Supply Chain Management with Overtime and Premium Freight

Eric Logan Huggins
hugginse@engin.umich.edu
Department of Industrial and Operations Engineering
University of Michigan
Ann Arbor, MI 48105

Tava Lennon Olsen
olsen@olin.wustl.edu
John M. Olin School of Business
Washington University in St. Louis
St. Louis, MO 63130

February, 2002
Abstract

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 either overtime production and/or premium freight shipments, both incurring per unit and setup costs. Overtime production occurs at the end of the period and incurs relatively high production costs; premium freight refers to building products at the beginning of the period they are needed and shipping them very quickly with relatively high shipping costs. Focusing primarily on the case where only one method of filling shortages is available, we 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 chain under centralized and decentralized control.
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 with either overtime production or with what we call ‘premium freight’. Premium freight consists of building parts the same day that they are required downstream and shipping them in an expedited fashion (e.g., by airplane or helicopter). In our problem, both methods of filling shortages are expensive, incurring 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 produces mostly engine parts used in vehicle assembly at one of the big three American automobile manufacturers. At PartCo, inventory levels are relatively low, yet they follow a policy of meeting all supply requests, frequently using overtime production and/or premium freight when shortages occur. Backordering is not considered an option because the parts they send downstream are essential in keeping the assembly lines moving, and the cost of shutting down the assembly lines at the automobile manufacturer is extremely high. We have heard a wide range of estimates for this cost, but all have been in 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 still 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 of 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 as well, where many manufacturers, notably Dell, 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 considering how and at what cost. Our results show that supply requests can always be met upstream using 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.

We have determined the optimal policies for this supply chain under centralized control when both options for filling shortages are available. However, for clarity of exposition, we will consider only one of the two mathematically equivalent options, overtime production, throughout most of this paper. The results are the same if we consider only premium freight. The structure of the slightly more complicated optimal policy for the combined model is included in Section 4 but the details and proof is left to Appendix 2.

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, replacing 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 examined 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, since 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 overtime production policy depends upon the problem data. In some cases, only one of the methods (either overtime production or premium freight) is used for all shortages. In other cases, there exists a threshold where premium freight is used for small shortages and overtime production is used for large shortages (or vice-versa). 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 forces the supplier to pay a possibly high fixed cost for either overtime production or premium freight. In the centralized case, we will show that the assembler is very sensitive to shortages and the optimal policy for the assembler will reflect this sensitivity.

In this paper, we consider a two-stage supply chain with two methods 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, Ganeshan, and Magazine [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 with setup costs at all stages. They establish lower bounds on the system costs under centralized control. Lawson and Porteus [1998] consider an \(m + 1\) stage supply chain where decisions must be made at each echelon about how many products to ship by regular means (which takes 1 period) and how many to expedite (which occurs instantaneously). Each expedited unit incurs a per unit cost higher than the per unit cost of regular shipping. They show that a “top-down” base-stock policy is optimal where the upstream managers ignore downstream decisions. 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. Although our paper is related to those just mentioned, our model has two methods of expediting, each incurring fixed costs, each with no lead time, and hence our results differ. Similarly, our paper is related to papers where outsourcing is used to meet shortages, but our results still differ. One reason is that we include a fixed cost when shortages occur, whereas most outsourcing assumptions only include a per unit cost. For example, in Gavirneni, Kapuscinski, and Tayur [1999], the cost of outsourcing is the same as it is from within the supply chain. Another difference is that we do not allow backorders at the upstream facility, which strongly affects our analysis. In Lee, So, and Tang [2000], the authors transform their outsourcing costs into backorder penalties, effectively allowing negative inventory levels upstream, even though supply requests are always met.

Finally, in our proof of the centralized model, we use results discussed by the following authors: Porteus [1990], Zheng [1991], Bertsekas [1995], and Rosling [1998]. In Zheng’s paper, he shows that \((s, S)\) policies are optimal given that the expected one period cost function is quasiconvex, which we show using a result from Porteus. We use this result to prove the optimal policy for the system inventory. The text by Bertsekas 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 discusses properties of logconcave functions.

The rest of this paper is organized as follows. In Section 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 Section 3. We prove that the optimal policies under the relaxed conditions are optimal for the original problem in Section 4. In Section 5, we conduct a numerical analysis
and discuss managerial insights. Section 6 concludes the paper and additional proofs are included in the appendices.

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 at stage 1 and at stage 2 each period. 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 run overtime production to meet the shortage and will incur high overtime production costs. If stage 1 orders everything stage 2 has on hand or less, stage 2 will avoid overtime production. The optimal policies for both stages will eventually reflect this relationship. Define the following variables:

\[ D_t = \text{the demand during period } t \]
\[ x_{1,t} = \text{the inventory level at stage 1 after demand has been experienced during period } t \]
\[ y_{1,t+1} = \text{the inventory position chosen for stage 1 for period } t + 1 \]
\[ x_{2,t} = \text{the inventory level at stage 2 before it experiences demand from stage 1 during period } t. \]
\[ y_{2,t+1} = \text{the inventory position chosen for stage 2 for period } t + 1. \]

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 produce \( (y_{1,t+1} - x_{1,t}) - x_{2,t} \) units with overtime production. Note that \( y_{1,t+1} \geq x_{1,t} \) and 
\[
\begin{align*}
y_{2,t+1} & \geq (y_{1,t+1} - x_{1,t}) - x_{2,t}^+.
\end{align*}
\]

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 \((c_2)\) and holding \((h_2)\), and overtime production incurs linear \((c_o)\) plus fixed \((K_o)\) costs (extensions of these to include premium freight are considered in Appendix 2). We also make the following assumptions about our model. First, the discount factor \( \alpha \) is assumed to be between 0 and 1. Second, we assume that demand is discrete, 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 we assume that demand is non-negative. Fourth, we assume that the per unit cost of overtime production 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:

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

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

(A3) \( p_d = 0 \) for \( d < 0 \).

