Combined Make-to-Order/Make-to-Stock Supply Chains

Philip Kaminsky
Dept. of Industrial Engineering and Operations Research, UC Berkeley
Onur Kaya
Faculty of Industrial Engineering, Koc University

October 2006
Revised March 2007, July 2007

Abstract

We consider a multi-item manufacturer served by a single supplier in a stochastic environment. The manufacturer and the supplier have to decide which items to produce to stock and which to produce to order. The manufacturer also has to quote due dates to arriving customers for make-to-order products. The manufacturer is penalized for long lead times, missing the quoted lead times and high inventory levels. We consider several variations of this problem, and design effective heuristics for the make-to-order/make-to stock decision, and to find the appropriate inventory levels for make-to-stock items. We also develop scheduling and lead time quotation algorithms for centralized and decentralized versions of the model. We perform extensive computational testing to assess the effectiveness of our algorithms, and to compare the centralized and decentralized models in order to quantify the value of centralized control in this supply chain. As centralized control is not always practical or cost-effective, we explore the value of limited information exchange for this system.

1 Introduction

Inventory costs make up a large portion of total costs in many supply chains, so effective inventory management is one of the most important issues facing supply chain managers. Indeed, in a multi-facility supply chain, a critical tactical decision involves the identification of intermediate stocking points in the supply chain, and the determination of appropriate inventory strategies at these stocking points. Managers must determine, in other words, which products (or components of products) will be made to stock, and which will be made to order. For those products or components that are made to stock, appropriate inventory strategies must be determined, and for those products that are made to order, approaches for customer lead time quotation must be developed. Furthermore, if different products or components utilize common production resources, operating strategies for those resources will impact system performance.

Traditionally, most companies utilized a “push” or make-to stock (MTS) system, holding inventory at the end of the supply chain. However, in an MTS system, firms need to be able to estimate demand to determine how much to produce and stock, and so these systems rely heavily on forecasts, which in many cases are not very accurate. Thus, many progressive companies have shifted to “pull” or make-to-order (MTO) systems, holding no inventory at all and producing to order. In these systems, companies produce based on actual customer demand instead of forecasts. Inventories are eliminated, but customers must now wait for delivery, perhaps leading to loss of competitiveness on the part of the firm. The decision to use either a push strategy or a pull strategy for a particular product therefore depends heavily on the characteristics of the system. Indeed, in a supply chain, using a push strategy for some products or components, and a pull strategy for others, might be much more effective then using either system exclusively. Because of this, firms are beginning to employ a hybrid approach, a “push-pull” strategy, or combined MTO-MTS system, holding inventory of some components, and producing others to order.

In addition to inventory decisions, the scheduling of the production of orders and the approach to lead time quotation to customers also have significant impact on the performance of supply chains, particularly in MTO supply chains. Companies need to quote short and reliable lead times to their customers to remain competitive in the market and to increase their profits. For a company that produces multiple products with different characteristics, the decision on when to produce each order affects the completion time of manufacturing and thus the lead time for that product. Thus, firms need a policy or approach that will help them to decide which items or components to produce to stock and which items or components to produce to order, what inventory levels to maintain for make-to-stock items, how to quote lead times to customers, and how to sequence orders to efficiently use limited resources. In this paper, we explore approaches to these four inter-related decisions for a two stage supply chain.

To the best of our knowledge, these four inter-related decisions have never been explored simultaneously in the context of supply chains. However, most of the components of these decisions have been considered separately or in pairs. MTO/MTS models have generally been studied for single stage systems. Williams [33], Federgruen and Katalan [12], Carr and Duenyas [5], and Youssef, Van Delft, and Dallery [34] assume that decisions regarding which items to produce-to-order and which to stock are made exogenously from the model, and they try to find the best way to operate the system given these decisions. In contrast, Li [21], Arreola-Risa and DeCroix [1] and Rajagopalan[26] all design models that focus on this decision as well as on associated inventory level decisions.

Researchers have introduced a variety of models in an attempt to understand effective due date quotation and sequencing. Keskinocak and Tayur[23], Kaminsky and Hochbaum [19], and Cheng and Gupta [7] survey due date quotation models in detail. The majority of earlier papers on due-date quotation are simulation based. For instance, Eilon and Chowdhury [11], Weeks [31], Miyazaki [24], Baker and Bertrand [2], and Bertrand [3] consider various due date assignment and sequencing policies, and in general demonstrate that policies that use estimates of shop congestion and job content information lead to better shop performance than policies based solely on job content. Some analytical results do exist for limited versions of these models. Primarily, these consist of deterministic, common due date models, where a single due date must be assigned for all jobs, and static models, where all jobs are available at time 0. For these simplified models, a variety of polynomial algorithms have been developed (see, for example, Brucker[4], Kahlbacher[15], Panwalkar et al.[25], Hall and Posner [13], Seidmann et al.[28], and Chand and Chhajed [6]); however, these results don't extend in an obvious way to more complex models.

A series of papers do use a variety of different approaches to consider more complex models. Kaminsky and Lee [16] derive analytical results for a model in which a series of jobs arrive at the system over time and in which all orders must be accepted. In other models, not all potential jobs must be accepted. Keskinocak, Ravi, and Tayur [22] consider a model in which the firm can decide whether or not to accept an order for the case in which revenue decreases with quoted lead time. Some researchers approach lead time quotation models within a queuing theoretic framework. For example, Wein [32] considers a multiclass M/G/1 queuing system under the objective of minimizing the weighted average lead time subject to the constraints of the maximum fraction of tardy jobs and the maximum average tardiness, Duenyas and Hopp [10] and Duenyas [9] consider models in which customers probability of placing an order decreases with increasing lead time, and Savasaneril, Griffin, and Keskinocak [27] also consider the possibility of holding inventory to reduce quoted lead times, thus increasing the probability that customers will place an order.

In Kaminsky and Kaya [18], we analyze pure MTO supply chains and design effective scheduling and due-date quotation algorithms for the centralized and decentralized versions of those systems. We show that these algorithms are probabilistically asymptotically optimal(i.e. the relative error of these heuristics when compared to the optimal solution goes to zero as $n \to \infty$) for the objective of minimizing $Z_n = \sum_{i=1}^{n} (c^d d_i + c^T T_i)$ where d_i is the quoted due-date for job i, $T_i = (C_i - d_i)^+$ is the tardiness of job i and c^d and c^T are the unit due date and tardiness costs for the model.

In this paper we integrate due-date quotation with combined MTO-MTS decision-making, consider several different approaches for sequencing jobs, and focus on two-stage supply chain models. Our analysis provides guidance for deciding when to use MTS and when to use MTO approaches in supply chains, and how to effectively operate the systems to minimize system wide-costs. We also quantify the value of centralization and information in these systems by building both decentralized and centralized models, obtaining good solutions to both of these classes of models, and designing computational experiments to explore the effectiveness of our algorithms and to compare the centralized and decentralized systems. We focus on finding answers to questions such as:

1. Which items should be produced MTO and which ones MTS at each step of the supply chain and what are the optimal levels of inventory for MTS items?
2. Which item should be produced next when a facility becomes available for production?
3. What due-date should be quoted to each customer at the time of arrival?
4. What is the benefit of a centralized supply chain as opposed to a decentralized system?
5. How much gain can be achieved through information exchange between supply chain members?

2 Models

We consider a serial supply chain with two stages, a supplier and a manufacturer. A stream of customer orders arrives at the manufacturer over time and the manufacturer quotes a lead time d_j for each order j when it arrives. In this context, the lead time represents the time until the order is expected to complete processing. An order can be for one of K distinct job types, each with different associated costs. Orders arrive at rate λ and each arriving order is for job type i with probability δ_i, $i = 1, 2, ...K$. To facilitate our analysis, we assume exponentially distributed, stationary and independent inter-arrival times, so each order for job type i arrive at rate $\lambda_i = \lambda\delta_i$. To start processing each order, the manufacturer requires a component specific to that particular job type manufactured by the supplier. The component for job type i requires a random processing time with mean μ_i^s at the supplier, and the order requires a random processing time with mean μ_i^m at the manufacturer. In general, inventory of some or all of the job types can be held by the manufacturer, and inventory of some or all of the component types can be held by the supplier. We assume no delivery lead time between the supplier and the manufacturer, so if the manufacturer needs to process a particular type and no inventory of the component of that type exists at the supplier, the manufacturer can begin processing immediately after the supplier has processed the necessary component.

Our objective is to minimize the total expected inventory, lead time, and tardiness cost in this system. Thus,

$$Z = \sum_{i=1}^{K} \{h_i E[I_i] + c_i^d E[d_i] + c_i^T E[W_i - d_i]^+\} \tag{2.1}$$

where h_i is the unit inventory holding cost of job type i, c_i^d is the unit lead time cost of job type i and c_i^T is the unit tardiness cost for job type i. $E[I_i]$ denotes the mean amount of inventory, $E[d_i]$ denotes the mean lead time of job, and $E[W_i - d_i]^+$ denotes the mean tardiness of job type i.

In general, the optimal inventory policy might be state dependent and quite complex. However, to facilitate our analysis, and consistent with traditional analysis of related models (see, for example, Arreola-Risa and DeCroix [1] and the references cited therein), we assume that a base-stock policy is used for inventory control of MTS items, so that inventory of job type i at the manufacturer is initially R_i, and whenever an order of type i arrives at the manufacturer, a production order is sent to replenish inventory. If $R_i = 0$, a make-to-order production system is employed for job type i. If demand is higher than the inventory level, then the extra demand is backlogged and satisfied later when it is produced. We also assume that orders arrive sufficiently slowly that the system is stable, that is we assume that for the manufacturer: $\sum_{i=1}^{K} \frac{\lambda_i}{\mu_i^m} < 1$, and for the supplier: $\sum_{i=1}^{K} \frac{\lambda_i}{\mu_i^s} < 1$.

We consider three versions of this model, which we briefly describe here and discuss in more detail in subsequent paragraphs. We analyze these three model versions in order to develop insight into their effective operation, and compare these three models to generate insights into the value of centralization. In the centralized version of the model, the entire system is operated by a single manager, who is aware of the inventory levels and processing times at both the supplier and the manufacturer. This manager decides on the inventory levels for each class of jobs, as well as the production schedule and the lead times that should be quoted to each customer. In the decentralized full information model, the manufacturer and the supplier are assumed to work independently, but the manufacturer has full information about both his processes and those of the supplier. Both parties attempt to minimize their own costs in a sequential manner. The supplier acts first to determine his optimal inventory levels and then the manufacturer acts to minimize his own costs using the optimal inventory levels for the supplier. Finally, in the simple decentralized model, the manufacturer and the supplier work independently from each other, but the manufacturer has very limited information about the supplier's status.

The objective of the manager or managers in each of these systems is to minimize holding, lead time, and tardiness costs, and to do this, due date quotation, sequencing, and inventory management have to be coordinated. Ideally, this would require simultaneous consideration of these three issues, but this is intractable. Thus, the approach we have elected to follow for this model (and throughout the paper) is slightly different. Observe that if management were able to quote lead times that precisely matched the waiting time of orders in the system, the objective would simplify to minimizing $\sum_{i=1}^{K}\{h_i E[I_i] + c_i^d E[W_i]\}$. Although it is not possible to be precisely accurate with lead time quotations in this system, this observation motivates our approach: we first determine a scheduling approach designed to reduce the sum of waiting times, and then based on that schedule, we find the optimal inventory levels to minimize $\sum_{i=1}^{K}\{h_i E[I_i] + c_i^d E[W_i]\}$ and design a due date quotation approach that presents due dates that are generally close to the completion times suggested by our scheduling approach. We note that this type of approach was applied in Kaminsky and Kaya[18] for pure MTO versions of this system.

In the next section, we introduce a preliminary single facility model and analyze this model. In Section

4, we present our supply chain models, algorithms, and results in detail, and in Section 5, we present our computational analysis.

3 Single Facility Model

3.1 The Model

Although our ultimate goal is to analyze multi-facility systems, we begin with a preliminary analysis of a single facility system. We focus on finding the optimal inventory levels for this system and the conditions under which an MTO strategy or an MTS strategy is preferred for this facility, as well as on designing effective scheduling and due date quotation heuristics for this system. We focus on a single facility identical to the manufacturer-supplier system described above, except that the manufacturer does not need to obtain a component from the supplier. This model can be considered a special case of the supply chain model described above where the processing times and costs at the supplier are all zero, so that the components are available to the manufacturer as soon as an order arrives. In other words, a single manufacturer faces a stream of customer orders that arrive over time and the manufacturer quotes a lead time d_j for each order j when it arrives. There are K distinct job types, each with different associated costs. Orders arrive at rate λ and each arriving order is for job type i with probability δ_i, i,$i = 1, 2, ..., K$. Inter-arrival times are exponentially distributed, stationary and independent, and each order of type i requires a random processing time (not known until processing is complete) with mean processing time μ_i^m. The facility can utilize both make-to-stock and make-to-order policies to minimize inventory and lead time related costs. Our goal is to find an effective operating policy to minimize these costs.

