M&SOM MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 10, No. 2, Spring 2008, pp. 278–287
ISSN 1523-4614 | EISSN 1526-5498 | 08 | 1002 | 0278
DOI 10.1287/msom.1070.0175

Strategic Inventory Placement in Supply Chains: Nonstationary Demand

Stephen C. Graves
Leaders for Manufacturing Program and A. P. Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, sgraves@mit.edu

Sean P. Willems
School of Management, Boston University, Boston, Massachusetts 02215, willems@bu.edu

The life cycle of new products is becoming shorter and shorter in all markets. For electronic products, life cycles are measured in units of months, with 6- to 12-month life cycles being common. Given these short product life cycles, product demand is increasingly difficult to forecast. Furthermore, demand is never really stationary because the demand rate evolves over the life of the product. In this paper, we consider the problem of where in a supply chain to place strategic safety stocks to provide a high level of service to the final customer with minimum cost. We extend our model for stationary demand to the case of nonstationary demand, as might occur for products with short life cycles. We assume that we can model the supply chain as a network, that each stage in the supply chain operates with a periodic review base-stock policy, that demand is bounded, and that there is a guaranteed service time between every stage and its customers. We consider a constant service time (CST) policy for which the safety stock locations are stationary; the actual safety stock levels change as the demand process changes. We show that the optimization algorithm for the case of stationary demand extends directly to determining the safety stocks when demand is nonstationary for a CST policy. We then examine with an illustrative example how well the CST policy performs relative to a dynamic policy that dynamically modifies the service times. In addition, we report on numerical tests that demonstrate the efficacy of the proposed solution and how it would be deployed.

Key words: base-stock policy; dynamic programming application; multiechelon inventory system; nonstationary demand; multistage supply chain application; safety stock optimization
History: Received: March 1, 2005; accepted: February 21, 2007. Published online in *Articles in Advance* January 4, 2008.

1. Introduction

Manufacturing firms are introducing new products at a higher frequency with increasingly shorter life cycles. For each new product, a firm must determine its supply chain and the associated operating policies to match supply to the demand to achieve the most revenue with the least cost. A major complication is that the demand process evolves over the product life cycle and is never stationary.

The product life cycle of a new product typically consists of four phases: (i) a product-launch phase in which the product is introduced to the market; (ii) a demand-ramp phase over which the demand rate grows rapidly; (iii) a peak-demand phase during which the product sells at its maximum rate; and (iv) an end-of-life phase during which the product demand declines as it is removed from the market. The demand rate is never stationary because the product moves from one life-cycle phase to another.

This research examines the problem of locating safety stocks in a supply chain in a way that accounts for uncertain, nonstationary demand processes. Given the inherent complexity of modeling nonstationary demand processes, we seek a pragmatic approach that requires approximations and compromises to get results that might apply in practice. We use the modeling framework from Graves and Willems (2000) (referenced as G-W) and introduce a nonstationary demand model. We show that the G-W safety stock placement optimization applies to this case of nonstationary demand.

In the remainder of this section, we briefly discuss related literature. In §2, we present the key assumptions for modeling a supply chain and its nonstationary demand. In §3, we extend the G-W model to accommodate nonstationary demand. In §4, we examine a simple example to explore the near optimality of a constant service time (CST) policy. An online appendix contains a more extensive computational experiment and real-world example that illustrate how the proposed policy might apply in practice.[1]

Related Literature. Relative to the stationary-demand inventory literature, there is much less work for nonstationary demand. We characterize this work by how the nonstationary demand is specified and whether the work focuses on optimization versus performance evaluation.

Morton and Pentico (1995) and Bollapragada and Morton (1999) focus on setting inventory policies for a single stage facing a general nonstationary demand process with proportional holding and backorder costs. When the order cost is zero, a time-varying base-stock policy is optimal; for a nonzero order cost, a time-varying (s, S) policy is optimal. This research develops computationally efficient upper and lower bounds on these optimal policies.

For short life cycle products, the challenge of accurately forecasting demand can be as important as determining inventory policies. Kurawarwala and Matsuo (1996) develop an integrated framework for forecasting and inventory management of short life-cycle products. Their approach estimates the parameters for a seasonal trend growth model and uses this as an input to a finite-horizon stochastic inventory model with time-dependent demands.

Nonstationary demand has also been modeled as a Markov-modulated Poisson demand process. One example is Chen and Song (2001), who show the optimality of echelon base-stock policies with state-dependent order-up-to levels for serial networks. A second example is Abhyankar and Graves (2001) who determine the optimal position of an inventory hedge in a two-stage serial supply chain that faces Markov-modulated demand with two states.

