

**DUE DATE SETTING AND SEQUENCING IN A
TIME COMPETITIVE ENVIRONMENT**

Izak Duenyas

Technical Report 92-40

**Department of Industrial and Operations Engineering
The University of Michigan
Ann Arbor, MI 48109**

June 1992

DUE DATE SETTING AND SEQUENCING IN A TIME COMPETITIVE ENVIRONMENT

IZAK DUENYAS

Department of Industrial and Operations Engineering
The University of Michigan
Ann Arbor, Michigan 48109

Abstract

We consider the interrelated problems of (1) quoting a due date to each customer arriving to a production system modeled as a single-server queue and (2) sequencing customer orders once they are in the system. We allow several different classes of customers, each with different preferences for leadtime and price. We first formulate the problem of quoting due dates under the assumption that customer orders are processed on a FCFS basis. Next, we consider the case where the firm has the option to schedule orders in other than FCFS order. For this case, we develop a heuristic for quoting due dates and sequencing orders. Simulation results suggest that policies that take into account customer price and due date preferences in scheduling and quoting due dates significantly outperform due date setting policies that do not.

1 Introduction

In many manufacturing environments, firms compete on the basis of delivery lead time as well as price. Recently, there has been a considerable amount of practitioner literature focusing on customer responsiveness (e.g., Blackburn 1991, Schemenner 1988, Stalk and Hout 1990, Thomas 1990, 1991). It has been observed that in many industries, companies that can promise shorter due date lead times and can deliver on time feel less pressured to compete on the basis of price. As Stalk points out (in Blackburn 1991, p. 69)

Response time is the final measure of how close a company is to its customers—how long must a customer wait from the point at which he or she asks for something until the company delivers it. Time-based competitors direct the benefits of their value-delivery systems towards the most attractive customers, the ones that are willing to pay for responsiveness, fast delivery and choice. The unattractive customers—those who will not pay for responsiveness and instead are willing to wait for the best price—are left for the competition.

The growing pressure on companies to perform on the basis of responsiveness has led to increased research on how to set due dates. Early research on scheduling assumed that due dates were exogenously set by the sales department and that the production schedulers had no input on setting these due dates. Hence, the responsibility of the production scheduler

was either to optimize the performance of the shop (e.g., minimize average flow times) without taking due dates into consideration, or to schedule jobs in order to minimize tardiness or lateness with respect to the given due dates. More recently, researchers have focused on the combined problem of quoting customer due dates and sequencing orders. (Baker 1984, Baker and Bertrand 1981, 1982, Bertrand, 1983, Eilon and Chowdhury 1976, Seidmann and Smith 1981, Shantikumar and Sumita 1988, Wein 1991), as well as on using leadtime estimates to improve the sequencing of orders (Vepsalainen and Morton, 1987, 1988).

The most common approach to the due date setting and sequencing problem is to minimize the average weighted due date lead time (the due date of a job minus its arrival time) while maintaining a certain service level (e.g., an upper bound on the proportion of tardy jobs or an upper bound on average tardiness). The results in the above cited papers indicate that due dates based on estimates of shop congestion and the estimated duration of the processing time of jobs lead to significantly lower average due dates. Baker and Bertrand (1982) suggest three parametric due date setting rules. The three rules quote customers due date lead times of a constant, a constant plus the expected processing time of the job and a constant times the expected processing time of the job, none of which depend on the amount of congestion in the system. Bookbinder and Noor (1985) derive a rule that takes into account the level of congestion in the system. Wein (1991) considers the problem of minimizing average weighted due date lead time for a multiclass $M/G/1$ queue, subject to constraints on the proportion of tardy jobs or the average tardiness. Wein uses the $c\mu$ rule (weighted shortest expected processing time rule) for sequencing the customers since the $c\mu$ rule minimizes the average weighted flow time. He then derives an expression for the conditional sojourn time distribution of jobs from each class. Using this distribution, the customer is quoted the minimum due date that satisfies the service constraint. This method leads to significant improvement in average weighted lead times. A particularly interesting conclusion of Wein's study is that proper due date setting offers a much larger improvement in performance than priority sequencing. Wein and Chevalier (1990) make the same observation for a two-station shop.

All of the above papers consider due date setting and priority sequencing from the scheduler's point of view. Due dates are selected to minimize holding costs, tardiness, fraction of late jobs etc., while trying to quote the customer the minimum possible due date. An implicit assumption in these papers is that each customer that is quoted a due date places an order. That is, whether the customer places an order does not depend on the date quoted him/her. As firms increasingly compete on the basis of lead time, the due date lead time quoted a customer will have a strong effect on whether he/she actually places an order. Hence, in quoting the customer a due date, the firm has to take into consideration the customers' preferences for due dates. An added complication is that a company always has a certain group of customers who are willing to wait longer in exchange for a lower price, and certain customers who are willing to pay more for quick service. A question that many companies face is what due dates to quote to different types of customers with different price and due date preferences. In an environment where the sales department and the customer negotiate the price and the due dates, this question is of real practical importance. The algorithms developed previously suggest an "optimal" due date which the sales department would quote to the customer, but do not indicate how the sales department should negotiate with the customer using this information. For example, suppose that our due date setting procedure indicates that an arriving customer should be assigned due date

lead time a_1 . Further suppose that the customer indicates that he/she will place an order only if the order can be delivered in $a_2 < a_1$ days. Should the sales department accept this offer? For what values of a_2 should an order still be accepted? Given that the offer was accepted with due date a_2 , what changes should be made in the sequence of orders to be processed? These questions are not answered by previous work on due date setting and motivate the work of this paper.

To our knowledge, the only two papers that address customer preferences in due date setting and sequencing are Wein et al. (1990) and Duenyas and Hopp (1992). Wein et al. (1990) consider an M/M/1 queue with a single customer class. The Poisson arrival rate of customers is a function of the customers' expected cost of obtaining service where the expected cost equals the quoted price plus a cost proportional to lead time plus a cost proportional to perceived expected tardiness. Wein et al. obtain optimal price and due date setting policies under five different scenarios under the assumption that the system operates under the FCFS discipline. Duenyas and Hopp (1992) also consider a system with a single customer class. They assume that customers arrive to the system (e.g., to the sales department to make inquiries) according to a random process, are quoted a due date lead time a , which may depend on the congestion level in the shop. The customer then places an order with probability $p(a)$, where $p(a)$ is a decreasing function of a . If the customer order is late, then the firm pays a penalty. They consider: (1) the case where orders are processed in a FCFS manner and (2) the case where the firm can choose the sequence of orders. For the second case, they prove that once customers are quoted a due date, orders are always processed according to the Earliest Due Date (EDD) protocol.

In this paper, we extend the framework of Duenyas and Hopp to the case where there are several customer classes with different price and lead time preferences with regard to a single product. We consider the problem of due date setting and order sequencing problems in the setting of a multiclass M/G/1 queue. Our intent is to demonstrate the importance of considering customer preference information in setting due dates and to develop a framework for utilizing this information when available. Our hope is to extend this framework in the future to the case where the facility has more than one machine and where different customer classes demand different products.