(A4) \( c_o > c_2 \).

(A5) \( b_1 \geq c_o + \alpha((1 - \alpha)c_1 - c_2) \) 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 \( F(x) \), \( \frac{F(x+1) - F(x)}{F(x)} \), is nonincreasing in \( x \in \mathcal{I} \). This is not a restrictive assumption as most common discrete distributions are in fact logconcave. For example, the discrete uniform, the Poisson and the binomial are all in this
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 variables representing the system inventory. Next, we relax some of the constraints on the cost per period to get our relaxed cost per period, \( g_r(\cdot) \). Lastly, 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 it’s 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 \( g(\text{period } k \text{ variables}) \), which consists of all the costs incurred by the system during decision period \( k \). For a given policy \( \pi \), 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 \( \pi \) out of the set of all feasible policies \( \Pi \) 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, define our cost per period as:

\[
g_1(x_1, y_1, x_2, y_2, D) \equiv 
\alpha c_1(y_1 - x_1) + K_o \delta((y_1 - x_1) - x_2) + c_o((y_1 - x_1) - x_2)^+ 
+ h_2(x_2 - (y_1 - x_1))^+ + \alpha c_2(y_2 - (x_2 - (y_1 - x_1))^+) + h_1(y_1 - D)^+ + b_1(y_1 - D)^-
\]

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 overtime production 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 (analogously to Veinott [1966]), we get:

$$g_2(x_1, y_1, x_2, y_2, D) \equiv \alpha(1 - \alpha)c_1 y_1 + \alpha^2 c_1 D + K_o \delta((y_1 - x_1) - x_2) + c_o((y_1 - x_1) - x_2)^+ +$$

$$h_2(x_2 - (y_1 - x_1))^+ + \alpha c_2(y_2 - (x_2 - (y_1 - x_1))^+) + h_1(y_1 - D)^+ + b_1(y_1 - D)^-$$

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:

$$g_3(x_1, y_1, x_s, y_s, D) \equiv \alpha(1 - \alpha)c_1 y_1 + \alpha^2 c_1 D + K_o \delta(y_1 - x_s) + c_o(y_1 - x_s)^+ +$$

$$h_2(x_s - y_1)^+ + \alpha c_2(y_s - y_1 - (x_s - y_1)^+) + h_1(y_1 - D)^+ + b_1(y_1 - D)^- = \alpha((1 - \alpha)c_1 - c_2)y_1 + \alpha^2 c_1 D + h_1(y_1 - D)^+ + b_1(y_1 - D)^- + K_o \delta(y_1 - x_s) + c_o(y_1 - x_s)^+ + (h_2 - \alpha c_2)(x_s - y_1)^+ + \alpha c_2 y_s$$

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,

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

In order 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 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:

(R1) \( y_1 \geq x_1 \).

(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 Appendix 1. \( \square \)

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

\[
f_r^*(x_s) = \min_{y_1, y_s \geq x_s} \mathbb{E}_D [g_r(x_1, y_1, x_s, y_s, D) + \alpha f_r^*(y_s - D)]
\]

\[
= \min_{y_1, y_s \geq x_s} \{ \mathbb{E}_D[L_1(y_1, D)] + L_2(y_1, x_s) + \alpha c_2 y_s + \alpha \mathbb{E}_D[f_r^*(y_s - D)] \}
\]

Now notice that \( y_1 \) has no effect on either \( y_s \) or the cost to go, \( \alpha \mathbb{E}_D[f_r^*(y_s - D)] \). Thus,

\[
f_r^*(x_s) = \min_{y_s \geq x_s} \left\{ \min_{y_1} \{ \mathbb{E}_D[L_1(y_1, D)] + L_2(y_1, x_s) \} + \alpha c_2 y_s + \alpha \mathbb{E}_D[f_r^*(y_s - D)] \right\}
\]

\[
= \min_{y_s \geq x_s} \{ m(x_s) + \alpha c_2 y_s + \alpha \mathbb{E}_D[f_r^*(y_s - D)] \}
\]

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

$$N_H(y_1) = (\alpha(1 - \alpha)c_1 - h_2)y_1 + E_D[\alpha^2c_1D + h_1(y_1 - D)^+ + b_1(y_1 - D)^-]$$

and

$$N_L(y_1) = (\alpha((1 - \alpha)c_1 - c_2) + c_0)y_1 + E_D[\alpha^2c_1D + h_1(y_1 - D)^+ + b_1(y_1 - D)^-].$$

We now have that

$$m(x_s) = \min_{y_1} \{ E_D[L_1(y_1, D)] + L_2(y_1, x_s) \}$$

$$= \min_{y_1} \begin{cases} 
E_D[L_1(y_1, D)] + (h_2 - \alpha c_2)(x_s - y_1) & \text{if } y_1 \leq x_s \\
E_D[L_1(y_1, D)] + K_o + c_o(y_1 - x_s) & \text{if } y_1 > x_s
\end{cases}$$

$$= \min_{y_1} \begin{cases} 
(h_2 - \alpha c_2)x_s + (\alpha(1 - \alpha)c_1 - h_2)y_1 + \\
E_D[\alpha^2c_1D + h_1(y_1 - D)^+ + b_1(y_1 - D)^-] & \text{if } y_1 \leq x_s \\
K_o - c_ox_s + (\alpha((1 - \alpha)c_1 - c_2) + c_0)y_1 + \\
E_D[\alpha^2c_1D + h_1(y_1 - D)^+ + b_1(y_1 - D)^-] & \text{if } y_1 > x_s
\end{cases}$$

$$= \min_{y_1} \begin{cases} 
(h_2 - \alpha c_2)x_s + N_H(y_1) & \text{if } y_1 \leq x_s \\
K_o - c_ox_s + N_L(y_1) & \text{if } y_1 > x_s
\end{cases}$$

$$= \min \begin{cases} 
(h_2 - \alpha c_2)x_s + \min_{y_1 \leq x_s} \{N_H(y_1)\} \\
K_o - c_ox_s + \min_{y_1 > x_s} \{N_L(y_1)\}
\end{cases}$$

Before continuing out 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) \( 0 \leq y_L \leq y_H \leq \infty \).