Within the bullwhip literature, several papers develop models for nonstationary demand. These papers generally assume that each stage follows an adaptive base-stock policy and then analyze the effect that different forecasting techniques and assumed demand distributions have on the inventory requirements at each stage. For instance, Lee et al. (1997) demonstrate that the adjustment of order-up-to levels at the retailer amplifies the variance of the order signal the retailer provides for the manufacturer. Two other examples are Graves (1999) and Chen et al. (2000).

Finally, there is a growing body of work on designing supply chains to handle nonstationary demand. Beyer and Ward (2000) use simulation to accurately model the inventory requirements in a two-echelon supply chain that utilizes two modes of distribution and is subject to nonstationary demand. Johnson and Anderson (2000) investigate the benefits of postponement in supply chains that introduce multiple products with short product life cycles. Ettl et al. (2000) minimize the total inventory in a multistage inventory system, where the key challenge is to approximate the replenishment lead times within the supply chain. To model nonstationary demand, they break the horizon into a set of stationary phases and adopt a rolling-horizon approach where the optimization is performed for each demand phase.

2. Assumptions

Our intent is to extend the G-W modeling framework to nonstationary demand. For completeness, we restate the key assumptions from G-W and introduce a few new assumptions to permit the consideration of nonstationary demand. We conclude with a short discussion of these new assumptions in the context of safety stock placement for nonstationary demand; we refer the reader to G-W for the justification of the original set of assumptions.

Multistage Network. We model a supply chain as a network with N nodes and arc set A, where nodes are stages in the supply chain and a directed arc denotes that an upstream stage supplies a downstream stage. A stage represents a major processing function such as raw material procurement, component production, finished-goods assembly and test, or the transportation of a product from a central distribution center to a regional warehouse. Each stage is a potential location for holding a safety stock of the item processed at the stage.

[1] An online appendix to this paper is available on the *Manufacturing & Service Operations* website (http://msom.pubs.informs.org/ecompanion.html).

We associate with each arc a scalar ϕ_{ij} to indicate how many units of the upstream component i are required per downstream unit j. If a stage is connected to several upstream stages, then its production activity is an assembly requiring inputs from each of the upstream stages. A stage that is connected to multiple downstream stages is either a distribution node or a production activity that produces a common component for multiple internal customers.

Production Lead Times. For each stage, we assume a deterministic production lead time and will call it T_i. When a stage reorders, the production lead time is the time from when all of the inputs are available until production is completed and available to serve demand. The production lead time includes the waiting and processing time at the stage, plus any transportation time to put the item into inventory. We assume that there are no capacity constraints that limit production at a stage.

Demand Process. External demand occurs only at nodes with no successors, which we term *demand nodes*. For each demand node j, we specify the end-item demand process over a planning horizon of length H. The average demand rate for demand node j at time t is $\mu_j(t)$ for $0 \le t \le H$. An internal stage serves only internal customers or successors. We find the average demand rate for an internal stage at time t from $\mu_i(t) = \sum_{(i,j)\in A} \phi_{ij}\mu_j(t)$.

As in G-W, we assume that demand at each stage j is *bounded* for the purposes of determining the safety stock placements. We specify the demand bound by the function $D_j(s,t)$ to represent the maximum demand for stage j for the time interval $(s,t]$. That is, we have for all demand realizations $d_j(s,t) \le D_j(s,t)$, where $d_j(s,t)$ represents the demand for stage j for the time interval $(s,t]$. To determine the safety stocks, we will use the demand bound net of the average demand, namely,

$$g_j(s,t) = D_j(s,t) - \int_{\tau=s}^{t} \mu_j(\tau)\,d\tau. \qquad (1)$$

We assume that the net demand bound $g_j(s,t)$ is a concave function over the planning horizon H. We discuss this assumption in the appendix.

Control Policy. Each stage operates with a periodic review base-stock replenishment policy with a common review period equal to a single time unit, say one day, and with guaranteed service times. We first detail our assumptions on how to structure and implement the base-stock policy for nonstationary demand and then describe the assumption of guaranteed service times.

Periodic Review Base-Stock Replenishment Policy. Each period, each stage j observes its demand from either its external customers or its downstream stages, and places orders on its suppliers to replenish the demand. That is, at time t, an external stage j observes its demand over the time interval $(t-1,t]$ and initiates its replenishment; to simplify notation, we denote the demand over the time interval $(t-1,t]$ by $d_j(t)$. Similarly, each internal stage i observes demand from its downstream customers given by $d_i(t) = \sum_{(i,j)\in A} \phi_{ij} d_j(t)$. There is no time delay in ordering, so during each period, each stage sees and initiates the replenishment of its customer demand.

