Cost Implications of Planned Lead Times in Supply Chain Operations Planning

J. M. Spitter, A. G. de Kok, and N. P. Dellaert
Eindhoven University of Technology, Department of Technology Management,
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
December 8, 2003

Abstract

In this paper we consider Supply Chain Operations Planning (SCOP) in a rolling horizon context under demand uncertainty. Production plans are calculated using a linear programming model. While in most mathematical programming models for SCOP the lead time is considered to be equal to zero or equal to the minimum processing time, we consider planned lead times. In this paper we concentrate on the optimal setting of the planned lead time. From queueing theory we know that the waiting time depends on the variance of the inter arrival times and on the utilization rate of the server. Since lead time consists for a part of waiting time, we consider the setting of planned lead times for various combinations of demand variance and utilization rates. A third factor we consider in the experiments is the difference in holding costs between the items produced by the capacitated resource and the end items. The results show that models with planned lead times are preferable above previous models without planned lead times.

Keywords: Supply Chain Operations Planning, linear programming, planned lead times, rolling horizon.

1 Introduction

In this paper we consider the Supply Chain Operations Planning (SCOP) of arbitrary supply chain structures. These arbitrary structures are characterized by the fact that the products sold to customers consist of multiple items, where each item may in turn consist of multiple items and may be used in multiple items as
In a discrete part manufacturing environment, this mixture of convergent and divergent relationships between items is most realistic, e.g. see Billington, McClain & Thomas [1983] and Shapiro [1993].

Figure 1: A typical mixture of convergent and divergent product structures. Boxes represent both the inventory and the assembly process of the item. In this example there are four end items, labeled 1, 2, 3, and 4.

The basic issue of planning in these systems is to ensure that, given the constraints of the system, i.e. resource and material availability constraints, the best possible quantity of the item is released at the best possible time, at the lowest costs.

For solving SCOP problems with capacity restrictions mathematical programming (MP) models are proposed, see e.g. Billington et al. [1983] and Chung & Krajewski [1984]. In their models they assume that the lead time of an item, the time between a request for an item and the delivery of that item, is zero or a fixed minimum lead time. The fixed minimum lead time is the time needed for, for example, the transfer of parts or for drying paint. So in these models, capacity is allocated at a fixed time offset. In Spitter [2002] a model with planned lead times is introduced. An released order can be produced during this planned lead time. So in this model, capacity can be allocated at any point in time within the planned lead time interval. Using planned lead times is not common, examples without planned lead times are numerous. In Belvaux & Wolsey [2001], for example, various mixed integer programming (MIP) models for lot sizing under capacity constraints are discussed. Stadtler [1996] presents five different MIP models for dynamic multi-item, multi-level lot sizing. In Hung & Chien [2000] MIP models are used for the study of multi-level capacitated lot sizing problems with multiple demand classes. Tempelmeier [1996] proposes a heuristic procedure for the dynamic multi-level, multi-item capacitated lot sizing problem, where the
production time is taken as the minimum lead time. The interaction between SCOP decisions and capacity acquisition decisions in environments with demand growth is explored in Rajagopalan & Swaminathan [2001]. In none of the discussed models there is a planned lead time. The actual lead time can be computed according to the demands on available capacity and are outputs of the model. Although the concept of planned lead time is not commonly used, it is used in methods for solving SCOP problems. In Gong, de Kok, & Ding [1994], e.g., it is shown that for an N-stage serial production line with random processing times of the orders, the optimal vector of planned lead times is obtained efficiently by solving an equivalent serial inventory model of the type considered in Clark & Scarf [1960].

The reason we use planned lead times is two-fold. First, by using planned lead times we can incorporate all kinds of queueing effects. As Karmarkar [1987] points out, on the scheduling level several other variables besides the capacity, affect the queueing behavior in complex production environments. These variables are, amongst others, lot sizes, release times of batches and the coordination of these release times, sequencing at machines, production mix, and the heterogeneity of items. Furthermore, linear programming models are applied in a rolling schedule setting, where periodically plans are derived based on forecasts of future demand and other exogenous information. The difference between the actual demand and its forecast induce the typical queueing behavior that arises when finite resource availability is confronted with stochastic resource requirements. Such queueing behavior cannot be made endogenous to a deterministic model.