Proof: The proof of (1) is straightforward. To prove (2), we examine the differential of both functions. See Appendix 1 for details.

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

\[
N(x_s) \equiv \alpha((1-\alpha)c_1-c_2)x_s + E_D[\alpha^2c_1D + h_1(x_s-D)^+ + b_1(x_s-D)^-],
\]

we have that

\[
m(x_s) = \min \begin{cases} (h_2 - \alpha c_2)x_s + \min_{y_1 \leq x_s} \{N_H(y_1)\} & \text{if } x_s \geq y_H \\ K_o - c_o x_s + \min_{y_1 > x_s} \{N_L(y_1)\} & \text{if } x_s < y_H \\ (h_2 - \alpha c_2)x_s + N_H(y_H) & \text{if } x_s \geq y_H \\ (h_2 - \alpha c_2)x_s + N_H(x_s) & \text{if } x_s < y_H \\ K_o - c_o x_s + N_L(x_s) & \text{if } x_s > y_L \\ K_o - c_o x_s + N_L(y_L) & \text{if } x_s \leq y_L \end{cases}
\]

\[
m(x_s) = \min \begin{cases} (h_2 - \alpha c_2)x_s + N_H(y_H) & \text{if } x_s \geq y_H \\ N(x_s) & \text{if } x_s < y_H \\ K_o + N(x_s) & \text{if } x_s > y_L \\ K_o - c_o x_s + N_L(y_L) & \text{if } x_s \leq y_L \end{cases}
\]

Define \( t_L \) as the smallest \( w \) such that \( N(w) \leq K_o - c_o w + N_L(y_L) \). We get that

\[
m(x_s) = \begin{cases} (h_2 - \alpha c_2)x_s + N_H(y_H) & \text{if } x_s \geq y_H \\ N(x_s) & \text{if } t_L \leq x_s < y_H \\ K_o - c_o x_s + N_L(y_L) & \text{if } x_s < t_L \end{cases}
\]

(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, \( x_s \geq y_H \), we order up to \( y_H \). If the system inventory is medium, \( t_L \leq x_s < y_H \), we use up the system inventory, \( x_s \). Finally, if 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

\[
y^*_1 = \begin{cases} 
  y_H & \text{if } x_s \geq y_H \\
  x_s & \text{if } t_L \leq x_s < y_H \\
  y_L & \text{if } x_s < t_L.
\end{cases}
\]  

(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

\[
f^*_r(x_s) = \min_{y_s \geq x_s} \{ m(x_s) + \alpha c_2 y_s + \alpha E_D[f^*_r(y_s - D)] \}.
\]

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

\[
f^*_{rm}(x_s) = \min_{y_s \geq x_s} \{ \alpha c_2 y_s + \alpha E_D[m(y_s - D)] + \alpha E_D[f^*_{rm}(y_s - D)] \}.
\]

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 a similar argument as Veinott [1966]. Second, to prove the existence of \( f^*_{rm}(\cdot) \), we must show that \( g(y_s) \equiv \alpha (c_2 y_s + m(y_s - D)) \geq 0 \) according to Proposition 1.1 of Bertsekas (1995, page 137). To prove \( g(y_s) \) is non-negative and to later prove that \( G(y_s) \) is quasiconvex, let us examine the function \( g^+(w) \equiv c_2 w + m(w) \). Graphically, the function looks as in Figure 1.

Starting from the left, \( g^+(\cdot) \) decreases at rate \(- (c_o - c_2)\) 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 $c_2w + N(w)$, decreasing at first, then increasing. From $y_H$ on, it increases at rate $h_2 + (1 - \alpha)c_2$.

![Graph of $g^+(w)$](image_url)

**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 Appendix 1.

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

**Proof:** Consider $g(y_s)$:

$$g(y_s) = \alpha(c_2y_s + m(y_s - D))$$
\[ \begin{align*}
\alpha(c_2y_s - c_2D + c_2D + m(y_s - D)) \\
\alpha(c_2D + g^+(y_s - D)) \geq 0
\end{align*} \]

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

\[ G(y_s) = \alpha E_D[c_2y_s + m(y_s - D)] \]
\[ = \alpha(c_2E_D[D] + E_D[g^+(y_s - D)]) \]

The first term is a constant, and the second term is a convolution of a quasiconvex function \( (g^+(y_s - D)) \) and a logconcave probability distribution by assumption (A2). Thus, according to Porteus [1990, page 619], \( G(\cdot) \) is a quasiconvex function. Since \( G(\cdot) \) is quasiconvex, the desired result follows from Zheng [1991].

\section{Optimal Policies for the Original Problem}

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

\[ f^*_r(x_s) = \min_{y_s \geq x_s} \{ m(x_s) + \alpha c_2y_s + \alpha E_D[f^*_r(y_s - D)] \} \]
\[ = \min_{y_s \geq x_s, y_1} E_D[g_r(x_1, y_1, x_s, y_s, D) + \alpha f^*_r(y_s - D)] \]

Recall that our fully constrained problem is

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

We must show that the optimal policies for the relaxed problem minimize the fully constrained problem and that both (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, \( y_H \).

\[ (A6) \ x_1 \leq y_H \]
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 when \( y_1 > x_s \) 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.

