THE ECONOMIC LOT AND DELIVERY SCHEDULING PROBLEM: MODELS FOR NESTED SCHEDULES

Juho Hahm
Department of Industrial Engineering
Seoul National University
Seoul, Korea

Candace Arai Yano
Department of
Industrial & Operations Engineering
University of Michigan
Ann Arbor, MI 48109-2117

Technical Report 91-13

April 1991
Revised November 1991
and
May 1992

THE ECONOMIC LOT AND DELIVERY SCHEDULING PROBLEM:
MODELS FOR NESTED SCHEDULES

Abstract

We investigate the problem of simultaneously determining schedules for the production of several assembly components at a captive supplier and delivery of those components to the customer. We consider situations in which production economies of scale in the form of setup costs and/or setup times make it desirable for the supplier to produce in batches that are larger than the desired order quantity of the customer. The objective is to minimize the average cost per unit time of transportation, inventory at both the customer and the supplier, and where applicable, setup costs.

We develop a heuristic solution procedure and a lower bounding approach for this problem. We also report experimental results which indicate that the heuristic provides solutions close to the lower bound in most instances. Our results provide a means to answer the often-asked question of whether just-in-time suppliers are (or should be) asked to hold inventory for their customers, and the question of how much setup costs and setup times need to be reduced so that the suppliers no longer need to hold that inventory.

THE ECONOMIC LOT AND DELIVERY SCHEDULING PROBLEM: MODELS FOR NESTED SCHEDULES

1. INTRODUCTION

Recent surveys of suppliers delivering parts on a "just-in-time" (JIT) basis have indicated that a majority of such suppliers believe that inventory reductions at their customers are being achieved by forcing the suppliers to hold the inventory instead. (See, for example, the annual automotive supplier survey in Ward's Auto World, July 1990.) Of course, these statistics may represent perceptions rather than fact, but they do raise the question of exactly how production and deliveries should be scheduled. We investigate the interface between a customer and one of its JIT suppliers and develop models that will provide insight into the appropriate degree of coordination, and the appropriate levels of inventory at each location.

In this paper, we focus on situations where a captive supplier produces several products for a single customer and makes direct deliveries to that customer. Such situations are not uncommon in the automobile industry, and they occur in other industries as well. Our model incorporates scheduling of the final manufacturing stage at the supplier, which, for simplicity, is assumed to consist of one machine or production line. Since the final manufacturing step at the supplier is the one most closely linked with the delivery process, we considered it most important to include in the model. We implicitly assume that the upstream manufacturing stages can supply input materials at the required rate.

The customer (or "assembly facility") uses the components at a nearly constant rate. This may be due to higher-level production smoothing, or various characteristics of the demand or manufacturing processes. For example, in the automotive industry applications that motivated this research, the customer is a final assembly facility that

uses a fixed-pace assembly line. To facilitate both workload smoothing and just-in-time arrangements, the assembly line is scheduled so as to smooth the usage rate of each component. Consequently, components required for every vehicle are used at a (perfectly) constant rate, and other components have usage rates with only small fluctuations. (See Schonberger (1983) for a discussion of the role of production smoothing within a just-in-time context.) Considering a situation with constant demand also allows us to gain insight into major tradeoffs that are less visible in models with time-varying demand.

The production line at the supplier can produce only one component at a time, and changeovers between components incur a setup cost and/or setup time. Since we are concerned with whether inventories are being "pushed back" to the supplier, we consider situations where the economic production quantities at the supplier are larger than the supplier's desired delivery quantities. If this were not true, the suppliers would not "complain" about holding inventory for the customer.

In this paper, we develop a procedure to simultaneously find a delivery interval and a common production cycle (interval) for all components which is an integer multiple of the delivery interval. Thus, during the production cycle, each component is produced once, while one or more deliveries are scheduled during the cycle. In a sequel (Hahm and Yano 1991b), we allow multiple production batches per cycle. The objective is to minimize the average cost per unit time of transportation, inventory at both the supplier and the customer, and setup costs at the supplier. In the next section, we provide a formal statement of the problem. The subsequent section contains a formulation. This is followed by development of a procedure to find a lower bound for a given production sequence. We then discuss sequencing issues, and develop a heuristic procedure. We report experimental results which indicate that the heuristic performs quite well. Finally, we conclude with a summary and a discussion of the implications of the model.

2. PROBLEM STATEMENT

We analyze a system with a single supplier at which the final production stage involves producing multiple components on a single machine or production line. The components are delivered directly to an assembly facility which uses these components at a constant rate. The supplier incurs a sequence-independent setup cost and/or setup time each time the production line is changed over from one component to another. On the other hand, setup costs and times for the assembly facility are assumed to be negligible, which is reflective of many assembly environments.

We assume that there is a fixed charge for each delivery. Later in the paper, we explain how this assumption can be relaxed to incorporate a fixed charge per truck. Deliveries to the customer occur at equal intervals (duration to be determined). The delivery lead time is assumed to be deterministic, and without loss of generality, equal to zero. (Incorporating in-transit inventory is straightforward.)

The supplier produces one batch of each component in each production interval (i.e., there is a common production cycle for all components). During the production interval, there are M (integer) equally spaced deliveries to the customer. In each delivery, the shipment quantities are exactly equal to the customer's requirements until the subsequent delivery arrives. Since the deliveries are equally spaced in time and the demand rates are constant, each shipment has the same composition. Inventory holding costs are charged on time-weighted average inventory levels, and the inventory holding cost of a component is assumed to be the same at the supplier as it is at the assembly facility. (Relaxation of this assumption is also straightforward.)

The objective is to minimize the average cost per unit time of setups, inventory, and transportation while ensuring that both demand and the supplier's capacity constraint are satisfied. We must decide the delivery and setup (or production) intervals and the detailed production schedule for the components.

A considerable amount of research has been done on multi-stage production systems with known constant demands. Nevertheless, little research has been done that considers both the cost of inventory accumulation prior to delivery and the cost of transportation in the determination of jointly optimal production and delivery schedules. One reason why the former has been ignored is that many models assume instantaneous production (e.g., Crowston et al. 1973, Blackburn and Millen 1982, Roundy 1985, Maxwell and Muckstadt 1985). Other papers incorporate capacity constraints but ignore some or all of the accumulation inventory (e.g., Caie and Maxwell 1981, Billington et al. 1983, Jackson et al. 1988), or treat transportation costs as fixed (Maxwell and Muckstadt 1981). Likewise, there is some research on procurement policies that include transportation costs, either as quantity discounts (e.g., Lee 1986) or fixed charge per shipment (e.g, Lippman 1971 and Lee 1989), but none explicitly considers the impact of the selected delivery schedule on the *supplier's* inventory levels. We are not aware of any papers that treat all of the issues in our problem. In the interest of brevity, we will review only those models most closely related to ours. The reader is referred to Schussel [1968], Taha and Skeith [1970], Jensen and Kahn [1972], Schwarz and Schrage [1975], Graves and Schwarz [1977], Bigham and Mogg [1979], Szendrovits [1981], Williams [1982], Blackburn and Millen [1982], and Moily [1986], among others, for a variety of results on continuous-time, multi-stage lot-sizing models.

Our problem might be viewed as a generalization of the "product cycling" problem first posed by Hanssmann (1962). The product cycling problem involves the scheduling of several items, each with a constant demand rate, that are produced on one machine. Each item is produced once in each cycle and the cycle is repeated. Setup times and costs may be incurred in changeovers between items. The objective is to minimize average setup and inventory costs per unit time. Hanssmann shows that a closed form solution exists for this problem. The product cycling problem is a special case of the well-known economic lot scheduling problem (ELSP). The ELSP allows multiple

production runs of each item within a cycle, whereas the product cycling problem allows only one. A review of the ELSP literature appears in Elmaghraby (1978), and more recent references can be found in Dobson (1987) and Zipkin (1991).

Our problem differs from Hanssmann's in several ways. First, usage of components occurs at a location other than where the components are produced. Thus, the supplier experiences periodic spikes in demand (at delivery times), and consequently, inventory of the components must be accumulated prior to each delivery.

Second, because of this inventory accumulation between deliveries, both (i) the sequence in which the components are produced and (ii) the exact timing of the "start" of the sequence relative to the times at which deliveries are made can affect the solution, even when setup times and costs are sequence-independent. If the production cost per unit time is similar across components, the production sequence between deliveries may not matter in practice, since the rate at which value is being added would be similar for all sequences. On the other hand, even with careful scheduling, the cost of holding the inventories accumulated prior to delivery may be greater than the cost of holding inventories at the assembly facility. The magnitude of the accumulation inventories can be controlled by appropriate choices of the delivery interval and the detailed production schedule.