The rest of this paper is organized as follows. In Section 2, we formulate the due date setting problem under the assumption that orders are processed in a FCFS manner. In Section 3, we formulate the problem of quoting due dates and sequencing orders under the assumption that the firm has the option to sequence customers in other than FCFS order. For that case, we suggest a heuristic based on the results of the FCFS case from Section 2. In Section 4, we present simulation results that compare the performance of our due date setting and sequencing procedure with other due date setting procedures that do not take customer preferences into account. The paper concludes in Section 5.

2 Setting Due Dates under FCFS Sequencing

In this section, we formulate the problem of quoting customer due dates when customer orders are serviced in a FCFS manner. We assume that there are n different classes of customers each demanding the same single product. Customers from different classes differ with respect to their preferences for price and leadtime. Customers from class i arrive to the sales department according to a Poisson process with rate λ_i to inquire about due dates.

The company quotes the customer a due date lead time. The probability that a customer from class i places an order when quoted due date lead time a is $p_i(a)$. (We assume that $p_i(a)$ is continuous and decreasing in a). In this way, we differentiate between the arrival rate of class i customers to the sales department to inquire about due dates and the *effective* arrival rate of customer orders to the factory, which depends on the quoted lead times to different customers. Once an order is placed, we assume that it is not cancelled. Orders are filled one at a time and we let F denote the c.d.f. of the production time, where F is continuous with derivatives f and f' . An order from a class i customer generates net revenue R_i . However, if the order is not filled on time, the firm incurs a penalty. We assume that an order that is x time units late incurs a penalty cx . This penalty may reflect the loss-of-goodwill, as well as extra shipping and handling costs due to the tardiness. As Wein et al. note, in some service industries, such as banking and pizza delivery, the server actually reimburses the customer for any tardiness experienced. In those cases, this penalty may reflect the cost of the reimbursement.

We note that it is possible to formulate an analogous due date setting, pricing and sequencing problem by assuming that: (1) a customer from class i places an order with probability $p_i(R, a)$, and (2) a customer can be quoted different lead times *and* different prices depending on the congestion in the system. It is not difficult to extend the analysis in this paper to that case. However, setting prices dynamically, based on the congestion level in the system is not possible in most situations, since this would require changing prices almost continuously. While quoting different lead times at different times is generally acceptable, most companies do not change their prices on a minute by minute basis.

Our assumption that the company faces several different customer classes applies in many practical situations. For example, the company may be able to apply price discrimination in different geographical regions, so long as customers in the same region are always quoted the same price, or customers in different regions may accept different lead times (thereby having different $p_i(a)$ functions). Similarly, the company may announce a price policy that depends on the urgency of the order. An example would be that customers that place orders to be delivered in a lead time of less than a time units would be charged a certain price, while those placing orders to be delivered later would be charged another (lower) price. However, if the congestion level in the system is low, and the new customer's order can be processed quickly, the firm may choose to promise a customer with an urgent order a lead time even less than a to get the customer's order. Our modeling assumptions are valid in all of these situations.

We denote the system state at customer arrival points by (k, t) , where k denotes the number of orders waiting to be processed, and t denotes the amount of time that the first order has already been processed. We denote by $a_{k,t}^i$, the optimal lead time quote to a customer from class i in state (k, t) . We can then formulate the problem of quoting optimal lead times as a Semi-Markov Decision Process. We let g be the average profit per customer arrival, and $v_{k,t}^i$ denote the relative value function. We can write the SMDP recursion as

$$g + v_{k,t}^i = \max_a p_i(a)(R_i - \varphi_{k+1}^t(a) + w_{k+1}^t) + (1 - p_i(a))(w_k^t) \quad (2.1)$$

where

$$\varphi_k^t(a) = \int_a^\infty c(x - a)dF_k^t(x)$$

and $F_k^t = F^t * F_{k-1}$ where F_k is the k -fold convolution of F and F^t denotes the c.d.f. of

the remaining processing time of a job that has already been processed for t time units. Clearly, $\varphi_k^t(a)$ represents the expected tardiness cost of a customer quoted lead time a in state (k, t) . Let Y represent the (random) interarrival time (which is exponential with rate $\sum_{i=1}^n \lambda_i$), and define

$$Q_t(k, \tau) = P\left\{\sum_{j=1}^k X_j \leq Y + t, \sum_{j=1}^{k+1} X_j > Y + t, Y + t - \sum_{j=1}^k X_j \leq \tau | X_1 > t\right\} \quad (2.2)$$

and

$$w_k^t = \sum_{i=1}^n s_i \left\{ \sum_{j=0}^{k-1} \int_{\tau=0}^{\infty} v_{k-j, \tau}^i dQ_t(j, \tau) + v_0^i \sum_{j=k}^{\infty} \int_{\tau=0}^{\infty} dQ_t(j, \tau) \right\} \quad (2.3)$$

and $s_i = \lambda_i / (\sum_{j=1}^n \lambda_j)$. $Q_t(k, \tau)$ is the joint conditional probability of completing exactly k services during an interarrival time and having the $k + 1^{\text{st}}$ customer in service for τ time units or less, given that we began with the current customer in service for t time units, assuming that we have an infinite supply of customers to serve.

We first prove that under the condition that $p_i(a)$ is concave in a for $i = 1, \dots, n$, the optimal lead time in (2.1) is higher than if we quote customer lead times myopically by disregarding the congestion effects of customers that place orders. For example, if we did not take into account future arrivals of customers, we would quote the due date $\hat{a}_{k,t}^i$ to a customer from class i in state (k, t) where

$$\hat{a}_{k,t}^i = \arg \max_a p_i(a) (R_i - \varphi_{k+1}^t(a))$$

However, quoting due dates in this manner does not take into account that by accepting this customer's order, we may also be forced to quote the next customer a longer lead time due to the increased congestion in the system. The following theorem states this result formally.

Theorem 1 *Assume that $p(a)$ is concave. Then, $\hat{a}_{k,t}^i \leq a_{k,t}^i$.*

Proof: It is straightforward to show recursively that $w_{k+1}^t \leq w_k^t$. By the definition of $\hat{a}_{k,t}^i$, we have

$$p_i'(\hat{a}_{k,t}^i) (R_i - \varphi_{k+1}^t(\hat{a}_{k,t}^i)) + p(\hat{a}_{k,t}^i) (c(1 - F_{k+1}^t(\hat{a}_{k,t}^i))) = 0$$

Hence, we have

$$p_i'(\hat{a}_{k,t}^i) (R_i - \varphi_{k+1}^t(\hat{a}_{k,t}^i) + w_{k+1}^t - w_k^t) + p(\hat{a}_{k,t}^i) (c(1 - F_{k+1}^t(\hat{a}_{k,t}^i))) \geq 0 \quad (2.4)$$

The result follows by the fact that substituting any $a < \hat{a}_{k,t}^i$ for $\hat{a}_{k,t}^i$ causes the left hand side of (2.4) to increase. \square .