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

\[
y_2^* = \begin{cases} 
x_s & \text{if } x_s > S^* \\
S^* & \text{if } x_s \leq S^*,
\end{cases}
\]

and \( y_2^* = y_s^* - y_1^* \).

These policies hold when only one of the two options for filling shortages (either overtime production or premium freight) is available. When both options are available, the proof of optimality is more complicated and is included in Appendix 2. However, the system and
stage 2 policies remain the same and the stage 1 policy becomes

\[
y_1^* = \begin{cases} 
  y_H & \text{if } x_s \geq y_H \\
  x_s & \text{if } t_M \leq x_s < y_H \\
  y_M & \text{if } t_L \leq x_s < t_M \\
  y_L & \text{if } x_s < t_L.
\end{cases}
\]

where \( t_M \) is the threshold between underordering and using premium freight, \( t_L \) is now the threshold between premium freight and overtime production, and \( y_M \) is the order-up-to level given premium freight will be utilized.

## 5 Numerical Analysis and Managerial Insights

The optimal policies for the decentralized and centralized supply chains, as discussed in the introduction, are quite different. 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. A natural question that arises from this difference is how much can be saved by using the centralized optimal policy rather than the decentralized optimal policy? To answer this question, we wrote a C++ program and performed numerical analyses for various parameters and demand distributions.

### 5.1 Numerical Analysis

In our experiment, we set \( c_1 = 10 \) and \( \alpha = 0.95 \). (We initially varied \( \alpha \), but the outcomes were similar. We chose \( \alpha = 0.95 \) as a trade-off between realism and fast convergence times.) We varied the other parameters as follows. For stage 1, we let \( h_1 = 2, 4, 8 \) and \( b_1 = 5, 10, 20 \). For stage 2, we let \( c_2 = 3, 5, 9 \) and \( h_2 = 2, 4, 8 \). For overtime production, we let \( c_o = 4, 6, 10 \) and \( K_o = 0, 25, 100 \). These variations lead to a total of \( 3^6 = 729 \) possible combinations.
However, many 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/overtime 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/overtime costs for the system under centralized control. Finally, we calculated three statistics comparing the centralized case and the decentralized case: the percentage savings in total costs, the percentage savings in inventory and overtime costs, and the percentage reduction in system inventory. We averaged these savings over all feasible combinations.

We compared the results for four different demand distributions: Poisson(mean), Uniform(lower bound, upper bound), Normal(mean, standard deviation), and Exponential(mean). Since we consider discrete demand, we used discrete approximations for the last three distributions. Also, we truncated each distribution below at zero and above at forty-nine to fit into our probability array, and adjusted the probabilities appropriately to ensure the total probability was one. (Again, we chose a probability array of size fifty as big enough to distinguish different distributions, but small enough to converge quickly.) The average results for several demand distributions are in the table below.

<table>
<thead>
<tr>
<th>Demand Distribution</th>
<th>Total Savings</th>
<th>Inv/OT Savings</th>
<th>Inventory Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform(25,25)</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Normal(25,1)</td>
<td>1.1%</td>
<td>10.9%</td>
<td>1.6%</td>
</tr>
<tr>
<td>Normal(25,5)</td>
<td>3.9%</td>
<td>20.5%</td>
<td>6.3%</td>
</tr>
<tr>
<td>Poisson(25)</td>
<td>4.0%</td>
<td>20.8%</td>
<td>6.2%</td>
</tr>
<tr>
<td>Uniform(0,49)</td>
<td>7.9%</td>
<td>23.1%</td>
<td>15.4%</td>
</tr>
<tr>
<td>Exponential(15)</td>
<td>11.2%</td>
<td>24.7%</td>
<td>15.7%</td>
</tr>
</tbody>
</table>
We ran the experiment for constant demand (Uniform(25,25)) as one way to check the accuracy of our computer code. When the demand is stochastic, observe that the three statistics can be quite significant, particularly when demand is exponentially distributed. Also, observe that as demand variance increases, so do the savings and inventory reduction. We will discuss these observations further in Section 5.2. We feel that insight can be gained by examining a typical numerical example along with an atypical numerical example.

As a typical example, consider a problem that has Poisson demand with mean 25. The per unit costs at stage 1 are 10 for production, 4 for holding, and 20 for backorders. The per unit costs at stage 2 are 5 for production, 2 for holding, and 6 for overtime production; the fixed cost for overtime production is 100. The discount factor is 0.95.

Under centralized control, the optimal policy at stage 1 is to order up to 28 if the system inventory is 18 or less, use up the available system inventory if the system inventory is between 19 and 31, and order up to 32 if the system inventory is 32 or more. The optimal policy for the system is to order up to a base-stock level of 57 and the optimal policy at stage 2 is to order the difference between the system inventory and the inventory at stage 1. Under decentralized control, the optimal policy at stage 1 is to order up to a base-stock level of 29 and the optimal policy at stage 2 is to order up to a base-stock level of 33. Note that the decentralized system carries a total of 62 units.