A third difference between our problem and the product cycling problem is that solutions for the latter typically include idle time only when setup costs are the dominant factor. In our problem, idle time may occur even when setup costs do not drive the production interval decision. By scheduling idle time, it may be possible to reduce accumulation inventories substantially.

A detailed review of articles that explicitly consider both transportation and inventory costs in a continuous-time setting appears in Hahm and Yano (1991a). These articles include models of single-item problems (Burns et al. 1985, Benjamin

5

1989, Hahm and Yano 1991c) and a multi-item model in which each item is shipped to a different customer (Blumenfeld et. al. 1991).

Hahm and Yano (1991a) model a problem similar to ours in which several components are produced on a single machine for a single customer by a captive supplier. They consider "just-in-time" schedules where one batch of each component is produced and these units are combined for delivery to the customer. Thus, the time between deliveries is equal to the duration of the overall production cycle. They develop an iterative heuristic procedure in which an initial production sequence is determined ignoring setup times. The corresponding optimal production (and delivery) cycle is then determined. An updated production sequence is found using this production cycle and incorporating setup times. The procedure iterates between updating the production sequence and the production cycle. The procedure is shown to converge, and to perform well in computational tests. They also present an associated lower bounding procedure.

Our model generalizes the Hahm and Yano (1991a) paper to allow for multiple deliveries within a production cycle. This generalization creates the need to partition the components into groups and to sequence the groups of components, as we will explain in more detail later.

3. FORMULATION

Notation

J : number of components.

D_j : demand for component j per unit time.

A: transportation cost per delivery.

S_j : setup cost for component j.

p_j : production time per unit of component j.

T : production interval (time between setups).

R: delivery interval (time between deliveries).

6

M: number of deliveries per production interval = (T/R).

h_j : inventory holding cost per unit of component j per unit time.

s_j : setup time for component j.

For the convenience of the reader, a list of these and other notation, along with their definitions, appears in Appendix 1.

We first develop the capacity constraint. During the production interval (T), each component is produced once. Thus we must have

$$\sum_{j=1}^{J}(s_j + p_jD_jT) \leq T ,$$

which is equivalent to $\qquad\qquad T \geq \tau$

where $\quad \tau = \dfrac{\sum\limits_{j}s_j}{1-\sum\limits_{j}p_jD_j}$.

Two terms in the objective function are evident: the average setup cost per unit time is $\dfrac{\sum\limits_{j}S_j}{T}$, and the average delivery cost per unit time is $\dfrac{A}{R}$. The inventory holding costs are somewhat more complicated. Inventory holding costs are incurred at both the supplier and the assembly facility. In our model, the average inventory of component j at the assembly facility is $\dfrac{1}{2}D_jR$ and the average inventory at the supplier is

$$\frac{1}{2}D_jT(1-p_jD_j) + p_jD_j^2R - \frac{1}{2}D_jR$$

if we ignore the time during which a component must wait for other components to be produced. Consider a component that does not have to wait for other components to be produced before delivery takes place. The system-wide inventory of this component is the same as in the standard economic production quantity (EPQ) model, except for the additional inventory that must accumulate to ensure availability of delivery quantities.

The first term in the above expression is simply the average inventory in the EPQ model. The second term reflects the fact that prior to *each* delivery, a quantity $D_j R$ must be available $p_j R$ time units earlier than would have been required in the EPQ model. Since this extra inventory must be incurred in each delivery interval, the additional inventory per unit time is $p_j D_j^2 R$. The last term captures the portion of the inventory that is present at the assembly facility rather than at the supplier.

With multiple components, any component j that is not produced immediately before a delivery occurs must wait an additional period of time equal to the duration of production of other components produced after j but prior to the *first* delivery of component j from the production batch in question. We define the *earliness* of component j, e_j, as the time between the start of the production of a batch and the *first* delivery of components from that batch. With this definition, we can re-express the total inventory holding cost per unit time (at the two locations) as the sum of the average inventory cost per unit time without consideration of earliness, $T \sum_j \alpha_j$, where $\alpha_j = \frac{1}{2} D_j h_j (1 - p_j D_j)$, and the average inventory cost per unit time caused by earliness, $\sum_j D_j h_j e_j$. (Note that the latter term includes the cost associated with the "accumulation" inventory, $\sum_j h_j (p_j D_j^2 R) = \sum_j D_j h_j (p_j D_j R)$.)

Although T and M are decision variables, to clarify earliness in the context of our problem, let us first analyze the scheduling problem with T and M given. We need to determine (i) which delivery point to assign to each component as the basis for its earliness and (ii) the sequence of components assigned to a common delivery point. We will refer to the set of components assigned to the same delivery point as a "group" and the procedure to form groups as "grouping."
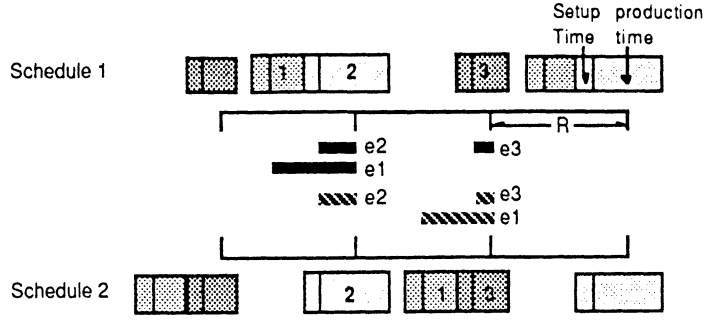
Figure 1. Effect of schedules (sequence and grouping) on earliness.

Figure 1 illustrates an example that shows the effects of scheduling on earliness. In this example, $T = 2R$ and $J = 3$. Two different schedules are compared in the figure. In each schedule, there are two groups: {1,2} and {3} in schedule 1 and {1,3} and {2} in schedule 2. Also, the sequences used in the two schedules differ. Components 2 and 3 have the same earliness in both schedules, but the earliness of component 1 differs between the two schedules.

We now present a formulation for simultaneous grouping and scheduling:

(P)

$$\text{Minimize} \quad \frac{1}{T} \sum_{j=1}^{J} S_j + T \sum_{j=1}^{J} \alpha_j + \sum_{j=1}^{J} D_j h_j e_j + \frac{A}{R} \tag{1}$$

$$\text{subject to} \quad T = M \cdot R \tag{2}$$

$$T \geq \tau \tag{3}$$

$$e_j \geq p_j D_j R \qquad \text{for } j=1 \text{ to } J \tag{4}$$

$$b_{[k]} + s_{[k]} + p_{[k]} D_{[k]} T \leq b_{[k+1]} \qquad \text{for } k=1 \text{ to } J \tag{5}$$

$$b_{[k]} + s_{[k]} + e_{[k]} = g_{[k]} R \qquad \text{for } k=1 \text{ to } J \tag{6}$$

$$b_{[J+1]} = b_{[1]} + T \tag{7}$$

$$b_{[k]} \geq 0 \qquad \text{for } k=1 \text{ to } J \tag{8}$$

9

$$g_{[k]} \geq 0 \text{ and integer} \qquad \text{for } k=1 \text{ to } J \qquad (9)$$

$$M \geq 1, \text{ integer} \qquad (10)$$

where $[k]$ is the index of the component which is produced in the k-th position in the sequence, b_j is the time when the setup for component j begins, g_j is the group index of component j, and g_jR is the delivery time to which component j is assigned (for computing earliness). Constraints (4) ensure that the delivery quantity of each component is available at each delivery time and constraints (5) ensure that only one component is produced at a time. Constraints (6) capture the assignment of components to delivery points. Constraint (7) ensures that the production interval has duration T. Constraints (8), (9) and (10) are nonnegativity and integrality constraints. In this formulation, constraint (3) is redundant because constraints (5) and (7) together imply constraint (3).

Conceptually, the problem could be solved as a dynamic program *if T and M were given*. One would proceed backward in time, starting at the end of the production interval. Since T is given, the setup plus processing time for each component is known. The state is represented by the starting time of the component with the earliest starting time (recall that we are scheduling backward in time), and the set of components that still need to be scheduled. The decision at each point is which component to schedule and whether it should be the last component scheduled in the previous (earlier) delivery interval, potentially requiring the scheduling of idle time after it. Since the total idle time is fixed for a given T, some assignments and some choices of idle times are not feasible and can be eliminated from consideration. Since all sequences are feasible, this problem is very difficult if J is large. Indeed, unless two or more components have identical characteristics, the solution of the dynamic program would involve complete enumeration in most instances. Since we also must determine M and T, such a

dynamic progam would need to be solved many times if the sequencing subproblem were solved in this fashion. For these reasons, we do not use this dynamic program in our solution procedure.

