Supply Chain Management with Guaranteed Delivery

Eric Logan Huggins • Tava Lennon Olsen
School of Business Administration, Fort Lewis College, Durango, Colorado 81301
John M. Olin School of Business, Washington University in St. Louis, St. Louis, Missouri 63130
huggins_e@fortlewis.edu • olsen@olin.wustl.edu

We consider a two-stage supply chain under centralized control. The downstream facility faces discrete stochastic demand and passes supply requests to the upstream facility. The upstream facility *always* meets the supply requests from downstream. If the upstream facility cannot meet the supply requests from inventory on hand, the shortage must be filled by expediting, which will incur per unit and setup costs. Such expediting may take the form of overtime production, which occurs at the end of the period and incurs relatively high production costs, or premium freight shipments, which involves building products at the beginning of the period they are needed and shipping them very quickly with relatively high shipping costs. We consider the case where one method of filling shortages is available and determine novel optimal inventory policies under centralized control. At both stages, threshold policies that depend only on the current inventory in the system are optimal; for the total inventory in the system, a base-stock policy is optimal. Numerical analysis provides insight into the optimal policies and allows us to compare the supply chains under centralized and decentralized control.
(*Expediting; Overtime; Supply Chain Coordination; Inventory; Markov Decision Processes*)

1. Introduction

In traditional supply chain situations, downstream facilities make decisions about their order quantities without regard to the actual inventory available upstream. If the upstream facilities do not have enough inventory on hand to fill the orders, it is often assumed that the downstream facility will take what it can get and backorder the rest. We consider a problem with stochastic demand where the downstream facility's supply requests are always met by the upstream facility. If the downstream facility orders more than the upstream facility has on hand, the upstream facility must meet the shortage by expediting. In practice, such expediting often consists of either overtime production or what we call "premium freight." A final option for expediting is outsourcing where the product is purchased from an outside supplier (at a premium).

Overtime production consists of building the required parts at the end of the day, at a higher cost than regular production, and shipping the parts by normal means, with the same shipping cost as would be the case for parts that were built during regular production. Premium freight consists of building parts at the beginning of the same day that they are required downstream, at the same cost as regular production, and expediting shipment, for example, shipping by airplane or helicopter with a higher shipping cost than would be the case for parts shipped the previous day. In our problem, we assume that only one of these mathematically equivalent options is available and that utilizing such expediting incurs

0025-1909/03/4909/1154
1526-5501 electronic ISSN

both fixed and per unit costs. Under these conditions, we examine how an upstream facility and a downstream facility can minimize system costs by working together; the upstream facility will always meet supply requests, and the downstream facility will avoid causing shortages upstream unless absolutely necessary.

We have modeled our problem after the actual inventory control problems faced by a large automobile-parts supplier in Michigan, which we will refer to as "PartCo". PartCo's principal business is to produce engine parts used in vehicle assembly at one of the big three U.S. automobile manufacturers. At PartCo, inventory levels are relatively low, yet PartCo follows a policy of meeting all supply requests, and frequently uses overtime production or premium freight when shortages occur. Backordering is not considered an option because the parts they send downstream are essential to the assembly line and the cost of shutting down the automobile manufacturer's assembly line is extremely high. We have heard a wide range of estimates for this cost, but all have been in the tens of thousands of dollars per hour! Therefore, overtime production and premium freight shipments are "commonly" used, according to our contacts at PartCo. We model a centrally controlled, two-stage supply chain where the upstream facility always meets supply requests from downstream.

In our model, we have attempted to capture the essence of the situation at PartCo, while keeping the analysis tractable. However, the model and results apply elsewhere in the automobile industry and in other industries. According to an article in *The Detroit News* (Smith 2001), Willow Run Airport outside Detroit has recently become the nation's third largest cargo airport due to shipment of automobile parts. The article states that "hardly a car or truck is made anywhere in the United States that doesn't include parts that have traveled through Willow Run Airport" and that "increasingly, Detroit's automakers are flying parts from city to city and from continent to continent." Clearly, shipping parts by air is a significant issue in the automotive supply chain.

Our model also applies to the computer and electronics industries where many manufacturers have reduced or even eliminated their requirements for warehousing, and receive parts in just-in-time fashion. Finally, our results yield new insight into a common assumption made in the inventory literature. In many single-location inventory models, it is assumed that supply requests upstream are always met, without consideration of how they are met, and at what cost. Our results show that supply requests can always be met upstream with some form of expediting, but that it may be much less expensive for the system if the downstream facility is sensitive to the inventory situation upstream and adjusts supply requests accordingly.

The proof of the optimal policies proceeds as follows. We define our cost per period in terms of variables representing the inventory levels and inventory positions at the assembler and the supplier. Note that we use the term assembler, even though no actual assembly may be taking place, to make it clear that this is the downstream partner. We then substitute variables representing the inventory level and position for the entire system for those of the supplier. To simplify the problem, we relax some of the constraints on the possible inventory levels for the assembler and the system, which leads to an optimal cost function that we can solve. Having relaxed some of the constraints, the optimal policy for the assembler becomes a myopic problem. We solve this myopic problem, which leaves us with an optimality equation that depends on the system variables only, so we derive the optimal policy for the system inventory. Finally, we show that the results of our relaxed problem meet the conditions of our original, fully constrained problem. All of our results are for the infinite horizon case, bypassing finite horizon results.

In our previous paper (Huggins and Olsen 2001), we examine the same supply chain, but under decentralized control. In other words, both the supplier and the assembler function independently without sharing any information or inventory decisions. In the decentralized case, the assembler ignores the situation at the supplier and follows a simple base-stock policy, because the assembler's supply requests are always met. At the supplier, we include a fixed cost for regular production and show that the optimal policy is an (s, S) policy for regular production. The form of the expediting policy depends upon the problem data.

An interesting result is that it may be optimal to use overtime production to not only fill a shortage but also to produce up to a positive inventory level. This case will not occur in our centralized model because we assume that there is no fixed cost for regular production at the supplier. The most important distinction to note between the decentralized and centralized cases is that in the decentralized case the assembler ignores the high expense to the supplier caused by shortages. A shortage of a single unit can force the supplier to pay a high fixed cost for expediting. In the centralized case, we will show that the assembler is very sensitive to shortages; the optimal policy for the assembler will reflect this sensitivity.

In this paper, we consider a two-stage supply chain with a single method of expediting. For a review of single-location models with expediting, we refer the reader to Chiang and Gutierrez (1998), Tagaras and Vlachos (2001), or our previous paper, Huggins and Olsen (2001). A thorough overview of supply chain literature may be found in the text edited by Tayur et al. (1998). The seminal work of Clark and Scarf (1960) first considered a multi-echelon inventory problem, and Federgruen and Zipkin (1984) showed that the same results hold over the infinite horizon. In their paper, they assume that if the supplier cannot meet supply requests from the assembler, the assembler is satisfied with as much as it can get. They prove that the optimal policy for the assembler is to ignore the supplier and follow an (s, S) policy (or a base-stock policy with no fixed cost for production). The optimal policy for the supplier is a base-stock policy, with an additional penalty for possibly not meeting supply requests that effectively increases the base-stock level.

The literature on supply chains with expediting or setup costs upstream is limited. Chen and Zheng (1994) consider supply chains with stochastic demand, constant lead times, and setup costs at all stages. They establish lower bounds on the system costs under centralized control. Parlar and Weng (1997) consider a two-period model for products with short life-cycles. Lawson and Porteus (1998) model an $m+1$-stage supply chain without setup costs where at each stage the options are to ship by regular means, expedite, or hold inventory. The authors show that a "top-down base-stock" policy is optimal, where the modified order-up-to decisions are made, in order, from the upstream stage to the downstream stage, and the regular shipment decisions are made before the expedited shipment decisions at each stage. The expediting costs at each stage are not shared with other stages, and the authors use decomposition to solve the problem. Moinzadeh and Aggarwal (1997) consider a two-stage system with one warehouse and several retailers. They assume modified one-for-one $(S-1, S)$ policies for both regular and expedited orders and develop a procedure to find optimal policy parameters. Finally, overtime, or a "vendoring option", in the context of inventory systems with production quotas was considered in Hopp et al. (1993), Duenyas et al. (1993), and Duenyas et al. (1997). These quotas are set in response to a downstream party's demand.