However, because demand is nonstationary, we need to adapt the base stock at each stage in concert with the evolution of the demand process. We let $B_j(t)$ denote the base stock for stage j at time t. In the next section, we pose an optimization problem from which we can find the base stock $B_j(t)$ for each time epoch. Because the base-stock level is not constant, we must also plan replenishments to adjust the base-stock levels for each stage in the supply chain. For a given set of base-stock levels $B_j(t)$, this is a deterministic planning problem. In the online appendix, we show how to determine the necessary replenishments to achieve the desired base-stock levels.

Guaranteed Service Times. We assume that each demand node j promises a *guaranteed outbound service time* S_j by which stage j will satisfy customer demand. That is, the customer demand observed at the review period t, $d_j(t)$, must be filled by time $t + S_j$. Furthermore, we assume that stage j provides *100% service* for the specified outbound service time: stage j delivers exactly $d_j(t)$ to the customer at time $t + S_j$.

Similarly, an internal stage i quotes and guarantees an outbound service time S_i for each downstream stage j, $(i,j) \in A$. At period t, the internal demand from stage j on stage i equals $\phi_{ij} d_j(t)$; then stage i delivers exactly this amount to stage j at time $t + S_i$. Graves and Willems (1998) describe how to extend the model to permit customer-specific service times.

The service times for the demand nodes and the internal stages determine where we place safety stocks in a supply chain and are *decision variables* for the optimization model in §3.

Discussion of Assumption of Guaranteed Service Times. G-W discuss and justify this assumption for the purposes of strategic safety stock placement for stationary demand.

The case of nonstationary demand raises an additional consideration. In stating the assumed control policy, we allow the base stock to vary over time; however, we do not vary the service times but restrict them to be *constant* over the planning horizon. We show later in this paper that this is not optimal. However, we defend this restriction with the following arguments.

First, we observe that constant service times translate into fixed placements for the safety stock in the supply chain. Changing the service times in a supply chain will change the locations for the safety stocks. From our experience in applying this model in practice, we know that firms are quite reluctant to switch the locations for their safety stocks unless there are very compelling reasons to do so. The choice of locations for the safety stocks dictates how the supply chain is controlled, so changing these buffer locations is a nontrivial exercise.

Second, we observe that the cost penalty from using constant service times is typically quite small; the online appendix provides evidence by means of a computational experiment.

Third, Graves and Willems (2002) show that the optimal service times for a supply chain with a single demand node and stationary demand are invariant to the parameters of the demand process. More specifically, suppose the demand bound for the demand node j is of the form

$$D_j(s,t) = (t-s)\mu_j + \sigma_j(t-s)^{\beta},$$

where μ_j, σ_j, and β are constants; then the optimal service times do *not* depend on the values for the parameters μ_j, σ_j. This suggests that a constant service time policy might perform quite well when the actual demand process is nonstationary and time varying. Graves and Willems (2002) extend this finding to supply chains with multiple demand nodes, with an additional assumption required for the calculus for setting the demand bound on the internal nodes.

Finally, we show in the next section that we can determine the best set of constant service times by solving an optimization problem that is comparable to that for the case of stationary demand. Hence, we have a practical approach for finding these service times.

3. Supply Chain Model

In this section, we present the multistage model of a supply chain and the optimization problem for determining safety stocks.

Inventory Model. The single-stage inventory model serves as the building block for modeling a multistage supply chain. We adapt the standard model of Kimball (1988) (see also Simpson 1958) to the case of nonstationary demand.

We define the inbound service time SI_j as the time for stage j to get supplies from its immediate suppliers. In each period t, stage j places an order equal to $\phi_{ij} d_j(t)$ on each upstream stage i for which $\phi_{ij} > 0$. The time for all orders to be delivered to stage j dictates when stage j can commence production to replenish its demand. This inbound service time is constrained by the maximum outbound service time from the upstream suppliers, i.e., $SI_j \geq \max_{(i,j)\in A}\{S_i\}$.

We assume we are given base stocks $B_j(t)$ for each period $t = 1, 2, \ldots, H$. For the stated assumptions, we can express the inventory at stage j at the end of period t as

$$I_j(t) = B_j(t) - d_j(t - SI_j - T_j, t - S_j), \tag{2}$$

where $d_j(s,t) = 0$ for $s \leq t \leq 0$ and $d_j(s,t) = d_j(0,t)$ for $s < 0 < t$.

We define the *net replenishment time* for stage j to be its replenishment time, net the stage's promised outbound service time, i.e., $SI_j + T_j - S_j$. This net replenishment time determines the safety stock at stage j. We always set the outbound service times so that the net replenishment time is nonnegative.[2]

[2] If the net replenishment time were negative, this implies that we replenish the stage's inventory before we need to serve a demand; we can reduce the stage's inventory by delaying the orders on its suppliers, which equates to increasing the inbound service time.