An operating policy for this problem consists of values for base-stock levels R_i for each job type as well as approaches for sequencing and lead time quotation. Recall that we will implement a three-phase approach to this problem. First, we will develop a sequencing approach, and then we will develop a due date quotation approach and an inventory setting approach. To analyze this system, we divide jobs in process and waiting to be processed into two classes, customer jobs, representing actual orders, and replenishment jobs, representing production orders whether they exist to replenish inventory or to meet external demand. Whenever an order arrives at the system, if that item exists in inventory, the demand is immediately satisfied from the inventory. Since that order is immediately satisfied, we don't add a customer job to the system, but we do add a replenishment job to the system, since it must be manufactured to replenish inventory. On the other hand, when an order arrives and there isn't any inventory of that item, a customer job is placed in the production queue, and a lead time is quoted. Thus, at any time the production queue may contain customer jobs representing currently unsatisfied orders, and replenishment jobs that are going to be produced to replenish inventory. We collectively refer to customer and replenishment jobs in the production queue as jobs in the production queue.

Observe that the production queue operates as in a pure make-to-order system because jobs are placed in the production queue to replenish inventory even though inventory is on hand to meet an order. Thus, the inventory level of an item (whether positive or zero) does not impact the production process, but does decrease the due date costs since we satisfy those orders immediately.

3.2 Analysis and Results

As discussed above, our approach to this problem starts with a sequencing rule. Motivated by the effectiveness of the SEPTA rule (the rule of choosing the job with the Shortest Expected Processing Time among Available jobs) to minimize the total completion times for the pure MTO systems analyzed in Kaminsky and Kaya [18], we sequence jobs in the production queue according to the SEPTA rule. Under this rule, each time a job completes processing, the shortest available job (in the expected sense) that has yet not been processed is selected for processing. To quote a due date, we utilize a modified version of the approach initially introduced in [18]. When an order j for job type i arrives at the system at time r_j, the following lead time d_j is quoted.

$$d_j = \begin{cases} 0 & if \;\; I_i > 0 \;\; at \;\; r_j \\ E[p_i] + E[M_j] + \frac{E[M_j]\lambda\psi_i\tau_i}{1-\lambda\psi_i\tau_i} & otherwise \end{cases}$$

Note that in general, if order j cannot be filled from inventory, some job in the production queue, job k, will be used to satisfy that order. So, M_j is the workload in front of job k at the time of arrival (that is, the total processing time of jobs to be processed before job k), ψ_i is the probability that an arriving job has processing time less than p_i and $\tau_i = E[p|p < p_i]$ is the expected processing time of a job given that it is less than p_i.

This lead time quotation rule is designed so that the following lemma is true:

Lemma 1. *For the model described above, for every job class i, the expected value of the quoted lead time for that class i is equal to the expected waiting time of a class i order to filled. In other words, for all i*

$$E[d_i] = E[W_i]$$

where W_i denotes the actual waiting time of order type i.

Proof. For any order j of class i, at the time of arrival r_j, if there is inventory at hand, then the lead time is 0 and the actual waiting time is also 0. However, if there is no inventory of that type of product on hand at the time of the arrival, then order j has to wait for job k, the job in the production queue that will be used to satisfy order j to complete processing. Thus, order j has to wait for the jobs in front of job k, plus all the jobs that arrive to the system before job k starts processing that have processing time less than $E[p_i]$. Note that, employing the standard terminology of queueing theory, the delayed busy period of a job k, at the time job k is added to the system when a workload of M_j already exists in the system, is defined as the amount of time spent processing jobs until the server becomes idle. So, the length of the delayed busy period consists of M_j plus the length of all the jobs that arrived while the server is busy. The expected length of a delayed busy period assuming Poisson arrivals is given in Conway et al. [8] as $E[T_j] = E[M_j]/(1-\rho)$ where $\rho = \lambda E[p]$. In our model, when an order arrives at the system and there is a workload of M_j in front of the corresponding job, this order needs to wait M_j plus all the jobs that are shorter than job k that arrive while job k is still in the queue. So, we can model this system as a delayed busy period model where the arrival rate is $\lambda\psi_i$, the arrival rate of jobs shorter than job k, and the expected processing time of such a job is τ_i. We don't consider the jobs longer than k since they are scheduled after k and don't impact its waiting time. Then the expected waiting time of order j is equal to $E[p_i]$ plus the delayed busy period of job k.

$$E[W_j] = E[p_i] + \frac{E[M_j]}{(1-\lambda\psi_i\tau_i)} = E[p_i] + E[M_j] + \frac{E[M_j]\lambda\psi_i\tau_i}{1-\lambda\psi_i\tau_i} = d_j$$

Since due dates are quoted such that $d_j = E[W_j]$ for all j, we conclude that for all types i, $E[d_i] = E[W_i]$ ∎

We call this sequencing/due date quotation approach SEPTA-LTQ, and based on the resulting schedule, we determine the optimal inventory levels for each class. We note, however, that our analysis is amenable to modification if different sequences are considered. In particular, we require the stationary distribution of the number of jobs of each type in the system, and this depends on the sequencing rule.

To find the optimal inventory levels, we minimize the following objective function:

$$\sum_{i=1}^{K}\{h_i E[I_i] + c_i^d E[W_i]\} \tag{3.1}$$

where $E[I_i]$ is the expected inventory level for job type i and $E[W_i]$ is the expected waiting time for job type i.

There is an obvious tradeoff between the inventory costs and the waiting time costs in this function. We can decrease the waiting times of class i jobs by holding additional inventory of that type but doing so will increase inventory costs. Observe that holding additional inventory for an item does not affect the production queue. Thus, the waiting times and the lead time quotation procedure for the other classes aren't impacted by inventory decisions.

Lemma 2. *Objective function* (3.1) *is equivalent to the following function:*

$$\sum_{i=1}^{K}\{h_i E[I_i] + c_i^d E[W_i]\} = \sum_{i=1}^{K}\{h_i E[I_i] + c_i E[N_i]\} = \sum_{i=1}^{K} h_i \sum_{x=0}^{R_i}(R_i - x) f_i(x) + c_i \sum_{x=R_i}^{\infty}(x - R_i) f_i(x) \tag{3.2}$$

where $f_i(x)$ denotes the probability of having x jobs of type i in the production queue, $c_i = c_i^d/\lambda_i$ and $E[N_i]$ is the expected number of orders for job type i waiting for production (or equivalently, the expected number of customer jobs of type i in the production queue).

Proof. By Little's Law, we can write $E[W_i] = E[N_i]/\lambda_i$ and the objective function becomes:

$$\sum_{i=1}^{K}\{h_i E[I_i] + c_i^d E[N_i]/\lambda_i\} = \sum_{i=1}^{K}\{h_i E[I_i] + c_i E[N_i]\} \tag{3.3}$$

Then, observe that, at any time, if we have x items of job type i in the production queue, we have $x - R_i$ customer jobs in the production queue if $x \geq R_i$, and thus we have waiting costs for those items. However, if $x \leq R_i$, then it means having $R_i - x$ units of inventory left at hand. Thus, $E[I_i] = \sum_{x=0}^{R_i}(R_i - x)f_i(x)$ and $E[N_i] = \sum_{x=R_i}^{\infty}(x - R_i)f_i(x)$. ▪

An expression of the form (3.3) was considered in Arreola-Risa and DeCroix [1] for a different model, and they derive results similar to those presented below for a FCFS queueing discipline. Below, we rederive their results for completeness and clarity of exposition using the notation of our model, and extend this analysis to SEPTA schedules.

Lemma 3. *Optimal inventory levels for each job type can be found by decomposing the problem into K subproblems, one for each type.*

Proof. Since the objective function (3.2) is written in terms of the functions related to the production queue only and since the production queue isn't affected by inventory levels, expected waiting time for a job type i only depends on the inventory level of that class. Thus, the inventory levels of the items other than class i have no effect on the distribution function $f_i(x)$, so we can divide the problem into K subproblems and solve each of them separately. ▪

The problem therefore reduces to minimizing

$$h_i \sum_{x=0}^{R_i}(R_i - x)f_i(x) + c_i \sum_{x=R_i}^{\infty}(x - R_i)f_i(x)$$

for each i.

Theorem 1. *The optimal level of inventory R_i is the minimum value $x \geq 0$ that satisfies*

$$F_i(x) \geq \frac{c_i}{c_i + h_i}.$$

Proof. Observe that the objective function

$$Z_i(R_i) = h_i E[I_i] + c_i E[N_i] = h_i \sum_{x=0}^{R_i} (R_i - x) f_i(x) + c_i \sum_{x=R_i}^{\infty} (x - R_i) f_i(x)$$

is exactly the discrete version of the newsboy problem and the optimal value of R_i will be the minimum value $x \geq 0$ that satisfies $F_i(x) \geq \frac{c_i}{c_i+h_i}$. ∎

Also, it follows that:

Corollary 1. *It is optimal to produce item i MTO if and only if* $F_i(0) \geq \frac{c_i}{c_i+h_i}$.

Corollary 2. *A job type's inventory level decreases (and thus it moves towards MTO) if its unit lead time cost, processing time or arrival rate decreases or its unit holding cost increases.*

Note that these results are not restricted to single server queues or the SEPTA schedule, and they are independent of the arrival or manufacturing process. However, these characteristics affect $F_i(x)$, the stationary distribution of the number of items of type i in the system. To assess the effectiveness of our scheduling algorithm and to explore how the scheduling approach impacts $F_i(x)$ and the objective function, we analyze two different schedules, SEPTA and FCFS, and present the following two results for these schedules. We employ these results in our computational analysis in Section 5.

Corollary 3. *If the FCFS scheduling rule is used in the production queue, and we have an M/G/1 queue, to minimize* (3.2) *it is optimal to produce product i MTO if and only if:*

$$\sum_{j=1}^{K} \delta_j E[e^{-\lambda_i \mu_j}] \leq \frac{(1-\delta_i) r_i}{r_i - (1-\rho)\delta_i}$$

where $r_i = \frac{c_i}{c_i+h_i}$, $\delta_i = \frac{\lambda_i}{\lambda}$ *and* $\rho = \sum_{i=1}^{K} \lambda_i \mu_i$.

Proof. For the FCFS scheduling rule, the pgf of the stationary distribution of the number in the system is given in Wallstrom [30] as:

$$E[z_1^{x_1} ... z_K^{x_K}] = (1-\rho) \frac{\sum_{i=1}^{K} \lambda_i (1-z_i) \beta_i\{L\}}{\sum_{i=1}^{K} \lambda_i [\beta_i\{L\} - z_i]} \tag{3.4}$$

where $L = \sum_{i=1}^{K} \lambda_i (1-z_i)$ and $\beta_i\{.\}$ is the Laplace-Stieltjes transform of the service time distribution for class i. Arreola-Risa and DeCroix [1] prove that

$$f_i(0) = \frac{(1-\rho)\lambda_i}{\lambda(1 - 1/\sum_{j=1}^{K} \delta_j E[e^{-\lambda_i \mu_j}]) + \lambda_i / \sum_{j=1}^{K} \delta_j E[e^{-\lambda_i \mu_j}]} \tag{3.5}$$

and $\lambda = \lambda_i/\delta_i$. Combining these results and employing Corollary 1 gives the desired result. ∎

Corollary 4. *For a class i, let a represent the set of job types with expected processing times less than that of type i, and let b represent the set of job types with expected processing times greater than that of type i. Then, if the SEPTA scheduling rule is used in the production queue, it is optimal to produce product type i to order if and only if:*

$$\frac{(1-\rho)(\lambda_i+\lambda_a-\lambda_a\nu_a(\lambda_i))+\lambda_b(1-\gamma_b(\lambda_i+\lambda_a-\lambda_a\nu_a(\lambda_i))}{\lambda_i\gamma_i(\lambda_i+\lambda_a-\lambda_a\nu_a(\lambda_i))}\geq r_i$$

where $r_i=\frac{c_i}{c_i+h_i}$, $\rho=\sum_{i=1}^{K}\lambda_i\mu_i$, $\lambda_a=\sum_{j\in a}\lambda_j$ is the total arrival rate of class a jobs, $\lambda_b=\sum_{j\in b}\lambda_j$ is the total arrival rate of class b jobs, $\gamma_i(z)=E[e^{-zp_i}]$ is the Laplace transform associated with the processing time of class i and $\nu_a(z)$ is the Laplace transform associated with the length of a busy period in which only class a jobs are processed (which can be found by solving the equation $\nu_a(z)=\gamma_a(z+\lambda_a-\lambda_a\nu_a(z))$).

Proof. Assume that the waiting time of a class i job in the system is w_i. If we assume that the jobs of different classes are scheduled according to SEPTA and the jobs of the same class are scheduled according to FCFS, then the number of class i jobs seen by a class i job at the time of its departure is equal to the number of class i jobs that arrived during w_i. Thus, the number of class i jobs seen at the time of the departure of a class i job, given that job has waited for w_i in the system, is poisson distributed with rate $\lambda_i w_i$. When we write the probability generating function, we get:

$$\xi(x_i)=E[z^{x_i}]=\int_{w_i=0}^{\infty}e^{-\lambda_i w_i}\sum_{n=0}^{\infty}z^n\frac{(\lambda_i w_i)^n}{n!}df_{SEPTA}(W_i)=E[e^{-(\lambda_i-\lambda_i z)w_i}]=\tilde{F}^{w_i}_{SEPTA}(\lambda_i(1-z)) \quad (3.6)$$

where $\tilde{F}^{w_i}_{SEPTA}(z)$ is the Laplace transform of the waiting time distribution of class i jobs under SEPTA schedule which is given in Conway et al.[8] as:

$$\tilde{F}^{w_i}_{SEPTA}(z)=\frac{(1-\rho)(z+\lambda_a-\lambda_a\nu_a(z))+\lambda_b(1-\gamma_b(z+\lambda_a-\lambda_a\nu_a(z))}{\lambda_i\gamma_i(z+\lambda_a-\lambda_a\nu_a(z))-\lambda_i+z}.$$

Then, the probability generating function of the number in the system for class i will be:

$$\xi_i(z)=\frac{(1-\rho)(\lambda_i-\lambda_i z+\lambda_a-\lambda_a\nu_a(\lambda_i-\lambda_i z))+\lambda_b(1-\gamma_b(\lambda_i-\lambda_i z+\lambda_a-\lambda_a\nu_a(\lambda_i-\lambda_i z))}{\lambda_i\gamma_i(\lambda_i-\lambda_i z+\lambda_a-\lambda_a\nu_a(\lambda_i-\lambda_i z))-\lambda_i z} \quad (3.7)$$

Using $\xi_i(0)=f_i(0)$ and combining this with Corollary 1 gives the result. ■

Similarly, the optimal inventory levels for MTS items can be obtained using the pgf as given above and employing Theorem 1.