The above theorem indicates that quoting due dates without taking into account the increased congestion that each order brings into the system leads to quoting lead times that are lower than the optimal lead times. We note that as formulated above, the due

date setting problem is a Semi-Markov Decision Process with uncountable state and action spaces. Hence, unless the processing times are exponentially distributed, we need to discretize the processing time distribution to obtain an approximate solution. One way to do this would be to discretize the processing time distribution so that at customer arrival points, the amount of time that the first order has been processed, t , is in some discrete set t_1, \dots, t_m . (For example, if the processing time is uniform between 0 and 1, then we assume that an arriving job finds the first job has had processing for $0, 1/n, 2/n, \dots, (n-1)/n$ time units). The transitions between these discrete states can be approximated from the original processing time distribution. We can then use a successive approximation algorithm where we start by setting all $w_k^{t_j}$ values equal to 0. We can calculate the optimal lead times and v_{k,t_j}^i values from (2.1) and update the $w_k^{t_j}$ values by using (2.3). We continue in this manner until a prespecified level of accuracy is reached.

For the special case of exponential service times, we can calculate the optimal due date lead times exactly since the amount of processing that the first job has already had does not affect its remaining processing time. In this case, (2.1) becomes

$$g + v_k^i = \max_a p_i(a)(R_i - \varphi_{k+1}(a) + w_{k+1}) + (1 - p_i(a))w_k \quad (2.5)$$

where $\varphi_{k+1}(a) = \int_a^\infty c(x-a)dF_{k+1}(x)$, $F_{k+1}(x)$ is the c.d.f. of the Erlang- $(k+1)$ distribution, and

$$w_k = \sum_{i=1}^n s_i \left\{ \sum_{j=0}^{k-1} q_j v_{k-j}^i + v_0^i \sum_{j=k}^{\infty} q_j \right\} \quad (2.6)$$

where q_j represents the probability that j customers are served between two successive customer arrivals. In this case, we can use the successive approximation algorithm described above to calculate the optimal lead times exactly. We can also use the SMDP formulation to show that the due date quoted to a customer from class i is increasing in the number of customer orders already in the system. We state the result in the following theorem:

Theorem 2 *If for all i , there exists a finite \hat{a}^i such that $p_i(a) = 0, a > \hat{a}^i$, and $p_i(a)$ is concave in a for all i , then the optimal solution to (2.5) has a_k^i increasing in k for each i .*

Proof: The proof is in the Appendix.

Theorem 2 is rather intuitive. It states that as the number of orders in the system increases, each type of customer is quoted increasing lead times. A second structural result of interest would be an ordering with respect to the due dates quoted to different customers. To show such a result, we need to make assumptions about the different types of customers. One reasonable set of assumptions is: (1) customers who are willing to pay more for quick service would have a lower probability of placing an order when quoted the same due date as those customers who would like to pay less and (2) the less patient customers are more sensitive to changes in lead times. Under these conditions, there is an ordering on the due dates quoted to different customers such that the customers who place the more profitable orders and are less willing to wait are quoted shorter due date lead times.

Theorem 3 *If $p_i(a)$ is concave in a for all i and if for $i \neq j$ and all a , $R_i \geq R_j$, $p_i(a) \leq p_j(a)$ and $p_i'(a) \leq p_j'(a)$ then $a_{k,t}^i \leq a_{k,t}^j$ in for any k and t .*

Proof: The proof is similiar to the proof of Theorem 1. \square .

3 Setting Due Dates and Sequencing Orders

In the previous section, we assumed that each customer order is filled in FCFS order. Even though in some environments, this may be a reasonable assumption (in cases where the customers have information about the status of orders and are displeased to learn that orders are not being processed in the order they arrived), in most situations this is not a reasonable assumption.

In most practical situations, the company is free to choose the sequence in which the orders are to be filled. It may be attractive to use a non-FCFS sequence. For instance, the company may choose to quote the more profitable customers shorter due dates and place their orders in front of existing orders from more patient customers. To accommodate this possibility, in this section we formulate the combined problem of quoting due dates and sequencing orders.

To model the case where the firm decides the sequence of orders, we must keep track of how much time is left on the due date of each order. To do this, we define states as $(i, u, k, t_1, \dots, t_k)$ where i denotes the class of the newly arrived customer, u denotes the amount of time that the order currently in process has already been processed, k denotes the number of orders in the system and t_1, \dots, t_k denote the remaining time until the due dates of these orders. We note that t_j could be negative, in which case t_j denotes how much time has passed since the due date of the j^{th} order. We assume that the first job cannot be preempted.

Letting $v_i(u, k, t_1, \dots, t_j, \dots, t_k)$ denote the relative value function, we can write the SMDP recursion as

$$g + v_i(u, k, t_1, \dots, t_j, t_k) = \max_{a, j=2, \dots, k+1} \{p_i(a)(R_i - \varphi_j^u(a) - \sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l^u(t_l) + w(u, k+1, t_1, \dots, t_{j-1}, a, t_j, \dots, t_k)) + (1 - p_i(a))w(u, k, t_1, \dots, t_k)\} \quad (3.7)$$

where

$$\varphi_j^u(a) = \int_a^\infty c(x - a) dF_j^u(x)$$

and F_j^u is defined as in Section 2.

Note that in this case the firm chooses not only the lead time to quote but also where to place the new order if the customer decides to place an order. If the order is placed in the j^{th} position, then the expected amount of time that orders j through k are delayed increases. Hence, the expected costs include the delay cost of the new order $\varphi_j^u(a)$, and the additional delay cost of the orders that are displaced $(\sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l^u(t_l))$. As in the previous section, the w values denote the expected profits after the new order is placed. However, in this case they are slightly more complicated. We let $h(t)$ denote the density of the inter-arrival time and define

$$Q_u^t(i, \tau) = P\left\{\sum_{j=1}^i X_j \leq t + u, \sum_{j=1}^{i+1} X_j > t + u, t + u - \sum_{j=1}^i X_j \leq \tau | X_1 > u\right\}$$

Then, we have

$$w(u, k, t_1, \dots, t_k) = \sum_i s_i \left\{ \int_{t=0}^\infty h(t) \left(\sum_{j=0}^{k-1} \int_{\tau=0}^\infty v_i(\tau, k-j, t_{j+1}-t, \dots, t_k-t) dQ_u^t(j, \tau) + v_i(0) \sum_{j=k}^\infty \int_{\tau=0}^\infty dQ_u^t(j, \tau) \right) \right\} \quad (3.8)$$

The above formulation of the due date setting and sequencing problem is a Semi-Markov Decision Process with uncountable state and action spaces. Since an exact solution seems out of the question, we now turn to the development of a heuristic for this problem. This heuristic will make use of the following structural results.

Theorem 4 *Customer orders are processed in an EDD manner.*

Proof: The proof is similar to the proof of the EDD result for a single customer class in Duenyas and Hopp (1992) and is omitted. \square

Theorem 5 *If $p_i(a)$ is concave in a for all i , then a lower bound on the optimal due date if the new customer of class i is placed in the j^{th} position is given by ψ_j where*

$$\psi_j = \arg \max_a \{p_i(a)(R_i - \varphi_j^u(a) - \sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l^u(t_l))\} \quad (3.9)$$

Proof: The proof is similar to the proof of Theorem 1. \square .