Similar to our assumption, the papers by Gavirneni et al. (1999) and Lee et al. (2000) consider two-stage supply chains where the demand from stage 1 is always met by stage 2. Gavirneni et al. study a capacitated two-stage supply chain under different levels of information and show that order-up-to policies are optimal, then discuss the value of the information to the supplier (stage 2). In their model, they assume that if the supplier faces a shortage, the retailer (stage 1) "acquires the missing part of the order elsewhere." Lee et al. attempt to quantify the value of information in a two-stage supply chain with correlated demand. They assume that when the manufacturer (stage 2) faces a shortage, the manufacturer "obtains units from an alternative source" which they resupply later. Although the manufacturer always meets demand from the retailer (stage 1), they effectively pay a backorder penalty. In both models, stage 1 does not suffer the consequences of expediting, unlike our centralized model. In fact, in a footnote, Lee et al. (2000) point out that:

> In the current paper, we make the assumption that the expedite cost is borne solely by the manufacturer so as to isolate the benefits of information sharing to the manufacturer. A similar assumption was made by most other researchers.... If this assumption is relaxed, then information sharing could bring benefits to both the manufacturer and the retailer, but this requires much more complex modeling of the contractual relationship between the manufacturer and the retailer.

Finally, in our proof of the centralized model, we use results discussed by Porteus (1990), Zheng (1991), Bertsekas (1995), and Rosling (2002). Zheng (1991) shows that (s, S) policies are optimal because the expected one-period cost function is quasi-convex, which we show using a result from Porteus (1990). We use this result to prove the optimal policy for the system inventory. Bertsekas (1995) has several useful propositions; specifically, one proposition states that the optimal cost function satisfies Bellman's equation under assumptions we show to be true in our model. Finally, we assume that our demand distribution is logconcave, and Rosling (2002) discusses properties of logconcave functions.

The rest of this paper is organized as follows. In §2, we define our model, develop cost functions, substitute system variables for supplier variables, and relax two constraints. Under these relaxed conditions, we determine the optimal policies for the assembler and for the system in §3. We prove that the optimal policies under the relaxed conditions are optimal for the original problem in §4. In §5, we conduct a numerical analysis and discuss managerial insights. Section 6 concludes the paper, and some proofs are included in the Appendix.

2. The Model and Cost Functions

We consider a two-stage supply chain where an upstream supplier (stage 2) must deliver products to a downstream assembler (stage 1). A single manager with perfect information about both stages makes all decisions in an effort to minimize total system discounted costs over the infinite horizon. This manager must decide how much to produce each period at stage 1 and at stage 2. Stage 1 experiences exogenous demand; the demand experienced by stage 2 is equal to the amount to be produced at stage 1 the next period. Thus, the production decision at stage 1 directly influences the costs incurred by stage 2. If stage 1 orders more than stage 2 has on hand, stage 2 is forced to expedite to meet the shortage and will incur high costs. If stage 1 orders everything stage 2 has on hand or less, stage 2 will avoid expediting. The optimal policies for both stages will eventually reflect this relationship. Define the following variables:

$D_t =$ the demand during period t,

$x_{1,t} =$ the inventory level at stage 1 after demand has been experienced during period t,

$y_{1,t+1} =$ the inventory position chosen for stage 1 for period $t+1$,

$x_{2,t} =$ the inventory level at stage 2 before it experiences demand from stage 1 during period t, and

$y_{2,t+1} =$ the inventory position chosen for stage 2 for period $t+1$.

The inventory decisions take place after demand is experienced at stage 1. At this point, the inventory at stage 1 is $x_{1,t}$ (which equals $y_{1,t} - D_t$) and the inventory at stage 2 is $x_{2,t}$ (which equals $y_{2,t}$). The manager must decide the inventory positions at both stages, $y_{1,t+1}$ and $y_{2,t+1}$. The decision for stage 1 ($y_{1,t+1}$) determines the demand experienced at stage 2, which is $y_{1,t+1} - x_{1,t}$. If $y_{1,t+1} - x_{1,t} > x_{2,t}$, there is a shortage at stage 2 and they must expedite $(y_{1,t+1} - x_{1,t}) - x_{2,t}$ units. Note that $y_{1,t+1} \geq x_{1,t}$, and $y_{2,t+1} \geq (x_{2,t} - (y_{1,t+1} - x_{1,t}))^+$.

Both of these decisions incur various costs, which we assume to be stationary. At stage 1, linear costs are assessed for production (c_1), holding (h_1), and backordering (b_1). At stage 2, linear costs are assessed for production (c_2), holding (h_2); and expediting incurs linear costs (c_e) plus fixed (K_e) costs. We also make the following assumptions about our model: First, the discount factor α is assumed to be between 0 and 1. Second, we assume that demand is discrete, independent and identically distributed in each period, stationary, and from a logconcave probability distribution (see below) and that the expected value of demand is positive and finite. Third, we let p_d be the probability that demand equals d, $F(d)$ be the probability that demand is less than or equal to d, and assume that demand is nonnegative. Fourth, we assume that the per unit cost of expediting is greater than the per unit cost of production at stage 2. Finally, to ensure that stage 1 policy is reasonable, we assume that the backordering cost at stage 1 is not too low and the holding cost at stage 2 is not too high. For later reference, we label our assumptions as follows:

Assumption 1 (A1). $0 < \alpha < 1$.

Assumption 2 (A2). *The demand distribution is logconcave and for all* t, $0 < E[D_t] < \infty$.

Assumption 3 (A3). $p_d = 0$ *for* $d < 0$.

Assumption 4 (A4). $c_e > c_2$.

Assumption 5 (A5). $b_1 \geq c_e + \alpha((1-\alpha)c_1 - c_2)$ *and* $h_2 \leq h_1 + \alpha(1-\alpha)c_1$.

Note that a distribution $F(x)$ is logconcave if $\log F(x)$ is concave in x. For discrete distributions this means that the differential of $F(x)$, $(F(x+1)-F(x))/F(x)$, is nonincreasing in $x \in \mathcal{I}$. This is not a restrictive assumption because most common discrete distributions are, in fact, logconcave. For example, the discrete uniform, the Poisson, and the binomial are all in this category. Similarly, most commonly assumed continuous distributions (such as the uniform, normal, and exponential distributions) are logconcave. We assume discretized versions of these distributions later in our numerical analysis.

We now proceed to define the various cost functions associated with the model described above. We originally define our cost per period in terms of stage 1 and stage 2 variables, but then make a substitution replacing the stage 2 variables with system inventory variables. Next, we relax some of the constraints on the cost per period to get our relaxed cost per period, $g_r(\cdot)$. Last, we formally define our relaxed optimal cost function $f_r^*(\cdot)$ and show that the stage 1 problem is now myopic. In later sections, we solve the relaxed problem and show that its results are also optimal for the original problem.