The explanation for Equation (2) follows that for the case of stationary demand. There are three transactions in period t: Stage j completes the replenishment of its demand from period $t - SI_j - T_j$; stage j fills its demand from period $t - S_j$; and stage j receives an additional replenishment equal to $\Delta B_j(t) = B_j(t) - B_j(t-1)$ so as to have the prescribed base-stock level. Hence, we can write an inventory balance equation:

$$I_j(t) = I_j(t-1) + d_j(t - SI_j - T_j) - d_j(t - S_j) + \Delta B_j(t). \quad (3)$$

We obtain (2) by applying (3) recursively and using the boundary condition $I_j(0) = B_j(0)$.

To derive Equation (2), we implicitly assume that we can always make the necessary adjustment $\Delta B_j(t)$ to the base-stock level. This need not be the case when the base-stock level decreases and $\Delta B_j(t) < 0$; in effect, we need to assume that $d_j(t - SI_j - T) + \Delta B_j(t) \geq 0$ so that the replenishment in period t is nonnegative. We expect this will typically be the case, and we assume this to be true so as not to overly complicate the presentation. (We note that when $d_j(t - SI_j - T) + \Delta B_j(t) < 0$, then Equation (2) provides a lower bound on the actual inventory level.)

Determination of Base Stock. For stage j to provide 100% service to its customers, we require that $I_j(t) \geq 0$; we see from Equation (2) that this requirement equates to

$$B_j(t) \geq d_j(t - SI_j - T_j, t - S_j).$$

Because demand is bounded, we satisfy the above requirement with the least inventory by setting the base stock as

$$B_j(t) = D_j(t - SI_j - T_j, t - S_j). \quad (4)$$

Thus, the base-stock level in period t is the maximum possible demand over a time interval $(t - SI_j - T_j, t - S_j]$ for which stage j filled its demand, but has yet to receive replenishments.

Safety Stock Model. We use Equations (2) and (4) to find the expected inventory level $E[I_j(t)]$:

$$\begin{aligned} E[I_j(t)] &= D_j(t - SI_j - T_j, t - S_j) - \int_{\tau = t - SI_j - T}^{t - S_j} \mu_j(\tau)\, d\tau \\ &= g_j(t - SI_j - T_j, t - S_j). \end{aligned} \quad (5)$$

The expected inventory represents the safety stock held at stage j and depends on the net replenishment time and the demand bound. We observe that stage j holds no safety stock whenever the net replenishment time is zero, i.e., $SI_j + T_j - S_j = 0$.

The supply chain will also have a work-in-process or pipeline inventory. This inventory corresponds to the replenishment of customer demand plus the planned adjustments to the base-stock levels. If we fix the base-stock levels for the start and end of the planning horizon, then we can show that the work-in-process does not depend on the choice of service times, but only on the average demand rates and the lead times at each stage. Hence, in posing an optimization problem, we ignore work-in-process and only model safety stock.

Multistage Model. To model the multistage system, we use Equation (5) for each stage where inbound service time is a function of the outbound service times for the upstream stages. We then formulate an optimization problem to find the optimal service times for the planning horizon:

$$\mathbf{P} \quad \text{Min} \sum_{t=1}^{H}\sum_{j=1}^{N} h_j E[I_j(t)] = \sum_{t=1}^{H}\sum_{j=1}^{N} h_j g_j(t - SI_j - T_j, t - S_j)$$

$$\begin{aligned} \text{subject to} \quad & S_j - SI_j \leq T_j \quad \text{for } j = 1, \ldots, N \\ & SI_j - S_i \geq 0 \quad \forall (i,j) \in A \\ & S_j = 0 \quad \forall \text{ demand nodes } j \\ & S_j, SI_j \geq 0, \text{ integer} \quad \text{for } j = 1, \ldots, N, \end{aligned}$$

where h_j denotes the holding cost per unit per time period for inventory at stage j. The objective of problem **P** is to minimize the safety stock holding cost over the planning horizon. The constraints assure that the net replenishment times are nonnegative, that each stage's inbound service time is no less than the maximum outbound service time quoted to the stage, and that the end-item stages satisfy their service guarantee.[3]

We define the planning model to end in period H and do not explicitly include any costs beyond this

[3] For ease of presentation, we require in **P** that each demand node provides a zero outbound service time to its customers. We can easily relax this assumption and permit nonzero service time requirements for the external customers; however, this unduly complicates the presentation, because we need to extend the planning horizon until all nonstationary end-item demand is served.

horizon. For instance, if H represents the end of the life cycle for a product, there would be disposal costs for the supply chain inventory left over at the end of the horizon. We could include a disposal cost by restating **P** with time-dependent holding costs.