Theorems 4 and 5 are useful in decreasing the region in which one must search to find the optimal solution. By Theorem 4, if we place the new customer in the j^{th} position, then the lead time quote must be in $[t_{j-1}, t_j]$. Furthermore it follows from Theorem 5 that (1) if $\psi_j < t_{j-1}$, then the lead time quote is in $[t_{j-1}, t_j]$; (2) if $t_{j-1} \leq \psi_j \leq t_j$, then the lead time quote is in $[\psi_j, t_j]$, and (3) if $\psi_j > t_j$, the best possible quote if we place the customer in the j^{th} position is t_j . However, Theorems 4 and 5 do not tell us where to place the customer's order. Nevertheless, Theorems 4 and 5 help in ruling out some positions. For example, if the immediate reward earned by a customer placing an order and being placed in position j and quoted lead time t_j is less than 0 (i.e., $R_i - \varphi_j^u(t_j) - \sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l^u(t_l) < 0$) then we can conclude that the customer order will not be placed in the j^{th} position under an optimal policy.

The above analysis immediately suggests a myopic heuristic for quoting customer due dates and sequencing orders:

Myopic Heuristic for Quoting Due Dates (H1)

1. For $j = 2, \dots, k+1$, compute $H_j = R_i - \varphi_j^u(t_j) - \sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l^u(t_l)$. If $H_j \leq 0$, then rule out position j . Let j' be the first position for which $H_j > 0$.
2. For $j = j', \dots, k+1$ compute ψ_j using (3.9)
3. for $j = j', \dots, k+1$, let

$$B_j = \begin{cases} p_i(t_j)(R_i - \varphi_{j+1}^u(t_j) - \sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l^u(t_l)) & \text{if } \psi_j > t_j \\ p_i(t_{j-1})(R_i - \varphi_{j+1}^u(t_{j-1}) - \sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l^u(t_l)) & \text{if } \psi_j < t_{j-1} \\ p_i(\psi_j)(R_i - \varphi_{j+1}^u(\psi_j) - \sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l^u(t_l)) & \text{if } t_{j-1} < \psi_j < t_j \end{cases}$$

4. Let $s = \arg \max_{j=j', \dots, k+1} B_j$. Place the customer's order in position s if customer decides to place an order. Quote the customer a lead time of ψ_s if $t_{s-1} < \psi_s < t_s$, a lead time of t_s if $\psi_s > t_s$ and a lead time of t_{s-1} otherwise.

The above heuristic is myopic because it does not take into account the future effects of the current decision. The purpose here is to maximize the expected revenue minus the displacement costs from the current customer. Hence, the major drawback of this heuristic policy is that it fails to account for the fact that if we accept a lower revenue customer now, the increased congestion in the system may lead us to lose a high revenue customer in the future. Hence, it is very important to consider the cost to the firm arising from increased congestion when an order is accepted. To do this, we first note that (3.7) can be rewritten as

$$g + v_i(u, k, t_1, \dots, t_j, t_k) = w(u, k, t_1, \dots, t_k) + \max_{a, j=2, \dots, k+1} \{p_i(a)(R_i - \varphi_j^u(a) - \sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l(t_l) + w(u, k+1, t_1, \dots, t_{j-1}, a, t_j, \dots, t_k) - w(u, k, t_1, \dots, t_k))\} \quad (3.10)$$

In our myopic heuristic, we were ignoring the term $w(u, k+1, t_1, \dots, t_{j-1}, a, t_j, \dots, t_k) - w(u, k, t_1, \dots, t_k)$. One way to think about this term is that it represents the congestion cost due to the arrival of a new order to the system. Since, there are an uncountable number of states and hence an uncountable number of different congestion costs, it is not possible to calculate this cost exactly. To incorporate this cost in our due date setting procedure, we develop a second heuristic for the due date setting and sequencing problem by using the congestion costs obtained from the FCFS case in the previous section. The assumption underlying this second heuristic is that when a new order joins the system, the major portion of the congestion cost is due to the fact that there is an extra order in the system. That is, we assume that the above term depends more on the number of orders already in the system than the due dates of the orders.

As an example for this kind of reasoning, consider the case where we are trying to decide whether to place the customer in the second or third position and quote him/her a lead time of a_2 or a_3 . In the first case, the increased congestion cost is given by $w(u, k+1, t_1, a_2, t_2, t_3, \dots, t_{k+1}) - w(u, k, t_1, \dots, t_k)$ while in the second case, it is given by $w(u, k+1, t_1, t_2, a_3, t_3, \dots, t_{k+1}) - w(u, k, t_1, \dots, t_k)$. However, note that if the next arrival occurs after three services, whether we placed this order in the second or third position will make no difference since the state of the system will be the same in both cases. If the next arrival occurs *before* three services, whether we placed this order in the second or third place will make a difference only if there is considerable slack in these orders' due dates. However, in all cases, the fact that there is an extra order in the system will make a difference when the next customer arrives. For example, if the next customer is a high revenue customer, the firm will incur extra displacement costs when placing his/her order in front of other orders. If the firm decides to place the new order last, the firm will have to quote a higher due date since it will now be placing the new order in the $k+1^{\text{st}}$ position instead of the k^{th} .

Using this assumption that a major portion of the increased congestion cost is due to the existence of an extra order in the system, we approximate the congestion costs $w(u, k+1, t_1, \dots, t_j, a, t_{j+1}, \dots, t_k) - w(u, k, t_1, \dots, t_k)$ by $w_{k+1}^u - w_k^u$ where $w_{k+1}^u - w_k^u$ is obtained by solving (2.1). We can now present our modified heuristic.

Modified Heuristic for Setting Due Dates (H2)

1. Solve SMDP (2.1) for the FCFS case to obtain $w_{k+1}^u - w_k^u$.
2. For $j = 2, \dots, k+1$, compute $H_j = (R_i - \varphi_j^u(t_j) - \sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l^u(t_l) + w_{k+1}^u - w_k^u)$. If $H_j < 0$, then rule out position j . Let j' be the first position for which $H_j > 0$.

3. For $j = j', \dots, k + 1$ compute ψ_j using

$$\psi_j = \arg \max_a p_i(a) (R_i - \varphi_j^u(a) - (\sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l^u(t_l)) + w_{k+1}^u - w_k^u) \quad (3.11)$$

4. For $j = j', \dots, k + 1$, let

$$B_j = \begin{cases} p_i(t_j) (R_i - \varphi_j^u(t_j) - (\sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l^u(t_l)) + w_{k+1}^u - w_k^u) & \text{if } \psi_j > t_j \\ p_i(t_{j-1}) (R_i - \varphi_j^u(t_{j-1}) - (\sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l^u(t_l)) + w_{k+1}^u - w_k^u) & \text{if } \psi_j < t_{j-1} \\ p_i(\psi_j) (R_i - \varphi_j^u(\psi_j) - (\sum_{l=j}^k \varphi_{l+1}^u(t_l) - \varphi_l^u(t_l)) + w_{k+1}^u - w_k^u) & \text{if } t_{j-1} < \psi_j < t_j \end{cases}$$

5. Let $s = \arg \max_{j=j', \dots, k+1} B_j$. Place the customer in position s if the customer decides to place an order. Quote the customer a lead time of ψ_s if $t_{s-1} < \psi_s < t_s$, a lead time of t_s if $\psi_s > t_s$, and a lead time of t_{s-1} otherwise.

