operations Research* 6, 863–873.

Song, J.-S., P. Zipkin (2003). Supply Chain Operations: Assemble-to-Order Systems, in: A. G. de Kok, S. C. Graves (eds.), *Handbooks in Oper. Res. And Management Sci. Supply Chain Management: Design, Coordination and Operation*, Vol. 11. Amsterdam, The Netherlands, North-Holland Publishing Company. Chapter 11.

Willems, S. P. (1999). Two papers in supply-chain design: supply-chain configuration and part selection in multigeneration products. MIT Ph.D. Dissertation, Cambridge MA.