To solve **P**, we first observe that we can rewrite the objective function as

$$\sum_{t=1}^{H}\sum_{j=1}^{N} h_j g_j(t - SI_j - T_j, t - S_j) = H \times \sum_{j=1}^{N} h_j G_j(SI_j, S_j),$$

where $G_j(SI_j, S_j) = H^{-1} \times \sum_{t=1}^{H} g_j(t - SI_j - T_j, t - S_j)$ is the average safety stock at node j as a function of its inbound and outbound service times. Thus, **P** is equivalent to the safety stock optimization problem for stationary demand in G-W, but with its objective function expressed in terms of the average safety stock $G_j(SI_j, S_j)$ over the planning horizon H. Furthermore, $G_j(SI_j, S_j)$ is a concave function, given the assumption that the net demand bound $g_j(s, t)$ is a concave function. As a consequence, we can solve **P** with the existing algorithms for stationary demand. G-W presents a dynamic programming algorithm for solving **P** for supply chains modeled as spanning trees; Humair and Willems (2006), Magnanti et al. (2006), and Lesnaia (2004) have each developed and tested algorithms for general acyclic networks.

4. Example

We examine the nonoptimality of a CST policy by means of an example chosen to exacerbate the underperformance of the CST policy. Consider a two-stage system, with upstream stage 1 serving stage 2, $\phi_{12} = 1$ and cost and time parameters in Table 1.

We assume that the planning horizon H is divided into two phases, with stationary demand in each phase. Phase 1 is the time window $0 < t \le Z$, whereas phase 2 is $Z < t \le H$. At time $t = Z + 1$, there is a step function shift in demand from a process with $\mu_1 = 100$, $\sigma_1 = 30$ to a process with $\mu_2 = 150$, $\sigma_2 = 50$. For any interval that spans the two phases, we model the demand bound by combining the maximum demands over the two phases as if they were independent random variables. Thus, we assume that the demand bound is as follows:

$$D(s, t) = (t - s)\mu_1 + k\sigma_1\sqrt{t - s} \quad \text{for } 0 < s < t \le Z, \quad (6)$$

$$D(s,t) = (Z-s)\mu_1 + (t-Z)\mu_2 + k\sqrt{(Z-s)\sigma_1^2 + (t-Z)\sigma_2^2} \quad \text{for } 0 < s < Z < t \le H, \quad (7)$$

$$D(s,t) = (t-s)\mu_2 + k\sigma_2\sqrt{t-s} \quad \text{for } Z \le s < t \le H, \quad (8)$$

where $\mu_1 = 100$, $\mu_2 = 150$, $\sigma_1 = 30$, $\sigma_2 = 50$, and the safety factor $k = 2$.

We assume a planning horizon of $H = 215$ days, with the transition point $Z = 115$ days. We ignore the initial 15-day start-up transient and solve **P** for a 200-day horizon, namely for days $t = 16, \ldots, 215$.

We note that by assumption the outbound service time for the end item is zero. Hence, there is a single decision variable, namely, the outbound service time at stage 1 (S_1) equal to the inbound service time for stage 2 (SI_2). The solution for **P** sets $S_1 = SI_2 = 0$. We hold a safety stock of intermediate product between stage 1 and stage 2, as well as an end-item safety stock at stage 2 to serve external demand. In Table 2 we report the base stocks from the solution to **P** for the transition from phase 1 to phase 2. We see that to accommodate the change in demand we need to increase the base stocks for both stages over this interval.

For this two-stage example, a CST policy is not optimal. The optimal policy is a dynamic service time (DST) policy that optimizes the service times in each period subject to the problem parameters and demand bounds for that period. For this example, the DST policy sets $S_1 = SI_2 = 0$ for all $t \le 115$ and $t \ge 130$, and then sets $S_1 = SI_2 = 10$ for $116 \le t \le 129$, which we term to be the transient time window.[4] During the

Table 1 Cost and Time Parameters for Example

	Stage 1	Stage 2
Holding cost h_j	0.5	1.0
Production lead time T_j	10	5

[4] The *transient time window* is the time over which the system inventory might depend on the demand from both phases. From Equations (2) and (7), we see that the inventory at stage j at time t depends on demand from both phases when $t - SI_j - T_j < Z$ and $t - S_j > Z$. Thus, the transient time window depends on the choice of service times. In the example, the maximum duration occurs when $S_1 = SI_2 = 10$ and, thus, the transient time window is $116 \le t \le 129$.