For a thorough review of infinite horizon, discounted total cost minimization problems, we refer the reader to Bertsekas (1995). In our problem, we first consider the cost per period g(period k variables), which consists of all the costs incurred by the system during decision period k. For a given policy π, the total, expected, discounted cost over the infinite horizon is $f_\pi(x_0)$, where x_0 is the initial inventory. Mathematically,

$$f_\pi(x_0) \equiv \lim_{N\to\infty} E\left[\sum_{k=0}^{N-1} \alpha^k g(\text{period } k \text{ variables})\right].$$

We are interested in finding the optimal policy π in the set of all feasible policies Π and hence the optimal total, expected, discounted cost over the infinite horizon, $f^*(x_0)$:

$$f^*(x_0) \equiv \min_{\pi\in\Pi} f_\pi(x_0).$$

Dropping the time subscripts for notational convenience, we define our cost per period as

$$\begin{aligned} & g_1(x_1, y_1, x_2, y_2, D) \\ &\quad \equiv \alpha c_1(y_1 - x_1) + K_e\delta((y_1-x_1)-x_2) + c_e((y_1-x_1)-x_2)^+ \\ &\qquad + h_2(x_2-(y_1-x_1))^+ + \alpha c_2(y_2-(x_2-(y_1-x_1))^+) \\ &\qquad + h_1(y_1-D)^+ + b_1(y_1-D)^- \end{aligned}$$

with $y_1 \geq x_1$ and $y_2 \geq (x_2-(y_1-x_1))^+$. The first term is the production cost at stage 1, the next two terms are expediting costs, the fourth term is the holding cost at stage 2, the fifth term is the production cost at stage 2, and the last two terms are holding and backordering costs at stage 1. Moving the $-\alpha c_1 x_1$ back to the previous period (analogous to Veinott 1966), we get

$$\begin{aligned} & g_2(x_1, y_1, x_2, y_2, D) \\ &\quad \equiv \alpha(1-\alpha)c_1 y_1 + \alpha^2 c_1 D + K_e\delta((y_1-x_1)-x_2) \\ &\qquad + c_e((y_1-x_1)-x_2)^+ + h_2(x_2-(y_1-x_1))^+ \\ &\qquad + \alpha c_2(y_2-(x_2-(y_1-x_1))^+) \\ &\qquad + h_1(y_1-D)^+ + b_1(y_1-D)^- \end{aligned}$$

under the same restrictions. Now we will define system variables and substitute for the stage 2 variables. Let the system inventory level be $x_s \equiv x_1 + x_2$ and the system inventory position be $y_s \equiv y_1 + y_2$ and substitute

$$\begin{aligned} & g_3(x_1, y_1, x_s, y_s, D) \\ &\quad \equiv \alpha(1-\alpha)c_1 y_1 + \alpha^2 c_1 D + K_e\delta(y_1-x_s) + c_e(y_1-x_s)^+ \\ &\qquad + h_2(x_s-y_1)^+ + \alpha c_2(y_s-y_1-(x_s-y_1)^+) \\ &\qquad + h_1(y_1-D)^+ + b_1(y_1-D)^- \\ &\quad = \alpha((1-\alpha)c_1-c_2)y_1 + \alpha^2 c_1 D + h_1(y_1-D)^+ \\ &\qquad + b_1(y_1-D)^- \qquad (1) \\ &\qquad + K_e\delta(y_1-x_s) + c_e(y_1-x_s)^+ \\ &\qquad + (h_2-\alpha c_2)(x_s-y_1)^+ + \alpha c_2 y_s \qquad (2) \end{aligned}$$

with $y_1 \geq x_1$ and $y_s \geq y_1 + (x_s - y_1)^+$. Note that the second restriction is equivalent to $y_s \geq \max\{y_1, x_s\}$. Also, we can rewrite $g_3(\cdot)$ as

$$g_3(x_1, y_1, x_s, y_s, D) = L_1(y_1, D) + L_2(y_1, x_s) + \alpha c_2 y_s$$

where $L_1(y_1, D)$ represents the terms on line (1) and $L_2(y_1, x_s)$ represents the terms on line (2). We can now write the fully constrained optimal cost function which we would like to solve, namely,

$$f^*(x_1, x_s) = \min_{y_s \geq \max\{x_s, y_1\},\, y_1 \geq x_1} E_D[g_3(x_1, y_1, x_s, y_s, D) + \alpha f^*(y_1 - D, y_s - D)].$$

To solve this equation, we relax some of the constraints. Later we will show that these constraints are always met by the optimal solution to the relaxed problem, and that they thus solve the original, fully constrained problem. First, we drop the constraint that $y_1 \geq x_1$. Second, we drop the constraint that $y_s \geq y_1$ in the case when $y_1 > x_s$. For later reference, we label the relaxed assumptions as:

Relaxed Assumption 1 (R1). $y_1 \geq x_1$.

Relaxed Assumption 2 (R2). $y_s \geq y_1$ *when* $y_1 > x_s$.

After relaxing the constraints, our relaxed cost per period has the same costs as $g_3(\cdot)$ but with only one constraint:

$$g_r(x_1, y_1, x_s, y_s, D) \equiv L_1(y_1, D) + L_2(y_1, x_s) + \alpha c_2 y_s,$$

with $y_s \geq x_s$. Now, we show that $g_r(\cdot) \geq 0$ and then apply a result from Bertsekas (1995) to obtain our relaxed optimal cost function.

Lemma 1. $g_r(x_1, y_1, x_s, y_s, D) \geq 0$.

Proof. The proof is in the Appendix. □

Because $g_r(x_1, y_1, x_s, y_s, D) \geq 0$, Proposition 1.1 of Bertsekas (1995, p. 137) holds, and the relaxed optimal cost function f_r^* satisfies

$$\begin{aligned} f_r^*(x_s) &= \min_{y_1,\, y_s \geq x_s} E_D\big[g_r(x_1, y_1, x_s, y_s, D) + \alpha f_r^*(y_s - D)\big] \\ &= \min_{y_1,\, y_s \geq x_s} \big\{E_D[L_1(y_1, D)] + L_2(y_1, x_s) + \alpha c_2 y_s \\ &\qquad + \alpha E_D[f_r^*(y_s - D)]\big\}. \end{aligned} \tag{3}$$

Notice that y_1 has no effect on either y_s or the cost to go, $\alpha E_D[f_r^*(y_s - D)]$. Thus,

$$\begin{aligned} f_r^*(x_s) &= \min_{y_s \geq x_s} \Big\{ \min_{y_1} \{E_D[L_1(y_1, D)] + L_2(y_1, x_s)\} \\ &\qquad + \alpha c_2 y_s + \alpha E_D[f_r^*(y_s - D)] \Big\} \\ &= \min_{y_s \geq x_s} \big\{ m(x_s) + \alpha c_2 y_s + \alpha E_D[f_r^*(y_s - D)] \big\}, \end{aligned} \tag{4}$$

where $m(x_s) = \min_{y_1}\{E_D[L_1(y_1, D)] + L_2(y_1, x_s)\}$. Finding the optimal inventory policy for stage 1 has become a myopic problem which we solve in the first part of §3.