Problem P with M and T as decision variables is a mixed-integer nonlinear program and convexity of the feasible region is not guaranteed. Therefore, direct solution of P does not seem practical. We suggest a hierarchical solution procedure, which disaggregates the above formulation into several manageable sub-problems. First, however, we describe a lower bounding procedure that also provides some insight into good values of R.

4. LOWER BOUNDS

For the moment, let us suppose that the sequence (q) is given and that the components are reindexed so that component j is produced in the j-th position in the sequence. Then, the problem can be formulated as:

(P1)

Minimize $\quad \frac{1}{T} \sum_{j=1}^{J} S_j + T \sum_{j=1}^{J} \alpha_j + \sum_{j=1}^{J} D_j h_j e_j + \frac{A}{R}$

subject to $\quad e_j \geq p_j D_j R$ $\qquad$ for $j=1$ to J $\qquad$ (11)

$\qquad\qquad b_j + s_j + e_j = g_j R$ $\qquad$ for $j=1$ to J $\qquad$ (12)

$\qquad\qquad b_j + s_j + p_j D_j T \leq b_{j+1}$ $\qquad$ for $j=1$ to J $\qquad$ (13)

$\qquad\qquad b_{J+1} = b_1 + T$ $\qquad\qquad$ (14)

$\qquad\qquad T = M \cdot R$ $\qquad\qquad$ (15)

$\qquad\qquad T \geq \tau$ $\qquad\qquad$ (16)

11

$$b_j \geq 0 \qquad\qquad \text{for } j=1 \text{ to } J \qquad\qquad (17)$$

$$g_j \geq 0 \text{ , integer} \qquad\qquad \text{for } j=1 \text{ to } J \qquad\qquad (18)$$

$$M \geq 1, \text{ integer} \qquad\qquad (19)$$

Let $TC(M, R, q)$ be the optimal objective value of P1 when the values of M, R and the sequence (q) are given. Suppose we have the corresponding solution. Components with the same g_j value are in the same group. (Some groups may be empty.) Let us now examine the earliness of components in an isolated group k ($1 \leq k \leq M$). Let n_k denote the number of components in group k. Also, let the superscript k index the group to which the component belongs. Suppose the components in this group are reindexed so that component j is produced in the j-th position in the sequence within that group. Then, we can characterize the earliness of the components in group k as follows:

$$e^k_{n_k} \geq p^k_{n_k} D^k_{n_k} R, \qquad\qquad (20)$$

$$e^k_{n_k\text{-}1} = e^k_{n_k} + s^k_{n_k} + p^k_{n_k\text{-}1} D^k_{n_k\text{-}1} T, \qquad\qquad (21)$$

$$\ldots$$

and $\qquad e^k_1 = e^k_2 + s^k_2 + p^k_1 D^k_1 T. \qquad\qquad (22)$

In general, we can express e^k_j, $j=1$ to n_k -1, as follows:

$$e^k_j = e^k_{j+1} + s^k_{j+1} + p^k_j D^k_j T. \qquad\qquad (23)$$

We now derive lower bounds on e^k_j. Let us define

$$\theta^k_{n_k} = p^k_{n_k} D^k_{n_k} R$$

and $\qquad \theta^k_j = \theta^k_{j+1} + p^k_j D^k_j R \qquad \text{for } j = 1 \text{ to } n_k \text{ -1}. \qquad\qquad (24)$

Then, because $s^k_j \geq 0$ and $T \geq R$, it is obvious that

$$e_j^k \geq \theta_j^k \quad \text{for } j = 1 \text{ to } n_k \quad \text{and all } k. \tag{25}$$

As a result,

$$\sum_{j=1}^{J} D_j h_j e_j = \sum_{k=1}^{M} \sum_{j=1}^{n_k} D_j^k h_j^k e_j^k \geq \sum_{k=1}^{M} \sum_{j=1}^{n_k} D_j^k h_j^k \theta_j^k \tag{26}$$

for any sequence and grouping. Note that the expression on the right hand side of the inequality in (26) provides a lower bound by ignoring setup times (essentially assuming that they are zero) and underestimating the production times by using R rather than T in the production time expressions.

The value of $\sum_{k=1}^{M} \sum_{j=1}^{n_k} D_j^k h_j^k \theta_j^k$ can be interpreted as the total weighted earliness

(considering time-weighted average inventory) for an identical processor scheduling problem where (1) the components in a group are assigned to the same machine, (2) all components are shipped at the end of the delivery interval, (3) setup times are zero, and (4) the production quantity of component j is $D_j R$ for all j. For this related scheduling problem, the processing time of component j is $p_j D_j R$ and the weight for component j is

$D_j h_j$. Since the grouping is constrained, $\sum_{k=1}^{M} \sum_{j=1}^{n_k} D_j^k h_j^k \theta_j^k$ is greater than or equal to the

total weighted earliness of the best schedule in which J components are assigned to M machines with no constraints on grouping. (With appropriate transformations, the problem can be posed as a multiple-machine total weighted flow time problem and related results can be used directly. Refer to Baker (1974) for results on this version of the problem.) Let $C(M, R)$ be the total weighted earliness of the best unconstrained schedule for this related scheduling problem. Then, the following theorem holds.

Theorem 1.

$$C(M, R) \geq \max\{ \beta R, \frac{1}{M} [Z(q^*)R + \frac{M-1}{2} \beta R] \}$$ (27)

where $\beta = \sum_{j=1}^{J} p_j D_j^2 h_j$,

$$Z(q^*) = \sum_{i=1}^{J} D_{[i]} h_{[i]} \sum_{j=i}^{J} p_{[j]} D_{[j]} ,$$ (28)

and q^* is the sequence in which components are sorted in non-increasing order of $\frac{p_j}{h_j}$.

Proof. See Appendix 2.

In Theorem 1, βR is the total weighted earliness of the schedule in which each of the J components is assigned a separate machine and $Z(q^*)R$ is that of the schedule in which all components are assigned to one machine. Therefore, βR and $Z(q^*)R$ represent lower and upper bounds on $C(M, R)$, respectively, and $\frac{1}{M} [Z(q^*)R + \frac{M-1}{2} \beta R]$ is a weighted sum of these two values.

Let $E(M, R, q)$ be the optimal value of $\sum_{j=1}^{J} D_j h_j e_j$ in P1 when the values of M and R, and the sequence (q) are given. Then, by Theorem 1 and (26), for any sequence q,

$$E(M, R, q) \geq \max\{ \beta R, \frac{R}{M} [Z(q^*) + \frac{M-1}{2} \beta] \} .$$ (29)

As a result,

$$TC(M, R, q) \geq \frac{1}{MR} \sum_{j=1}^{J} S_j + MR \sum_{j=1}^{J} a_j + \max\{ \beta R, \frac{R}{M} [Z(q^*) + \frac{M-1}{2} \beta] \} + \frac{A}{R} .$$ (30)

For simplicity, let $LB(M, R)$ represent the right-hand-side of the above inequality. Then it is obvious that $LB(M, R)$ does not depend on the sequence. Therefore, if we define $TC(M, R)$ as $\min_q \{TC(M, R, q)\}$, we have

$$TC(M, R) \geq LB(M, R). \tag{31}$$

We now determine a lower bound on $LB(M,R)$. Consider the following problem:

(P2) minimize $LB(M, R)$

      subject to $MR \geq \tau$

               $R \geq 0$

               $M > 0$, integer

Let M^* and R^* be optimal values of M and R, respectively, for P2. Then, $LB(M^*, R^*)$ is

a lower bound on $LB(M, R)$ under the constraints of P2, and therefore also on $TC(M, R)$.

We will use M^* and R^* as the initial values of M and R. As we will discuss later, these

initial values yield a reasonably good solution.

$LB(M^*, R^*)$ is calculated using the sequence q^*. However, it does not necessarily

mean that q^* will yield the minimum $TC(M, R)$. Therefore, we need to find out what

characterizes a good sequence.

5. SEQUENCING ISSUES AND A HEURISTIC PROCEDURE

As mentioned earlier, the earliness of the components in group k can be expressed

as in (20) through (22). If the total setup and processing time for each group is less than

or equal to R, then $e_{n_k}^k = p_{n_k}^k D_{n_k}^k R$ (i.e., the minimum possible) for all k. If not, the

value of $e_{n_k}^k$ may depend on the schedules of the other groups. However, once the value of

$e_{n_k}^k$ is determined, the earliness of each of the other components in the same group is

determined so as to satisfy (21) and (22). It is easy to show that if we know which

component produced last in a group, the other components in that group should be

arranged in non-increasing order of $\frac{1}{h_j D_j}(s_j + p_j D_j T)$. (For fixed T, this is a single-

machine weighted earliness problem. By making appropriate transformations, results

15

in Baker (1974) for the single-machine weighted flow time problem can be applied directly.)