Secondly, by using planned lead times there is the possibility to produce earlier. While in models without planned lead times production starts as late as possible, the production can already start at the beginning of the planned lead time in the model of Spitter et al. [2002]. In Spitter et al. [2003], we have shown that producing early, especially when utilization rate are high, is beneficial. By producing early, not only the amount of unused capacity is reduced, but also the safety stocks. Since early production increases the work-in-process costs, the correct setting of planned lead times is important.

In this paper we investigate the setting of planned lead times when using LP models. Our main question in this research is: Which factors influence the planned lead times, and how? Our approach to this research question is both analytical and experimental. We first identify factors influencing the planned lead times by using queueing theory and earlier research. Then we formulate testable hypotheses about interesting experimental settings we created. We use the LP model to test these hypotheses.

The remainder of this paper is organized as follows. In the next section we identify the factors influencing the planned lead times. In section 3, the LP model for solving SCOP problems using planned lead times is given. In section 4, we describe some experimental settings, pose our hypotheses and test them on the experimental settings. The last section consists of the conclusions and some
suggestions for further research.

2 Factors influencing the planned lead time

In this section we are looking for the factors influencing the planned lead time. The waiting time plays an important role in the lead time, therefore we first look at the expected waiting time for a single server queue when using queueing theory.

We consider a GI/G/1 queue, where the service time and the inter-arrival time between the orders have a general distribution. When an order arrives and the server is busy, the order waits in a queue. When all earlier arrived orders are finished, the order is produced. So, we have a first in, first out (FIFO) system. The service time also has a general distribution. Let the inter-arrival distribution $A$ have mean $1/\lambda$ and variance $\sigma_a^2$, and let the service distribution $G$ have mean $1/\mu$ and variance $\sigma_g^2$. For a stable system we must have $0 < \rho = \lambda/\mu < 1$. Let $c_a^2 = \lambda^2 \sigma_a^2$ and $c_g^2 = \mu^2 \sigma_g^2$. $c$ is the coefficient of variation. We are interested in the average waiting time of this system. The approximation for the waiting time, which was first investigated by Kingman [1962], is given by:

$$W_q = \frac{\lambda(\sigma_a^2 + \sigma_g^2)}{2(1 - \rho)}$$

(1)

In our SCOP model it is assumed that demand occurs once per period, the demand distribution is general and we only know the expected demand $E[D]$ and the squared coefficient of variation $c_d^2$. The orders derived from the demand are produced in order of arrival, and the utilization rate $\rho$ of the server is known. There is a demand for products at the beginning of each period, so the inter-arrival time of orders is deterministic and equal to one period. Thus we have a variance of $\sigma_a^2 = 0$ and, since $E[A] = 1/\lambda$, $\lambda = 1$. The service time per unit is fixed, so the service time of an order only depends on the size of the order. Hence, the coefficients of variation of the distribution functions of both the demand and the server are equal, thus $c_g^2 = c_d^2$. Further we know that $\sigma_g^2 = c_g^2/\mu^2 = c_d^2/\mu^2$ and $\rho = \lambda/\mu = 1/\mu$, thus $\sigma_g^2 = \rho^2 c_d^2$. The expected waiting time is now given by:

$$W_q = \frac{\rho^2 c_d^2}{2(1 - \rho)}$$

(2)

In (2), we see that the waiting time is dependent on the utilization rate of the machine and on the coefficient of variation. Higher utilization rates as well as more variation in the demand increase the waiting time. Hence, the utilization rate and the coefficient of variation are two factors influencing the planned lead time.

So far we have identified two factors influencing the planned lead time, a third factor can be found in Spitter et al. [2003]. In this paper we studied the timing of production. We considered early and late production within the planned lead
time. The timing is amongst other factors dependent on the holding costs of the items. Producing early is only beneficial if the safety stocks savings are larger than the increased work-in-process costs. For planned lead times a similar reasoning is applicable. Longer planned lead times can reduce the safety stocks and thus the involved costs, but on the other hand the work-in-process inventory and thus the involved costs increase with longer planned lead times. So the difference between the holding costs of the end items and the holding costs of the items produced on capacitated resources is also a factor influencing the setting of planned lead times.

3 SCOP model

We look at the factors influencing the planned lead time in capacitated SCOP using the LP model proposed in Spitter et al. [2002]. This LP model is introduced in the second subsection, but first the production dynamics are discussed and the used notation is defined. Since the work-in-process inventory is not included in the LP model, this is separately discussed in section 3.3.