4 Supply Chain Models

Building on our analysis in Section 3, we analyze the inventory decisions, scheduling and due-date quotation issues for two-stage supply chains. As we did for the single stage system, we develop effective heuristics to determine optimal inventory levels at both facilities and design effective algorithms for scheduling and due-date quotation for both the centralized and decentralized versions of these systems. These algorithms allow us to compare the value of centralization and information exchange in supply chains under a variety of different conditions.

Observe that when a manufacturer is working with a supplier, required components obtained from this supplier may not be immediately available to the manufacturer at the time an order arrives at the manufacturer. If this is the case, the manufacturer has to wait for these components before initiating production of that order. Thus, the supplier-manufacturer relationship impacts inventory levels, as well as scheduling and lead time quotation decisions.

We model this system as a two facility serial supply chain with a manufacturer and a supplier where both parties can choose to stock some of the items and use a make-to-order strategy for the others in a multi-item, stochastic environment. As before, we assume that the supplier and the manufacturer employ a one-to-one replenishment strategy and a base-stock policy for inventory control of their items. The manufacturer starts with an inventory of R_i^m units of finished goods and the supplier starts with an inventory of R_i^s units of semi-finished goods that the manufacturer needs to complete his production.

We again define production replenishment jobs as we did for the single facility model. When an order arrives, if that item is in the manufacturer's inventory, the order is immediately satisfied. At the same time, a replenishment job of that class is started in the system, although the order is not placed in the manufacturer's production queue until the component required for that job is received from the supplier. If the supplier has an inventory of the appropriate component, it is sent immediately to the manufacturer, and so the replenishment job appears immediately in the manufacturer's production queue. If the supplier does not have a positive inventory level of that component, the supplier must process the component before sending it to the manufacturer, at which point the component appears in the manufacturer's production queue. Note that regardless of whether or not the component is in inventory, the supplier places that job in his production queue, either to ship to the manufacturer or to replenish the supplier inventory.

If that item is neither in the manufacturer's nor the supplier's inventory, then a lead time is quoted to the customer and a customer job order is sent to both facilities to satisfy this order. This customer job appears immediately in the supplier's production queue, and after it is delivered from the supplier, it appears in the manufacturer's production queue.

If the item is in the supplier's inventory but not in the manufacturer's inventory, then a customer job immediately appears in the manufacturer's production queue, and a replenishment job is placed in the supplier's production queue. A shorter lead time is quoted in this case since the customer only needs to wait for the production at the manufacturer, and waiting time at the supplier is 0.

Let x_i^s and x_i^m denote the number of jobs of type i in the supplier's and manufacturer's production queues, respectively. Then, let N_i^s denote the number of jobs of class i waiting in the supplier's production queue that should be delivered to the manufacturer immediately upon completion, I_i^s denote the amount of components inventory of class i, N_i^m denote the total number of customers of class i waiting in the system for their orders to be filled and I_i^m denote the amount of finished goods inventory at the manufacturer of class i. Then,

$$\begin{aligned}
N_i^s &= max\{x_i^s - R_i^s, 0\} \\
I_i^s &= max\{R_i^s - x_i^s, 0\} \\
N_i^m &= max\{x_i^m + max(x_i^s - R_i^s, 0) - R_i^m, 0\} \\
I_i^m &= max\{R_i^m - x_i^m - max(x_i^s - R_i^s, 0), 0\}
\end{aligned} \tag{4.1}$$

In the following subsections, we develop heuristics for the centralized and decentralized supply chains for scheduling, inventory and lead time quotations. Conceptually, each of our heuristics has three main steps:

Step 1: Determine the scheduling rule to be used in each facility and find the stationary distributions of the number of jobs in the system.

Step 2: Based on the stationary distributions and the objective function to be minimized, determine the base-stock inventory levels.

Step 3: Quote lead times to the arriving customers based on the schedule used, the inventory levels and the current state of the system at the time of the arrival.

4.1 The Centralized Supply Chain Model

4.1.1 The Model

When the manufacturer and the supplier belong to the same firm, the system can be modeled with a central agent that has complete information about both stages and makes all the decisions about both stages – clearly, this type of centralized control will be more effective in minimizing the total costs in the system. In this section, we consider this type of model, in which we assume that the manufacturer and the supplier work as a single entity and they are both controlled by the same agent that has all the information about both sides. The decisions about the scheduling at both facilities and lead time quotation for the customer as well as the inventory levels for each party are made by this agent.

In the centralized model, our objective function to minimize is:

$$\sum_{i=1}^{K} \{h_i^s E[I_i^s] + h_i^m E[I_i^m] + c_i^d E[d_i] + c_i^T E[W_i - d_i]^+\} \tag{4.2}$$

where h_i^s is the unit holding cost of semi-finished goods at the supplier and h_i^m is the unit holding cost of finished goods at the manufacturer.

4.1.2 Analysis and Results

In this system, we use an approach similar to the one we used for the single facility case. We first find the optimal inventory levels for any schedule that is independent of the workload or inventory levels in the system, in order to minimize the objective function

$$Z(R^s, R^m) = \sum_{i=1}^{K} \{h_i^s E[I_i^s] + h_i^m E[I_i^m] + c_i^d E[W_i]\}. \tag{4.3}$$

Then, we present an effective scheduling algorithm consistent with this model with the goal of minimizing the total waiting times of the jobs in the system and a lead time quotation algorithm that matches these waiting times.

Using the definitions (4.1) and the equations $c_i^d = \lambda c_i$ and $E[W_i] = E[N_i^m]/\lambda_i$ due to Little's law, we rewrite the objective function in terms of inventory levels and jobs in the production queues:

$$\begin{aligned} Z(R^s, R^m) &= \sum_{i=1}^{K} Z(R_i^s, R_i^m) = \sum_{i=1}^{K} \{h_i^s E[I_i^s] + h_i^m E[I_i^m] + c_i E[N_i^m]\} = \sum_{i=1}^{K} \{h_i^s \sum_{y_i^s=0}^{R_i^s} (R_i^s - y_i^s) P(x_i^s = y_i^s) \\ &+ h_i^m [\sum_{y_i^s=0}^{R_i^s-1} \sum_{y_i^m=0}^{R_i^m} (R_i^m - y_i^m) P(x_i^s = y_i^s, x_i^m = y_i^m) + \sum_{y_i^s=R_i^s}^{R_i^s+R_i^m} \sum_{y_i^m=0}^{R_i^s+R_i^m-y_i^s} (R_i^s + R_i^m - y_i^s - y_i^m) P(x_i^s = y_i^s, x_i^m = y_i^m)] \\ &+ c_i [\sum_{y_i^s=0}^{R_i^s-1} \sum_{y_i^m=R_i^m}^{\infty} (y_i^m - R_i^m) P(x_i^s = y_i^s, x_i^m = y_i^m) + \sum_{y_i^s=R_i^s}^{\infty} \sum_{y_i^m=R_i^s+R_i^m-y_i^s}^{\infty} (y_i^s + y_i^m - R_i^s - R_i^m) P(x_i^s = y_i^s, x_i^m = y_i^m)]\} \end{aligned} \tag{4.4}$$

where R^s and R^m are the array of inventory levels at the supplier and the manufacturer and x_i^s and x_i^m are the number of class i jobs in the supplier's and manufacturer's production queue, respectively.

Observe that for any schedule independent of inventory levels and workloads, both the supplier and the manufacturer's production queues operate independent of the values of R^m, and the supplier's production queue is independent of R^s. Unfortunately, the state of the manufacturer's production queue, and thus the stationary distribution of number of jobs at the manufacturer $f_m(x)$, is a function of R^s.

Theorem 2. *For fixed inventory levels R_i^s for each class i at the supplier, the optimal levels of inventory for the manufacturer are the minimum R_i^m values that satisfy:*

$$P(x_i^s \leq R_i^s, x_i^m \leq R_i^m) + P(x_i^s > R_i^s, x_i^s + x_i^m \leq R_i^s + R_i^m) \geq \frac{c_i}{c_i + h_i^m}. \tag{4.5}$$

Proof. Since both the supplier's and manufacturer's production queues operate independent of R^m values, for fixed R^s, the optimal R_i^m is independent of R_j^m for any pair of classes i and j. Thus, we can separate the objective function (4.4) into K subproblems and analyze each class separately. Then, for each class i,

$$\begin{aligned}\frac{\Delta Z(R_i^s, R_i^m)}{\Delta R_i^m} &= Z(R_i^s, R_i^m + 1) - Z(R_i^s, R_i^m) \\ &= h_i^m[P(x_i^s \leq R_i^s, x_i^m \leq R_i^m) + P(x_i^s > R_i^s, x_i^s + x_i^m \leq R_i^s + R_i^m)] \\ &\quad + c_i[-1 + P(x_i^s \leq R_i^s, x_i^m \leq R_i^m) + P(x_i^s > R_i^s, x_i^s + x_i^m \leq R_i^s + R_i^m)]\end{aligned}$$

and $\frac{\Delta^2 Z(R_i^s, R_i^m)}{\Delta^2 R_i^m} = \frac{\Delta Z(R_i^s, R_i^m+1)}{\Delta R_i^m + 1} - \frac{\Delta Z(R_i^s, R_i^m)}{\Delta R_i^m} \geq 0$ for every R_i^s, R_i^m pair.

Thus, the optimal inventory level is the minimum R_i^m that satisfies $\frac{\Delta Z(R_i^s, R_i^m)}{\Delta R_i^m} \geq 0$ since the second order conditions show that our objective function decreases up to a level and then starts increasing. The result follows from this relationship. ∎

Thus, if the supplier is MTO (inventory zero) or the supplier's inventory level is otherwise known we can characterize the optimal manufacturer inventory levels.

Corollary 5. *If the manufacturer is working with a pure MTO supplier, then the manufacturer's optimal inventory levels for each class are the minimum R_i^m values that satisfy:*

$$P(x_i^s + x_i^m \leq R_i^m) \geq \frac{c_i}{c_i + h_i^m}. \tag{4.6}$$

Proof. This is a special case of Theorem 2 where the inventory levels for every class at the supplier are all fixed and 0. Using $R_i^s = 0$ in (4.5), gives the result. ∎

If this is not the case, we need to start to make approximating assumptions to begin to characterize the manufacturer's inventory levels:

Lemma 4. *If we assume that the change in the stationary distributions of the number of jobs at the manufacturer is negligible w.r.t. a unit increase in the amount of the inventory levels at the supplier, the optimal level of R_i^m is nonincreasing w.r.t. R_i^s.*

Proof. The optimal R_i^m is the minimum value that satisfies (4.5) for given R_i^s and observe that the left hand side of (4.5) is an increasing function of both R_i^s and R_i^m. Thus, as R_i^s increases, smaller values of R_i^m can satisfy the relation. ∎

Unfortunately, in general, finding optimal inventory levels is difficult, because as mentioned above, $f_m(x)$, the distribution of the number of jobs at the manufacturer's production queue, depends on the inventory levels R^s at the supplier. If, for the centralized, model $h_i^s \geq h_i^m$ for a product type i, we can show that an MTO strategy for type i at the supplier is optimal, but typically we would not expect this condition to hold (that is, typically $h_i^s < h_i^m$). Thus, we are motivated to approximate the optimal R^s values. To do this, we assume that the change in the stationary distributions of the number of jobs at the manufacturer is negligible w.r.t. a change in the amount of the inventory levels at the supplier. In this case, we can divide the problem into K subproblems and analyze each class separately. In the remaining part of this section, we state our results under this assumption. Unfortunately, still, we can't prove a result similar to Theorem 2 for the inventory values at the supplier, since the objective function (4.4) doesn't possess the convexity structure in R_i^s for fixed R_i^m. (i.e. $Z(R_i^s+1, R_i^m) - Z(R_i^s, R_i^m)$ is not nondecreasing in R_i^s for every R_i^m.)

In order to determine the optimal levels of R^s, we therefore employ a one-dimensional search on R_i^s. For each R_i^s, we calculate the optimal values of R_i^m, calculate the total cost using the objective function (4.4) and pick the pair with minimum cost for each class i. We can decrease the size of the search space, however, using some properties of the objective function.