Since a value of $e_{n_k}^k$ greater than $p_{n_k}^k D_{n_k}^k R$ has an adverse impact on the earliness of all the components in the group, it might be advantageous to equalize the total setup and production time across groups so as to minimize the interactions among the groups. If this is not possible, it would be desirable to alternate groups with long and short production times. This has two beneficial effects. First, it helps to eliminate the "domino effect" that would occur if groups with long production times were placed sequentially. Second, it helps to reduce unnecessary idle time in the schedule, which may permit a shorter production interval (and therefore less inventory). These considerations lead us to the following sequencing algorithm.

Algorithm SEQ

Step 1. Let $t_j = s_j + p_j D_j T$ and $w_j = h_j D_j$ for all j.

Let M be the number of groups.

Step 2. Select the unassigned component with the largest t_j and assign it to the group whose total production time is the shortest thus far.

Repeat until all components are assigned.

Let PT_k be the total production time of group k (i.e., $PT_k = \sum\limits_{\substack{\text{component } j \\ \text{in group } k}} t_j$), k=1 to M.

Step 3. Arrange the components in each group in non-increasing order of $\dfrac{t_j}{w_j}$.

Step 4. Arrange the groups in increasing order of PT_k , k=1 to M.

Let $\{k\}$ be the k-th group in the sequence.

Sequence the groups in following order: $\{1\}$, $\{M\}$, $\{2\}$, $\{M$-$1\}$, $\{3\}$, $\{M$-$2\}$,... . •

Step 2 in algorithm SEQ tends to minimize the maximum PT_k and thereby also attempts to equalize the PT_ks. Step 3 determines the sequence within each group and Step 4 determines the sequence of the groups. The resulting sequence usually is not optimal for a given M and R. Because of the combinatorial complexity of the problem, however, it is usually impractical to find the optimal sequence. In practice, the number of components produced on one production line typically falls in the range from a few to several dozen depending upon the application. At the low end of this spectrum, optimal solutions can be obtained by complete enumeration. In the upper portion of this range, however, the problems of partitioning and sequencing become difficult because the number of deliveries per cycle (M) tends to increase with the number of components, and the number of possible partitions increases dramatically with M. We believe that the procedure described above is a practical and logical means to find a good sequence. Another alternative would be to find the optimal sequence once the final values of M and R are determined.

We now have a way to obtain initial values of M and R and a procedure to construct a sequence given M and R. We now examine how to improve the initial solution. If the production sequence and the values of M and R were given, we would be able to solve problem P1 to find the best grouping using a variant of the dynamic program mentioned earlier. (The only decisions would be which component should be last in the cycle and when to start a new group.) The grouping resulting from solving this problem may differ from the original grouping. If so, we can adjust the sequence of the components in each of the new groups to achieve a further reduction in the inventory cost. However, we want to do this without affecting the schedules of the other groups. One way to do this while ensuring an improvement in the objective function is to allow changes in the sequence within a group with the exception of the components in the last position of each group. By retaining the identity of the groups, we prevent them from

affecting the schedule of the preceding or succeeding groups. The reason for retaining the position of the last component in each group is to ensure that we maintain $e_{n_k}^k = p_{n_k}^k D_{n_k}^k R$. If another component is placed in the last position in the group, the value of $e_{n_k}^k$ may increase and we then cannot guarantee an improvement. With these restrictions, it is optimal to rearrange the remaining components in each group in non-increasing order of $\frac{1}{h_j D_j}(s_j + p_j D_j T)$, as mentioned earlier. Of course, it would be possible to consider retaining the groups and evaluate all components in each group as a candidate for the last position in the group, but preliminary results indicated that this would not provide much additional benefit.

This adjustment will result in a new sequence. However, the value of g_j for each component will remain the same, as will the value of M. At this point, we could take the new sequence and the associated g_j values as given, and improve the solution further by allowing R and the e_js to change. Let $\alpha = \sum_{j=1}^{J} \alpha_j$ and $S = \sum_{j=1}^{J} S_j$. The sum of terms related to R in the objective function, $\frac{S}{MR} + \alpha MR + \frac{A}{R}$, is convex in R. We now consider the e_js and define the related problem P3 as follows.

(P3) minimize $\sum_{j=1}^{J} D_j h_j e_j$

subject to (11) through (17).

Given the values of M and the g_js, P3 can be reformulated as a parametric right-hand-side linear program where the parameter is R. The parametric right-hand-side linear minimization problem is well-known to be convex in the parameter. Therefore, the objective function in P3 is convex in R and the problem can be solved optimally. Let $f(R)$ be the objective function value of P3 for a given R. Also let $F(R) = f(R) + \alpha MR$. Then, the problem is to find the value of R which minimizes $F(R) + \frac{S}{MR} + \frac{A}{R}$. The function $F(R)$ is easily verified to be convex in R, and $\frac{S}{MR} + \frac{A}{R}$ is convex decreasing

18

in R. Therefore, we can restrict our attention to values of $R \geq r$, where r

$= \underset{R}{\text{argmin}} \{F(R)\}$. Note that the value of r is easily found by linear programming.

Discussions thus far suggest the following algorithm to solve problem P when initial values of M and R are given.

Algorithm BESTR

Step 1. Apply algorithm SEQ to provide an initial sequence q.

Step 2. Solve P1 using a mixed integer programming technique. The resulting objective function value is $TC(M, R, q)$.

Step 3. Arrange all except the last component in each group in non-increasing order of $\frac{1}{h_j D_j}(s_j + p_j D_j T)$, retaining the position of the last component in each group.

Step 4. Find $r = \underset{R}{\text{argmin}} \{F(R)\}$ using a LP technique.

Step 5. For $R \geq r$, find the value of R which minimizes $F(R) + \dfrac{S}{MR} + \dfrac{A}{R}$ using a parametric LP technique. •

Step 3 in algorithm BESTR changes the initial sequence and Step 4 and 5 changes the value of R. These steps are designed to produce a further cost reduction. Let $TC^B(M, R)$ be the total cost resulting from algorithm BESTR for a given M and initial R. Then, $TC^B(M, R)$ will be less than or, at most, equal to $TC(M, R, q)$ in Step 2.

Algorithm BESTR solves problem (P) when values of M and R are given. We will use M^* and R^* as initial values in algorithm BESTR. However, in the special the case where $M = 1$, we use the specialized algorithm developed by Hahm and Yano (1991a).

Before we describe the entire algorithm, let us define some notation and introduce

a preliminary result. For any positive integer M, let

$$R(M) = \underset{R \geq 0}{\text{argmin}}\{ LB(M, R) \mid MR \geq \tau \} .$$

Lemma 1.

$LB(M, R(M))$ is unimodal in M .

Proof. See Appendix 3.

Suppose M^* is 1. Then, we can use the algorithm developed by Hahm and Yano

(1991a) for the special case of $M = 1$ (henceforth referred to as the HY Algorithm). Let us

refer to the associated lower bound as $LB1$. Then, the new lower bound for the case of M

$= 1$ is max$\{LB1, LB(1, R(1))\}$. Furthermore, if $LB1$ is greater than $LB(2, R(2))$, the lower

bound for problem (P) no longer occurs at $M=1$ but occurs at $M=2$ by Lemma 1. (Figure 2

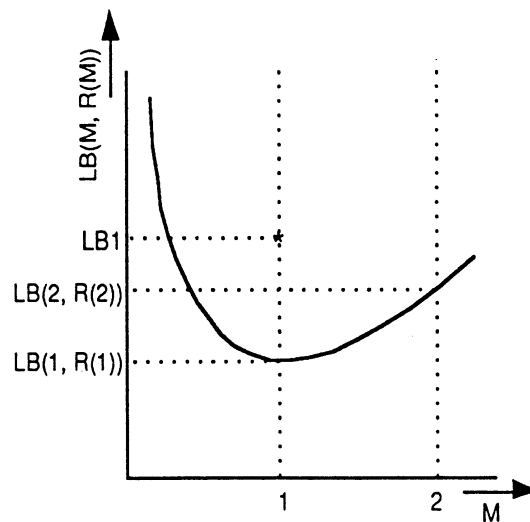illustrates an example.) In this case, we will reset M^* to 2.

Figure 2. Lower Bound for M = 1 and 2.

We now present the entire algorithm.

Heuristic Algorithm

Step 1. Find M^* and R^*.

Step 2. If $M^* \geq 2$, go to Step 5.

Using the HY algorithm, find the total cost and the lower bound and let us refer to these as TC^* and LB^*, respectively. Also let M^* be 1 and R^* be the resulting value of R.