3.1 Production dynamics

We consider the production of an item $i$. At time $t$ there is no work in process and an order $R_{it}$ for item $i$ is received. Each item has its own planned lead time $\tau_i$, and an order received at time $t$ will only be available for further production at time $t + \tau_i$. Execution of the order consists of several consecutive steps during the time interval $(t, t + \tau_i]$.

First, after receiving the order, all the respective amounts of components are gathered. If item $j$ is a component of $i$ and we need $h_{ji}$ units of $j$ for the production of a single unit of item $i$, there must be $h_{ji}R_{it}$ units of component $j$ available at time $t$.

During the time period $(t, t + \tau_i]$ the order can be executed. For all resources, in all the periods $(t, t+1], (t+1, t+2], \ldots, (t+\tau_i-1, t+\tau_i]$ the capacity constraints should be met. The order $R_{it}$ can be split up and each part may be produced on several resources that can operate in parallel. The amount of work for one unit of item $i$ equals $p_i$, so that the total amount of work for order $R_{it}$ equals $p_iR_{it}$.

Finally, at time $t + \tau_i$, $R_{it}$ units of items $i$ become available. The $R_{it}$ units are either stored at the inventory of $i$, or directly passed to fulfill dependent or independent demand.

We now define the notation used in the rest of the paper.

**Input parameters**

$T : T \in \mathbb{N}$, the planning horizon. The planning period $[-\tau_i, T]$ is divided in a starting time $-\tau_i$ and $T+\tau_i$ time slots $t$, defined as $(t-1, t], t = 1-\tau_i, \ldots, T$. 

$n : n \in \mathbb{N}$, the number of different items. The items are labeled $1, \ldots, n$.

$p_i : p_i > 0$, the capacity required for producing an unit of item $i$, $i = 1, \ldots, n$.
We assume this is the same for each of the facilities that could assemble $i$.

$H : H = (h_{ij})_{1 \leq i, j \leq n}, h_{ij} \in \mathbb{N}$, is the matrix representing the bills of material, i.e., $h_{ij}$ is the number of units of item $i$ needed for the production of a single unit of item $j$. We assume that the matrix $H$ is lower triangular and that the diagonal of $H$ contains only zero elements. This implies that no item can be (directly or indirectly) a component of itself.

$\tau_i : \tau_i \in \mathbb{N}, \tau_i \geq 1$, $i = 1, \ldots, n$, the lead time of item $i$.

$\alpha_{it} : \alpha_{it} > 0$, $i = 1, \ldots, n$, $t = 1, \ldots, T$, the inventory costs of holding one unit of item $i$ during a time slot $t$.

$\beta_{it} : \beta_{it} \geq 0$, $i = 1, \ldots, n$, $t = 1, \ldots, T$, the costs due to backordering at time $t$, i.e., the unit penalty costs induced by backorders of item $i$ during a time slot $t$.

$D_{it} : D_{it} \in \mathbb{N}, i = 1, \ldots, n, t = 0, \ldots, T - 1$, the exogenously determined demand for item $i$ during time slot $t$.

$k : k \in \mathbb{N}$, the number of different resources. The resources are labeled $1, \ldots, k$.

$c_{ut} : c_{ut} \in \mathbb{N}, u = 1, \ldots, k$, $t = 1, \ldots, T$, the maximum assembly time available on resource $u$ in time slot $t$.

$\mathcal{R}_i :$ the set of resources that are used by item $i$, $i = 1, \ldots, n$.

$I_{i,-1} :$ the inventory of item $i$ at time $-1$, $i = 1, \ldots, n$.

$B_{i,-1} :$ the amount of backorders of item $i$ at time $-1$, $i = 1, \ldots, n$.

$\bar{R}_{it} :$ the planned work order release of item $i$ already determined in the past, $i = 1, \ldots, n$, $t = -\tau_i, \ldots, -1$. These orders become available at time $0, 1, \ldots, \tau_i - 1$, respectively.

$\bar{V}_{ijt} :$ the amount of work already carried out at resource $u$ during time slot $t$, for producing some items $i$, $i = 1, \ldots, n$, $j = 1, \ldots, k$, $t = -\tau_i + 2, \ldots, 0$.