Lemma 5. *For this model, the function $Z(R_i^s, R_i^m)$ as given in (4.4) is supermodular for every class i.*

Proof.

$$\begin{aligned}
Z(R_i^s+1, R_i^m+1) - Z(R_i^s, R_i^m+1) &= h_i^s P(x_i^s \le R_i^s) + h_i^m[P(x_i^s > R_i^s, x_i^s + x_i^m \le R_i^s + R_i^m + 1)] \\
&\quad + c_i[-1 + P(x_i^s \le R_i^s) + P(x_i^s > R_i^s, x_i^s + x_i^m \le R_i^s + R_i^m + 1)] \\
&\ge h_i^s P(x_i^s \le R_i^s) + h_i^m[P(x_i^s > R_i^s, x_i^s + x_i^m \le R_i^s + R_i^m)] \\
&\quad + c_i[-1 + P(x_i^s \le R_i^s) + P(x_i^s > R_i^s, x_i^s + x_i^m \le R_i^s + R_i^m)] \\
&= Z(R_i^s+1, R_i^m) - Z(R_i^s, R_i^m) \qquad (4.7)
\end{aligned}$$

By transivity, $Z(\acute{R}_i^s, \acute{R}_i^m) - Z(R_i^s, \acute{R}_i^m) \ge Z(\acute{R}_i^s, R_i^m) - Z(R_i^s, R_i^m)$ for every $\acute{R}_i^s \ge R_i^s$ and $\acute{R}_i^m \ge R_i^m$. ■

Theorem 3. *Let $\bar{R}_i^s \ge 0$ be the minimum value that satisfies*

$$P(x_i^s \le \bar{R}_i^s) \ge \frac{c_i}{c_i + h_i^s}.$$

Then, the optimal level of inventory $R_i^{s*} \le \bar{R}_i^s$.

Proof. For every R_i^s, R_i^m pair

$$\begin{aligned}
\frac{\Delta Z(R_i^s, R_i^m)}{\Delta R_i^s} &= Z(R_i^s+1, R_i^m) - Z(R_i^s, R_i^m) \\
&= h_i^s P(x_i^s \le R_i^s) + h_i^m[P(x_i^s > R_i^s, x_i^s + x_i^m \le R_i^s + R_i^m)] \\
&\quad + c_i[-1 + P(x_i^s \le R_i^s) + P(x_i^s > R_i^s, x_i^s + x_i^m \le R_i^s + R_i^m)] \qquad (4.8) \\
\frac{\Delta^2 Z(R_i^s, R_i^m)}{\Delta^2 R_i^s} &= \frac{\Delta Z(R_i^s+1, R_i^m)}{\Delta R_i^s + 1} - \frac{\Delta Z(R_i^s, R_i^m)}{\Delta R_i^s} \\
&= h_i^m[P(x_i^s > R_i^s+1, x_i^s + x_i^m = R_i^s + R_i^m + 1) - P(x_i^s = R_i^s + 1, x_i^m < R_i^m)] \qquad (4.9) \\
&\quad + h_i^s P(x_i^s = R_i^s + 1) + c_i[P(x_i^s = R_i^s + 1, x_i^m \ge R_i^m) + P(x_i^s > R_i^s + 1, x_i^s + x_i^m = R_i^s + R_i^m + 1)
\end{aligned}$$

Observe that $\frac{\Delta^2 Z(R_i^s,0)}{\Delta^2 R_i^s} \geq 0$ for every R_i^s, so $\frac{\Delta Z(R_i^s,0)}{\Delta R_i^s}$ is increasing. Thus, the minimum R_i^s value that satisfies $\frac{\Delta Z(R_i^s,0)}{\Delta R_i^s} \geq 0$ minimizes the cost function for fixed $R_i^m = 0$.

$$\frac{\Delta Z(R_i^s,0)}{\Delta R_i^s} = Z(R_i^s+1,0) - Z(R_i^s,0) = h_i^s P(x_i^s \leq R_i^s) + c_i[-1 + P(x_i^s \leq R_i^s)] \geq 0$$

Observe that $\bar{R}_i^s$ is the minimum value that satisfies this relation, and therefore the minimizer of the objective function when $R_i^m = 0$. Thus $Z(R_i^s,0) \geq Z(\bar{R}_i^s,0)$ for every R_i^s. Also, recall that our objective function is supermodular by Lemma 5, thus for $\acute{R}_i^s \geq \bar{R}_i^s$ and $\acute{R}_i^m \geq 0$:

$$\begin{aligned} Z(\acute{R}_i^s,\acute{R}_i^m) - Z(R_i^s,\acute{R}_i^m) &\geq Z(\acute{R}_i^s,R_i^m) - Z(R_i^s,R_i^m) \\ \Rightarrow Z(\acute{R}_i^s,\acute{R}_i^m) - Z(\bar{R}_i^s,\acute{R}_i^m) &\geq Z(\acute{R}_i^s,0) - Z(\bar{R}_i^s,0) \geq 0 \Rightarrow Z(\acute{R}_i^s,\acute{R}_i^m) \geq Z(\bar{R}_i^s,\acute{R}_i^m). \end{aligned}$$

∎

We conclude that there is no need to search for R_i^{s*} beyond $\bar{R}_i^s$. Also, this bound allows us to observe that:

Corollary 6. *It is optimal for the supplier to use a MTO strategy to produce product type i if* $F_i^s(0) \geq \frac{c_i}{c_i+h_i^s}$.

Proof. Due to Theorem 3. ∎

In general, for the fixed inventory value R_i^m at the manufacturer, if for every R_i^s, $\frac{\Delta^2 Z(R_i^s,R_i^m)}{\Delta^2 R_i^s} \geq 0$, then the following theorem holds. An example of this case occurs when $h_i^s \geq h_i^m$.

Theorem 4. *For fixed inventory levels R_i^m at the manufacturer, if $\frac{\Delta^2 Z(R_i^s,R_i^m)}{\Delta^2 R_i^s} \geq 0$, the optimal levels of inventory at the supplier are the minimum value R_i^s that satisfies:*

$$(h_i^s + c_i)P(x_i^s \leq R_i^s) + (h_i^m + c_i)P(x_i^s > R_i^s, x_i^s + x_i^m \leq R_i^s + R_i^m) \geq c_i. \tag{4.10}$$

Proof. Since $\frac{\Delta^2 Z(R_i^s,R_i^m)}{\Delta^2 R_i^s} \geq 0$ for every R_i^s, the optimal inventory level will be the minimum R_i^s that satisfies $\frac{\Delta Z(R_i^s,R_i^m)}{\Delta R_i^s} \geq 0$. The result follows from equation (4.8). ∎

Once inventory levels are determined, we turn to sequencing and lead time quotation decisions. Once again, we generalize the approaches in Kaminsky and Kaya [18] to MTS/MTO systems. In particular, we assume that processing times at the supplier and the manufacturer are exchangeable, and we process the jobs according to $SEPTA_p$ (Shortest Expected Processing Time Available rule, based on total expected processing time $p_i = p_i^s + p_i^m$) at the supplier, and in FCFS order at the manufacturer. [1]

[1] In Kaya [20], we consider scheduling and lead time quotation if processing times at the supplier and the manufacturer are not exchangeable. In essence, we focus on the bottleneck facility and develop a scheduling algorithm similar to $SEPTA_p$ and a lead time quotation algorithm for that schedule.

To quote a due date, we utilize a modified version of the approach initially introduced in [18]. This approach is shown to be effective in [18] to minimize the objective function $E[d_i] + E[W_i - d_i]^+$. This lead time quotation algorithm is designed assuming that jobs are scheduled according to $SEPTA_p$ at the supplier and $FCFS$ at the manufacturer, and using the inventory values and state of the system at the time of job arrivals.

When an order j for job type i arrives at the system at time r_j, the following lead time d_j is quoted:

$$d_j = \begin{cases} 0 & if \ \ I_i^m > 0 \ \ at \ \ r_j \\ E[p_j^m] + t_j^{mm} & if \ \ I_i^m = 0, \ \ I_i^s > 0 \ \ at \ \ r_j \\ d_j^s + E[p_j^m] + max\{t_j^{ms} + t_j^{mm} + slack_j^m - d_j^s, 0\} & otherwise \end{cases}$$

where $d_j^s = E[p_j^s] + E[M_j^s] + \frac{E[M_j^s]\lambda pr\{p<p_j\}E\{p^s|p<p_j\}}{1-\lambda pr\{p<p_j\}E\{p^s|p<p_j\}}$. Let k denote the job in the production queue used to satisfy order j. Then, M_j^s is the workload at the supplier in front of job k at time r_j, $t_j^{ms} = \sum_{l \in A} E[p_l^m]$ where A=set of jobs in supplier queue scheduled before job k and to be sent to manufacturer immediately after being processed at the supplier and $t_j^{mm} = \sum_{l \in B} E[p_l^m]$ where B=set of jobs in manufacturer queue at time r_j, $slack_j^m = (d_j^s - p_j^s)\lambda pr\{p < p_j\}E\{p^m|p < p_j\} + \sum_{l \in L} min\{(d_j^s - p_j^s)\lambda_l, I_l^s\}E[p_l^m]$ where L is the set of job types that will be scheduled after k at the supplier. Note that if we have inventory of job type $l \in L$, those jobs will be sent to the manufacturer before job k even though they are longer than k and may cause an increase in the completion time of k, thus we add the last component in the slack calculation to account for this.

As in the single stage case, the schedule used in the system only impacts the stationary distributions of the number of jobs in the system (i.e. $f(x)$). With the sequencing rule we presented above, the supplier is using the SEPTA rule with respect to total processing time of the jobs and the manufacturer is scheduling jobs FCFS. Since the production queue at the supplier is independent of R^s and R^m, it operates just like the single facility described in Section 3, so we can find the stationary distributions of the number of jobs in the supplier's production queue using the pgf given in (3.7).

However, the inventory levels at the supplier affect the processes at the manufacturer, making exact determination of the stationary distributions analytically difficult. Thus, we are motivated to use the usual decomposition approach to approximate the stationary distributions of the number of jobs at the manufacturer. Note that when the interarrival times are exponentially distributed in a queueing model (like the well known Jackson network model presented first in Jackson[14]), the departure process is poisson distributed. Since the inter-arrival time to our system is exponentially distributed, we approximate the departure process from the supplier with a poisson distribution and thus we assume that the arrivals to the manufacturer are poisson. So, we treat the processes at the manufacturer as a single facility with multiple classes with poisson arrivals scheduled FCFS. The probability generating function for the number of customers at steady state in this system is given in equation (3.4). The required probabilities to solve this model can be obtained from the pgf (3.4).

4.2 The Decentralized Supply Chain, Full Information Model

While some supply chains are relatively easy to control in a centralized fashion, often this is not the case. Even if the stages in a supply chain are owned by a single firm, information systems, control systems, and local performance incentives need to be designed and implemented in order to facilitate centralized control. In many cases, of course, the supplier and manufacturer are independent firms, with relatively limited information about each other. Implementing centralized control in these supply chains is typically even more difficult and costly, since the firms need to coordinate their processes, agree on a contract, implement an information technology system for their processes, etc. Thus, for either centrally owned or independent firms, centralization might not be worth the effort if the gains from centralization are not substantial.

Although centralization in supply chains may be difficult and costly to implement, decentralized systems may be substantially less profitable. In some cases, rather than agreeing to centralized controls, firms may elect to share information in order to (hopefully) increase profits. We are therefore motivated to explore the gains from information exchange, as these gains may approach those of centralization. In this model, the manufacturer has complete information about the entire system, but has no control over the supplier's decisions. For this system, we find the optimal inventory levels for the manufacturer and the supplier, as well as an effective scheduling and due-date quotation algorithm. We also computationally assess the differences between the centralized model and this decentralized model in the next section.

4.2.1 The Model

In this decentralized supply chain model, we assume that the two parties work independently from each other and aim to minimize their own costs. However, the manufacturer has full information about the processes at the supplier as well as his own processes. We assume that, the supplier also incurs a unit lead time cost of c_i^s in addition to his unit holding cost of h_i^s. So, the supplier works as a single facility (see section 3) where the manufacturer is the customer of the supplier. The manufacturer has a unit lead time cost of c_i and a unit holding cost of h_i^m per unit time for his inventories. Since the supplier works independently from the manufacturer and tries to minimize his own costs, the results from the single facility case (Section 3) apply for optimal inventory level determination and sequencing at the supplier. Then, for that sequence and set of inventories at the supplier, we find the optimal inventory levels and design an effective scheduling and lead time quotation algorithm for the manufacturer.

4.2.2 Analysis and Results

As in Section 3, the supplier's objective is to minimize $\sum_{i=1}^{K}\{h_i^s E[I_i^s]+c_i^s E[N_i^s]\}$. Thus, the optimal inventory level for class i jobs at the supplier, R_i^s, is the minimum $x \geq 0$ that satisfies

$$F_i^s(x) \geq \frac{c_i^s}{c_i^s + h_i^s}$$

Using the approach in Section 3, the optimal inventory levels for the supplier can be obtained.