Table 2 Comparison of Policies for Two-Stage Example, with Optimal Policy in Bold

Time (t)	Holding cost for $S_1 = 0$ (\$)	Stage 1 base stock for $S_1 = 0$	Stage 2 base stock for $S_1 = 0$	Holding cost for $S_1 = 10$ (\$)	Stage 1 base stock for $S_1 = 10$	Stage 2 base stock for $S_1 = 10$
115	**229**	**1,189**	**634**	232	0	1,732
116	259	1,256	706	**246**	**0**	**1,796**
117	286	1,321	775	**258**	**0**	**1,858**
118	310	1,385	843	**271**	**0**	**1,921**
119	333	1,448	909	**282**	**0**	**1,982**
120	354	1,511	974	**293**	**0**	**2,043**
121	360	1,573	974	**304**	**0**	**2,104**
122	366	1,634	974	**314**	**0**	**2,164**
123	371	1,695	974	**324**	**0**	**2,224**
124	377	1,756	974	**334**	**0**	**2,284**
125	382	1,816	974	**344**	**0**	**2,344**
126	382	1,816	974	**353**	**0**	**2,403**
127	382	1,816	974	**362**	**0**	**2,462**
128	382	1,816	974	**370**	**0**	**2,520**
129	382	1,816	974	**379**	**0**	**2,579**
130	**382**	**1,816**	**974**	387	0	2,637

transient time window, the optimal safety stock policy is to eliminate the intermediate inventory between stages 1 and 2 and enlarge the base stock of the end item to cover the total lead time of 15 days. In Table 2, we compare the two policies over the transient time window to examine how nonoptimal a CST policy might be. We see that the cost difference can be 20% in a single day. The cost penalty for using the CST policy ($S_1 = 0$) during the transient time window is 11%.

To explain why the optimal policy switches, we must consider the structure of the two policies. In Figure 1, we plot the safety stock costs for both policies over the transient period.

During the transient time window, the CST policy increases the safety stock at both stages to protect against the increased variability in phase 2. In contrast, the DST policy maintains only an end-item safety stock; when we transition to phase 2, the DST safety stock increases, but less dramatically than that for the CST policy. The DST policy is better able to pool the increased demand variability from phase 2 with the lower demand variability of phase 1, due to having a longer net replenishment time at its single safety stock location.

Although not optimal, an open question is how well the CST policy performs in the transient time window relative to the DST policy. In Table 3, we report the percentage cost penalty for using the CST policy during the time periods $116 \le t \le 129$ relative to the DST policy for various values of the ratio of holding costs h_1/h_2. We see that there is only a cost penalty when $0.26 < h_1/h_2 < 0.52$, because otherwise the CST policy is also optimal in the transient time window. We also see that the maximum cost penalty is 12.0% and occurs when $h_1/h_2 = 0.51$.

Figure 1 Expected Safety Stock Cost as a Function of Time

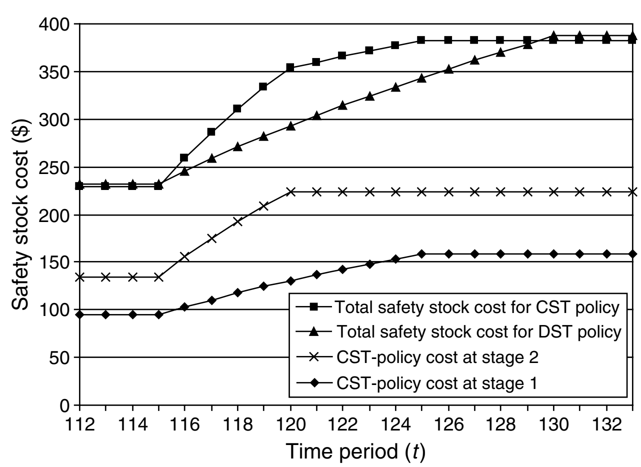

Table 3 Cost Penalty for CST Optimal Policy During Transient Time Window

Holding cost ratio h_1/h_2	≤ 0.26	0.3	0.4	0.5	0.51	≥ 0.52
Cost penalty (%)	0	0.3	4.0	11.1	12.0	0

Table 4 Cost Penalty for CST Optimal Policy During Transient Time Window

Variability parameter σ_2	30	40	50	60	70
Cost penalty (%)	0	6.1	11.1	14.4	16.8

In Table 4, we report the percentage cost penalty for using the CST policy during the time periods $116 \leq t \leq 129$ relative to the optimal policy for various values of the demand variability parameter σ_2 with $h_1/h_2 = 0.5$. The cost penalty grows with the demand variability in phase 2, albeit at a declining rate. If the demand variability parameters are the same in both phases, then the CST policy is optimal for the transient time window.

To assess the significance of the cost penalty in this example, we make three comments. First, from Table 3, we see that there is a nonnegligible cost penalty for a fairly limited range of choices for the cost ratio h_1/h_2. Second, from Table 4, we see that doubling the demand variability results in only a 14.4% cost penalty. Third, if the transient time window is a fraction f of the overall planning horizon H, then the cost penalty for following a CST policy is $11\% \times f$ for the base case in Table 2. In our base example, $f = 0.07$ for a cost penalty of less than 1%.