3. Optimal Policies for the Relaxed Problem

In this section, for the relaxed problem, we determine the optimal inventory policies for stage 1 and for the system. We study the function $m(x_s)$ and show that the stage 1 policy depends only on the system inventory level x_s. Next, we show that the optimal inventory policy for the system is a base-stock policy. Define $N_H(y_1)$ and $N_L(y_1)$ as

$$\begin{aligned} N_H(y_1) &= (\alpha(1-\alpha)c_1 - h_2) y_1 \\ &\quad + E_D\big[\alpha^2 c_1 D + h_1 (y_1 - D)^+ + b_1 (y_1 - D)^-\big] \end{aligned}$$

and

$$\begin{aligned} N_L(y_1) &= (\alpha((1-\alpha)c_1 - c_2) + c_e) y_1 \\ &\quad + E_D\big[\alpha^2 c_1 D + h_1 (y_1 - D)^+ + b_1 (y_1 - D)^-\big]. \end{aligned}$$

We now have that

$$\begin{aligned} & m(x_s) \\ &= \min_{y_1} \big\{E_D[L_1(y_1, D)] + L_2(y_1, x_s)\big\} \\ &= \min_{y_1} \begin{cases} E_D[L_1(y_1, D)] + (h_2 - \alpha c_2)(x_s - y_1) \\ \quad \text{if } y_1 \leq x_s \\ E_D[L_1(y_1, D)] + K_e + c_e(y_1 - x_s) \\ \quad \text{if } y_1 > x_s \end{cases} \\ &= \min_{y_1} \begin{cases} \left\{ \begin{array}{l} (h_2 - \alpha c_2) x_s + (\alpha(1-\alpha)c_1 - h_2) y_1 \\ \quad + E_D[\alpha^2 c_1 D + h_1(y_1 - D)^+ \\ \qquad + b_1 (y_1 - D)^-] \end{array} \right\} & \text{if } y_1 \leq x_s \\ \left\{ \begin{array}{l} K_e - c_e x_s + (\alpha((1-\alpha)c_1 - c_2) + c_e) y_1 \\ \quad + E_D[\alpha^2 c_1 D + h_1(y_1 - D)^+ \\ \qquad + b_1 (y_1 - D)^-] \end{array} \right\} & \text{if } y_1 > x_s \end{cases} \\ &= \min_{y_1} \begin{cases} (h_2 - \alpha c_2) x_s + N_H(y_1) & \text{if } y_1 \leq x_s \\ K_e - c_e x_s + N_L(y_1) & \text{if } y_1 > x_s \end{cases} \end{aligned}$$

$$= \min \begin{cases} (h_2 - \alpha c_2)x_s + \min_{y_1 \le x_s}\{N_H(y_1)\} \\ K_e - c_e x_s + \min_{y_1 > x_s}\{N_L(y_1)\}. \end{cases}$$

Before continuing our study of $m(x_s)$, we derive properties about $N_L(y_1)$ and $N_H(y_1)$ in the following lemma.

Lemma 2. *Define* $y_H = \arg\min_{y_1}\{N_H(y_1)\}$ *and* $y_L = \arg\min_{y_1}\{N_L(y_1)\}$.
(1) $N_L(y_1)$ *and* $N_H(y_1)$ *are convex in* y_1.
(2) $0 \le y_L \le y_H \le \infty$.

Proof. The proof of part (1) is straightforward. To prove part (2), we examine the differential of both functions; see the Appendix for details. □

Returning to our study of $m(x_s)$ and defining

$$N(x_s) \equiv \alpha((1-\alpha)c_1 - c_2)x_s + E_D\big[\alpha^2 c_1 D + h_1(x_s - D)^+ + b_1(x_s - D)^-\big],$$

we have that

$$m(x_s) = \min \begin{cases} (h_2 - \alpha c_2)x_s + \min_{y_1 \le x_s}\{N_H(y_1)\} \\ K_e - c_e x_s + \min_{y_1 > x_s}\{N_L(y_1)\} \end{cases}$$

$$= \min \begin{cases} (h_2 - \alpha c_2)x_s + N_H(y_H) & \text{if } x_s \ge y_H \\ (h_2 - \alpha c_2)x_s + N_H(x_s) & \text{if } x_s < y_H \\ K_e - c_e x_s + N_L(x_s) & \text{if } x_s > y_L \\ K_e - c_e x_s + N_L(y_L) & \text{if } x_s \le y_L \end{cases}$$

$$= \min \begin{cases} (h_2 - \alpha c_2)x_s + N_H(y_H) & \text{if } x_s \ge y_H \\ N(x_s) & \text{if } x_s < y_H \\ K_e + N(x_s) & \text{if } x_s > y_L \\ K_e - c_e x_s + N_L(y_L) & \text{if } x_s \le y_L \end{cases}$$

$$= \min \begin{cases} (h_2 - \alpha c_2)x_s + N_H(y_H) & \text{if } x_s \ge y_H \\ N(x_s) & \text{if } x_s < y_H \\ K_e - c_e x_s + N_L(y_L) & \text{if } x_s \le y_L. \end{cases}$$

Define t_L as the smallest w such that $N(w) \le K_e - c_e w + N_L(y_L)$. We get that

$$m(x_s) = \begin{cases} (h_2 - \alpha c_2)x_s + N_H(y_H) & \text{if } x_s \ge y_H, \\ N(x_s) & \text{if } t_L \le x_s < y_H, \\ K_e - c_e x_s + N_L(y_L) & \text{if } x_s < t_L. \end{cases} \tag{5}$$

So, we have defined $m(x_s)$ explicitly and in the process we have determined the optimal inventory control policy at stage 1. If the system inventory is large, $x_s \ge y_H$, we order up to y_H. If the system inventory is medium, $t_L \le x_s < y_H$, we use up the system inventory, x_s. Finally, if the system inventory is small, $x_s < t_L$, we order up to y_L.

Theorem 1. *Let* y_1^* *be the optimal inventory position at stage* 1 *for the relaxed problem. Then,*

$$y_1^* = \begin{cases} y_H & \text{if } x_s \ge y_H, \\ x_s & \text{if } t_L \le x_s < y_H, \\ y_L & \text{if } x_s < t_L. \end{cases} \tag{6}$$

Proof. By definition of $m(x_s)$. □

Given $m(x_s)$, we now have that the optimal relaxed cost function is in terms of system variables only. From Equation (4), we have

$$f_r^*(x_s) = \min_{y_s \ge x_s}\big\{m(x_s) + \alpha c_2 y_s + \alpha E_D[f_r^*(y_s - D)]\big\}.$$

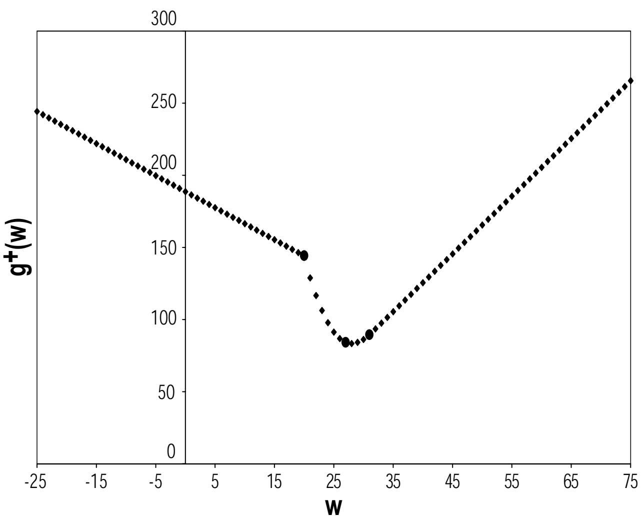
Now, we move the $m(x_s)$ term back to the previous period and get

$$f_{rm}^*(x_s) = \min_{y_s \ge x_s}\big\{\alpha c_2 y_s + \alpha E_D[m(y_s - D)] + \alpha E_D[f_{rm}^*(y_s - D)]\big\}$$
$$= \min_{y_s \ge x_s}\big\{G(y_s) + \alpha E_D[f_{rm}^*(y_s - D)]\big\}$$

where $G(y_s) = \alpha c_2 y_s + \alpha E_D[m(y_s - D)]$. We need to justify two steps here. First, we can move $m(x_s)$ back a period and $f_r^*(\cdot)$ will have the same optimal policy as $f_{rm}^*(\cdot)$ using an argument similar to Veinott (1966). Second, to prove the existence of $f_{rm}^*(\cdot)$, we must show that $g(y_s) \equiv \alpha(c_2 y_s + m(y_s - D)) \ge 0$ according to Proposition 1.1 of Bertsekas (1995, p. 137). To prove $g(y_s)$ is nonnegative and to later prove that $G(y_s)$ is quasi-convex, let us examine the function $g^+(w) \equiv c_2 w + m(w)$. Graphically, the function looks similar to Figure 1.