Output parameters

$\hat{I}_{it} :$ the work-in-process inventory of item $i$ at time $t$ waiting after being produced, $i = 0, \ldots, n$, $t = 0, \ldots, T - 1$.

$\breve{I}_{it} :$ the work-in-process inventory of item $i$ at time $t$ waiting to be used in production, $i = 1, \ldots, n$, $t = 0, \ldots, T - 1$. 
Decision variables

$I_{it}$: the inventory of item $i$ at time $t$, $i = 0, \ldots, n$, $t = 0, \ldots, T - 1$.

$G_{it}$: the gross requirements of item $i$ at time $t$, $i = 1, \ldots, n$, $t = 0, \ldots, T - 1$.

$R_{it}$: the planned work order release of item $i$ at time $t$, $i = 1, \ldots, n$, $t = -\tau_i, \ldots, T - 1$.

$B_{it}$: the backorders of item $i$ at time $t$, $i = 1, \ldots, n$, $t = 0, \ldots, T - 1$.

$V_{iut}$: the capacity allocated to some items $i$ at resource $u$ in time slot $t$, $i = 1, \ldots, n$, $u = 1, \ldots, k$, $t = -\tau_i + 2, \ldots, T$.

Note that the characterization of the variables $I_{it}$, $G_{it}$, $R_{it}$, $B_{it}$, and $V_{iut}$ depends on $t$. If $t < 0$ these variables are inputs to our model, if $t \geq 0$ they are outputs of our model.

3.2 Linear programming model

A linear programming formulation of the problem, where the production variables $R_{it}$, $i = 1, \ldots, n$, $t = 0, \ldots, T - 1$, are decision variables, is as follows:

$$\min \sum_{t=1}^{T} \sum_{i=1}^{n} \alpha_{it} I_{i,t-1} + \sum_{t=1}^{T} \sum_{i=1}^{n} \beta_{it} B_{i,t-1} + \varepsilon \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{u=1}^{k} t V_{iut}$$  \hspace{1cm} (3)

subject to

$$I_{it} = I_{i,t-1} + R_{i,t-\tau_i} - D_{it} - G_{it} + B_{it} - B_{i,t-1}, \quad i = 1, \ldots, n, \quad t = 0, \ldots, T - 1$$  \hspace{1cm} (4)

$$G_{it} = \sum_{j=1}^{n} h_{ij} R_{jt}, \quad i = 1, \ldots, n, \quad t = 0, \ldots, T - 1$$  \hspace{1cm} (5)

$$\sum_{s:-\tau_i < s \leq t} p_i R_{is} \geq \sum_{u \in R_i} \sum_{s:-\tau_i + 1 < s \leq t+1} V_{ius}, \quad i = 1, \ldots, n, \quad t = -\tau_i + 1, \ldots, T - 1$$  \hspace{1cm} (6)

$$\sum_{s:-\tau_i < s \leq t} p_i R_{is} \leq \sum_{u \in R_i} \sum_{s:-\tau_i + 1 < s \leq t+\tau_i} V_{ius}, \quad i = 1, \ldots, n, \quad t = -\tau_i + 1, \ldots, T - 1$$  \hspace{1cm} (7)

$$\sum_{i \in u \in R_i} V_{iut} \leq c_{ut}, \quad u = 1, \ldots, k, \quad t = 1, \ldots, T$$  \hspace{1cm} (8)

$$B_{it} - B_{i,t-1} \leq D_{it}, \quad i = 1, \ldots, n, \quad t = 0, \ldots, T - 1$$  \hspace{1cm} (9)

$$R_{it} = \bar{R}_{it}, \quad i = 1, \ldots, n, \quad t = -\tau_i + 1, \ldots, -1$$  \hspace{1cm} (10)

$$V_{iut} = \bar{V}_{iut}, \quad i = 1, \ldots, n, \quad u = 1, \ldots, k, \quad t = -\tau_i + 2, \ldots, 0$$  \hspace{1cm} (11)

$$R_{it}, B_{it}, I_{it} \geq 0, \quad i = 1, \ldots, n, \quad t = 0, \ldots, T - 1$$  \hspace{1cm} (12)
\[ V_{iut} \geq 0, \quad i = 1, \ldots, n, \quad u = 1, \ldots, k, \quad t = 1, \ldots, T \quad (13) \]

The objective function (3) consists of three terms. The first two terms correspond with the inventory and the backorder costs. We want to minimize these costs. The last term, where \( \varepsilon \) is a small positive number, is added to force the model to produce as early as possible. In Spitter et al. [2003], we have seen that early production is cheaper for high utilization rates. In the experiments in this research we have utilization rates of 85% or higher, so early production is desirable.