Finally, the optimal policy in Table 2 would be difficult to implement. It requires the supply chain to consolidate the component inventory into the end-item inventory at the start of the transient time window, and then to reverse this consolidation when the transition ends; the supply chain must also change its internal service times during the transient time window. It would be much easier to follow the CST policy by which the base stocks are gradually increased over the course of the transient time window, and there is no change to the internal service times.

The online appendix contains two more numerical examples to illustrate the differences between the CST and DST policies. The online appendix also describes and documents the planning problem for adjusting the base stocks that would be prescribed by a CST policy.

5. Conclusion

In this paper, we introduce and develop a model for positioning safety stock in a supply chain subject to nonstationary demand. We show how to extend the G-W model to find the optimal placement of safety stocks under the assumption of a CST policy. In particular, we show that if the nonstationary demand bound is a concave function, then this optimization problem for the nonstationary demand case is equivalent to that for the stationary demand case. Thus, it is no more difficult to solve the nonstationary problem than it is to solve for stationary demand, for which there are highly efficient algorithms. We also show that the CST policy need not be optimal; rather, an optimal policy might entail changing the service times, and thus the locations of safety stocks, as demand evolves over time. Nevertheless, we argue and provide limited evidence that applying the CST policy is near optimal and has obvious implementation advantages. In the online appendix, we provide additional evidence with a more extensive numerical test as well as an application to a consumer-packaged goods company.

There are several interesting areas in which to extend this research. First, one might further explore the effectiveness of using the CST policy to better understand when it works well and when not. This might provide some insight or guidelines as to when to deviate from a CST policy. Second, we would hope to examine how best to model real demand processes and their demand bounds in the context of safety stock placement. Finally, we want to investigate what happens when we relax some of the key assumptions to the model: No capacity constraints, deterministic lead times, and a common review period.

Acknowledgments
This research has been supported in part by the Massachusetts Institute of Technology (MIT) Leaders for Manufacturing Program, a partnership between MIT and major 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. The authors thank the editors and referees for their very helpful and constructive feedback on earlier versions of the paper.

Appendix. Concavity of Safety-Stock Function

The intent of the appendix is to provide a justification for the assumption that the net demand bound $g_j(s, t)$ is a concave function. For ease of presentation, we drop the stage subscript j in this appendix.

In G-W, we suggest that for the purposes of strategic safety stock placement, a typical demand bound for stationary demand is given by:

$$D(s, t) = (t-s)\mu + k\sigma\sqrt{t-s},$$

where μ and σ correspond to the mean and standard deviation of demand in each period, and k is a safety factor that is chosen to reflect the percentage of time that the safety stock covers the demand variation. We have found that this bound is applicable in many contexts where the assumption of stationary demand is reasonable.

By analogy, for nonstationary demand we suggest the following specification of the bound:

$$D(s, t) = \int_{\tau=s}^{t} \mu(\tau)\, d\tau + k\sqrt{\int_{\tau=s}^{t} \sigma^2(\tau)\, d\tau},$$

where $\mu(\tau)$ is the average rate of demand at time τ and $\sigma^2(\tau)$ is the instantaneous variance at time τ. We assume independence over time. The net demand function is then:

$$g(s, t) = k\sqrt{\int_{\tau=s}^{t} \sigma^2(\tau)\, d\tau}.$$

To examine the shape of this function, it will be helpful to rewrite it as:

$$g(s, t) = k\sqrt{f(t) - f(s)} \quad \text{for } f(t) = \int_{\tau=0}^{t} \sigma^2(\tau)\, d\tau.$$

Thus, the function $f(t)$ represents the demand variance over the time interval $(0, t]$ and is an increasing, positive function over $0 \le t \le H$. We assume that $f(t)$ is continuous and twice differentiable.

The concavity of $g(s, t)$ depends upon the shape of $f(t)$. In particular, by examining the Hessian, we find that $g(s, t)$ is concave if the following two conditions hold:

$$2f''(t)f(t) \le (f'(t))^2 \quad \forall\, t, \tag{9}$$

$$(f'(t))^2 f''(s) - (f'(s))^2 f''(t) \ge 0 \quad \forall\, s, t,\ s < t. \tag{10}$$

The first condition (9) says that the variance function cannot increase faster than a quadratic function at any point in time. This seems reasonable for most contexts, because we expect that the variability in demand will change gradually over time.