The above procedure can also be used for negotiating with customers with respect to due dates and price. Suppose that a customer of class i arrives to the sales department and asks for a due date quote. Assume that the firm quotes due date a_1 , and the heuristic indicates that if the customer places an order, the order should be placed in position s_1 . Let $z_1 = B_{s_1}$. However, once quoted the due date, the customer states that he/she will place an order if the company can deliver before a_2 , ($a_2 < a_1$) time units. Now, assuming that the customer will indeed place an order if the firm accepts this offer, this means that the customer has revealed his/her due date preference. Then, using $p_i(a) = 1$ if $a < a_2$, with $p_i(a) = 0$ otherwise, the modified heuristic can be used to compute the position, s_2 , for the new order. Let $z_2 = B_{s_2}$ obtained with this new probability function. If $z_2 > z_1$, then the firm should accept the customer's order. Otherwise, the firm can make a counter offer, for example, asking the customer to pay $z_1 - z_2$ more than the usual price since other orders already in the system will be delayed for the customer.

In the next section, we conduct a simulation experiment where we test the performance of the two heuristics developed in this section and of the due date setting rules in the literature which do not take into account customer preferences in setting due dates.

4 A Simulation Study

In this section, we describe the due date setting rules that we tested in our simulation study. We tested six different due date setting rules on a set of problems with two customer classes. Customer arrivals to the sales department were Poisson distributed while processing times were exponential with rate μ .

The first two heuristics that we tested were the modified heuristic and the myopic heuristic from Section 3. We compared these heuristics to four other heuristics from the literature. One problem that we have to address before comparing these different heuristics is that all the due date setting rules in the literature require the specification of a service level, either average tardiness or the proportion of late orders. Hence, to compare the heuristics in this paper with those in the literature, using simulation, we computed the average tardiness of jobs obtained by using the modified heuristic. (Of the two heuristics developed in this paper, the modified heuristic always gave better solutions) For the average

tardiness obtained by the modified heuristic, we computed, by simulation, the performance of the due date setting rules in the literature. Hence, the modified heuristic and the four due date setting rules from the literature are all compared under the same average tardiness constraint.

The first of the heuristics from the literature (CONS) quotes a constant lead time to each customer. We note that since, in this paper, we assume that the firm produces a single product, all the due date setting rules described in Baker and Bertrand would quote a constant lead time to customers. The constant lead time that results in the same average tardiness as the modified heuristic is computed by simulation.

The second heuristic from the literature that we test in this paper is a policy that depends on the congestion level in the system. This heuristic, developed by Eilon and Chowdhury, and modified by Bookbinder and Noor quotes all customers a due date of D_k when there are k orders in the system, where

$$D_k = 2\mu^{-1} + c[\mu^{-1}(k + 1)] \quad (4.12)$$

where c is the constant to be determined by simulation to satisfy the average tardiness constraint. We refer to this policy as the (STATE DEPENDENT) policy as in Wein (1991).

The final two policies that we tested are two policies that we modified from Wein (1991). For a multiclass $M/G/1$ queue where customers from different classes have different priorities, Wein derived one non-parametric and two parametric rules for setting due dates. However, we note that Wein (1991) assumes that the effective arrival rate of customer orders to the system is constant and does not depend on the due dates quoted. The first parametric rule is a function of the mean sojourn time of an order, $ES_i(k_1, k_2)$, where the mean sojourn time depends on the class of the customer as well as the number of orders from each class already in the system. This parametric rule quotes the customer the lead time $\alpha ES_i(k_1, k_2)$, where α is set by simulation. We note that the $c\mu$ rule is used in sequencing customers. Hence, the expected sojourn time of high priority customers depends only on how many high priority customers there are in the system, (along with the amount of time left until the job currently being processed is finished), while the sojourn time of low priority customers also depends on the high priority jobs that will arrive in the future. The second parametric rule quotes an arriving customer from class i , the lead time $E[S_i(k_1, k_2)] + \beta\sigma[S_i(k_1, k_2)]$ where $\sigma[S_i(k_1, k_2)]$ is the standard deviation of a class i order that joins the system when there are k_1 high priority and k_2 low priority orders already in the system, and β is set by simulation. Finally, Wein's non-parametric rule quotes a due date a to a customer from class i , where

$$\tau = \int_{x=a}^{\infty} (x - a) dS_i(k_1, k_2)$$

where τ is the average tardiness.

We chose to test modifications of the parametric rules in Wein in our simulation study. We can not use Wein's parametric rules directly, since the arrival of orders into the system in our problem depends on the lead times quoted to the customers. However, we can modify Wein's due date setting rules by taking this dependence into account. To do this, we must define which class of customers has a high priority. It is natural to assign a higher priority to those customers that are willing to pay more and wait less. We must also rederive the expected sojourn times of customers. It is clear that the sojourn time distributions for the

high priority customers are not affected by the assumptions of our problem. However, a major difficulty of deriving sojourn times for the low priority class is that the arrival rates of the high priority customers (who will be placed in front of the low priority ones) actually depend on the lead times quoted to them. Hence, to be able to carry out an analysis, we need to fix the lead times quoted to the high priority customers. One way of doing this is to quote them the lead times in Wein's non-parametric rule. That is, a high priority customer arriving when there are already k_1 high priority orders in the system is quoted lead time $a_1(k_1)$ where

$$\tau = \int_{a_1(k_1)}^{\infty} (x - a_1(k_1)) dF_{k+1}(x) \quad (4.13)$$

where τ is the average tardiness, and F_{k+1} is the Erlang-($k+1$) distribution. (Note that since orders that are currently being processed can not be preempted, we assume that a low priority order becomes high priority as soon as its processing begins). Next, we can utilize the lead times obtained in (4.13) to derive the mean and variance of sojourn times for the low priority customers. To do this, we first let $\lambda_1(k_1)$ denote the *effective* arrival rate of high priority orders to the system when there are already k_1 high priority orders in the system (i.e., $\lambda_1(k_1) = \lambda_1 p_1(a_1(k_1))$). Note that, in order to compute the sojourn time of a particular low priority order, we need only keep track of future processing completions and future arrivals of high priority orders (since other low priority orders arriving in the future will be processed after this particular low priority order). Assume that a low priority customer places an order when there are already k_1 high priority orders, and k_2 low priority orders in the system. We let $S_2(k_1, k_2)$ denote the (random) sojourn time of this order. The next event will either be a service completion or arrival of a high priority order. We denote the (random) time until the next event as $Z(k_1)$. $Z(k_1)$ is exponential with rate $\mu + \lambda_1(k)$. Let I_a denote the identity function which equals 1 if the next event is an arrival and 0 if it is the completion of an order. Then we can write

$$S_2(k_1, k_2) = Z(k_1) + I_a S_2(k_1 + 1, k_2) + (1 - I_a) S_2(k_1 - 1, k_2) \text{ for } k_1 > 1 \quad (4.14)$$

and

$$S_2(1, k_2) = Z(1) + I_a S_2(2, k_2) + (1 - I_a) S_2(1, k_2 - 1) \text{ for } k_2 > 1 \quad (4.15)$$

with

$$S_2(1, 0) = Z(1) + I_a S_2(2, 0) + (1 - I_a) S_2(0, 0) \quad (4.16)$$

and $S_2(0, 0)$ is exponential with rate μ .