Starting from the left, $g^+(\cdot)$ decreases at rate $-(c_e - c_2)$ until point $t_L - 1$. (The big dot on the left is $t_L - 1$, the big dot in the middle is y_L, and y_H is the big dot on the right.) From t_L to $y_H - 1$, it follows $c_2 w + N(w)$, decreasing at first, then increasing. From y_H on, it increases at rate $h_2 + (1-\alpha)c_2$.

Figure 1 **Graph of $g^+(w)$**



Lemma 3. *The function $g^+(\cdot)$ has exactly one minimum which occurs between t_L and $y_H - 1$ and is positive.*

Proof. The proof is in the Appendix. □

Theorem 2. *For the relaxed problem, the optimal inventory control policy for the system inventory is a base-stock policy.*

Proof. Consider $g(y_s)$:

$$\begin{aligned} g(y_s) &= \alpha(c_2 y_s + m(y_s - D)) \\ &= \alpha(c_2 y_s - c_2 D + c_2 D + m(y_s - D)) \\ &= \alpha(c_2 D + g^+(y_s - D)) \geq 0, \end{aligned}$$

where the inequality holds because $g^+(\cdot) \geq 0$. Note from Figure 1 or Lemma 3 that $g^+(\cdot)$ is a quasi-convex function with a minimum point. Now consider $G(y_s)$:

$$\begin{aligned} G(y_s) &= \alpha E_D[c_2 y_s + m(y_s - D)] \\ &= \alpha(c_2 E_D[D] + E_D[g^+(y_s - D)]). \end{aligned}$$

The first term is a constant, and the second term is a convolution of a quasi-convex function ($g^+(y_s - D)$) and a logconcave probability distribution by Assumption (A2). Thus, according to Porteus (1990, p. 619), $G(\cdot)$ is a quasi-convex function. Because $G(\cdot)$ is quasi-convex, the desired result follows from Zheng (1991). □

4. Optimal Policies for the Original Problem

In the previous section, we determined the optimal policies for the relaxed problem

$$\begin{aligned} f_r^*(x_s) &= \min_{y_s \geq x_s} \{m(x_s) + \alpha c_2 y_s + \alpha E_D[f_r^*(y_s - D)]\} \\ &= \min_{y_s \geq x_s,\, y_1} E_D\big[g_r(x_1, y_1, x_s, y_s, D) + \alpha f_r^*(y_s - D)\big]. \end{aligned}$$

Recall that our fully constrained problem is

$$\begin{aligned} f^*(x_1, x_s) = \min_{y_s \geq \max\{x_s, y_1\},\, y_1 \geq x_1} & E_D\big[g_r(x_1, y_1, x_s, y_s, D) \\ & + \alpha f^*(y_1 - D, y_s - D)\big]. \end{aligned}$$

We must show that the optimal policies for the relaxed problem minimize the fully constrained problem and that both Relaxed Assumptions (R1) and (R2) are met. To do so, we need one additional assumption that our initial inventory at stage 1 does not exceed the maximum order level at stage 1, y_H.

Assumption 6 (A6). $x_1 \leq y_H$.

Theorem 3. $f^*(x_1, x_s) = f_r^*(x_s)$.

Proof. The optimal policies for the relaxed problem minimize the costs for the fully constrained problem because both relaxed constraints are met and y_1 does not affect y_s or the costs-to-go. If $x_s < t_L$, $y_1^* = y_L \geq t_L \geq x_s \geq x_1$. If $t_L \leq x_s < t_H$, $y_1^* = x_s \geq x_1$. Finally, if $x_s \geq t_H$, $y_1^* = y_H \geq x_1$ by Assumption (A6). Thus, the first relaxation (R1) is satisfied. To show that (R2) is satisfied, define y_s^* to be the optimal system inventory position, S^* to be the optimal system base-stock level, and y_2^* to be the optimal inventory position for stage 2. We must show that $y_s \geq y_1$ when $y_1 > x_s$. The only time that $y_1 > x_s$ is when $x_s < t_L$ (otherwise, $y_1 = x_s$ or $y_1 = y_H \leq x_s$). In this case, $y_1^* = y_L \leq S^* = y_s^*$. The inequality holds because $0 \leq y_2^* = y_s^* - y_1^* = S^* - y_L$. □

For the original, fully constrained problem, we now know the optimal policies for stage 1, stage 2, and for the system. The order-up-to levels are as follows:

$$y_1^* = \begin{cases} y_H & \text{if } x_s \geq y_H, \\ x_s & \text{if } t_L \leq x_s < y_H, \\ y_L & \text{if } x_s < t_L, \end{cases}$$

$$y_s^* = \begin{cases} x_s & \text{if } x_s > S^*, \\ S^* & \text{if } x_s \le S^*, \end{cases} \qquad \text{and}$$

$$y_2^* = y_s^* - y_1^*.$$

5. Numerical Analysis and Managerial Insights

The optimal policies for the decentralized and centralized supply chains, as discussed in the introduction, are quite different from each other. In the decentralized case, stage 1 and stage 2 both follow base-stock policies. In the centralized case, as we have just shown, stage 1 and stage 2 follow interesting policies that only depend on the system inventory level. Natural questions arise from this difference, such as how much can be saved by using the centralized optimal policy rather than the decentralized optimal policy? Also, how do changes in the various parameters affect inventory, savings, and frequency of expediting? To answer these questions, we wrote a C++ program and performed numerical analyses for various parameters and demand distributions.

5.1. General Numerical Results

In our experiment, we set $c_1 = 10$ and varied all the other parameters. We let $\alpha = 0.95$, 0.99, and 0.995. For stage 1, we let $h_1 = 0.01$, 0.05, and 0.10 and $b_1 = 20$, 30, and 40. For stage 2, we let $c_2 = 3$, 5, and 9 and $h_2 = 0.005$, 0.01, and 0.05. For expediting, we let $c_e = 4$, 6, and 10 and $K_e = 0$, 50, and 200. These variations led to a total of $3^7 = 2{,}187$ possible combinations. However, one-third of the combinations violate either Assumption (A4) or (A5) and these results were not considered.

For each combination, we made several calculations. We calculated the optimal base-stock levels, the total costs, and the inventory/expediting costs for stage 1 and stage 2 in the decentralized case. For the centralized case, we calculated the optimal inventory control parameters for stage 1: t_L, y_L, and y_H. Using these parameters, we calculated the system base-stock level S^* and the total cost and inventory/expediting costs for the system under centralized control. We calculated four statistics comparing the centralized case and the decentralized case: the percentage savings in total costs (TS%), the percentage savings in inventory and expediting costs (I/ES%), the percentage reduction in system inventory (IR%), and the probability of using expediting under both decentralized (P(E)D) and centralized (P(E)C) control. We averaged these statistics over all feasible combinations. Finally, we divided the probability of expediting under decentralized control by the probability of expediting under centralized control to develop a ratio (D/C) that reflects how much more frequently expediting is used under decentralized control.

We compared the results for four different demand distributions: Poisson(mean), Uniform(lower bound, upper bound), Normal(mean, standard deviation), and Exponential(mean). Because we considered discrete demand, we used discrete approximations for the last three distributions. Also, we truncated each distribution below at zero and above at 49 to fit into our probability array, and adjusted the probabilities appropriately to ensure that the total probability was one. (We chose a probability array of size 50 as being large enough to distinguish different distributions, but small enough to converge quickly. Also, note that each unit of demand could represent a batch of 10, 100, 1,000, or any number of parts.) The average results for several demand distributions are in Table 1.