Step 3. If $LB^* \leq LB(2, R(2))$, then stop. •

Set LB^* to $LB(2, R(2))$ and find the value of $TC^B(2, R(2))$ using algorithm BESTR.

Step 4. If $TC^* \leq TC(2, R(2))$, then stop. •

Set M^* to 2, R^* to the value of R resulting from algorithm BESTR in Step 3, and TC^* to $TC^B(2, R(2))$. Stop. •

Step 5. Using algorithm BESTR, find the value of $TC^B(M^*, R^*)$.

Set R^* to the value of R resulting from algorithm BESTR and TC^* to $TC^B(M^*, R^*)$.

Stop. •

A flow chart of the algorithm is represented in Figure 3. Experimental results are discussed in the next section.
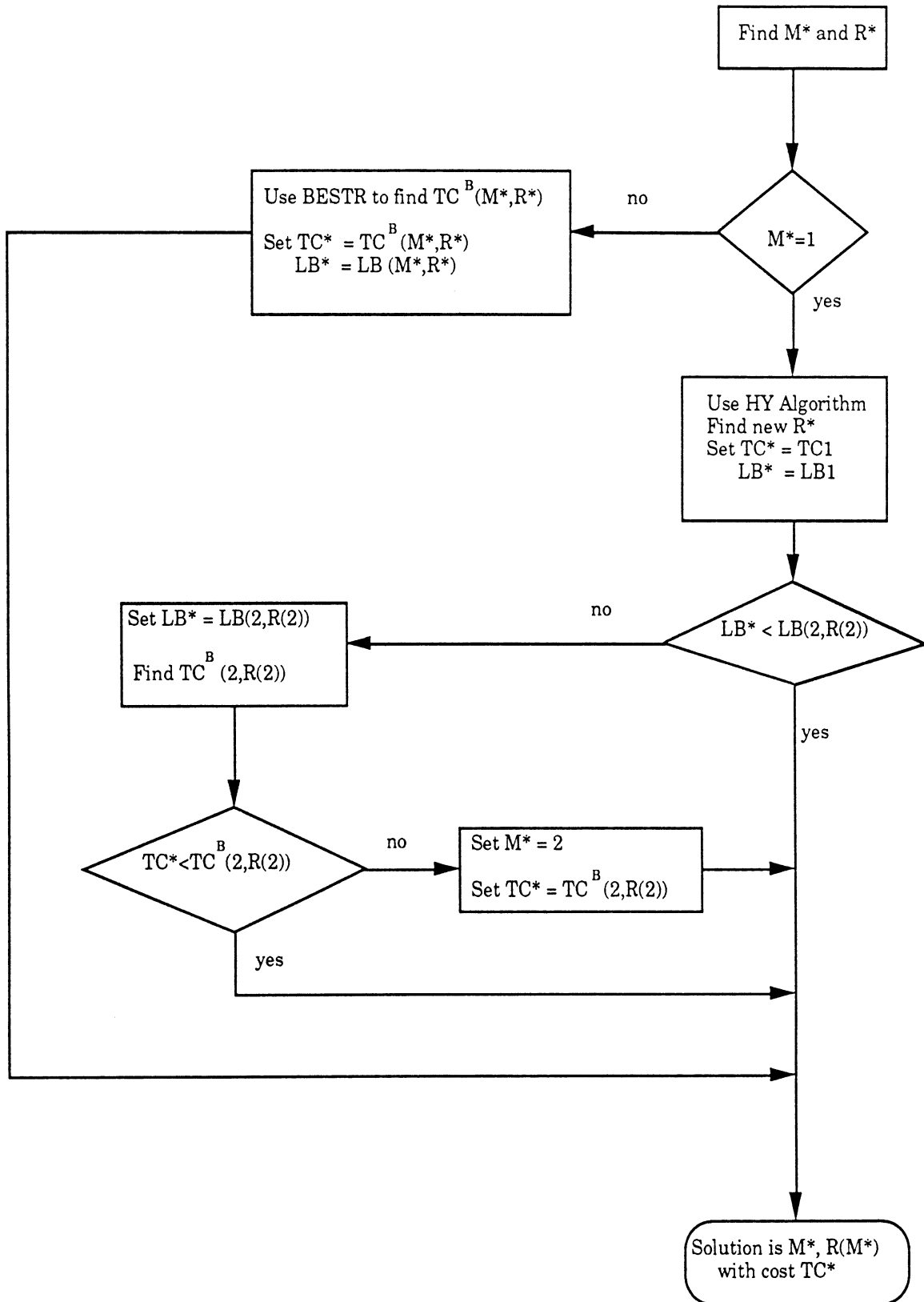
Figure 3. Flow Chart of the Heuristic Algorithm.

6. EXPERIMENTAL RESULTS

We tested the heuristic algorithm on a set of 72 problems which was designed to evaluate the impact of various problem parameters on the performance of the algorithm. Three major factors were considered. The first factor is the tightness of the capacity constraint. We consider two levels of tightness: $\sum_j p_j D_j$ drawn from a uniform distribution on [0.3,0.5] (not tight) and from [0.7,0.9] (tight).

The second factor is the spread of the natural production intervals among the components. Since our problem requires $T \geq R$, we determined constrained natural production and delivery intervals by solving the following problem, which is a variation of the original problem in which the accumulation inventories are ignored:

Minimize $\quad \sum_{j=1}^{J} \dfrac{S_j}{T_j} + \sum_{j=1}^{J} \alpha_j T_j + bR + \dfrac{A}{R}$

subject to $\quad \sum_{j=1}^{J} \dfrac{s_j}{T_j} \leq 1 - \sum_{j=1}^{J} p_j D_j \qquad\qquad\qquad (32)$

$$T_j \geq R \qquad\qquad\qquad \text{for } j = 1 \text{ to } J \qquad (33)$$

where T_j is the production interval of component j, and R is the delivery interval. Let T_j' and R_j' be the optimal values of T_j and R_j for this problem, which we refer to as the constrained natural production intervals and delivery interval, respectively. These values can be expressed as:

$$T_j' = \sqrt{\dfrac{S_j + \lambda_o s_j}{\alpha_j - \lambda_j}}, \quad 1 \leq j \leq J \qquad (34)$$

and $\qquad\qquad\qquad R' = \sqrt{\dfrac{A}{\beta + \sum_{j=1}^{J} \lambda_j}} \qquad\qquad\qquad (35)$

where λ_o is chosen so that constraint (32) is satisfied as an equality, or $\lambda_o = 0$ if the constraint is not binding; and for each j, λ_j is chosen so that constraint (33) is satisfied

as an equality or $\lambda_j = 0$ if it is not binding. We measure the spread of the constrained natural production intervals using the ratio of the maximum to the minimum T_j'. This ratio is generated from a uniform distribution on [1.5,2.5] for half of the problems (low variance) and on [4,8] for the other half (high variance).

The third factor is the magnitude of the *unconstrained optimal* delivery interval (i.e, the optimal delivery interval considering only the assembly facility's inventory costs and the transportation cost) relative to the natural production intervals. We say that the optimal unconstrained delivery interval is relatively small if it is less than or equal to the minimum T_j', medium if it is strictly between the minimum T_j' and maximum T_j', and large if it is greater than or equal to the maximum T_j'. In the first case, we might expect T^* to be constrained by the minimum T_j'; in the second case, we might expect T^* to be close to the unconstrained optimal R; and in the third case, economic tradeoffs may dictate the solution.

The two levels of capacity tightness, two levels of natural production interval spreads, and three levels of unconstrained optimal delivery intervals give twelve (2x2x3) combinations of levels. For each combination, we generated a set of six problems: two problems with three components, another two with six components, and two with nine components.

Summary statistics appear in Table 1. The average deviation of the cost of the solutions from their respective lower bounds is 5.93%. The lower bound that we use here is $LB(M^*,R^*)$ as described earlier. The maximum error is 16.8%, and we were able to confirm that 9 of the 72 solutions are optimal (solution equal to lower bound). The maximum error occurs in a problem with 9 components and a medium unconstrained optimal delivery interval. For this problem, the value of M^* is 2. In this case, only 2 delivery points are shared by 9 components and therefore, earliness costs are significant. Note that earliness of a component is incurred during the production times as well as the setup times of the subsequent components. The lower bound does not

account for setup times at all, and includes only the effects of R/T (= $1/M$) of the processing times. Therefore, if setup times are fairly large, $M \gg 1$, and/or many components share a small number of delivery points, the lower bound may be quite loose. Nearly all of the problems with large gaps between the cost of the heuristic solution and the lower bound had these characteristics. Thus, the larger gaps appear to reflect the poor quality of the lower bounds in these instances rather than large deviations from optimality of the heuristic solutions.