So, under the centralized policy, the system carries 5 less units of inventory. This is an inventory reduction of 8.1%, which is one factor that contributes to cost savings. The other factor that contributes is how often stage 2 is forced to run overtime production. In the centralized case, stage 2 must run overtime when the system inventory after demand is less than the low threshold, or when $57 - D < 19$. The probability that stage 2 must run overtime equals the probability that $D \geq 39$, which is 0.0034. In the decentralized case, stage 2 must run overtime when demand is greater than 33; the probability that $D > 33$ is 0.034. So, for
this particular example, stage 2 must run overtime production ten times as often under the decentralized optimal policy! These two factors lead to a total cost savings of 3.2% and an inventory/overtime cost savings of 17.5%. These results are typical for the majority of our experimental outcomes.

Now, as an atypical example, consider a problem that has exponential demand with mean 15. The per unit costs at stage 1 are 10 for production, 8 for holding, and 10 for backorders. The per unit costs at stage 2 are 3 for production, 8 for holding, and 6 for overtime production; the fixed cost for overtime production is 25. The discount factor is 0.95.

Under centralized control, the optimal policy at stage 1 is to order up to 6 if the system inventory is -1 or less and to use up the available system inventory if the system inventory is 0 or greater. Stage 1 will never order up to it’s high inventory level of 48, because that value is greater than the optimal base-stock level for the system inventory, which is 21. Under decentralized control, the optimal policy at stage 1 is to order up to a base-stock level of 10 and the optimal policy at stage 2 is to order up to a base-stock level of 9. Note that the decentralized system carries a total of 19 units.

So, under the centralized policy, the system actually carries 2 more units of inventory than under the decentralized policy, or an inventory reduction of -10.5%. However, the total cost savings is 13.4% and the inventory/overtime cost savings is 23.0%. The reason for such significant savings despite the inventory increase is that under decentralized control, the probability that stage 2 uses overtime production is 0.50, so stage 2 incurs the fixed overtime cost half of the time! Under centralized control, this probability decreases to 0.20. Note again that this type of result is uncommon; an inventory increase for the centralized policy occurred in only 2% of our experimental outcomes.
5.2 Managerial Insights

As shown in Table 1, the centralized optimal policy generally affects significant savings over 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/overtime savings are typically over 20%. In particular, if the demand experienced by stage 1 is highly variable, following the centralized optimal policy seems worth the effort.

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” in order to save significant overtime production 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]. 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, the centralized optimal policy offers a way to reduce inventories and costs simultaneously. Not only that, but the centralized optimal policy generally reduces the likelihood of overtime production and/or premium freight, an outcome with which most managers would be 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 get a myopic problem for stage 1 that we solve for the optimal policy which depends on two thresholds and the system inventory. Next, we solve the optimality equation for the system and show that a base-stock policy is optimal. Finally, we show that the solutions for the relaxed problem solve the original, fully constrained problem, and hence we have found the overall optimal policies.

We performed an experiment for several different demand distributions and parameter values. The results of this experiment clearly indicate that the centralized optimal policy affects significant savings over the decentralized optimal policy. Numerical examples yield insight into where the savings occur. Our main managerial insights are that savings are significant, that to cut costs stage 1 must occasionally underorder, and that inventory levels can generally be reduced by following the centralized policy.

Some of our analyses and results are distinctive when compared to 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 (or three) 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 in order to create the two thresholds that determine inventory positions at both stages.

Acknowledgements

We would like to thank Brad Johnson and Michael Knox, our 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 1

Proof of Lemma 1

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

When $y_1 \leq x_s$,

$$g_r(x_1, y_1, x_s, y_s, D) \geq h_2(x_s - y_1) - \alpha c_2 x_s + \alpha c_2 y_s \geq \alpha c_2 (y_s - x_s) \geq 0.$$ 

The first inequality holds because we drop non-negative terms from equations (1) and (2) and cancel the $\alpha c_2 y_1$ terms. The second inequality holds since $y_s \geq x_s$.

When $y_1 > x_s$,

$$g_r(x_1, y_1, x_s, y_s, D) \geq -\alpha c_2 y_1 + c_o (y_1 - x_s) + \alpha c_2 y_s$$

$$= c_o (y_1 - x_s) - \alpha c_2 (y_1 - y_s)$$

$$\geq c_o (y_1 - x_s) - \alpha c_2 (y_1 - x_s)$$

$$= (c_o - \alpha c_2)(y_1 - x_s) \geq 0.$$ 

The first inequality holds because we drop non-negative terms from equations (1) and (2), the second inequality holds since $y_s \geq x_s$ and the third inequality holds since $c_o \geq \alpha c_2$ by assumptions (A1) and (A4).

Proof of Lemma 2, part (2)

Note that the middle inequality is satisfied because $-h_2 < 0 < c_o - \alpha c_2$ 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

\[ \Delta N_L(y_1) = N_L(y_1 + 1) - N_L(y_1) \]

\[ = (\alpha((1 - \alpha)c_1 - c_2) + c_o)(y_1 + 1) + E_D[\alpha^2c_1D + h_1(y_1 + 1 - D)^+ + b_1(y_1 + 1 - D)^-] - \]

\[ \alpha((1 - \alpha)c_1 - c_2) + c_o)y_1 - E_D[\alpha^2c_1D + h_1(y_1 - D)^+ + b_1(y_1 - D)^-] \]

\[ = (\alpha((1 - \alpha)c_1 - c_2) + c_o) + \]

\[ E_D[h_1(y_1 + 1 - D)^+ - h_1(y_1 - D)^+ + b_1(y_1 + 1 - D)^- - b_1(y_1 - D)^-] \]

\[ = (\alpha((1 - \alpha)c_1 - c_2) + c_o) + h_1F(y_1) - b_1(1 - F(y_1)) \]

\[ = \alpha((1 - \alpha)c_1 - c_2) + c_o - b_1 + (h_1 + b_1)F(y_1) \]