We ran the experiment for constant demand (Normal(25,0)) as one way to check the accuracy of our computer code. First, notice that the total savings percentages are relatively small, less than 1%. We believe this result is due to the high production costs relative to the low holding costs; with the holding costs we assumed, it is relatively inexpensive to hold buffer inventory. On the one hand, in an earlier experiment we assumed much higher holding costs

Table 1 **Average Savings of the Centralized Optimal Policy**

Demand distribution	TS%	I/ES%	IR%	P(E)D%	P(E)C%	D/C
Normal(25, 0)	0.00	0.0	0.0	0.00	0.00	N/A
Normal(25, 1)	0.04	18.8	2.3	0.70	0.28	2.5
Normal(25, 5)	0.20	18.4	8.9	1.64	0.67	2.4
Normal(25, 10)	0.27	13.3	9.8	1.70	0.67	2.5
Poisson(25)	0.22	19.1	9.7	1.64	0.65	2.5
Uniform(0, 49)	0.09	3.5	3.2	1.41	0.61	2.3
Exponential(15)	0.78	13.4	15.9	1.85	0.69	2.7

(to model a lean inventory system) and found much higher total savings percentages. On the other hand, notice that both the inventory and expediting savings percentages and the inventory reduction percentages are quite significant. Finally, observe that under decentralized control expediting is about two and one-half times more likely than under centralized control. Avoiding expediting and reducing system inventories are significant factors that lead to the observed savings under centralized control. We feel that some insight can be gained by examining a typical numerical example, as follows.

Consider a problem that has Poisson demand with mean 25 and has the middle value for all the parameters. (That is, $\alpha = 0.99$, $h_1 = 0.05$, $b_1 = 30$, $c_2 = 5$, $h_2 = 0.025$, $c_e = 6$, and $K_e = 50$.) Under decentralized control, the optimal policy at both stages is a base-stock policy, coincidentally with a base-stock level of 39 at both stages. Under centralized control, the optimal policy at stage 1 is

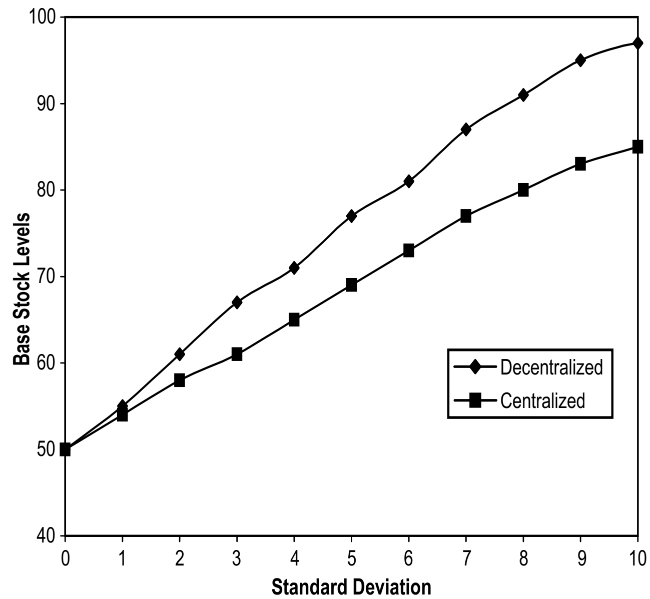
$$y_1^* = \begin{cases} 39 & \text{if } x_s \geq 39, \\ x_s & \text{if } 25 \leq x_s < 39, \\ 34 & \text{if } x_s < 25, \end{cases}$$

and the optimal policy for the system is to order up to a base-stock level of 70. So, under the centralized policy, the system carries eight fewer units of inventory, or an inventory reduction of 10.3%. This inventory reduction leads to a savings on inventory and expediting costs of 21.3% and a total savings of 0.16%. These results are typical for the majority of our experimental outcomes.

5.2. Parametric Changes

To further understand the differences between decentralized and centralized control, we varied the problem parameters for the typical example mentioned in the previous section. First, we assumed a normal distribution with mean 25, but varied the standard deviation from 0 to 10. Second, we assumed a Poisson distribution with mean 25, then varied the holding costs, the backorder cost, the setup cost for expediting, and the discount factor. We varied h_1 from 0.01 to 1.01 (letting $h_2 = h_1/2$), b_1 from 10 to 50, K_e from 0 to 200, and α from 0.899 to 0.999.

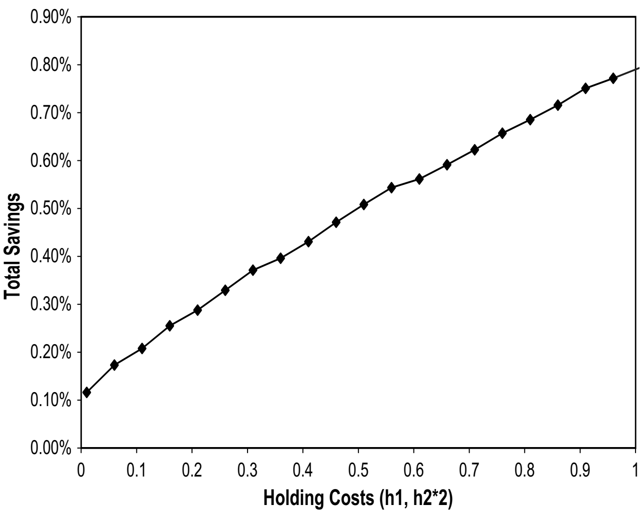
Figure 2 Base-Stock Levels vs. Standard Deviation



As we increased the standard deviation of the demand distribution, all the inventory levels increased as expected, as seen in Figure 2. The lower line is the system base-stock level under centralized control and the upper line is the sum of the base-stock levels at both stages under decentralized control. The base-stock levels appear to diverge.

For our next experiment, we varied the holding costs. As holding costs increased, all the inventory levels decreased, as expected. As seen in Figure 3, the total savings percentage of centralized control over decentralized control increased almost linearly as the holding costs increased (whereas the inventory and expediting production savings percentage and the inventory reduction percentage decreased). The probability of using expediting increased under both kinds of control as the holding costs increased (shown in Figure 4), as expected, while the ratio (D/C) comparing decentralized and centralized expediting utilization decreased. In Figure 4, the lower line represents the probability of using expediting under centralized control and the upper line represents the same probability under decentralized control. Note that this probability is always lower under centralized control.

Figure 3 Total Savings Percentage vs. Holding Costs



Next, we varied the backorder cost and the setup cost for expediting. As the backorder cost increases, the percentage savings on inventory and expediting tends to decrease. High backorder costs make the centralized policy less effective. As the expediting setup cost increases, all statistics increase except for the probability of expediting which decreases under both forms of control. Figure 5 shows the increase in inventory and expediting cost savings percentage (the top line) and in inventory reduction percentage (the bottom line). Under centralized control, the expediting