When we look at the constraints of the model we see that equation (4) stipulates that the inventory at period \( t \) equals the inventory at the previous period plus what is coming in minus what is going out. Demand can be determined exogenously, represented by \( D_{it} \), or endogenously, represented by \( G_{it} \). The amount \( B_{i,t-1} \) equals the outstanding backorders during time slot \( t \). Equations (5) relate the gross requirements of item \( i \) to planned orders of items of which \( i \) is a component. Note that for any end item \( i \) we have \( G_{it} = 0 \). In the cumulative requirements (6) and (7), the required capacity for request \( R_{it} \) of item \( i \) at time \( t \) is divided over the so-called capacity claims \( V_{iut} \). Capacity can only be claimed from resources that are able to produce item \( i \). Inequality (6) ensures that capacity is not claimed before the request is made and inequality (7) sees to it that the requests are produced within their planned lead time. We assume here that the production of an order does not start before all earlier requests have been fulfilled. We explain these inequalities in example 3.1. The capacity claim \( V_{iut} \) is the capacity allocated to some items \( i \) at resources \( u \) during time slot \( t \). The total capacity claimed on a certain resource during a time slot cannot exceed the capacity of that resource, which is ensured by inequality (8). Equations (9) enforce that backordering can only be effectuated on actual exogenously determined demand. Finally, equations (10) and (11) give respectively the requests and capacity claims from the past which still have influence on the future.

Note that although the planning horizon equals \( T \), solutions are only obtained up to time \( T - 1 \). This is due to the fact that values of \( R_{iT} \) only depend on future inputs determined after time \( T \).

**Example 3.1** We explain inequalities (6) and (7) with an example. Assume that for a certain item the planned lead time is equal to 3 and is produced on resource \( u \). Assume furthermore that there are no request or capacity claims from the past, and \( p_i = 1 \). Then we have

\[
V_0 \leq R_0 \leq V_0 + V_1 + V_2 \quad (14)
\]
\[
V_0 + V_1 \leq R_0 + R_1 \leq V_0 + V_1 + V_2 + V_3 \quad (15)
\]
\[
V_0 + V_1 + V_2 \leq R_0 + R_1 + R_2 \leq V_0 + V_1 + V_2 + V_3 + V_4 \quad (16)
\]
For simplicity only the time indices are used. In (14) we see that capacity for request $R_0$ can be claimed in the periods 0, 1, and 2. In (15) it is immediately clear that capacity for request $R_1$ can be claimed up to period 3, it is not so clear that capacity for this request cannot be claimed before period 1. Assume that capacity is claimed for $R_1$ in period 0. Since production of an order does not start before all earlier orders are produced, all the capacity needed for request $R_0$ is also claimed in period 0. Thus the capacity claim $V_0$ consist of the whole request $R_0$ plus (a part of) request $R_1$. This is in contradiction with the left-hand-side of inequality (14). Hence capacity for request $R_1$ cannot be claimed before period 1.

3.3 Work-in-process inventory

The production of an item takes place during its planned lead time. For an item $i$ this means that it is in stock until there is a request for it. At that time this item is taken out of the stock and waits to be used. At a certain moment during its planned lead time the item is transformed into its parent item $j$ together with the other child items of $j$. This parent item has to wait during the remaining planned lead time before it is put into the stock of items $j$. Waiting as a child item before production is less expensive than waiting as a parent item after production, due to the fact that production adds value in the form of labor, resource usage etc. Thus producing early reduces the backorder costs, but it increases the work-in-process inventory costs. In this section we give expressions for the amount of work-in-process inventory.