Taking the expectation in (4.14), (4.15), and (4.16) gives us a system of equations for calculating mean sojourn times. Similarly, squaring both sides in (4.14), (4.15) and (4.16) and taking expectations gives us the second moment of sojourn times. Hence the first heuristic (W1) modified from Wein (1991) quotes the low priority customers the due dates $\alpha E S_2(k_1, k_2)$ and the second one (W2) quotes the due dates $E S_2(k_1, k_2) + \beta \sigma[S_2(k_1, k_2)]$, where α and β are set by simulation. Both heuristics quote the high priority customers the due dates obtained by solving (4.13).

We tested our heuristic on several different test problems. We considered a variety of different conditions such as different levels of utilization, and different types of $p_i(a)$ functions (linear, concave, convex). In this section, we report the results on 6 representative cases. For each of the test problems, we simulated the system for 2000 time units 10 times

Case	Class 1				Class 2			
	$p_1(a)$		R_1	λ_1	$p_2(a)$		R_2	λ_2
1	1	if $a < 0$	2	0.5	1	if $a < 2$	1	0.9
	$1 - a/3$	if $0 \leq a < 3$			$1 - (a - 2)/4$	if $2 \leq a < 6$		
	0	if $a \geq 3$			0	if $a \geq 6$		
2	1	if $a < 0$	2	1.0	1	if $a < 2$	1	1.5
	$1 - a/3$	if $0 \leq a < 3$			$1 - (a - 2)/4$	if $2 \leq a < 6$		
	0	if $a \geq 3$			0	if $a \geq 6$		
3	1	if $a < 0$	2	0.5	1	if $a < 0$	1	0.9
	$e^{-a/2}$	if $a \geq 0$			$e^{-a/4}$	if $a \geq 0$		
4	1	if $a < 1$	1.5	0.7	1	if $a < 4$	1	0.7
	$1 - (a - 1)^2/9$	if $1 \leq a < 4$			$1 - (a - 4)^2/9$	if $4 \leq a < 7$		
	0	if $a \geq 4$			0	if $a \geq 7$		
5	1	if $a < 1$	1.1	0.8	1	if $a < 4$	1	0.8
	$1 - (a - 1)^2/9$	if $1 \leq a < 4$			$1 - (a - 4)^2/9$	if $4 \leq a < 7$		
	0	if $a \geq 4$			0	if $a \geq 7$		
6	1	if $a < 0$	2	0.8	1	if $a < 3$	1	0.8
	$1 - a/5$	if $0 \leq a < 5$			$1 - (a - 3)/5$	if $3 \leq a < 8$		
	0	if $a \geq 5$			0	if $a \geq 8$		

Table 1: Input Data for Cases 1-6.

and averaged the total profits accumulated until time 2000 under each due date setting rule. Each simulation run took less than 5 minutes on a 486 computer. Since, the arrival of nearly 3000 customers was simulated in each run, this indicates that all the heuristics including the modified heuristic, generated due dates very rapidly. In all of our problems $\mu = 1$, and for cases 1-5, the tardiness penalty per unit time was $c = 1$. The arrival rates, the revenues and the $p_i(a)$ functions were changed in each problem. The input data for our test problems are displayed in Table 1. For each test problem, we report the total profits obtained in 2000 time units (π), the utilization (ρ), and the average tardiness per order (τ) in Table 2. We also report the standard deviation for the profits, utilization and average tardiness in parantheses.

The first test case is one with medium utilization and linear, downward sloping $p_i(a)$ curves. The arrival rate of the high revenue customers was lower than the arrival rate of the low revenue customers. Case 2 was the same as case 1 except that the arrival rates of both classes of customers were higher, resulting in higher utilization for the system. In both cases, the modified heuristic gave the best results. This is not surprising, since the modified and the myopic heuristics use more information about customer preferences than the other heuristics for setting due dates. However, note that the myopic heuristic did not do particularly well, especially in the case with higher rates of arrival. The reason for this is that since the myopic heuristic is not taking into account the increased cost of congestion, it is accepting more customers into the system, and in many cases displacing them to place a new order in front of one already in the system. For these reasons, the myopic heuristic resulted in larger average tardiness per order, and hence lower profits. The two heuristics based on Wein's work gave the next best results. This is also expected since, in their modified versions, these two heuristics are utilizing some information about

customer preferences. In particular, they are using the information on the arrival rates of the higher priority customer class at different congestion levels. We note that in all our test problems, we found no significant difference between the two heuristics (W1) and (W2).

Case 3 was similar to case 1. The $p_i(a)$ functions were changed to see the effect of convex functions. There was very little difference in the results, with the modified heuristic giving the best results. In Case 4, we tested a case with strictly concave lead time preference functions. Again, the modified heuristic gave the best results, with heuristics (W1) and (W2) giving the next best results.

In Cases 1-4, the net revenue obtained from a higher revenue order was much higher than the revenue obtained from a low revenue order. In Case 5, we tested an example where the difference between the revenues obtained from different classes of customers was low. The modified heuristic still gave the best results. However, the difference between the modified heuristic and the two heuristics (W1) and (W2) was lower in this case.

In all of the cases we have described so far, the tardiness cost was the same for all customer classes. This may be a reasonable assumption in many cases. For example, today's patient customer may have an urgent order the next time he/she places an order. If this customer experiences some tardiness now, he/she may be less likely to place an order next time with an urgent order. Similarly, the company may have a policy of reimbursing all customers at the same rate for any tardiness experienced. However, even though having the same tardiness cost for all customer classes may be a reasonable assumption, we would still like to know how the different heuristics compare under the assumption of different tardiness costs.

A major difficulty of analyzing the case with different tardiness costs is that Theorem 4 which states that customer orders are processed in an EDD manner once a due date is set for an order is not valid for the case where the different classes of customers have different tardiness costs. In particular, whereas in the case with identical tardiness costs, we need to consider the sequence of the orders in the system only when we are quoting a due date to a new customer, with different tardiness costs, it may be optimal to resequence after each service completion. We also note that the $c\mu$ rule for sequencing orders is not optimal for this case, since the $c\mu$ rule minimizes the average weighted flow time for orders, and not the tardiness costs (i.e., the $c\mu$ rule would be optimal if all orders had a due date of 0, or all orders were already late). However, the results in Wein (1991) and the cases that we have tested indicate that due date setting has a much larger impact on profits than sequencing. Hence, for this problem, we ignored the fact that it may be profitable to resequence after each service, and still considered the two heuristics described in Section 3. The only difference is that, we now keep track of not only the remaining time of each order until its due date, but also its class. Hence, in step 3 of the modified algorithm, when we compute the ψ_j values, we modify the $\varphi_j^u(t_i)$ values so that they now depend on the class of the customers displaced, (i.e., we replace $\varphi_j^u(a)$ by $\varphi_{i,j}^u(a)$ where $\varphi_{i,j}^u(a) = \int_a^\infty c_i(x-a)dF_j^u(x)$) so that displacing an order with a high tardiness penalty costs more than displacing one with a low tardiness penalty if both have the same due date. We used the modified and the myopic heuristics described in Section 3 with this slight change for the cases where the different customer classes have different tardiness costs.