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, \( \Delta N_L(y_L) = \alpha((1 - \alpha)c_1 - c_2) + c_o - b_1 + (h_1 + b_1)F(y_L) \approx 0 \) 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, \( y_L \approx F^{-1}\left(\frac{b_1 - \alpha((1 - \alpha)c_1 - c_2) - c_o}{h_1 + b_1}\right) \) 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 is non-negative and that the second fraction is less than or equal to one. So, we require that \( b_1 - \alpha((1 - \alpha)c_1 - c_2) - c_o \geq 0 \) and \( b_1 + h_2 - \alpha(1 - \alpha)c_1 \leq h_1 + b_1 \) which both hold by assumption (A5). Under these conditions, we have that \( 0 \leq y_L \leq y_H \leq \infty \). \( \square \)

**Proof of Lemma 3**

To the left of \( t_L \), the slope of \( g^+(\cdot) \) is \(-(c_o - c_2) < 0\) and to the right of \( y_H - 1 \), the slope of \( g^+(\cdot) \) is \( h_2 + (1 - \alpha)c_2 > 0 \). Also note that \( g^+(t_L) \leq g^+(t_L - 1) \) 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 \( c_2w + N(w) \), a convex function, and thus there is exactly one minimum. The minimum value is positive since

\[ c_2w + N(w) = c_2w + \alpha((1 - \alpha)c_1 - c_2)w + E_D[\alpha^2c_1D + h_1(w - D)^+ + b_1(w - D)^-] \]

\[ \geq (1 - \alpha)(\alpha c_1 + c_2)w + E_D[b_1(w - D)^-] \geq 0. \]
The first inequality is true because we drop two non-negative terms. The second inequality depends on whether $w$ is negative or not.

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


Appendix 2 (for inclusion online only)

In this appendix, we study the problem where both overtime production and premium freight are viable options. We follow the previous proof but add additional comments or proofs when necessary. In this problem, the manager must actually make three decisions: how much to produce at stage 1, how much to produce during overtime at stage 2, and how much to produce during regular production at stage 2. Now, if stage 1 orders more than stage 2 has on hand, stage 2 must fill the shortage with some combination of overtime production and/or premium freight shipments. Define a new variable \( z_{2,t} \) = the amount produced with overtime production at stage 2 during time \( t \).

Again, the inventory decisions take place after demand is experienced at stage 1. The manager must decide the inventory position at stage 1, \( y_{1,t+1} \), the overtime production quantity at stage 2, \( z_{2,t} \), and the inventory position at stage 2, \( y_{2,t+1} \). As before, stage 2 begins overtime with \( x_{2,t} - (y_{1,t+1} - x_{1,t}) \) units on hand; stage 2 then produces \( z_{2,t} \) units during overtime, bringing the inventory level up to \( x_{2,t} - (y_{1,t+1} - x_{1,t}) + z_{2,t} \). If this quantity is negative, premium freight must be used to fill the remaining shortage incurring linear \( c_p \) and fixed \( K_p \) costs. Note that \( y_{1,t+1} \geq x_{1,t}, z_{2,t} \geq 0, \) and \( y_{2,t+1} \geq (x_{2,t} - (y_{1,t+1} - x_{1,t}) + z_{2,t})^+ \).

When both options are viable, we require three additional assumptions. First, we modify assumption (A5) slightly. Second, to simplify the analysis, we assume that the fixed cost for overtime production is not less than the discounted fixed cost for premium freight and that the per unit cost of overtime is not more than the discounted per unit cost of premium freight plus regular production at stage 2. Similar results hold when these inequalities are reversed. We label these assumptions as:

(A5) \( b_1 \geq \alpha((1 - \alpha)c_1 + c_p) \) and \( h_2 \leq h_1 + \alpha(1 - \alpha)c_1 \).

(A6) \( K_o \geq \alpha K_p \)

(A7) \( c_o \leq \alpha(c_p + c_2) \)

Here we develop similar cost functions to those of Section 2, using the same notation.
where possible.

\[ g_1(x_1, y_1, x_2, y_2, z_2, D) \equiv \alpha c_1(y_1 - x_1) + K_0 \delta(z_2) + c_o z_2 + h_2(x_2 - (y_1 - x_1) + z_2)^+ \\
+ \alpha K_p \delta((x_2 - (y_1 - x_1) + z_2)^-) + \alpha c_p(x_2 - (y_1 - x_1) + z_2)^-
\]

\[ + \alpha c_2(y_2 - (x_2 - (y_1 - x_1) + z_2)) + h_1(y_1 - D)^+ + b_1(y_1 - D)^- 
\]

with \( y_1 \geq x_1, z_2 \geq 0, \) and \( y_2 \geq (x_2 - (y_1 - x_2) + z_2)^+ \). Again, we move the \(-\alpha c_1 x_1\) term back to get:

\[ g_2(x_1, y_1, x_2, y_2, z_2, D) \equiv \alpha(1 - \alpha)c_1 y_1 + \alpha^2 c_1 D + K_0 \delta(z_2) + c_o z_2 + h_2(x_2 - (y_1 - x_1) + z_2)^+
\]

\[ + \alpha K_p \delta((x_2 - (y_1 - x_1) + z_2)^-) + \alpha c_p(x_2 - (y_1 - x_1) + z_2)^-
\]

\[ + \alpha c_2(y_2 - (x_2 - (y_1 - x_1) + z_2)) + h_1(y_1 - D)^+ + b_1(y_1 - D)^- 
\]

under the same restrictions. Substituting system variables, we get:

\[ g_3(x_1, y_1, x_s, y_s, z_2, D) \equiv \alpha(1 - \alpha)c_1 y_1 + \alpha^2 c_1 D + h_1(y_1 - D)^+ + b_1(y_1 - D)^- 
\]