Figure 4 Probability of Expediting vs. Holding Costs

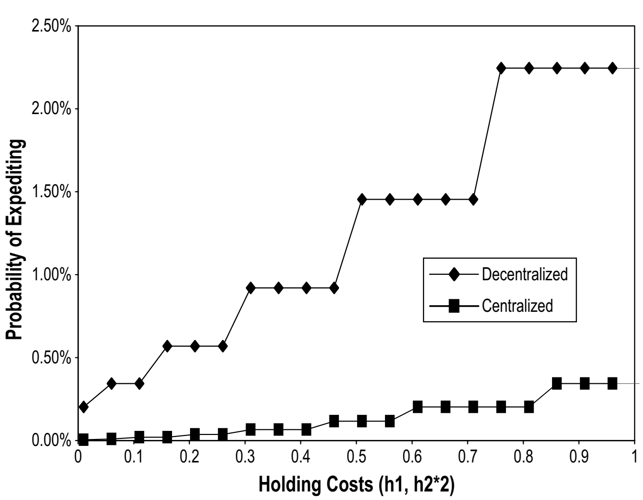
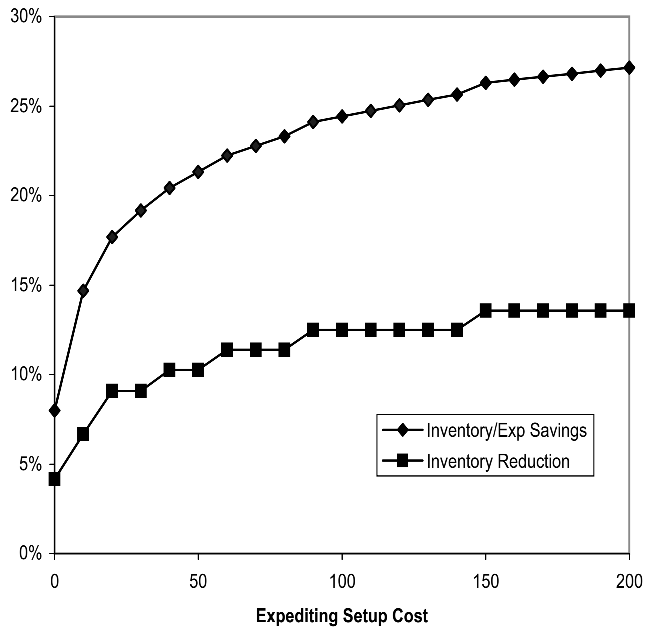


Figure 5 Inventory/Expediting Costs and Inventory Reduction vs. Expediting Setup Cost



setup cost can be avoided if stage 1 “underorders” from stage 2.

Finally, we varied the discount factor. As the discount factor increased, so did all the inventory levels. On the other hand, as the discount factor increased, the total savings percentage and the probability that expediting will be used (under both centralized and decentralized control) generally decreased. A graph of the probability of expediting versus the discount factor is not shown as it looks very similar to Figure 4 mirrored so that the probabilities are decreasing rather than increasing.

5.3. Managerial Insights

As shown in Table 1, the centralized optimal policy generally affects significant savings on inventory and expediting costs and reduces system inventory when compared with the decentralized optimal policy. In real situations it could be costly to coordinate the two stages and share information, but it may well be cost effective, considering that the inventory/expediting savings are typically more than 10%. In particular, if the discount factor is low or the holding costs or expediting costs are high, following the centralized optimal policy seems worth the effort. Specifically, companies operating under a lean inventory paradigm may consider their holding and expediting costs to be very high and hence may want to consider the benefits of centralized control.

Clearly, to cut costs in this kind of supply chain, stage 1 must be sensitive to the amount of inventory available at stage 2. Stage 1 must be willing to occasionally "underorder" to save significant expediting costs at stage 2. By the same token, stage 2 must be willing to produce extra units when stage 1 underorders, trusting that stage 1 will want those additional units the next period. Here, it is interesting to compare our centralized results with those of Federgruen and Zipkin (1984) (the infinite horizon extension of Clark and Scarf 1960). In their model, stage 1 completely ignores stage 2 and follows a base-stock model dependent on only stage 1 cost parameters; stage 2 also follows a base-stock policy, but with a higher base-stock level to reduce the chance of not filling supply requests from stage 1. In our model, stage 1 is sensitive to the costs and inventory available at stage 2, and orders accordingly; stage 2 orders more when stage 1 underorders, bringing the system inventory up to a base-stock level.

Finally, in most scenarios, the centralized policy is an effective way to reduce the total inventory in the system. For managers interested in following a lean inventory paradigm, not only does the centralized optimal policy offer a way to reduce inventories and costs simultaneously but the centralized optimal policy reduces the likelihood of expediting, an outcome which would make most managers very happy.

6. Conclusion and Extensions

In this paper, we have modeled a two-stage supply chain where supply requests are always met by the upstream facility. We have shown that the optimal inventory control policies for both stages depend only on the system inventory, and that the optimal policy for the system inventory is a base-stock policy.

We solved the problem by substituting variables for system inventory and then relaxing two constraints. After this relaxation, we got a myopic problem for stage 1 that we solved for the optimal policy which depends on two thresholds and the system inventory. Next, we solved the optimality equation for the system and showed that a base-stock policy is optimal. Finally, we showed that the solutions for the relaxed problem solved the original, fully constrained problem, and hence we found the overall optimal policies.

We performed an experiment for several different demand distributions and parameter values. The results of this experiment indicate that the centralized optimal policy saves only a small amount on total costs, but significantly reduces the inventory and expediting costs, the system inventory level, and the probability of using expediting. Numerical examples illustrate where the savings occur. Our main managerial insights are that to cut costs stage 1 must occasionally underorder, and that by underordering, inventory and expediting costs as well as inventory levels can generally be reduced by following the centralized policy.

Some of our analyses and results are distinctive when compared with traditional inventory literature. Traditional two-echelon proofs proceed by separating variables and then solving two independent problems. Our separation is a little different. We found that by substituting system variables and relaxing some constraints, we could first solve a myopic problem, then solve a straightforward dynamic program. Our optimal policies also vary from traditional optimal inventory policies. Our stage 1 policy of ordering up to two separate inventory levels and occasionally underordering is quite different from traditional inventory policies. Hence, we feel that our base-stock result for the system is also interesting.

The most obvious extension to this model is channel coordination. Is there a way to induce both stage 1 and stage 2 to follow the centralized optimal policy? If so, how will the two stages share the various costs involved? We are currently working on this problem and can make a few observations about possible solutions. First, the two stages must share information to achieve the centralized results. The centralized policies depend only on the system inventory, which is the sum to the inventories at both stages, and thus at least one of the stages must know the system inventory to order appropriately. Second, for a cost structure to coordinate both stages, the cost structure will very likely be two-tiered to create the two thresholds that determine inventory positions at both stages.

Other possible extensions to this include multiple types of expediting, capacity constraints, and the inclusion of lead times. Capacity constraints may apply to either regular production, overtime production, or both. However, if both types of production are constrained, we cannot guarantee that demand will always be met. In our problem, we assume that deliveries either occur overnight or almost instantaneously with premium freight. This situation is a reasonable reflection of reality at PartCo because they deliver predominantly to neighboring manufacturing plants. However, some of their shipments travel farther, even outside the country, and in these cases lead times would apply. Incorporating other types of expediting, capacity, or lead times are left as the subjects of future work.

Acknowledgments

The authors thank two anonymous referees and the associate editor for their extremely helpful suggestions that improved this paper. They also thank Brad Johnson and Michael Knox, their contacts at "PartCo," for sharing their insights about the real inventory challenges faced by their employer. This research was funded in part by NSF grant DMI-9713727.

Appendix

Proof of Lemma 1. We consider the case where $y_1 \le x_s$ and the case where $y_1 > x_s$. Note that in both cases it is possible that $y_1 < 0$, because backorders are allowed at stage 1.

When $y_1 \le x_s$,

$$\begin{aligned} g_r(x_1, y_1, x_s, y_s, D) &\ge h_2(x_s - y_1) - \alpha c_2 x_s + \alpha c_2 y_s \\ &\ge \alpha c_2 (y_s - x_s) \ge 0. \end{aligned}$$

The first inequality holds because we drop nonnegative terms from Lines (1) and (2) and cancel the $\alpha c_2 y_1$ terms. The second inequality holds because $y_s \ge x_s$.