<div align="center">Table 1</div>

The practical benefits of the heuristic are more dramatic. We considered a variety of alternative benchmarks. One apparently simple alternative is to use independent solutions for the production and delivery intervals. Unfortunately, adapting the independent solutions to construct a feasible solution is not easy. If the independently-determined value of T is not an integer multiple of the independently-determined R, it would be possible to construct a schedule in which the overall cycle duration is equal to the least common multiple of T and R. One could, instead, choose M to be close to T/R and then choose T and R accordingly. Unfortunately, both of these approaches ignore the grouping and sequencing issues, and consequently do not specify a complete solution. Furthermore, even for an arbitary grouping and sequencing decision, sufficient inventory must be interjected into the system to eliminate shortages, and determining how much inventory to interject is not a simple matter. Moreover, with arbitrary grouping and sequencing, we might expect the aggregate level of earliness to be quite large. For these reasons, we decided to use a different benchmark obtained from a constrained version of the problem.

Our benchmark involves setting $M = 1$ and finding the constrained optimal solution. This solution might be viewed as the best "just-in-time" schedule given current system costs and parameters, and the difference between the cost for $M = 1$ and

<div align="center">25</div>

M unconstrained can be viewed as the cost of using just-in-time when just-in-time is not optimal. We report the ratio of the cost of the heuristic solution to the cost of this "just-in-time" schedule in Table 1. If the ratio is 1.0, then $M = 1$ is optimal. From the table, it is clear that this happens infrequently in this problem set. The average penalty from using the best just-in-time schedule versus the heuristic is over 20%. Thus, it may be desirable in many instances to allow production quantities to be larger than the delivery quantities until setup costs and times are reduced.

7. CONCLUSIONS

We have developed a heuristic procedure and a lower bounding approach for the problem of simultaneously scheduling the final production stage of a captive supplier and deliveries from that supplier to the customer. Experimental results indicate that the heuristic performs well in an absolute sense. The results also show that a "just-in-time" schedule with equal production and delivery quantities may not necessarily be optimal.

It is important to point out that the procedure can be adapted to a "fixed-cost-per-truck" transportation cost structure with minor modifications. It is easy to show that with such a cost structure, the optimal solution has a delivery cycle less than or equal to the time supply of one truckload of goods, i.e., there is one truck per shipment. The unimodality and convexity results presented here imply that the optimal value of the delivery cycle (for a given production sequence) is either the unconstrained optimal value (if it is less than the time supply of one truckload), or exactly equal to the time supply of one truckload.

The heuristic can be used in a variety of ways. First, it can provide first-cut guidance on the frequency of deliveries and the relative frequency of production. Second, it can be used to study the system-wide effects of setup cost and setup time reductions, and the effects of transportation cost reductions that might be achieved through co-location or a change of suppliers. Finally, solutions from the heuristic can be

26

used to obtain estimates of the benefits from better supplier-customer coordination through the joint determination of production and delivery schedules.

Acknowledgement

REFERENCES

BAKER, K. R., 1974, Introduction to Sequencing and Scheduling, John Wiley and Sons, New York, New York, pp. 20-22.

BENJAMIN, J., 1989, "An Analysis of Inventory and Transportation Costs in a Constrained Network," *Transportation Science*, Vol. 23, 177-183.

BIGHAM, P.E. AND MOGG, J.M., 1979, "Converging Branch Multi-Stage Production Schedules with Finite Production Rates and Start-up Delays," *J. of Opl. Res. Soc.*, Vol. 30, 737-745.

BILLINGTON, P., MCCLAIN, J.O., AND THOMAS, L.J., 1983, "Mathematical Programming Approaches to Capacity-Constrained MRP Systems: Review, Formulation and Problem Reduction," *Management Science*, Vol. 29, 1126-1141.

BLUMENFELD, D.E., BURNS, L.D. AND DAGANZO, C.F., 1991, "Synchronizing Production and Transportation Schedules," *Transportation Research*, Vol. 25B, 23-37.

BLACKBURN, J. D. AND MILLEN, R. A., 1982, "Improved Heuristics for Multi-Stage Requirements Planning Systems," *Management Science*, Vol. 28, 44-56.

BURNS, L.D., HALL, R.W., BLUMENFELD, D.E., AND DAGANZO, C.F., 1985, "Distribution Strategies that Minimize Transportation and Inventory Costs," *Operations Research*, Vol. 33, 469-490.

CAIE, J.P. AND MAXWELL, W.L., "Hierarchical Machine Load Planning," in Multilevel Production/Inventory Control Systems: Theory and Practice, L.B. Schwarz (Ed.), North Holland, New York, 1981.

CROWSTON, W. B., WAGNER, H. M. AND WILLIAMS, J.F., 1973, "Economic Lot Size Determination in Multi-Stage Assembly Systems," *Management Science*, Vol. 19, 517-527.

DOBSON, G., 1987, "The Economic Lot-Scheduling Problem: Achieving Feasibility using Time-Varying Lot Sizes," *Operations Research*, Vol. 35, 764 -771.

EASTMAN, S., EVEN, S. AND ISAACS, I., 1964, "Bounds for the Optimal Scheduling of n Jobs on m Processors," *Management Science*, Vol. 11,

ELMAGHRABY, S. E., 1978, "The Economic Lot Scheduling Problem(ELSP): Review and Extensions.," *Management Science*, Vol. 24, 587-598.

GRAVES, S.C. AND SCHWARZ, L.B., 1977, "Single Cycle Continuous Review Policies for Arborescent Production/Inventory Systems," *Management Science*, Vol. 23, 529-540.

HAHM, J. AND YANO, C.A., 1991a, "Economic Lot and Delivery Scheduling Problem: The Common Cycle Case," Working Paper, Department of Industrial & Operations Engineering, University of Michigan, Ann Arbor, MI (to appear in *IIE Transactions*).

28

HAHM, J. AND YANO, C.A., 1991b, "Economic Lot and Delivery Scheduling Problem: Powers of Two Policies," Working Paper, Department of Industrial & Operations Engineering, University of Michigan, Ann Arbor, MI.

HAHM, J. AND YANO, C.A., 1991c, "The Economic Lot and Delivery Scheduling Problem: The Single Item Case," *Journal of Manufacturing and Operations Management* (to appear).

HANSSMANN, F., 1962, Operations Research in Production and Inventory Control, John Wiley & Sons, New York. 158-160.

JACKSON, P.L., MAXWELL, W., AND MUCKSTADT, J.A., 1988, "Determining Optimal Reorder Intervals in Capacitated Production-Distribution Systems," *Management Science*, Vol. 34, No. 8, 938-958.

JENSEN P. A. AND KHAN, H. A., 1972, "Scheduling in a Multistage Production System with Set-up and Inventory Costs," *AIIE Transactions*, Vol. 4, 126-133.

LEE, C.Y., 1986, "The Economic Order Quantity for Freight Discounts," *IIE Transactions*, Vol. 18, 318-320.

LEE, C.Y., 1989, "A Solution to the Multiple Setup Problem with Dynamic Demand," *IIE Transactions*, Vol. 21, 266-270.

LIPPMAN, S.A., 1971, "Economic Order Quantities and Multiple Setup Costs," *Management Science*, Vol. 18, 39-47.

MAXWELL, W. AND MUCKSTADT, J.A. 1981, "Coordination of Production Schedules and Shipping Schedules, " in Multilevel Production/Inventory Control Systems: Theory and Practice, L.B. Schwarz (Ed.), North Holland, New York, 1981.

MAXWELL, W. AND MUCKSTADT, J.A., 1985, "Establishing Consistent and Realistic Reorder Intervals in Production-Distribution Systems," *Operations Research*, Vol. 33, No. 6, 1316-1341.

MOILY, J.P., 1986, "Optimal and Heuristic Procedures for Component Lot-Splitting in Multi-Stage Manufacturing Systems," *Management Science*, Vol. 32, 113-125.

ROUNDY, R., "A 98%-Effective Lot-Sizing Rule for a Multi-Product, Multi-Stage Production/Inventory System," *Math. of OR*, Vol. 11, No. 4, 699-727.

SCHONBERGER, R., 1983, Japanese Manufacturing Techniques, The Free Press, New York, New York.

SCHUSSEL, G., 1968, "Job-Shop Lot Release Sizes," *Management Science*, Vol. 14, B449-B472.

SCHWARZ, L.G. AND SCHRAGE, L., 1975, "Optimal and System Myopic Policies for Multi-Echelon Production/Inventory Assembly Systems," *Management Science*, Vol. 21, 1285-1294.

SZENDROVITS, A., 1981, "Comments on the Optimality in 'Optimal and System Myopic Policies for Multi-Echelon Production/Inventory Assembly Systems'," *Management Science*, Vol. 27, 1081-1087.