\[ + K_0 \delta(z_2) + (c_o - \alpha c_2) z_2 + h_2(x_s - y_1 + z_2)^+ \quad (7) \]

\[ + \alpha K_p \delta((x_s - y_1 + z_2)^-) + \alpha c_p(x_s - y_1 + z_2)^- - \alpha c_2 x_s \quad (8) \]

\[ + \alpha c_2 y_s \]

with \( y_1 \geq x_1, z_2 \geq 0, \) and \( y_s \geq y_1 + (x_s - y_1 + z_2)^+ = \max\{y_1, x_s + z_2\} \). We can rewrite \( g_3(\cdot) \) as

\[ g_3(x_1, y_1, x_s, y_s, z_2, D) = L_1(y_1, D) + L_2(y_2, x_s, z_2) + \alpha c_2 y_s \]

where \( L_1(y_1, D) \) represents the terms on line \((7)\) and \( L_2(y_2, x_s, z_2) \) represents the terms on lines \((8)\) and \((9)\). At this point, our fully constrained optimal cost function is

\[ f^*(x_1, x_s) = \min_{y_s \geq \max\{y_1, x_s + z_2\}, z_2 \geq 0, y_1 \geq x_1} E_D [g_3(x_1, y_1, x_s, y_s, z_2, D) + \alpha f^*(y_1 - D, y_s - D)] . \]

Again, to solve this equation we relax some constraints. First, we drop the constraint that \( y_1 \geq x_1 \). Second, we relax the constraint on the system inventory position so that \( y_s \geq x_s \). For later reference, we label the relaxed assumptions as:
(R1) $y_1 \geq x_1$.

(R2) $y_s \geq \max\{y_1, x_s + z_2\} \rightarrow y_s \geq x_s$.

After relaxing the constraints, our cost per period becomes

$$g_r(x_1, y_1, x_s, y_s, z_2, D) \equiv L_1(y_1, D) + L_2(y_1, x_s, z_2) + \alpha c_2 y_s$$

with $z_2 \geq 0$ and $y_s \geq x_s$. The function $g_r(\cdot)$ can be shown to be non-negative by analysis similar to Lemma 1 and we can thus use the same result from Bertsekas (1995, page 137) for the optimal cost function $f^*_r$.

$$f^*_r(x_s) = \min_{y_s \geq x_s, z_2 \geq 0} E_D [g_r(x_1, y_1, x_s, y_s, z_2, D) + \alpha f^*_r(y_s - D)]$$

$$= \min_{y_s \geq x_s, z_2 \geq 0} \left\{E_D[L_1(y_1, D)] + L_2(y_1, x_s, z_2) + \alpha c_2 y_s + \alpha E_D[f^*_r(y_s - D)]\right\}$$

Now notice that $y_1$ and $z_2$ have no effect on either $y_s$ or the cost to go, $\alpha E_D[f^*_r(y_s - D)]$.

Thus,

$$f^*_r(x_s) = \min_{y_s \geq x_s} \left\{\min_{z_2 \geq 0} \left\{E_D[L_1(y_1, D)] + L_2(y_1, x_s, z_2)\right\} + \alpha c_2 y_s + \alpha E_D[f^*_r(y_s - D)]\right\}$$

$$= \min_{y_s \geq x_s} \left\{m(x_s) + \alpha c_2 y_s + \alpha E_D[f^*_r(y_s - D)]\right\}$$

where $m(x_s) = \min_{z_2 \geq 0} \{E_D[L_1(y_1, D)] + L_2(y_1, x_s, z_2)\}$. Again, finding the optimal inventory policy for stage 1 has become a myopic problem that depends only on $y_1$ and $z_2$.

Now consider $m(x_s)$ under two cases, when stage 1 does not order more than the system inventory on hand ($y_1 \leq x_s$) and when stage 1 does order more than the system inventory on hand ($y_1 > x_s$). In the first case, we get that

$$L_2(y_1, x_s, z_2) = K_o \delta(z_2) + (c_o - \alpha c_2)z_2 + h_2(x_s - y_1 + z_2) - \alpha c_2 x_s$$

$$= K_o \delta(z_2) + (h_2 + c_o - \alpha c_2)z_2 + h_2(x_s - y_1) - \alpha c_2 x_s$$

which is minimized when $z_2 = 0$ and thus $L_2(y_1, x_s, z_2) = h_2(x_s - y_1) - \alpha c_2 x_s$ when $y_1 \leq x_s$. 

31
In the second case, there are four options.

\[ L_2(y_1, x_s, z_2) = \begin{cases} 
\alpha K_p + \alpha c_p (y_1 - x_s) - \alpha c_2 x_s & \text{if } z_2 = 0 \\
K_o + \alpha K_p + (c_o - \alpha (c_p + c_2)) z_2 + \alpha c_p (y_1 - x_s) - \alpha c_2 x_s & \text{if } 0 < z_2 < y_1 - x_s \\
K_o + (c_o - \alpha c_2) (y_1 - x_s) - \alpha c_2 x_s & \text{if } z_2 = y_1 - x_s, \\
K_o + (h_2 + c_o - \alpha c_2) z_2 + h_2 (x_s - y_1) - \alpha c_2 x_s & \text{if } z_2 > y_1 - x_s. 
\end{cases} \]

It is easy to show that the third option is less expensive than both the second and fourth options using assumption (A7). Thus, either the first option or the third option are optimal.

Either \( z_2 = 0 \) (use premium freight) with \( L_2(y_1, x_s, z_2) = \alpha K_p + \alpha c_p (y_1 - x_s) - \alpha c_2 x_s \) or \( z_2 = y_1 - x_s \) (use overtime production up to 0) with \( L_2(y_1, x_s, z_2) = K_o + (c_o - \alpha c_2) (y_1 - x_s) - \alpha c_2 x_s \).