Corollary 7. *The optimal inventory levels for the supplier in the decentralized, full information case will be greater than or equal to the optimal level in the centralized case when $c_i^s = c_i$ and if the same schedules are used for both cases.*

Proof. Due to Theorem 3. ∎

Assuming that the supplier uses the optimal inventory levels for his facility, the manufacturer tries to minimize his own costs for those fixed R^s values. So, we attempt to minimize the objective function:

$$\sum_{i=1}^{K}\{h_i^m E[I_i^m] + c_i E[N_i^m]\}$$

Using the definitions (4.1) and writing them explicitly for fixed R^s, we get the objective function that the manufacturer wishes to minimize:

$$\begin{aligned} Z(R^s, R^m) &= \sum_{i=1}^{K}\{h_i^m E[I_i^m] + c_i E[N_i^m]\} \\ = & \sum_{i=1}^{K}\{h_i^m[\sum_{y_i^s=0}^{R_i^s-1}\sum_{y_i^m=0}^{R_i^m}(R_i^m - y_i^m)P(x_i^s = y_i^s, x_i^m = y_i^m) + \sum_{y_i^s=R_i^s}^{R_i^s+R_i^m}\sum_{y_i^m=0}^{R_i^s+R_i^m-y_i^s}(R_i^s + R_i^m - y_i^s - y_i^m)P(x_i^s = y_i^s, x_i^m = y_i^m)] \\ + & c_i[\sum_{y_i^s=0}^{R_i^s-1}\sum_{y_i^m=R_i^m}^{\infty}(y_i^m - R_i^m)P(x_i^s = y_i^s, x_i^m = y_i^m) + \sum_{y_i^s=R_i^s}^{\infty}\sum_{y_i^m=R_i^s+R_i^m-y_i^s}^{\infty}(y_i^s + y_i^m - R_i^s - R_i^m)P(x_i^s = y_i^s, x_i^m = y_i^m)]\} \end{aligned} \tag{4.11}$$

Observe that both the supplier's and the manufacturer's production queues operate independent of R^m for fixed R^s.

Then, for fixed inventory levels R_i^s for each class i at the supplier, the optimal levels of inventory R_i^m for the manufacturer can be found by employing Theorem 2. Also, the optimal inventory levels for a manufacturer working with a pure MTO supplier satisfy Corollary 5.

Corollary 8. *The optimal inventory level for the manufacturer in the decentralized, full information case will be less than or equal to the optimal level in the centralized case if $c_i^s = c_i$ and the same schedules are used in both cases assuming that the change in the stationary distributions of the number of jobs at the manufacturer is negligible with respect to a unit increase in inventory level at the supplier.*

Proof. Due to Corollary 7 and Lemma 4. ∎

Observe that the optimal inventory level for class i at the supplier decreases as c_i^s decreases. If we can choose the unit lead time costs c_i^s to charge the supplier for backlogs of type i, we can coordinate the decentralized supply chain by charging $c_i^s \leq c_i^d$ such that the supplier's inventory amounts will be decreased and the manufacturer's inventory levels will be increased to the levels that they carry in the centralized model for the same schedule. In this case, the total cost of the centralized supply chain can be achieved in this decentralized model.

However, note that it might not be in the best interest of the manufacturer to decrease c^s in the decentralized model, because decreasing c^s will decrease the supplier's costs by decreasing the manufacturer's gains, and the benefits gained by the manufacturer from the decrease in the total supply chain costs might not be worth the decrease in his gains. It would be possible though to design contracts between the supplier and the manufacturer that will coordinate the supply chain and achieve the minimum total cost in the supply chain by dividing the gains between the parties in an efficient manner or by making the supplier to pay a certain percentage of his gains to the manufacturer such that both parties will benefit from this coordination and will agree to take part in it. We leave the details of these contracts for future research.

Corollary 9. *Using a MTO strategy for class i jobs at the manufacturer is optimal if and only if*

$$P(x_i^s \leq R_i^s, x_i^m \leq 0) \geq \frac{c_i}{c_i + h_i^m} \tag{4.12}$$

where R_i^s is the optimal inventory level for class i at the supplier.

Proof. Substituting $R_i^m = 0$ in Theorem 2 gives the result. ∎

Since the manufacturer is working independently from the supplier, he also uses an SEPTA schedule according to his own processing times to sequence his jobs, and employs a lead time quotation algorithm similar to the centralized model since he has full information about the supplier. Thus, he will quote lead time for order j of job type i as follows:

$$d_j = \begin{cases} 0 & if \;\; I_i^m > 0 \;\; at \;\; r_i \\ E[p_j^m] + t_j^{mm} + \frac{t_j^{mm}\lambda\psi_j^m\tau_j^m}{1-\lambda\psi_j^m\tau_j^m} & if \;\; I_i^m = 0, \;\; I_i^s > 0 \;\; at \;\; r_j \\ d_j^s + E[p_j^m] + M_j^m + slack_j^m & otherwise \end{cases}$$

where $d_j^s = E[p_j^s] + E[M_j^s] + slack_j^s$, $slack_j^s = \frac{E[M_j^s]\lambda\psi_j^s\tau_j^s}{1-\lambda\psi_j^s\tau_j^s}$ and $\psi_j^s = pr\{p^s < p_j^s\}$, $\tau_j^s = E\{p^s|p^s < p_j^s\}$, $\psi_j^m = pr\{p^m < p_j^m\}$, $\tau_j^m = E\{p^m|p^m < p_j^m\}$

For the equations above, let k denote the job in the production queue used to satisfy order j (note that k and j are jobs of the same type thus have equal expected processing times), A denote the set of jobs l

in supplier queue with $E[p_l^m] < E[p_k^m]$ scheduled before job k and to be sent to manufacturer immediately after being processed instead of being kept as inventory at the supplier, and B denote the set of jobs at manufacturer queue at time r_j with $E[p_l^m] < E[p_k^m]$. Then, M_j^s is the estimated workload at the supplier in front of job k at time r_j, $t_j^{ms} = \sum_{l \in A} E[p_l^m]$ is the estimated workload for the manufacturer because of the current jobs at the supplier queue in front of job k, and $t_j^{mm} = \sum_{l \in B} E[p_l^m]$ is the estimated workload for the manufacturer because of the current jobs at the manufacturer queue that are shorter than job k.

Let L be the set of jobs of class l with $E[p_l^m] < E[p_k^m]$ and $E[p_l^s] \geq E[p_k^s]$, and then $sl_j^m = (slack_j^s + E[M_j^s])\lambda pr\{p^s < p_j^s, p^m < p_j^m\}\tau_j^m + \sum_{l \in L} min\{(slack_j^s + E[M_j^s])\lambda_l, I_l^s\}E[p_l^m]$ is the estimated workload for the manufacturer because of jobs with shorter processing time at the manufacturer than job k that haven't arrived to the system yet but are expected to arrive before job k is finished processing at the supplier and are expected to move to the manufacturer before job k. Finally, $M_j^m = \max\{t_j^{ms} + t_j^{mm} + sl_j^m - d_j^s, 0\}$ is the estimated workload at the manufacturer when job k is moved to the manufacturer from the supplier to satisfy order j and $slack_j^m = \frac{M_j^m \lambda \psi_k^m \tau_k^m}{1 - \lambda \psi_k^m \tau_k^m}$ is the estimated workload for the manufacturer because of future jobs with shorter processing times at the manufacturer that will arrive at the manufacturer while job k is still in the manufacturer's queue.

In this case, the stationary distributions of the number of jobs at the supplier can be found using the pgf given in (3.7) in Section 3. The results from the single facility case hold for the supplier.

The manufacturer is also scheduling jobs according to SEPTA. As in the centralized case, we use the decomposition approach to approximate the stationary distributions of the number of jobs at the manufacturer, and assume that the arrivals to the manufacturer are exponentially distributed. Thus, the stationary distribution of the number of jobs at the manufacturer can also be found by equation (3.7). Note that since the manufacturer and supplier have different processing times for each class and thus different schedules, their stationary distributions will also be different (although both are found through the same pgf).

4.3 Simple Decentralized Model

In many real systems, the supplier and the manufacturer have very limited information about each other. The manufacturer is unaware of the processes at his supplier and needs to make his own decisions without any information from the supplier. The manufacturer may only be aware, for example, of the average time it takes for the supplier to process an order and send it to him. For our model of this kind of decentralized supply chain, we determine optimal inventory levels for the manufacturer and an effective way to schedule the orders and to quote due-dates to the customers with little information from the supplier. In the next section, we compare the performance of this simple decentralized model to that of the centralized model and the decentralized model with information exchange to provide insight as to whether or not centralization and information exchange are worth the effort.

4.3.1 The Model

In this model, we again assume that the manufacturer and the supplier work independently from each other, but in this case, the manufacturer has no information about the supplier, and thus he can't deduce the production schedule or inventory levels at the supplier. The supplier still behaves as in the information sharing model above, and thus those results still hold.

However, in this case, the manufacturer only knows the average time it takes for the supplier to deliver a job type i, denoted by $E[d_i^s]$. Thus, for each job, the manufacturer acts as if each job of type i is going to be delivered to him from the supplier after $E[d_i^s]$ time units. Based on this assumption, we determine the optimal inventory levels for the manufacturer and design an effective scheduling and lead time quotation algorithm.

4.3.2 Analysis and Results

Since the supplier acts as a single facility as in the previous section, the optimal inventory levels for class i jobs, R_i^s are the minimum values that satisfy

$$F_i^s(R_i^s) \geq \frac{c_i^s}{c_i^s + h_i^s}$$

However, in this case, since the manufacturer only knows the average time it takes for an order of type i to be delivered by the supplier, when an order arrives, he assumes, that order will be delivered by the supplier to him after $E[d_i^s]$ time units. We model the supplier as an $M/D/\infty$ queue with deterministic processing times $E[d_i^s]$ for type i and without any inventories, so each job of type i will take exactly $E[d_i^s]$ time units to be delivered from the supplier to the manufacturer as assumed by the manufacturer; we use analysis similar to that of previous sections, so the manufacturer attempts to minimize the following function:

$$\begin{aligned} Z(R^m) = \sum_{i=1}^{K}\{h_i^m E[I_i^m] + c_i E[N_i^m]\} = \sum_{i=1}^{K} \quad & \{h_i^m[\sum_{y_i^s+y_i^m=0}^{R_i^m}(R_i^m - y_i^s - y_i^m)P(x_i^s + x_i^m = y_i^s + y_i^m)] \\ & +c_i[\sum_{y_i^m+y_i^s=R_i^m}^{\infty}(y_i^s + y_i^m - R_i^m)P(x_i^s + x_i^m = y_i^s + y_i^m)]\} \end{aligned}$$

Using Corollary 5, the optimal levels of inventory are the minimum R_i^m values that satisfy:

$$P(x_i^s + x_i^m \leq R_i^m) \geq \frac{c_i}{c_i + h_i^m}.$$

We find the stationary distributions of the jobs at the supplier using the $M/D/\infty$ queue model with processing times $E[d_i^s]$. Since we assume that each order is delivered to the manufacturer at time $r_i + E[d_i^s]$ by the supplier, the processes at the supplier don't affect the arrival process to the manufacturer and the distribution of the interarrival time of jobs to the manufacturer follows the same exponential distribution as the interarrival time of jobs to the system. Thus, the arrival process of the jobs to the manufacturer is also poisson distributed, so the stationary distributions of the number of jobs at both facilities are calculated using equation (3.7). Thus, using the relations given in Section 3 for a single facility, we find the stationary distributions of the number of jobs and the inventory levels for the manufacturer and the supplier.

Since the manufacturer focuses on his own objectives, he uses an SEPTA schedule according to his own processing times. We assume that the manufacturer has no information about the supplier and is unaware of the schedule or inventory levels there, but from his previous experience, he can deduce an average delivery time for each class of products. So, he uses this piece of information to quote due dates to his customers. We use a similar approach to the previous cases to quote due dates, but now, using only the available information that the manufacturer has, we estimate some of the values that were known by the manufacturer in the previous cases we considered. Thus, we quote lead time for order j of job type i as follows:

$$d_j = \begin{cases} 0 & if \;\; I_i^m > 0 \;\; at \;\; r_j \\ E[p_j^m] + t_j^{mm} + \frac{t_j^{mm}\lambda\psi_j^m\tau_j^m}{1-\lambda\psi_j^m\tau_j^m} & if \;\; I_i^m = 0, \;\; I_i^s > 0 \;\; at \;\; r_j \\ E[d_i^s] + E[p_j^m] + M_j^m + slack_j^m & otherwise \end{cases}$$

where k is the job in the production queue used to satisfy order j and $E[d_i^s]$ is the average delivery time of class i products from the supplier. We use the same notation as in the previous section, but in this case we only have information about the manufacturer. B denotes the set of jobs in the manufacturer's queue at time r_j with $E[p_l^m] < E[p_k^m]$ and $t_j^{mm} = \sum_{l \in B} E[p_l^m]$. $sl_j^m = \frac{E[d_i^s]\psi_k^m\tau_k^m}{E[p^s]}$ is the estimated workload of the manufacturer because of the jobs that will arrive from the supplier before job k and have shorter processing time than p_k^m at the manufacturer. Then, $M_j^m = max\{t_j^{mm} + sl_j^m - E[d_i^s], 0\}$ is the estimated workload that job k sees when it arrives at the manufacturer, and finally, $slack_j^m = \frac{M_j^m\lambda\psi_k^m\tau_k^m}{1-\lambda\psi_k^m\tau_k^m}$ is the estimated workload at the manufacturer because of future jobs with shorter processing time than p_k^m at the manufacturer that will arrive at the manufacturer while job k is still in the manufacturer's queue.