TAHA, H. A. AND SKEITH R.W., 1970, "The Economic Lot Sizes in Multistage Production Systems," *IIE Transactions*, Vol. 2, 157-162.

WILLIAMS, J. F., 1982, "On the Optimality of Integer Lot Size Ratios in 'Economic Lot Size Determination in Multi-Stage Systems,'" *Management Science*, Vol.28, 1341-1349.

ZIPKIN, P., 1991, "Computing Optimal Lot Sizes in the Economic Lot Scheduling Problem," *Operations Research*, Vol. 39, 56-63.

"1990 Supplier Survey," *Ward's Auto World*, Vol. 26, No. 7, 37-39, July 1990.

Table 1. Summary of Experimental Results

J : Number of Components.
C: 1 if capacity is not tight and 2 if capacity is tight.
D: 1 if low spread of natural setup cycles, 2 if high.
V: 1 if unconstrained del. interval is small, 2 if medium, 3 if large.

TC: Total Cost from Heuristic.
LB: Lower Bound.
JIT: Cost of Solution with
   M = 1

J	C	D	V	TC / LB	TC / JIT	J	C	D	V	TC / LB	TC / JIT
3	1	1	1	101.69%	109.25%	3	1	2	1	100.00%	112.38%
3	1	1	1	102.56%	103.96%	3	1	2	1	100.00%	110.94%
3	1	1	2	102.81%	100.67%	3	1	2	2	100.00%	100.00%
3	1	1	2	102.11%	104.06%	3	1	2	2	100.00%	100.00%
3	1	1	3	102.55%	101.41%	3	1	2	3	100.00%	100.00%
3	1	1	3	102.69%	102.16%	3	1	2	3	100.00%	100.00%
3	2	1	1	103.17%	146.55%	3	2	2	1	101.39%	138.61%
3	2	1	1	101.47%	136.70%	3	2	2	1	103.59%	132.84%
3	2	1	2	113.80%	121.47%	3	2	2	2	100.19%	100.00%
3	2	1	2	114.72%	119.45%	3	2	2	2	103.26%	103.59%
3	2	1	3	102.80%	100.00%	3	2	2	3	100.00%	100.00%
3	2	1	3	111.72%	107.25%	3	2	2	3	102.35%	100.00%
6	1	1	1	106.20%	125.38%	6	1	2	1	101.46%	121.88%
6	1	1	1	105.89%	130.16%	6	1	2	1	101.07%	118.01%
6	1	1	2	104.84%	106.80%	6	1	2	2	106.44%	107.17%
6	1	1	2	106.47%	105.75%	6	1	2	2	104.58%	104.40%
6	1	1	3	105.59%	106.22%	6	1	2	3	100.00%	100.00%
6	1	1	3	105.78%	106.78%	6	1	2	3	100.00%	100.00%
6	2	1	1	106.08%	169.99%	6	2	2	1	102.28%	129.10%
6	2	1	1	112.42%	152.22%	6	2	2	1	107.69%	143.96%
6	2	1	2	110.10%	138.69%	6	2	2	2	105.66%	108.16%
6	2	1	2	109.55%	148.24%	6	2	2	2	107.37%	107.43%
6	2	1	3	105.90%	109.91%	6	2	2	3	109.22%	108.33%
6	2	1	3	112.14%	119.61%	6	2	2	3	106.28%	100.12%
9	1	1	1	108.86%	127.60%	9	1	2	1	105.91%	122.02%
9	1	1	1	108.58%	125.32%	9	1	2	1	101.90%	118.81%
9	1	1	2	110.42%	114.89%	9	1	2	2	106.88%	106.60%
9	1	1	2	116.80%	118.09%	9	1	2	2	106.33%	107.11%
9	1	1	3	111.69%	111.19%	9	1	2	3	106.56%	102.23%
9	1	1	3	110.79%	107.14%	9	1	2	3	102.74%	100.00%
9	2	1	1	106.08%	182.58%	9	2	2	1	108.36%	134.07%
9	2	1	1	109.48%	158.09%	9	2	2	1	105.34%	166.48%
9	2	1	2	107.68%	171.91%	9	2	2	2	113.80%	137.68%
9	2	1	2	110.04%	136.28%	9	2	2	2	109.78%	165.94%
9	2	1	3	109.11%	125.12%	9	2	2	3	111.65%	107.64%
9	2	1	3	110.43%	148.25%	9	2	2	3	112.18%	132.81%

APPENDIX 1

LIST OF NOTATION

(in order of presentation)

T :   production interval (time between setups).

J :   number of components.

s_j :   setup time for component j.

p_j :   production time per unit of component j.

D_j :   demand for component j per unit time.

τ :   minimum production interval $= \dfrac{\sum\limits_{j} s_j}{1 - \sum\limits_{j} p_j D_j}$.

S_j :   setup cost for component j.

A: transportation cost per delivery.

R: delivery interval (time between deliveries).

e_j: earliness of component j (time between the start of the production of a batch and the *first* delivery of components from that batch.

h_j :   inventory holding cost per unit of component j per unit time.

α_j : $\frac{1}{2} D_j h_j (1 - p_j D_j)$.

M: number of deliveries per production interval $= (T/R)$.

$[k]$:   index of the component which is produced in the k-th position in the sequence.

b_j: time at which the setup for component j begins.

g_j: group index for component j (for computing earliness).

q: a sequence.

32

$TC(M,R,q)$: optimal objective value of P1 when M, R and q are given.

n_k: number of components in group k.

e_j^k : earliness of the jth component in group k.

θ_j^k : lower bound on e_j^k .

$C(M,R)$: total weighted earliness of the best unconstrained schedule for the scheduling problem with M and R given.

β: $\sum_{j=1}^{J} p_j D_j^2 h_j$,

q^*: sequence in which components are sorted in non-increasing order of $\dfrac{p_j}{h_j}$.

$Z(q^*)$: $\sum_{i=1}^{J} D_{[i]} h_{[i]} \sum_{j=i}^{J} p_{[j]} D_{[j]}$.

$E(M,R,q)$: optimal value of $\sum_{j=1}^{J} D_j h_j e_j$ in P1 when the values of M and R, and the sequence (q) are given.

$LB(M,R)$: right-hand-side of (30).

$TC(M,R)$: $\min_{q}\{TC(M,R,q)\}$.

M^* : optimal value of M for P2.

R^*: optimal value of R for P2.

t_j: $s_j + p_j D_j T$.

w_j: $h_j D_j$.

PT_k : total production time of group k (i.e., $PT_k = \sum\limits_{\substack{\text{component } j \\ \text{in group } k}} t_j$).

α: $\sum_{j=1}^{J} \alpha_j$.

S : $\sum_{j=1}^{J} S_j$.

$f(R)$:   objective function value of P3 for a given R.

$F(R)$: $f(R) + \alpha MR$.

r: $\underset{R \geq 0}{\text{argmin }} F(R)$

$TC^B(M, R)$:   total cost resulting from algorithm BESTR for a given M and initial R.

$R(M)$: $\underset{R \geq 0}{\text{argmin}}\{ LB(M, R) \mid MR \geq \tau \}$.

$LB1$:   lower bound on solution from HY algorithm $(M=1)$.

$TC1$:   objective value of solution from HY algorithm $(M=1)$.

APPENDIX 2

PROOF OF THEOREM 1

Theorem 1:

$$C(M, R) \geq \max\{\ \beta R, \frac{R}{M}\ [Z(q^*) + \frac{M-1}{2}\ \beta\ ]\ \}$$ (A2-1)

where $\beta = \sum_{j=1}^{J} p_j D_j^2 h_j$,

$$Z(q^*) = \sum_{i=1}^{J} D_{[i]} h_{[i]} \sum_{j=i}^{J} p_{[j]} D_{[j]} ,$$

and q^* is the sequence in which components are sorted in non-increasing order of $\frac{p_j}{h_j}$.

Proof.

Note that $C(M, R)$ has been defined as the total weighted earliness of J components when the processing time of component j is $p_j D_j R$ for all j and the components are assigned to M identical machines. Eastman et al. (1964) proved the following result.

$$C(M, R) \geq \max\{\ C(J, R), \frac{1}{M}\ C(1, R) + \frac{M-1}{2M}\ C(J, R)\ \}.$$ (A2-2)

For our problem, $C(J, R)$ and $C(1, R)$ are defined as follows.

$$C(J, R) = \sum_{i=1}^{J} D_{[i]} h_{[i]} \cdot p_{[j]} D_{[j]} R = \beta R.$$ (A2-3)

$$C(1, R) = \min_{q} Z(q) \cdot R = Z(q^*) \cdot R.$$ (A2-4)

Substitution of (A2-3) and (A2-4) into (A2-2) gives the result in (A2-1). ∎

35

APPENDIX 3

PROOF OF LEMMA 1

Lemma 1:

$$LB(M, R(M)) \text{ unimodal in } M.$$

<u>Proof.</u>

Throughout this proof, we use the term unimodal to refer to a function that initially decreases in the argument and then increases (i.e., first order conditions define a local minimum, not a local maximum). We want to show that $LB(M, R(M))$ is unimodal for positive real values of M and thus, also unimodal for positive integer M. Therefore, initially, let us assume that M is a positive real number.