Before returning to \( m(x_s) \), define \( N_M(y_1) \) as

\[ N_M(y_1) = \alpha ((1 - \alpha) c_1 + c_p) y_1 + E_D [\alpha^2 c_1 D + h_1(y_1 - D)^+ + b_1(y_1 - D)^-]. \]

Similar to the results of Lemma 2, define \( y_M = \arg \min_{y_1} \{N_M(y_1)\} \). As before, \( N_M(y_1) \) is convex and we have that \( 0 \leq y_M \leq y_L \leq y_H < \infty \). \([0 \leq y_M \text{ by assumption (A5), } y_M \leq y_L \text{ by assumption (A7), } y_L \leq y_H \text{ by algebra, and } y_H < \infty \) as before.]

We have

\[
m(x_s) = \min_{z_2 \geq 0, y_1} \{E_D [L_1(y_1, D)] + L_2(y_1, x_s, z_2)\}
\]

\[
= \min \left\{ \begin{array}{ll}
\min_{y_1 \leq x_s} \{(h_2 - \alpha c_2) x_s + N_H(y_1)\} & \text{if } y_1 \leq x_s \\
\min_{y_1 > x_s} \{\alpha K_p - \alpha (c_p + c_2) x_s + N_M(y_1)\} & \text{if } y_1 > x_s \\
\min_{y_1 > x_s} \{K_o - c_o x_s + N_L(y_1)\} & \text{if } y_1 > x_s, \\
(h_2 - \alpha c_2) x_s + \min_{y_1 \leq x_s} \{N_H(y_1)\} & \\
\alpha K_p - \alpha (c_p + c_2) x_s + \min_{y_1 > x_s} \{N_M(y_1)\} & \\
K_o - c_o x_s + \min_{y_1 > x_s} \{N_L(y_1)\} & \end{array} \right\}
\]
\[
\begin{align*}
&= \min \left\{ \begin{array}{ll}
(h_2 - \alpha c_2)x_s + N_H(y_H) & \text{if } x_s \geq y_H \\
N(x_s) & \text{if } x_s < y_H \\
\alpha K_p + N(x_s) & \text{if } x_s > y_M \\
\alpha K_p - \alpha(c_p + c_2)x_s + N_M(y_M) & \text{if } x_s \leq y_M \\
K_0 + N(x_s) & \text{if } x_s > y_L \\
K_0 - c_0x_s + N_L(y_L) & \text{if } x_s \leq y_L
\end{array} \right. \\
&= \left\{ \begin{array}{ll}
(h_2 - \alpha c_2)x_s + N_H(y_H) & \text{if } x_s \geq y_H \\
N(x_s) & \text{if } t_M \leq x_s < y_H \\
\alpha K_p - \alpha(c_p + c_2)x_s + N_M(y_M) & \text{if } t_L \leq x_s < t_M \\
K_0 - c_0x_s + N_L(y_L) & \text{if } x_s < t_L
\end{array} \right.
\end{align*}
\]

where \( t_M \) is defined as the smallest \( w \) such that \( N(w) \leq \alpha K_p - \alpha(c_p + c_2)w + N_M(y_M) \) and \( t_L \) is redefined as

\[
t_L \equiv \left[ \frac{K_0 + N_L(y_L) - \alpha K_p - N_M(y_M)}{c_0 - \alpha(c_p + c_2)} \right].
\]

We have now defined \( m(x_s) \) explicitly and again determined the optimal inventory control policy at stage 1. The optimal policy for stage 1 is now

\[
y_1^* = \left\{ \begin{array}{ll}
y_H & \text{if } x_s \geq y_H \\
x_s & \text{if } t_M \leq x_s < y_H \\
y_M & \text{if } t_L \leq x_s < t_M \\
y_L & \text{if } x_s < t_L.
\end{array} \right. 
\quad (12)
\]

From equation (11), we have that

\[
f^*_r(x_s) = \min_{y_s \geq x_s} \{ m(x_s) + \alpha c_2y_s + \alpha E_D[f^*_r(y_s - D)] \}
\]

As before, we move the \( m(x_s) \) term back to get

\[
f^*_{rm}(x_s) = \min_{y_s \geq x_s} \{ G(y_s) + \alpha E_D[f^*_{rm}(y_s - D)] \}
\]

where \( G(y_s) = \alpha c_2y_s + \alpha E_D[m(y_s - D)] \). We justify this step as before. The only difference is that now the \( g^+(\cdot) \) function has an extra kink on the left. From the left, \( g^+(\cdot) \) starts
out decreasing at rate \(-(c_o - c_2)\) up to point \(t_L\); next, the function decreases at the steeper rate \(-\alpha(c_p + c_2) - c_2\) up to the point \(t_M\); after this point, the function behaves as before. Analogous results to Lemma 3 hold, \(G(\cdot)\) is a quasiconvex function, and the optimal system inventory control policy is a base-stock policy.

We have just shown that the optimal policies for the relaxed problem minimize the fully constrained problem, assuming the original constraints are met. The first relaxed constraint (R1) is satisfied by similar analysis as before. To show that the second relaxed constraint (R2) is satisfied, note that in the relaxed optimal solution, \(z_2 = 0\) or \(z_2 = y_1 - x_s\). Thus, the original constraint \(y_s \geq \max\{y_1, x_s + z_2\}\) is equivalent to \(y_s \geq \max\{y_1, x_s\}\) under optimality and similar analysis as before yields that (R2) is satisfied. Thus, we have solved our original problem when both methods of resupply are available.