5 Computational Analysis

Using the approaches for inventory values determination, scheduling, and lead time quotation developed in the previous sections, we designed several computational experiments to assess the effectiveness of our algorithms, how the performance of combined MTS/MTO systems differs from that of pure MTO or MTS systems, and the differences in performance between centralized and decentralized supply chains. For the centralized system, we assume that both parties cooperate and use the heuristics explained in Section 4.1 for finding the inventory levels, scheduling and lead time quotation. For the decentralized systems, we assume that each facility acts independently and uses the heuristics for setting inventory levels, scheduling, and lead time quotation intended to minimize its own costs as explained in Sections 4.2 and 4.3. We implemented our heuristics in C++, and present the results in the following subsections.

5.1 Single Facility

We performed a simulation study to assess the effectiveness of our heuristics for single facility models, and to determine the value of implementing a combined MTS/MTO system. We generated $n = 1000$ jobs for each simulation, and considered several scenarios with different numbers of job types K and different multipliers h, c^d and c^T for cost functions to evaluate the effect of these parameters on our system (for this computational study, we use the same h, c^d and c^T values among different job types). For each scenario, we considered 10 instances with different exponentially distributed arrival and processing rates. In Table 1, the ratios of the average of total costs $Z = \sum_{i=1}^{K} \{h_i E[I_i] + c_i^d E[d_i] + c_i^T E[W_i - d_i]^+\}$ are shown, where $E[I_i]$ denotes the average inventory, $E[d_i]$ denotes the average quoted lead time and $E[W_i - d_i]^+$ denotes the average tardiness in a simulation run for product type i.

5.1.1 The Impact of MTS/MTO and Sequencing Decisions

The first and second columns compare combined MTS/MTO system using the SEPTA-LTQ algorithm with pure MTS and MTO systems, respectively. For the pure MTO system, we assume that no inventory is held for any of the product types. For the pure MTS system, we assume that the company carries enough inventory to satisfy at least 95% of the customers from inventory and we find the inventory levels that satisfy this constraint. Recall that for the combined system, we utilize Theorem 1 to find the optimal inventory levels. For the pure MTS system, using an approach similar to that of Theorem 1, we find the minimum values of x that satisfy the relation $F_i(x) \geq 0.95$ for product type i (instead of $F_i(x) \geq \frac{c_i}{c_i + h_i}$) and use these x values as the inventory levels for the pure MTS system. The combined system, on average, provides a 20% decrease in costs as opposed to pure MTS systems and a 15% decrease as opposed to the pure MTO systems.

In the third column, we compare the costs of applying the SEPTA-LTQ algorithm for this combined system with the costs of applying the appropriate lead time quotation algorithm with a FCFS schedule. We see that the SEPTA-LTQ algorithm leads to about 15% lower costs on average than the FCFS-LTQ algorithm. Note that with the FCFS schedule, since future arrivals don't effect the delivery time of the previous orders, the lead times quoted with FCFS schedule have error only due to the stochasticity of the processing times of jobs already at the queue and thus, in general, have less error than the SEPTA schedule. Thus, if anything, lead time quotation decreases the gains achieved by a SEPTA schedule.

In the last column, we analyze the effectiveness of our lead time quotation algorithm. We compare the objective function Z using the scheduling and lead time quotation algorithm SEPTA-LTQ with the cost function $Z' = \sum_{i=1}^{K}\{h_i E[I_i] + c_i^d E[W_i]\}$ using the same schedule SEPTA but without quoting lead times. Instead we use the actual waiting times, W, of jobs in the system in the calculation of Z'. Note that the optimal off-line LTQ algorithm quotes lead times that are exactly equal to the waiting times of the jobs in the system, and thus Z' is a lower bound for the LTQ algorithm with a SEPTA-based schedule. We see that the total cost with the lead time quotation algorithm, Z, is about 4% more than Z'. Thus, we conclude that our lead time quotation algorithm is effective for minimizing the objective function Z. Also, we see that the performance of the lead time quotation algorithm increases as h or c^T decreases. When h is lower, more inventory will be held so when an order arrives, it is more likely that the order is satisfied from the inventory and there is less error in lead time quotation. When c^T is lower, the impact of tardiness costs on total costs decrease, and thus Z is closer to Z'.

5.1.2 The Impact of System Parameters

We also see the effect of the parameters on the system performance in Table 1. Unsurprisingly, as the unit inventory holding cost increases while all other parameters remain constant, the performance of the combined system moves toward the performance of a pure MTO system and gives much better results than pure MTS systems since holding inventory becomes much more costly as h increases. The system is affected in the same way as the unit due-date cost, c^d, decreases because in that case lead times become less important and a MTO system becomes much more attractive as c^d decreases. We also see that the SEPTA-LTQ algorithm gives much better results than the FCFS-LTQ algorithm as h increases (alternatively c^d decreases) because as h increases, the system moves towards a MTO strategy and the schedule has an important effect on the objective in MTO systems; meanwhile, for MTS systems, since the orders are satisfied from the inventory, the effect of minimizing the total completion times of the jobs on the system performance is not too significant. On the other hand, as the unit tardiness cost, c^T, increases while all other parameters remain constant, MTO systems perform worse and pure MTS systems become much more attractive. Also, we see that the performance difference between SEPTA-LTQ algorithm and FCFS-LTQ algorithm decreases as c^T increases and the FCFS-LTQ algorithm starts to give better results as c^T becomes very high. This happens because, with the FCFS-LTQ algorithm, the future arrivals don't affect the completion times of the previous jobs, and thus we can quote the lead times much closer to the actual completion times so the tardiness decreases. However, with the SEPTA-LTQ algorithm, tardiness cost will be higher and as c^T increases, the cost of tardiness overcomes the gains in total completion times with SEPTA-LTQ algorithm, and thus FCFS-LTQ leads to lower total costs than SEPTA-LTQ.

Table 1: Comparison of combined MTO-MTS system with pure systems and with different schedules for a single facility with $c^d = 2$ and different combinations of h, c^T and K

$\frac{h=0.5}{c^T=2.5}$	$\frac{Z}{Z_{MTS}}$	$\frac{Z}{Z_{MTO}}$	$\frac{Z_{SEPTA-LTQ}}{Z_{FCFS-LTQ}}$	$\frac{Z'}{Z}$	$\frac{h=1}{c^T=2.1}$	$\frac{Z}{Z_{MTS}}$	$\frac{Z}{Z_{MTO}}$	$\frac{Z_{SEPTA-LTQ}}{Z_{FCFS-LTQ}}$	$\frac{Z'}{Z}$
K=3	0.963	0.526	0.942	0.989	K=3	0.829	0.880	0.894	0.991
K=5	0.972	0.613	0.973	0.984	K=5	0.798	0.844	0.850	0.993
K=10	0.947	0.734	0.926	0.977	K=10	0.778	0.831	0.724	0.987
$\frac{h=1}{c^T=2.5}$	$\frac{Z}{Z_{MTS}}$	$\frac{Z}{Z_{MTO}}$	$\frac{Z_{SEPTA-LTQ}}{Z_{FCFS-LTQ}}$	$\frac{Z'}{Z}$	$\frac{h=1}{c^T=3}$	$\frac{Z}{Z_{MTS}}$	$\frac{Z}{Z_{MTO}}$	$\frac{Z_{SEPTA-LTQ}}{Z_{FCFS-LTQ}}$	$\frac{Z'}{Z}$
K=3	0.832	0.906	0.913	0.971	K=3	0.847	0.873	0.922	0.966
K=5	0.814	0.827	0.897	0.962	K=5	0.825	0.829	0.908	0.958
K=10	0.795	0.872	0.762	0.959	K=10	0.808	0.820	0.795	0.954
$\frac{h=2}{c^T=2.5}$	$\frac{Z}{Z_{MTS}}$	$\frac{Z}{Z_{MTO}}$	$\frac{Z_{SEPTA-LTQ}}{Z_{FCFS-LTQ}}$	$\frac{Z'}{Z}$	$\frac{h=1}{c^T=5}$	$\frac{Z}{Z_{MTS}}$	$\frac{Z}{Z_{MTO}}$	$\frac{Z_{SEPTA-LTQ}}{Z_{FCFS-LTQ}}$	$\frac{Z'}{Z}$
K=3	0.724	0.936	0.846	0.963	K=3	0.880	0.845	0.931	0.956
K=5	0.719	0.934	0.823	0.957	K=5	0.896	0.796	0.911	0.952
K=10	0.683	0.967	0.695	0.952	K=10	0.913	0.753	0.924	0.946
$\frac{h=5}{c^T=2.5}$	$\frac{Z}{Z_{MTS}}$	$\frac{Z}{Z_{MTO}}$	$\frac{Z_{SEPTA-LTQ}}{Z_{FCFS-LTQ}}$	$\frac{Z'}{Z}$	$\frac{h=1}{c^T=10}$	$\frac{Z}{Z_{MTS}}$	$\frac{Z}{Z_{MTO}}$	$\frac{Z_{SEPTA-LTQ}}{Z_{FCFS-LTQ}}$	$\frac{Z'}{Z}$
K=3	0.573	1	0.837	0.958	K=3	0.925	0.743	1.032	0.948
K=5	0.551	1	0.776	0.951	K=5	0.946	0.719	1.131	0.943
K=10	0.472	0.987	0.672	0.943	K=10	0.953	0.651	1.096	0.937
average	0.754	0.858	0.838	0.964		0.866	0.799	0.926	0.961

5.1.3 The Impact of Arrival and Processing Rates on Algorithm Performance

The results in Table 1 are averaged across arrival and processing rates, but the performance of our algorithm clearly depends on these rates. We analyze the performance of our algorithm for varying arrival and processing rate pairs in Table 2 using two different product types (i.e $K = 2$). We use the same parameter values for both product types, $h = 1$, $c^d = 2$ and $c^T = 2.5$, and generate $n = 3000$ jobs. We fix the arrival rate of product type 1 as $\lambda_1 = 1$ and use varying arrival rates for product type 2, λ_2 and varying processing rates μ_1 and μ_2. In Table 2, we list the congestion level $(\lambda_1/\mu_1 + \lambda_2/\mu_2)$ for each run in the first column and the optimal inventory values for both product types, for each run, under SEPTA scheduling (R_i^{SEPTA}) and under FCFS scheduling (R_i^{FCFS}) in columns 2 to 5. We also present the ratios of the objective functions of the combined system as opposed to pure MTO and MTS systems and see that the combined system performs significantly better than the pure systems. As the congestion level increases, the pure MTS system performs relatively closer to the combined system, but still the difference is about 10% in the best case. Forand when we look at the pure MTO system, for high congestion levels, the difference increases to more than 60%. We also see in the last column that our lead time quotation algorithm performs effectively in all cases.

We also see in Table 2 that under the SEPTA schedule, since the shorter job type gets precedence, we carry less inventory of the shorter type and more inventory of the longer type in comparison with the FCFS schedule, and that inventory levels under FCFS schedule are proportional to relative arrival rates. Also, comparing objective functions under the two scheduling regimes, SEPTA and FCFS have similar performance when the processing rates μ_1 and μ_2 are close, and there are even some cases where FCFS performs better than SEPTA. This lack of significant distinction between the performance of the two policies is due to the fact that similar processing rates means that the sequences generated by the two policies are similar. However, as the difference between processing rates increase, SEPTA performs significantly better than FCFS.

5.2 Effect of Inventory Decisions and Lead-Time Quotation on Supply Chains

We next explore the effect of inventory decisions and lead-time quotation on our system. In the first four columns of Table 3, without considering lead time quotation and using the SEPTA scheduling algorithm, we compare the objective functions $Z' = \sum_{i=1}^{K}\{h_i^s E[I_i^s] + h_i^m E[I_i^m] + c_i^d E[W_i]\}$ to explore only the effect of inventory decisions in this system. Recall that we made some assumptions about the interaction between the supplier and the manufacturer regarding the stationary distributions of the number of jobs at the manufacturer and we also assumed that inventory values of different types at the supplier don't effect the stationary distributions of other types at the manufacturer. We explore how these assumptions effect the optimal inventory levels and the objective function with this computational study. We also compare the centralized and decentralized versions of the combined MTS/MTO supply chain with this objective function. In the last column of Table 3, we also analyze the effectiveness of the lead-time quotation algorithm for the

Table 2: Comparison of combined MTO-MTS system with pure systems and with different schedules for a single facility with $h = 1$, $c^d = 2$, $c^T = 2.5$, $K = 2$, $\lambda_1 = 1$ and different combinations of λ_2, μ_1 and μ_2

congestion	λ_2	μ_1	μ_2	R_1^{SEPTA}	R_2^{SEPTA}	R_1^{FCFS}	R_2^{FCFS}	$\frac{Z}{Z_{MTS}}$	$\frac{Z}{Z_{MTO}}$	$\frac{Z_{SEPTA}}{Z_{FCFS}}$	$\frac{Z'}{Z}$
0.992	10	12	11	0	38	3	35	0.892	0.373	0.969	0.942
0.924	10	11	12	5	6	1	12	0.826	0.665	1.033	0.952
0.976	10	7	12	11	7	2	21	0.885	0.544	0.955	0.931
0.952	10	2.1	21	5	5	3	27	0.867	0.561	0.426	0.927
0.995	5	6.2	6	0	40	6	34	0.913	0.372	0.973	0.959
0.973	5	6	6.2	18	6	4	21	0.852	0.546	1.062	0.935
0.969	5	5	6.5	13	5	3	17	0.841	0.552	0.994	0.943
0.714	5	2.1	21	2	2	1	4	0.817	0.754	0.621	0.946
0.989	2	3.1	3	1	35	11	24	0.897	0.437	0.988	0.940
0.978	2	3	3.1	24	3	9	18	0.869	0.569	0.990	0.949
0.992	2	3.6	2.8	1	42	14	31	0.904	0.384	0.948	0.957
0.571	2	2.1	21	1	0	1	1	0.794	0.852	0.743	0.962
0.976	1	2.1	2	2	26	13	15	0.858	0.601	1.012	0.973
0.990	1	2.3	1.8	1	42	21	25	0.886	0.564	0.968	0.946
0.992	1	1.1	12	12	1	11	11	0.894	0.561	0.568	0.937
0.523	1	2.1	21	1	0	1	0	0.802	0.828	0.813	0.953

centralized case by comparing our original objective function Z_{Cen} which includes lead-time quotation with Z'_{Cen}. Throughout Table 3, we consider different numbers of product types K and use the same parameter values over each of the product types, $c^d = 2$ and $c^T = 2.5$ for both the supplier and the manufacturer, and different combinations of h^s and h^m.