By definition,

$$LB(M, R) = \frac{1}{MR} \sum_{j=1}^{J} S_j + MR \sum_{j=1}^{J} \alpha_j + \max\{ \beta R, \frac{R}{M} [Z(q^*) + \frac{M\text{-}1}{2} \beta] \} + \frac{A}{R}$$

$$= \max \{ \frac{1}{MR} \sum_{j=1}^{J} S_j + MR \sum_{j=1}^{J} \alpha_j + \beta R + \frac{A}{R} ,$$

$$\frac{1}{MR} \sum_{j=1}^{J} S_j + MR \sum_{j=1}^{J} \alpha_j + \frac{R}{M} [Z(q^*) + \frac{M\text{-}1}{2} \beta] + \frac{A}{R} \}.$$

Let us define

$$LB'(M, R) = \frac{1}{MR} \sum_{j=1}^{J} S_j + MR \sum_{j=1}^{J} \alpha_j + \beta R + \frac{A}{R}$$

and $LB''(M, R) = \frac{1}{MR} \sum_{j=1}^{J} S_j + MR \sum_{j=1}^{J} \alpha_j + \frac{R}{M} [Z(q^*) + \frac{M\text{-}1}{2} \beta] + \frac{A}{R}$.

Then,

$$LB(M, R) = \max\{ LB'(M, R), LB''(M, R) \}.$$

Let

$$R'(M) = \underset{R \geq 0}{\mathrm{argmin}}\{ LB'(M, R) \mid MR \geq \tau \}$$

and

$$R''(M) = \underset{R \geq 0}{\mathrm{argmin}}\{ LB''(M, R) \mid MR \geq \tau \}.$$

Then, it is easily verified that

$$LB(M, R(M)) = \max\{ LB'(M, R'(M)), LB''(M, R''(M)) \}.$$

As a result, if both $LB'(M, R'(M))$ and $LB''(M, R''(M))$ are unimodal in M, $LB(M, R(M))$ which is the maximum of these, is also unimodal in M. (This property is easy to prove for univariate unimodal functions of the type considered here.)

First, let us show $LB'(M, R'(M))$ is unimodal in M.

It is easy to show that

$$R'(M) = \max \{ R_L(M), \frac{\tau}{M} \}$$

where $R_L(M) = \sqrt{\dfrac{S + MA}{M(\alpha M + \beta)}}$.

Let m be the value of M which satisfies $R_L(M) = \dfrac{\tau}{M}$. Then, it is easily verified that the value of m is unique and it is expressed in the following form.

$$m = \frac{\alpha \gamma^2 - S + \sqrt{(\alpha \gamma^2 - S)^2 + 4A\beta\gamma^2}}{2A} . \tag{A3-1}$$

Also, because m is unique, it is clear that

$$R'(M) = \begin{cases} \dfrac{\tau}{M} & \text{for } M \leq m \\ R_L(M) & \text{for } M \geq m \end{cases}$$

37

As a result,

$$LB'(M, R'(M)) = \begin{cases} LB'(M, \dfrac{\tau}{M}) = \dfrac{S}{\tau} + \alpha\gamma + \dfrac{\beta\gamma}{M} + \dfrac{AM}{\tau} & \text{for } M \le m \\[3mm] LB'(M, R_L(M)) = 2\sqrt{\dfrac{1}{M}(S + MA)(\alpha M + \beta)} & \text{for } M \ge m \end{cases}$$

Note that

$$\frac{\tau}{m} = R_L(m) \text{ and } LB'(m, \frac{\tau}{m}) = LB'(m, R_L(m)). \tag{A3-2}$$

To show that $LB'(M, \dfrac{\tau}{M})$ and $LB'(M, R_L(M))$ are unimodal, we use the following lemmas.

Lemma 2: $LB(M, R_L(M))$ is increasing in M for $M \ge \sqrt{\dfrac{S\beta}{A\alpha}}$,

decreasing in M for $M \le \sqrt{\dfrac{S\beta}{A\alpha}}$.

Proof:

$$LB(M, R_L(M)) = 2\sqrt{\frac{(S+M\cdot A)(\alpha M+\beta)}{M}} .$$

It is easily verified that $(S+M\cdot A)(\alpha M+\beta)/M$ is increasing in M for $M \ge \sqrt{\dfrac{S\beta}{A\alpha}}$ and

decreasing in M for $M \le \sqrt{\dfrac{S\beta}{A\alpha}}$. Therefore, Lemma 2 follows.•

Lemma 3: $LB(M, \dfrac{\tau}{M})$ is a convex function of M and achieves its minimum at $M = \dfrac{\tau}{R'}$.

Proof:

$$LB(M, \frac{\tau}{M}) = \frac{S}{\tau} + \alpha\gamma + \frac{\beta\gamma}{M} + M\cdot\frac{A}{\tau} .$$

It is easily verified that $\dfrac{\partial^2}{\partial M^2} LB(M, \dfrac{\tau}{M}) > 0$ and that $\dfrac{\partial}{\partial M} LB(M, \dfrac{\tau}{M}) = 0$ at $M = \dfrac{\tau}{R'}$.

•

38

Let M^1 and M^2 be the values of M that minimize $LB'(M, \frac{\tau}{M})$ and $LB'(M, R_L(M))$ respectively. From Lemmas 2 and 3,

$$M^1 = \tau \sqrt{\frac{\beta}{A}} \quad \text{and } M^2 = \sqrt{\frac{S\beta}{A\alpha}} \ . \tag{A3-3}$$

Before we prove Lemma 1, we want to show that

(i) if $M^1 \geq m$, then $M^2 \geq m$

and (ii) if $M^1 \leq m$, then $M^2 \leq m$.

Proof of (i).

First we want to show that, if $M^1 \geq m$, $\sqrt{\frac{S}{\alpha}} \geq \tau$.

Suppose $\sqrt{\frac{S}{\alpha}} < \tau$, which means $\alpha\gamma^2 > S$. Then,

$$m = \frac{\alpha\gamma^2 - S + \sqrt{(\alpha\gamma^2 - S)^2 + 4A\beta\gamma^2}}{2A}$$

$$> \frac{\sqrt{4A\beta\gamma^2}}{2A}$$

$$= \tau\sqrt{\frac{\beta}{A}}$$

$$= M^1$$

which contradicts the assumption $M^1 \geq m$.

Therefore, $\sqrt{\frac{S}{\alpha}} \geq \tau$.

As a result,

$$M^2 = \sqrt{\frac{S\beta}{A\alpha}} \geq \tau\sqrt{\frac{\beta}{A}} = M^1 \geq m \ .$$

The proof of (ii) is similar to that of (i) and is omitted here.

Case 1: $M^1 \geq m$.

If $M^1 \geq m$, $LB'(M, \frac{\tau}{M})$ is monotonically decreasing in M for $M \leq m$ by Lemma 3.

By (A3-2), $LB'(m, \frac{\tau}{m}) = LB'(m, R_L(m))$. Moreover, by (i) and Lemma 2, $LB'(M,$

$R_L(M))$ is monotonically decreasing in M for $m \leq M \leq M^2$ and monotonically

increasing for $M > M^2$. As a result, $LB'(M, R'(M))$ is unimodal in M.

Case 2: $M^1 \leq m$.

If $M^1 \geq m$, $LB'(M, \frac{\tau}{M})$ is monotonically decreasing in M for $M \leq M^1$ and

monotonically increasing in M for $M^1 \leq M \leq m$ by Lemma 3. By (A3-2), $LB'(m, \frac{\tau}{m})$

$= LB'(m, R_L(m))$. Moreover, by (ii) and Lemma 2, $LB'(M, R_L(M))$ is monotonically

increasing in M for $M \geq m$. As a result, $LB'(M, R'(M))$ is unimodal in M.

Consequently, $LB'(M, R'(M))$ is unimodal in M.

The proof that $LB''(M, R''(M))$ is unimodal in M parallels the proof for $LB'(M, R'(M))$.

Since $LB'(M, R'(M))$ and $LB''(M, R''(M))$ are unimodal in M, $LB(M, R(M))$ is

unimodal in M.

40