When $y_1 > x_s$,

$$\begin{aligned} g_r(x_1, y_1, x_s, y_s, D) &\ge -\alpha c_2 y_1 + c_e(y_1 - x_s) + \alpha c_2 y_s \\ &= c_e(y_1 - x_s) - \alpha c_2 (y_1 - y_s) \\ &\ge c_e(y_1 - x_s) - \alpha c_2 (y_1 - x_s) \\ &= (c_e - \alpha c_2)(y_1 - x_s) \ge 0. \end{aligned}$$

The first inequality holds because we drop nonnegative terms from Lines (1) and (2), the second inequality holds because $y_s \ge x_s$, and the third inequality holds because $c_e \ge \alpha c_2$ by Assumptions (A1) and (A4). □

Proof of Lemma 2, Part (2). Note that the middle inequality is satisfied because $-h_2 < 0 < c_e - \alpha c_2$ by Assumptions (A1) and (A4). To prove the other inequalities, we define the differential of each function as $\Delta N_i(y_1) = N_i(y_1 + 1) - N_i(y_1)$ for $i = L, H$. To calculate y_L, we must solve $\Delta N_L(y_1) = 0$. If the solution to this equation is not integer, y_L will be either the ceiling or the floor of the solution to this equation. Consider

$$\begin{aligned} \Delta N_L(y_1) &= N_L(y_1 + 1) - N_L(y_1) \\ &= (\alpha((1-\alpha)c_1 - c_2) + c_e)(y_1 + 1) \\ &\quad + E_D[\alpha^2 c_1 D + h_1(y_1 + 1 - D)^+ + b_1(y_1 + 1 - D)^-] \\ &\quad - (\alpha((1-\alpha)c_1 - c_2) + c_e) y_1 \\ &\quad - E_D[\alpha^2 c_1 D + h_1(y_1 - D)^+ + b_1(y_1 - D)^-] \\ &= (\alpha((1-\alpha)c_1 - c_2) + c_e) \\ &\quad + E_D[h_1(y_1 + 1 - D)^+ - h_1(y_1 - D)^+ \\ &\quad + b_1(y_1 + 1 - D)^- - b_1(y_1 - D)^-] \\ &= (\alpha((1-\alpha)c_1 - c_2) + c_e) + h_1 F(y_1) - b_1(1 - F(y_1)) \\ &= \alpha((1-\alpha)c_1 - c_2) + c_e - b_1 + (h_1 + b_1) F(y_1). \end{aligned}$$

Similarly,

$$\Delta N_H(y_1) = \alpha(1-\alpha)c_1 - h_2 - b_1 + (h_1 + b_1) F(y_1).$$

Thus, at each respective minimum, $\Delta N_L(y_L) = \alpha((1-\alpha)c_1 - c_2) + c_e - b_1 + (h_1 + b_1)F(y_L) \approx 0$ and $\Delta N_H(y_H) = \alpha(1-\alpha)c_1 - h_2 - b_1 + (h_1 + b_1)F(y_H) \approx 0$. Or,

$$y_L \approx F^{-1}\left(\frac{b_1 - \alpha((1-\alpha)c_1 - c_2) - c_e}{h_1 + b_1}\right) \quad \text{and}$$

$$y_H \approx F^{-1}\left(\frac{b_1 + h_2 - \alpha(1-\alpha)c_1}{h_1 + b_1}\right).$$

For y_L and y_H to exist, we require that the first fraction be nonnegative and that the second fraction be less than or equal to one. So, we require that $b_1 - \alpha((1-\alpha)c_1 - c_2) - c_e \ge 0$ and $b_1 + h_2 - \alpha(1-\alpha)c_1 \le h_1 + b_1$, which both hold by Assumption (A5). Under these conditions, we have that $0 \le y_L \le y_H \le \infty$. □

Proof of Lemma 3. To the left of t_L, the slope of $g^+(\cdot)$ is $-(c_e - c_2) < 0$ and to the right of $y_H - 1$, the slope of $g^+(\cdot)$ is $h_2 + (1-\alpha)c_2 > 0$. Also note that $g^+(t_L) \le g^+(t_L - 1)$ by definition of t_L. Thus, any minima of the function occur between t_L and $y_H - 1$. Between these values, $g^+(\cdot)$ follows $c_2 w + N(w)$, a convex function, and thus there is exactly one minimum. The minimum value is positive because

$$\begin{aligned} c_2 w + N(w) &= c_2 w + \alpha((1-\alpha)c_1 - c_2) w \\ &\quad + E_D[\alpha^2 c_1 D h_1(w - D)^+ + b_1(w - D)^-] \\ &\ge (1-\alpha)(\alpha c_1 + c_2) w + E_D[b_1(w - D)^-] \ge 0. \end{aligned}$$

The first inequality is true because we drop two nonnegative terms. The second inequality depends on whether or not w is negative.

If $w \geq 0$, the second inequality is obvious.

If $w < 0$, the second inequality is true because we get $c_2 w + N(w) \geq (1-\alpha)(\alpha c_1 + c_2)w - b_1 w + E_D[D] \geq 0$ by Assumptions (A2) and (A5). $\square$

References

Bertsekas, D. 1995. *Dynamic Programming and Optimal Control*, Vol. 2. Athena Scientific, Belmont, MA.

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

Chiang, C., G. Gutierrez. 1998. Optimal control policies for a period review inventory system with emergency orders. *Naval Res. Logist.* **45** 187–204.

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

Duenyas, I., W. J. Hopp, Y. Bassok. 1997. Production quotas as bounds on interplant JIT contracts. *Management Sci.* **43** 1372–1386.

——, ——, M. L. Spearman. 1993. Characterizing the output process of a CONWIP line with deterministic processing and random outages. *Management Sci.* **39** 975–988.

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

Gavirneni, S., R. Kapuscinski, S. Tayur. 1999. Value of information in capacitated supply chains. *Management Sci.* **45** 16–24.

Hopp, W. J., M. L. Spearman, I. Duenyas. 1993. Economic production quotas for pull manufacturing systems. *IIE Trans.* **25** 71–79.

Huggins, E., T. Olsen. 2001. Inventory control with generalized expediting. Working paper, Fort Lewis College, Durango, CO.

Lawson, D., E. Porteus. 1998. Multi-stage inventory management with expediting. *Oper. Res.* **48** 878–893.

Lee, H., K. So, C. Tang. 2000. The value of information sharing in a two-level supply chain. *Management Sci.* **46** 626–643.

Moinzadeh, K., P. Aggarwal. 1997. An information-based multi-echelon inventory system with emergency orders. *Oper. Res.* **45** 694–701.

Parlar, M., Z. K. Weng. 1997. Designing a firm's coordinated manufacturing and supply decisions with short product life cycles. *Management Sci.* **43** 1329–1344.

Porteus, E. 1990. Stochastic inventory theory. D. Heyman, M. Sobel, eds. *Stochastic Models*. Elsevier Science Publishers, New York.

Rosling, K. 2002. Inventory cost rate functions with nonlinear shortage costs. *Oper. Res.* **50** 1007–1017.

Smith, J. 2001. Willow Run third in the nation for cargo, but auto slowdown might dim outlook. *Detroit News* (January 21), http://detnews.com/2001/business/0101/21/e01-177947.htm.

Tagaras, G., D. Vlachos. 2001. A periodic review inventory system with emergency replenishments. *Management Sci.* **47** 415–429.

Tayur, S., R. Ganeshan, M. Magazine, eds. 1998. *Quantitative Models for Supply Chain Management*. Kluwer Academic Publishers, Boston, MA.

Veinott, A. 1966. On the optimality of (s, S) inventory policies: New conditions and a new proof. *SIAM J. Appl. Math.* **14** 1067–1083.

Zheng, Y. 1991. A simple proof for optimality of (s, S) policies in infinite-horizon inventory systems. *J. Appl. Probab.* **28** 802–810.

Accepted by Paul Glasserman; received March 4, 2002. This paper was with the authors 4 months for 1 revision.