Case 6 is an example where the two customer classes have different tardiness costs. Each order by a customer from class 1 brings in a net revenue of \$2, while each order from a class 2 customer brings in a net revenue of \$1. If a high revenue order is late, the tardiness cost

Case		Modified	Myopic	CONS	STATE D.	W1	W2
1	π	1484 (25)	1330 (32)	712 (44)	1019 (38)	1204 (20)	1208 (29)
	ρ	0.70 (0.01)	0.80 (0.01)	0.62 (0.02)	0.65 (0.01)	0.69 (0.01)	0.68 (0.01)
	τ	0.24 (0.01)	0.44 (0.02)	0.24 (0.01)	0.24 (0.01)	0.24 (0.01)	0.24 (0.01)
2	π	1892 (43)	1192 (29)	801 (31)	1226 (67)	1398 (50)	1407 (67)
	ρ	0.80 (0.01)	0.94 (0.01)	0.55 (0.02)	0.81 (0.02)	0.87 (0.01)	0.87 (0.02)
	τ	0.28 (0.01)	0.65 (0.03)	0.28 (0.04)	0.27 (0.03)	0.28 (0.02)	0.28 (0.03)
3	π	1608 (25)	1525 (37)	917 (34)	1093 (29)	1366 (10)	1358 (26)
	ρ	0.61 (0.01)	0.76 (0.02)	0.50 (0.02)	0.65 (0.02)	0.66 (0.03)	0.65 (0.02)
	τ	0.37 (0.02)	0.51 (0.03)	0.37 (0.03)	0.37 (0.03)	0.37 (0.01)	0.37 (0.01)
4	π	1721 (40)	1327 (35)	901 (51)	1392 (44)	1446 (43)	1439 (49)
	ρ	0.80 (0.02)	0.88 (0.02)	0.52 (0.01)	0.76 (0.02)	0.75 (0.02)	0.76 (0.02)
	τ	0.15 (0.01)	0.44 (0.02)	0.15 (0.02)	0.15 (0.02)	0.15 (0.02)	0.15 (0.01)
5	π	1568 (28)	1212 (54)	926 (55)	1408 (27)	1416 (17)	1404 (37)
	ρ	0.84 (0.01)	0.90 (0.02)	0.52 (0.01)	0.77 (0.02)	0.77 (0.02)	0.76 (0.02)
	τ	0.10 (0.01)	0.36 (0.02)	0.10 (0.01)	0.10 (0.01)	0.10 (0.01)	0.10 (0.01)
6	π	1948 (41)	1398 (64)	925 (66)	1589 (42)	1639 (36)	1645 (40)
	ρ	0.80 (0.02)	0.89 (0.02)	0.50 (0.01)	0.73 (0.01)	0.72 (0.02)	0.73 (0.02)
	τ	0.11 (0.01)	0.41 (0.03)	0.11 (0.01)	0.11 (0.01)	0.11 (0.01)	0.11 (0.01)

Table 2: Results for Cases 1-6.

is \$2 per unit time, while if a low revenue order is late, then the tardiness cost is \$1 per unit time. The results for this case were similar to the results for Cases 1-5. The modified heuristic gave the best results, resulting in nearly 20% percent higher profit than the best among the other heuristics.

These six cases are representative of our experience with the due date setting rules developed in this paper and those from the literature. Our observations can be summarized as follows:

1. The modified heuristic gave the best results, easily outperforming all other due date setting rules. The difference in profits obtained by using the modified heuristic and the other heuristics varied between 11% and 136%. The difference in performance between the modified heuristic, and the other due date setting rules become more pronounced at high utilization levels (as much as 35 % between the modified heuristic and W2 in Case 2). The computational requirements of the modified heuristic were very minor. In contrast, the myopic heuristic performed poorly demonstrating the importance of considering the effects of increased congestion if an order is accepted, when quoting due dates.
2. The two heuristics modified from Wein (1991) also performed very well. We note that these heuristics require less information about customer preferences than the modified heuristic. To implement these due date setting rules, the only information necessary is the arrival rate of higher priority orders at different congestion levels. There was very little difference between (W1) and (W2) suggesting that it may be enough to

calculate the expected sojourn times. Hence, in cases where it may be difficult to estimate the customers' lead time preferences, these two heuristics requiring minor data collection are very good alternatives to the modified heuristic.

3. The state dependent rule did not perform as well as the three rules mentioned above. However, if there is no data on customer preferences, it seems to be the best alternative.

5 Conclusions and Further Research

In this paper, we formulated the problem of quoting due dates and sequencing customer orders in a manufacturing system. We have demonstrated the importance of considering customer preferences for price and due dates in the due date setting process. We have developed a heuristic that performed very well in our simulation study relative to other due date setting rules in the literature. We suggested a way to use this heuristic in bargaining with customers with respect to due dates and price. We also found that in cases where the company does not have much information on customer due date preferences, two heuristics modified from Wein (1991) performed very well. The difference in the performance of the various heuristics is due to how explicitly each heuristic considers the dependence of demand to quoted lead times in setting due dates and sequencing orders. The results in this paper indicate that companies that collect data on how their demand depends on the lead times they quote customers and utilize this data in their lead time quoting procedures may be able to realize an important competitive advantage.

Much remains to be done in the area of setting lead times and sequencing orders. A set of research topics that seem promising are:

1. Developing an understanding of the dependence of customer demand on the quoted lead times. Considerable research has been done on characterizing the price sensitivity of demand in different industries. The recent practitioner literature indicates that in many industries demand is also sensitive to lead times. Empirical work on the dependence of demand on quoted lead times is necessary to make use of due date setting models in practical settings.
2. Incorporating the long-term consequences of failing to deliver orders on time. In this paper, we made the assumption common in the literature, that the tardiness cost captures the consequences of failing to deliver orders on time. In particular, each tardiness cost corresponds to an average tardiness so that if a firm uses average tardiness as a service constraint, a simple search can be used to find the corresponding tardiness cost. However, in some cases it may be reasonable to assume that a customer whose order is tardy several times, will have a lower probability of placing an order (or even requesting a due date quote) next time. More modeling and empirical work is necessary to address this issue. This problem appears to be difficult both due to the difficulty of collecting empirical data (e.g., is a customer whose order was late one time, and early two times previously, more or less likely to place an order?) and also due to modeling difficulties.
3. Incorporating multiple machines and/or multiple products. In this paper, we treated a simple case, with one machine and one product type. Addressing the realistic

situations of multiple machines and/or multiple products will require more complex queuing models and larger state spaces. The resulting analysis is likely, therefore, to be a challenging topic for further work.

Appendix

In this appendix, we prove Theorem 2. We first note that there exists a level of orders in the system \hat{k} such that for all $k > \hat{k}$, the optimal decision is to reject customers from any class (i.e., to quote lead times such that customers from all classes will place an order with probability 0). Obviously an upper bound on \hat{k} is $\{\sup_i \inf_k, \varphi_k, (\hat{a}_i) > R_i\}$, which is the point where the expected tardiness cost associated with any customer is larger than the revenue generated by his/her order. Hence, we can restrict our problem to a finite state space problem.