To understand the second condition, we note that it is trivially true when $f''(s) \ge 0$ and $f''(t) \le 0$, and it fails when $f''(s) < 0$ and $f''(t) > 0$. Thus, the second condition can hold only if the variance function is convex up to some point and then concave thereafter. This does not seem overly restrictive because it permits the variance function to be either concave or convex over its entire range, or to have an S-shape. In the following discussion, we assume that the variance function is convex for $0 \le t \le Z$ and concave for $Z \le t \le H$ for some Z, $0 \le Z \le H$.

We need condition (10) to hold in both the convex region and the concave region of the variance function. In the convex region, the condition (10) is true if the second derivative is nonincreasing, i.e., if $f''(s) \ge f''(t) \ge 0$, $\forall\, s, t \ni 0 \le s < t \le Z$. For instance, in the convex region, the variance function could be a quadratic or a linear function, or some smooth convex function that grows no faster than a quadratic. This is quite similar to what is assumed for (9) to hold; however, (9) does not imply (10). For instance, if the variance function initially grows linearly and then transitions to quadratic growth, it will satisfy (9) but not (10). Nevertheless, a more realistic scenario (say, for an emerging product) would be for the variance function to initially increase at a quadratic rate and then slow down to linear growth; such a variance function satisfies both (9) and (10).

In the concave region, (9) is always true and condition (10) holds if the second derivative is nonincreasing, i.e., if $0 \ge f''(s) \ge f''(t)$, $\forall\, s, t \ni Z \le s < t \le H$. Thus, we satisfy (10) in the concave region if the growth rate of the variance function decreases at an increasing rate. For instance, suppose the concave region starts with the variance function growing linearly and then transitions to (concave) quadratic growth, as might be expected at the end of life for a product; it will satisfy (10).

In summary, we have found conditions for the net demand bound $g(s, t)$ to be a concave function. We have interpreted these conditions and contend that it is reasonable to assume that these conditions hold for many supply chains of interest.

References

Abhyankar, H. S., S. C. Graves. 2001. Creating an inventory hedge for Markov-modulated Poisson demand: An application and a model. *Manufacturing Service Oper. Management* **3** 306–320.

Beyer, D., J. Ward. 2000. Network server supply chain at HP: A case study. HP Labs Technical Report HPL-2000-84, Hewlett Packard, Palo Alto, CA.

Bollapragada, S., T. E. Morton. 1999. A simple heuristic for computing nonstationary (s, S) policies. *Oper. Res.* **47** 576–584.

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

Chen, F. Y., J. K. Ryan, D. Simchi-Levi. 2000. The impact of exponential smoothing forecasts on the bullwhip effect. *Naval Res. Logist.* **47** 269–286.

Coughlin, R. L. 1998. Optimization and measurement of a worldwide supply chain. S.M. thesis, Leaders for Manufacturing Program, MIT, Cambridge, MA.

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

Graves, S. C. 1999. A single-item inventory model for a nonstationary demand process. *Manufacturing Service Oper. Management* **1** 50-61.

Graves, S. C., S. P. Willems. 1998. Optimizing strategic safety stock placement in supply chains. Working paper, http://web.mit.edu/sgraves/www/papers/.

Graves, S. C., S. P. Willems. 2000. Optimizing strategic safety stock placement in supply chains. *Manufacturing Service Oper. Management* **2** 68–83.

Graves, S. C., S. P. Willems. 2002. Strategic inventory placement in supply chains: Nonstationary demand. Working paper, http://web.mit.edu/sgraves/www/papers/.

Humair, S., S. P. Willems. 2006. Optimal inventory placement in networks with clusters of commonality. *Oper. Res.* **54** 725–742.

Johnson, M. E., E. Anderson. 2000. Postponement strategies for channel derivatives. *Internat. J. Logist. Management* **11** 19–35.

Kimball, G. E. 1988. General principles of inventory control. *J. Manufacturing Oper. Management* **1** 119–130.

Kurawarwala, A., H. Matsuo. 1996. Forecasting and inventory management of short life-cycle products. *Oper. Res.* **44** 131–150.

Lee, H. L., V. Padmanabhan, S. Whang. 1997. Information distortion in a supply chain: The bullwhip effect. *Management Sci.* **43** 546–558.

Lesnaia, E. 2004. Optimizing safety stock placement in general network supply chains. PhD thesis, MIT Operations Research Center, Cambridge, MA.

Magnanti, T. L., Z.-J. M. Shen, J. Shu, D. Simchi-Levi, C.-P. Teo. 2006. Inventory placement in acyclic supply chain networks. *Oper. Res. Lett.* **34** 228–238.

Morton, T. E., D. W. Pentico. 1995. The finite horizon nonstationary stochastic inventory problem: Near-myopic bounds, heuristics, testing. *Management Sci.* **41** 334–343.

Simpson, K. F. 1958. In-process inventories. *Oper. Res.* **6** 863-873.