To explore the effectiveness of our algorithms, we compare the objective function using our inventory levels for the centralized model with the minimum objective function for that case. In the simulations, we find the optimal solution by considering all of the possible combinations of inventories of each type for the supplier and the manufacturer and selecting the best one. We use these minimum objective values as lower bounds in our simulations. However, this process takes a significant amount of time, especially when there are a large number of types. Although we can find the optimal inventories by trying all possible solutions for the cases analyzed in this experiment since they are of relatively small size, note that this method becomes almost impossible to apply as the problem size gets bigger. For example consider the case when there are 10 different jobs and the inventory of each job type at the supplier and the manufacturer can take a value from 0 to 4. In that case we need to evaluate $(5 * 5)^{10}$ different possible solutions for the trial and error method and if each of them takes a nanosecond(10^{-9} seconds), the whole process takes more than a day. In our heuristic, however, there is no interaction between the different types and we consider them separately. Also, we do only a one-dimensional search over the supplier's inventory levels up to an upper bound and find the corresponding inventory level for the manufacturer using Theorem 2. Thus, for the same case, we would only need to make $5 * 10$ comparisons in the worst case, which can be completed almost instantaneously, and thus this heuristic approach and can be easily applied to even larger problems.

We construct the lower bound by using the optimal inventory values found by exhaustive search, and compare it with the objective value obtained by using the inventory values found by our heuristic. The first column in Table 3 compares these two objective functions for the centralized supply chain model. We see that the inventory values found by our heuristic are very close to the optimal inventory values and there is only a 5% difference, on average, between the minimum costs and the costs obtained by using our inventory values. The difference is due to our assumption that having inventory of type i at the supplier doesn't affect the stationary distributions of the number of jobs at the manufacturer.

Comparing the centralized and decentralized models, we see that the costs of the centralized model are, on average, 10% less than those of the decentralized model with full information. The cost savings due to inventory decisions increases to more than 15% when we compare the centralized model with the simple decentralized model. We also see that, without centralization, if the manufacturer has full information about the supplier, he can decrease the costs by about 7% by effectively adjusting his own inventory levels without changing anything else. We see that if the manufacturer had control over the supplier, he could cut his costs significantly. However, even if the manufacturer didn't have control over the supplier but had full information about the whole system, he could still cut his costs and increase his profits.

In the last column of Table 3, we also analyze the effectiveness of our lead time quotation algorithm for the centralized model by comparing Z'_{Cen}, which excludes lead time quotation, with the original cost function Z_{Cen}. Note that the difference between Z_{Cen} and Z'_{Cen} is due to errors in lead time quotation. Observe that the difference between Z_{Cen} and Z'_{Cen} is about 5%, demonstrating that our lead time quotation algorithm is effective for minimizing the objective function Z_{Cen}.

We can also observe the impact of the parameters h^s, h^m, c^d and K on system performance. As the inventory holding cost at the supplier, h^s, increases while everything else remains the same, our heuristic gives results closer to the lower bound. Intuitively, as h^s increases, the optimal inventory levels at the supplier and their effect on the manufacturer decreases, as we assumed in our approximation. In addition, if the supplier uses a pure MTO strategy, then our heuristic finds the optimal solution since our approximation is exact for this case. We see the same effect as h^m decreases, because in this case it is better to carry inventories at the manufacturer than at the supplier, and the inventory levels at the supplier and their effects on the manufacturer decrease. Similarly, as c^d decreases, the inventory levels at the supplier also decrease, and our heuristic gives results closer to those of the optimal solution.

When we compare supply chain models, we see that as h^s decreases, the decentralized models give results similar to those of the centralized model. This is because as h^s decreases, the optimal inventory levels at the supplier for the centralized model become close to the upper bound we presented in Theorem 3, which is also the optimal inventory level for the supplier in the decentralized models assuming that the same schedule is used in both cases. Thus, as h^s decreases, the inventory levels at the supplier and the inventory levels at the manufacturer for the centralized and decentralized models, approach each other and the objective values for the decentralized models become closer to the centralized model objective value. We can observe the same effect as h^m increases, because in the centralized model as h^m increases, the inventory levels at the manufacturer decreases and the inventory levels at the supplier moves closer to the upper bound. Thus, the decentralized models' results get closer to those of the centralized model.

For the lead time quotation algorithm, observe that the performance of the algorithm increases as h^s or h^m decreases. This is because as the inventory holding cost decreases, more inventory will be held and when an order arrives, it is more likely that the order is satisfied from the inventory and there is less error in the lead time quotation.

In Table 3, we use the same unit lead time cost c^d for the supplier and the manufacturer. In this case, assuming the same schedule is used in the centralized and decentralized models, the supplier carries more inventory and the manufacturer carries less inventory in the decentralized model with full information (DFI model) than in the centralized model (see Corollary 7 and Corollary 8). If we have the option to choose the unit lead time costs c_i^s to charge the supplier for backlog for each product type i, we can coordinate the decentralized supply chain by charging $c_i^s \leq c_i^d$ such that the supplier's inventory amounts will decrease to the levels carried in the centralized model, and thus the total cost of the centralized supply chain can be

Table 3: Analysis of inventory decisions on centralized and decentralized systems with $c^d = 2$, $c^T = 2.5$ and different combinations of h^s, h^m and K

$\frac{h^s=0.5}{h^m=2}$	$\frac{Z'_{LB}}{Z'_{Cen}}$	$\frac{Z'_{Cen}}{Z'_{DFI}}$	$\frac{Z'_{Cen}}{Z'_{SD}}$	$\frac{Z'_{DFI}}{Z'_{SD}}$	$\frac{Z'_{Cen}}{Z_{Cen}}$	$\frac{h^s=0.5}{h^m=0.6}$	$\frac{Z'_{LB}}{Z'_{Cen}}$	$\frac{Z'_{Cen}}{Z'_{DFI}}$	$\frac{Z'_{Cen}}{Z'_{SD}}$	$\frac{Z'_{DFI}}{Z'_{SD}}$	$\frac{Z'_{Cen}}{Z_{Cen}}$
K=3	0.916	0.963	0.896	0.930	0.965	K=3	0.964	0.834	0.794	0.952	0.980
K=5	0.933	0.925	0.849	0.917	0.953	K=5	0.959	0.877	0.808	0.921	0.975
K=10	0.929	0.923	0.832	0.901	0.947	K=10	0.967	0.892	0.836	0.937	0.978
$\frac{h^s=1}{h^m=2}$	$\frac{Z'_{LB}}{Z'_{Cen}}$	$\frac{Z'_{Cen}}{Z'_{DFI}}$	$\frac{Z'_{Cen}}{Z'_{SD}}$	$\frac{Z'_{DFI}}{Z'_{SD}}$	$\frac{Z'_{Cen}}{Z_{Cen}}$	$\frac{h^s=0.5}{h^m=1}$	$\frac{Z'_{LB}}{Z'_{Cen}}$	$\frac{Z'_{Cen}}{Z'_{DFI}}$	$\frac{Z'_{Cen}}{Z'_{SD}}$	$\frac{Z'_{DFI}}{Z'_{SD}}$	$\frac{Z'_{Cen}}{Z_{Cen}}$
K=3	0.921	0.935	0.868	0.928	0.938	K=3	0.943	0.945	0.890	0.942	0.973
K=5	0.949	0.914	0.855	0.935	0.925	K=5	0.909	0.856	0.788	0.920	0.964
K=10	0.952	0.922	0.862	0.934	0.929	K=10	0.918	0.875	0.808	0.923	0.956
$\frac{h^s=1.9}{h^m=2}$	$\frac{Z'_{LB}}{Z'_{Cen}}$	$\frac{Z'_{Cen}}{Z'_{DFI}}$	$\frac{Z'_{Cen}}{Z'_{SD}}$	$\frac{Z'_{DFI}}{Z'_{SD}}$	$\frac{Z'_{Cen}}{Z_{Cen}}$	$\frac{h^s=0.5}{h^m=5}$	$\frac{Z'_{LB}}{Z'_{Cen}}$	$\frac{Z'_{Cen}}{Z'_{DFI}}$	$\frac{Z'_{Cen}}{Z'_{SD}}$	$\frac{Z'_{DFI}}{Z'_{SD}}$	$\frac{Z'_{Cen}}{Z_{Cen}}$
K=3	0.969	0.856	0.802	0.937	0.921	K=3	0.922	0.961	0.881	0.917	0.947
K=5	1	0.839	0.784	0.934	0.913	K=5	0.911	0.975	0.874	0.896	0.939
K=10	1	0.803	0.771	0.960	0.908	K=10	0.936	0.953	0.841	0.882	0.934
average	0.952	0.898	0.835	0.931	0.933		0.939	0.908	0.836	0.921	0.960

achieved by the DFI model. This can be seen in Table 4, in which we consider changing c_i^s and c_i^d values. For the centralized version of the model with parameters described in the table caption, the optimal inventory levels are $R_1^s = 27$, $R_2^s = 1$, $R_1^m = 19$ and $R_2^m = 3$. Since less inventory is carried at the supplier and more inventory is carried at the manufacturer as c_i^s decreases, by carefully selecting c_1^s and c_2^s, we achieve centralized inventory levels and costs and thus coordinate the DFI system.

However, if the models use different schedules, we are not necessarily able to achieve the total cost of the centralized system by choosing appropriate c_i^s values in the DFI system. Depending on the system characteristics, arrival and processing rates, and the schedules used, the decentralized model might give higher costs, but it can also give lower costs than the centralized model. For example, for the case discussed above, by using the $SEPTA_{p^s} - SEPTA_{p^m}$ schedule in the DFI system and $SEPTA_p - FCFS$ in the centralized system, the optimal inventory levels are $R_1^s = 27$, $R_2^s = 1$, $R_1^m = 15$ and $R_2^m = 6$ with the total cost 5568.3 for the centralized model, which is slightly larger than the minimum cost achieved in the coordinated DFI system, 5473.6.

In Table 5, we present a characteristic example of how the optimal inventory levels change in the centralized and decentralized models depending on the arrival and processing rates for different number of classes using $h^s = 0.5$ and $h^m = 1$. In the first trial we use 3 different product types; in trial 2, we use 5 different product types and in trial 3, we use 10 different product types. We observe that (at least in this example) more inventory is carried for a product type as the arrival rate of that type increases or processing rate decreases. Also, we see that the processing rates play a more important role than the arrival rates on the

Table 4: Comparison of the centralized model with DFI model with $K = 2$, $h^s = 0.5$, $h^m = 1$, $c^m = 2$,$\lambda_1 = 1$, $\lambda_2 = 1$, $\mu_1^s = \mu_1^m = 2$, $\mu_2^s = \mu_2^m = 2.1$ and changing values for c_i^s

c_1^s	c_2^s	R_1^S	R_2^S	R_1^M	R_2^M	$\frac{Z'_{Cen}}{Z'_{DFI}}$		c_1^s	c_2^s	R_1^S	R_2^S	R_1^M	R_2^M	$\frac{Z'_{Cen}}{Z'_{DFI}}$
0.2	0.2	19	0	24	4	0.847		1.2	1.2	37	2	18	3	0.886
0.4	0.6	23	1	21	3	0.923		1.6	1.6	41	3	19	2	0.839
0.6	0.8	27	1	19	3	1		2	2	45	3	19	2	0.789
0.8	0.8	32	1	18	3	0.953		2.4	2.4	45	4	19	2	0.783
1	1	36	2	18	3	0.899		2.8	2.8	47	4	19	2	0.759

optimal inventory levels when an SEPTA schedule is used since the sequence, and thus the waiting times, depend highly on the processing times; however, for a FCFS schedule, the effect of the processing rates is smaller since the processing times are not considered for sequencing and the effect of the arrival rates become larger for this schedule. We also see from the optimal inventory levels for the centralized and decentralized versions that the optimal inventory levels at the manufacturer and supplier are closely related to each other and that the manufacturer has to consider the supplier as well as his own production facility when setting inventory levels to minimize his costs. We also see that the optimal inventory levels at the supplier are lower for the centralized case than for the decentralized cases and the manufacturer carries more inventory in the centralized model than he carries in the decentralized models.