The problem of choosing optimal lead times at arrival points is equivalent to the problem of choosing arrival rates to an M/M/1 queue (since by quoting lead time a_k^i to customers from class i , we are setting their *effective* arrival rate in state k to $\lambda_i p_i(a_k^i)$). The state 0 is reachable under any stationary policy, hence by Ross (1983) a stationary policy is optimal.

We let $\pi_\sigma(k, j)$ be the expected profit between one visit from state k to state j under stationary policy σ , and $t_\sigma(k, j)$ be the expected number of customers that arrive to the system (but do not necessarily place an order) until that visit to state j . Letting g_σ be the long run average profit per customer who inquires about due dates, we have $g_\sigma = \pi_\sigma(k, k)/t_\sigma(k, k)$. If g^* is the optimal average profit per inquiring customer, then we have $\pi_\sigma(k, k) - g^* t_\sigma(k, k) \leq 0$ with equality obtained by the optimal policy σ^* . We use a result due to Stidham and Weber (1989) that for any g , the quantity $\pi_\sigma(k, k) - g t_\sigma(k, k)$ equals the total g -revised profit (the profit obtained by subtracting g from expected revenues in each state) obtained between any two visits to state k . That is, solving the problem of maximizing g^* -revised profits between any two visits to state k is equivalent to solving (2.5). Letting $\pi(k, 0)$ denote the optimal expected g^* -revised profit starting in state k until state 0 is reached for the first time, and a_k^i denote the lead time quote to class i customers in state k , by renewal-reward theory we can write:

$$\pi(k, 0) = \max_{a_k^1, \dots, a_k^n} \frac{-g^* \Lambda + \sum_{i=1}^n \{\lambda_i p_i(a_k^i)(R_i - \varphi_{k+1}(a_k^i) + \pi(k+1, 0)) + \lambda_i(1 - p_i(a_k^i))\pi(k, 0)\} + \mu \pi(k-1, 0)}{\Lambda + \mu}$$

where $\Lambda = \sum_{i=1}^n \lambda_i$. Arranging terms, and using the fact that $\pi(k, 0) = \pi(k, k-1) + \pi(k-1, 0)$, we get

$$\pi(k, k-1) = \max_{a_k^1, \dots, a_k^n} \frac{-g^* \Lambda + \sum_{i=1}^n \lambda_i p_i(a_k^i)(R_i - \varphi_{k+1}(a_k^i) + \pi(k+1, k))}{\mu}$$

It is straightforward to show by induction on k (starting from \hat{k} and continuing until 0) that $\pi(k, k-1)$ is decreasing in k . Moreover, the first-order condition for the optimality of the lead time quoted to a customer from class i in state k is

$$\lambda_i p_i'(a_k^i)(R_i - \varphi_{k+1}(a_k^i) + \pi(k+1, k)) + \lambda_i p_i(a_k^i)(c(1 - F_{k+1}^i(a_k^i))) = 0$$

Now, $\pi(k+2, k+1) \leq \pi(k+1, k)$ and $\varphi_{k+2}(a_k^i) \geq \varphi_{k+1}(a_k^i)$, hence $R_i - \varphi_{k+2}(a_k^i) + \pi(k+2, k+1) \leq R_i - \varphi_{k+1}(a_k^i) + \pi(k+1, k)$. Also, $c(1 - F_{k+2}^i(a_k^i)) \geq c(1 - F_{k+1}^i(a_k^i))$. Hence, we have

$$\lambda_i p_i'(a_k^i)(R_i - \varphi_{k+2}(a_k^i) + \pi(k+2, k+1)) + \lambda_i p_i(a_k^i)(c(1 - F_{k+2}^i(a_k^i))) \geq 0$$

Since substituting any $a < a_k^i$ will cause the left hand side to increase, we conclude that $a_{k+1}^i \geq a_k^i$. Since the problem of maximizing expected g -revised profits until reaching state 0 is equivalent to problem (2.5), Theorem 2 has been proven. \square

6 References

- Baker, K., "Sequencing Rules and Due-Date Assignments in a Job Shop," *Management Science* **30**, (1984), 1093-2004.
- Baker, K., and J. Bertrand, "A Comparison of Due-Date Selection Rules," *AIIE Transactions* **13**, (1981), 123-131.
- Baker, K., and J. Bertrand, "An Investigation of Due-Date Assignment Rules with Constrained Tightness," *Journal of Operations Management* **3**, (1982), 109-120.
- Bertrand, J., "The Effect of Workload Dependent Due-Dates on Job Shop Performance," *Management Science*, **29** (1983), 799-816.
- Bertrand, J., "The Use of Workload Information to Control Job Lateness in Controlled and Uncontrolled Release Production Systems," *Journal of Operations Management*, **3**, (1983), 67-78.
- Blackburn, Joseph, *Time Based Competition*, Richard D. Irwin, Homewood, Illinois, (1990).
- Bookbinder, J.H., and A.I. Noor, "Setting Job-Shop Due Dates with Service Level Constraints," *Journal of the Operational Research Society*, **36**, (1985), 1017-1026.
- Duenyas, I., and W.J. Hopp, "Quoting Customer Lead Times," Technical Report 91-24, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, MI 48109-2117 (1992).
- Eilon, S., and I.G. Chowdhury, "Due Dates in Job Shop Scheduling," *International Journal of Production Research*, **14**, (1976) 223-237.
- Ross, S., *Introduction to Stochastic Dynamic Programming*, Avademic Press, New York, 1983.
- Schmenner, Roger W., "The Merit of Making Things Fast," *Sloan Management Review*, Fall Quarter, 11-17, (1988).
- Seidmann, A., and M. Smith, "Due Date Assignment for Production Systems," *Management Science* **27**, (1981), 401-413.
- Shanthikumar, J., and U. Sumita, "Approximations for the Time Spent in a Dynamic Job Shop with Applications to Due Date Assignment," *International Journal of Production Research* **26**, (1988), 1329-1352.
- Stalk, George and Thomas M. Hout, *Competing Against Time: How Time- Based Competition is Reshaping Global Markets*, The Free Press, New York, (1990).
- Stidham, S., and R. Weber, "Monotonic and Insensitive Optimal Policies for Control of Queues with Undiscounted Costs," *Operations Research*, **37**, (1989), 611-625.
- Thomas, Philip R., *Competitiveness Through Total Cycle Time: An Overview for CEO's*, McGraw-Hill, New York, (1990).



Thomas, Philip R., *Getting Competitive: Middle Managers and the Cycle Time Ethic*, McGraw-Hill, New York, (1991).

Vepsalainen A., and T. Morton, "Priority Rules for Job Shops with Weighted Tardiness Costs," *Management Science*, **33**, (1987) 1035-1047.

Vepsalainen, A., and T. Morton, "Improving Local Priority Rules with Global Lead Time Estimates," *Journal of Manufacturing and Operations Management* **1**, (1988), 102-118.

Wein, L.M., "Due-Date Setting and Priority Sequencing in a Multiclass M/G/1 Queue," *Management Science* **37**, (1991), 834-850.

Wein, L.M., and P. Chevalier, "A Broader View of the Job-Shop Scheduling Problem," to appear in *Management Science*.

Wein, L.M., S. Whang, and L.J. Lemire, "Due Date Setting and Pricing in a Single-Server Queue," Sloan School of Management, M.I.T., Technical Report, September 1990.