5.3 Effectiveness of a Combined System and Efficiency of Heuristics for Supply Chains

We also completed a similar study to the one discussed above, except that lead time quotation is included in our analysis, so the objective function becomes $Z = \sum_{i=1}^{K} \{h_i^s E[I_i^s] + h_i^m E[I_i^m] + c_i^d E[d_i] + c_i^T E[W_i - d_i]^+\}$. We use the multipliers $h^s = 0.5$, $h^m = 1$, $c^d = 2$ and $c^T = 2.5$ for this case. We are comparing the costs for the entire system for the centralized and decentralized models. In the first column in Table 6, we compare the total costs of the centralized model using the algorithm $SEPTA_p - LTQ_C$ with the total cost of scheduling jobs according to FCFS in both systems and quoting lead times appropriately. We conclude that the schedule used to produce the jobs has an important effect on the total cost and we see that, on average, our heuristic performs about 20% better than the commonly used schedule FCFS. Recall from Table 1 that this was also the case for the single facility case.

Also, when we compare the values in the second column of Table 6 with the values in the second column of Table 3 for parameters $h^s = 0.5$, $h^m = 1$, $c^d = 2$ and $c^T = 2.5$, which compare the centralized case and the decentralized case with full information, we see that there is a small difference between these values. This difference is due to the inclusion of lead time quotation in Table 6. When the manufacturer has full

Table 5: Optimal inventory levels for centralized and decentralized supply chains with $h^s = 0.5$, $h^m = 1$ and different combinations of λ, μ and K

Product Type	Arrival rate	Proc.rate of supp.	Proc. rate of manuf.	Centralized model		Decen. with full info		Simple dec. model		Centralized with FCFS	
				Sup.	Man.	Sup.	Man.	Sup.	Man.	Sup.	Man.
Trial 1											
1	0.9501	6	3	0	2	2	1	2	1	7	3
2	0.2311	2	1	1	1	1	1	1	1	3	1
3	0.6068	0.9	5	10	1	11	1	11	0	6	3
Trial 2											
1	0.9501	8	6	0	1	1	1	1	1	0	5
2	0.2311	5	4	0	1	0	1	0	0	0	1
3	0.6068	1.8	1.5	0	5	3	5	3	4	0	3
4	0.486	3	5.5	1	1	1	1	1	0	0	2
5	0.8913	4	5	1	1	1	1	1	1	0	3
Trial 3											
1	0.9501	4	3	0	3	3	2	3	2	0	2
2	0.2311	11	6	0	0	0	0	0	0	0	1
3	0.6068	13	12	0	0	0	0	0	0	1	1
4	0.486	16	15	0	0	0	0	0	0	0	1
5	0.8913	8	9	0	1	1	1	1	0	1	2
6	0.7621	6	11	1	0	1	0	1	0	1	1
7	0.4565	14	5	0	0	0	0	0	0	0	1
8	0.0185	2	4	0	0	0	0	0	0	0	0
9	0.8214	4.5	4.5	0	1	1	1	1	0	1	1
10	0.4447	3	4	0	1	1	0	1	0	0	1

Table 6: Comparison of centralized and decentralized supply chains for combined MTO-MTS system with $h^s = 0.5$, $h^m = 1$, $c^d = 2$ and $c^T = 2.5$

	$Z_{SEPTA_p-LTQ_C}/Z_{FCFS-LTQ}$	Z_{Cen}/Z_{DFI}	Z_{Cen}/Z_{SD}	Z_{DFI}/Z_{SD}
K=3	0.859	0.954	0.679	0.712
K=5	0.775	0.857	0.575	0.671
K=10	0.682	0.893	0.584	0.654
Average	0.772	0.902	0.613	0.680

information in the DFI case, he can quote lead times as effectively as in the centralized case and we can conclude that including lead time quotation doesn't have much effect on the cost ratio between the centralized case and the decentralized case with full information.

However, when we compare the ratios in the third column of Table 6 with the ratios in the third column of Table 3 for parameters $h^s = 0.5$, $h^m = 1$, $c^d = 2$ and $c^T = 2.5$, which compare the simple decentralized model with the centralized model, we see that there is a large difference. Since the manufacturer has very little information about the supplier, he can no longer quote reliable due dates, so costs increase dramatically in the simple decentralized case due to poor lead time quotation. The costs of the simple decentralized model are about 40% worse than those of the centralized model. Observe that although there wasn't much difference between the decentralized model with full information and the simple decentralized model in Table 3, in Table 6, this difference increases to about 30% due to the difficulty of lead time quotation.

In Table 7, we compare the cost Z'_{Cen} which excludes lead time quotation with the original cost function Z_{Cen} for the centralized model using different number of jobs, n. Recall that the difference between these two costs is simply due to the errors in lead time quotation. For small number of jobs, the difference is about 25% but the performance of the algorithm increases significantly as the number of jobs increases and there is only about 6% difference on average between these two costs for $n = 1000$ jobs. This demonstrates the effectiveness of our lead time quotation approach for minimizing Z, at least for relatively large numbers of jobs.

We conclude this subsection with the observation that having full information about the supplier is critical to the manufacturer for lead time quotation. In addition to full information, if the manufacturer has complete control over the supplier, total costs can be substantially decreased with better inventory allocation.

6 Conclusion

We have considered a variety of stylized models of a supply chain with a single manufacturer and a single supplier in order to find effective inventory values that should be carried at each facility and to assess the

Table 7: Analysis of the effectiveness of the lead time quotation algorithm for the centralized model

Z'_{Cen}/Z_{Cen}	$n = 10$	$n = 50$	$n = 100$	$n = 500$	$n = 1000$
K=3	0.853	0.911	0.937	0.948	0.944
K=5	0.692	0.879	0.924	0.931	0.935
K=10	0.738	0.865	0.916	0.940	0.932
Average	0.761	0.885	0.926	0.939	0.937

impact of manufacturer-supplier relations on inventory decisions and effective lead time quotation. In our models, we consider several variations of inventory, scheduling and lead time quotation problems in combined MTO-MTS supply chains in order to minimize a function of the total inventory, lead times and tardiness. We derive the conditions for both the centralized and decentralized versions under which an MTO or MTS system should be used for each product at each facility and we present algorithms to find the optimal inventory levels. We also present effective lead time quotation and scheduling algorithms for centralized and decentralized versions of this model. Computational tests demonstrate the effectiveness of these approaches.

Using our algorithms, we investigate the value of coordination schemes involving information sharing between supply chain members for this system. We see that costs can be cut dramatically by using a combined system instead of pure MTO or MTS systems and that information exchange between the supplier and the manufacturer is critical for effective lead time quotation. We also discover that if centralization is not possible, information exchange in the decentralized model can improve the level of performance dramatically (particularly with respect to lead time quotation), although not as significantly as centralized control. In addition, we observe that when the manufacturer has significant information about the status of the supplier, he can improve the performance of the decentralized system, and indeed move close to the performance of a centralized system, by selecting the appropriate long lead time penalty to charge the supplier.

Of course, these are stylized models, and real world systems have many more complex characteristics that are not captured by these models. Nevertheless, this is to our knowledge one of the first papers that analytically explores inventory decisions, scheduling and lead time quotation together in the context of a supply chain, and that explores the impact of the supplier-manufacturer relationship in these systems.

In the future, we hope to evaluate more complex supply chain systems and assess how supply chain architecture can impact scheduling and due-date quotation decisions. We also intend to expand this research to consider different functions of lead time in the objective function. In some systems, the manufacturer doesn't have to accept all orders and has the option to reject certain orders. Pricing and capacity decisions can also be incorporated into these models. In all of these models and variants, the manufacturer needs to develop strategies for system design, and for scheduling and and due-date quotation.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grants DMI-0092854 and DMI-0020439.

References

[1] Arreola-Risa, A. and G. A. DeCroix (1998). Make-to-Order versus Make-to-Stock in a Production-Inventory System with General Production Times. *IIE Transactions* **30**, no.8, pp.705-713.

[2] K.R. Baker and J.W.M. Bertrand (1981), A Comparison of Due-Date Selection Rules. *AIIE Transactions* **13**, pp. 123-131.

[3] Bertrand, J.W.M. (1983), The Effect of Workload Dependent Due-Dates on Job Shop Performance. *Management Science* **29**, pp. 799-816.

[4] Brucker, P. (1998), *Scheduling Algorithms*. Springer, New York.

[5] Carr, S. and I. Duenyas (2000). Optimal Admission Control and Sequencing in a Make-to Stock, Make-to-Order Production System. *Operations Research* **48**, no.5, pp.709-720.

[6] Chand, S. and D. Chhajed (1992), A Single Machine Model for Determination of Optimal Due Date and Sequence. *Operations Research* **40**, pp. 596-602.

[7] Cheng, T.C.E. and M.C. Gupta (1989), Survey of Scheduling Research Involving Due Date Determination Decisions. *European Journal of Operational Research* **38**, pp. 156-166.

[8] Conway R., W. Maxwell and L. Miller (1967). Theory of Scheduling.Addison-Wesley Publishing Company, Reading, MA.

[9] Duenyas, I. (1995), Single Facility Due Date Setting with Multiple Customer Classes. *Management Science* **41**, pp. 608-619.

[10] Duenyas, I and W.J. Hopp (1995), Quoting Customer Lead Times. *Management Science* **41**, pp. 43-57.

[11] Eilon, S. and I.G. Chowdhury (1976), Due Dates in Job Shop Scheduling. *International Journal of Production Research* **14**, pp. 223-237.

[12] Federgruen, A. and Z. Katalan (1999). The Impact of Adding a Make-to-Order Item to a Make-to-Stock Production System. *Management Science* **45**, no.7, pp.980-994.

[13] Hall, N.G. and M.E. Posner (1991), Earliness-Tardiness Scheduling Problems, I: Weighted Deviation of Completion Times about a Common Due Date. *Operations Research* **39**, pp. 836-846.

[14] Jackson, J.R. (1963) Job Shop-like Queuing Systems. *Management Science* **10** pp.131-142.

[15] Kahlbacher, H. (1992), Termin-und Ablaufplanung - ein analytischer Zugang, *Ph.D. thesis*, University of Kaiserslautern, cited in [4]

[16] Kaminsky, P. and Z-H Lee (2004) Analysis of On-line Algorithms for Due Date Quotation. *Submitted for publication.*

[17] Kaminsky, P. and D. Simchi-Levi (2001). Probabilistic Analysis of an On-line Algorithm for the Single Machine Completion Time Problem with Release Dates. *Operations Research Letters* **29** pp. 141-148.

[18] Kaminsky, P. and O. Kaya (2006). Scheduling and Due-Date Quotation in a MTO Supply Chain. *Submitted for Publication.*

[19] Kaminsky, P and D. Hochbaum (2004) Due Date Quotation Models and Algorithms. *To be published as a chapter in the forthcoming book Handbook on Scheduling Algorithms, Methods and Models, Joseph Y. Leung (ed.), Chapman Hall/CRC.*

[20] Kaya, O. (2006). MTO-MTS Production Systems in Supply Chains. PhD Thesis, University of California, Berkeley, USA.

[21] Li, L. (1992) The Role of Inventory in Delivery Time Competition *Management Science* **38**, no.2, pp.182-197.

[22] Keskinocak, P., R. Ravi, and S. Tayur (2001), Scheduling and Reliable Lead Time Quotation for Orders with Availability Intervals and Lead Time Sensitive Revenues. *Management Science* **.47** pp. 264-279.

[23] Keskinocak, P., and S. Tayur (2004), Due-Date Management Policies. *In Handbook of Quantitative Supply Chain Analysis: Modeling in the E-Business Era, edited by David Simchi-Levi, S. David Wu, and Z. Max Shen* Kluwer Academic Publishers, Norwell, MA, pp. 485-553

[24] Miyazaki, S. (1981), Combined Scheduling System for Reducing Job Tardiness in a Job Shop. *International Journal of Production Research* **19**, pp. 201-211.

[25] Panwalkar, S.S., M.L. Smith, and A. Seidmann (1982), Common Due Date Assignment to Minimize Total Penalty for the One Machine Scheduling Problem. *Operations Research* **30**, pp. 391-399.

[26] Rajagopalan, S. (2002) Make-to-Order or Make-to-Stock: Model and Application. *Management Science* **48**, no.2, pp.241-256.

[27] S. Savansaneril, P.M. Griffin, and P. Keskinocak (2006) Analysis of Lead Time Quotation Policies for Base-Stock Inventory Systems. *Proceedings of the Manufacturing and Services Operations Management Conference*, Atlanta, GA.

[28] Seidmann, A., S.S. Panwalker, and M.L. Smith (1981), Optimal Assignment of Due Dates For a Single Processor Schedluing Problem. *International Journal of Production Research* **19**, pp. 393-399.

[29] Simchi-Levi, D., P. Kaminsky and E. Simchi-Levi. (2003) Designing and Managing the Supply Chain. New York: McGraw-Hill.

[30] Wallstrom, B. (1980). On the M/G/1 queue with several classes of customers having different service time distributions, Report 1-19, Lund Institute of Technology.

[31] Weeks, J.K. (1979), A Simulation Study of Predictable Due-Dates. *Management Science* **25**, pp. 363-373.

[32] Wein, L. M. (1991) Due-date Setting and Priority Sequencing in a Multiclass M/G/1 Queue. *Management Science* **37**, pp. 834-850.

[33] Williams, T.M. (1984). Special Products and Uncertainty in Production/Inventory Systems. *European Journal of Operations Research* **15**, pp.46-54.

[34] Youssef, K.H., C. Van Delft, and Y. Dallery (2004) Efficient Scheduling Rules in a Combined Make-To-Stock and Make-to-Order Manufacturing System. *Annals of Operations Research,* **126**, pp. 103-134.