In figure 2 the three inventory points of an item $i$ are shown. We see the inventory $\hat{I}_{it}$ is equal to the inventory $\hat{I}_{i,t-1}$ plus the items that have been produced during time slot $t$ minus the items going to inventory $I_{it}$. The items that are going to inventory $I_{it}$ are those items requested $\tau_i$ periods before, thus $R_{i,t-\tau_i}$. In equations (18) the work-in-process inventory balances of $\hat{I}_{it}$ are given.

\begin{equation}
\hat{I}_{it} = \hat{I}_{i,t-1} + \sum_{u \in R_i} V_{iut} - R_{i,t-\tau_i}, \quad i = 1, \ldots, n, \quad t = 0, \ldots, T - 1
\end{equation}
The inventory $\bar{I}_{it}$ is equal to the inventory of the period before plus the requests of parent items for the particular item at time $t$ minus that part of the request that had gone into production at time $t$. In equations (19) the work-in-process inventory balances of $\bar{I}_{it}$ are given. Note that these equations are non-existent for the end items. They do not have parent items, thus they cannot get a request from them. For end items the inventory remains on the same level as the initial value.

\[
\bar{I}_{it} = \bar{I}_{i,t-1} + G_{it} - \sum_{j=1}^{n} \sum_{u \in R_j} h_{ij} V_{j,u,t+1}, \quad i = 1, \ldots, n, \quad t = 0, \ldots, T - 1 \quad (19)
\]

4 Numerical experiments

In section 2 we have seen that the utilization rate of the resource and the coefficient of variation of the demand structure have influence on the waiting time. The waiting time is a large part of the lead time. Hence, our hypothesis is that the best value for the planned lead time is dependent on the utilization rate of the resource and the variance of the demand. In Spitter et al. [2003], we have seen that also the difference between the holding costs of the end items and the holding costs of the items produced on capacitated resources influences the best value for the planned lead time. In this section we look at these three influencing factors in an experimental setting. The chosen supply chain structures for the experiments are discussed in section 4.1. In section 4.2, we pose hypotheses about the setting of the planned lead time and we look at the results in section 4.3.

4.1 Experimental setting

As we have seen in equation (2), long waiting times only occur if the utilization rate of a machine is sufficiently high. We call the machines with high utilization rates capacitated. In the experiments we use three different supply chain structures, see figure 3. The items produced on capacitated resources are indicated with a circle. We vary the planned lead times of these items to find an optimal value. The planned lead time of the other items are set to one period.

The three chosen supply chain structures are representative for all kind of supply chain structures. Structure 1 represents supply chains with only one capacitated resource. If items, produced on several capacitated resources, are used in the same end item, the optimal planned lead times of these items are influenced by each other because production of end items is only possible if all necessary items are available. Structure 2 is representative for this situation. Structure 3 represents supply chains with two or more capacitated resources where the demand of the item produced on the upstream capacitated resource is influenced by the production rate of the downstream item. The production rate
of the downstream item is influenced by the length of its planned lead time. So
the optimum planned lead times of upstream items is influenced by the optimum
planned lead time of downstream items.

We assume a one-to-one production ratio, where one unit of each child item
is needed for the production of a single unit of a parent item. The capacity
requirement $p_i$ for producing one item $i$ is equal to one for all items. The raw
material is delivered by suppliers. We assume that these suppliers do not deliver
before the end of the lead time, so we do not have work-in-process inventory ($\hat{I}_{it}$)
costs for these items. During the experiments, we vary the planned lead times
of the items that are produced on the capacitated resources to find the optimum
value, the lead times of the other items are equal to one.

We study these supply chain structures in a rolling horizon setting. For a
certain time horizon, the requests for orders are calculated based on a forecast
of the demand. After a period is ended, the actual demand is known and the
inventory and backorder parameters are adjusted to this demand. With the
adjusted parameters a plan is calculated for the coming time horizon. In the
experiments, the forecast is equal to the mean actual demand. The actual demand
of the end item follows a gamma distribution with a mean of 100 items per
period. The time horizon is chosen such that it exceeds the cumulative planned
lead times. The determination of the safety stock of end items is done as follows.
First a complete simulation without safety stocks has been carried out. Next all
the realized periodic inventory levels of the end items are considered and safety stocks are selected such that in 95% of the time no backorders occur.

4.2 Hypotheses

For different planned lead times, utilization rates, demand variations, and holding costs, the SCOP model is simulated. For a number of periods, the time horizon, the optimal production plan is calculated based on the mean demand. After one period the actual demand of that period is known. The parameters are adjusted to this actual demand and the production plan for the coming time horizon is calculated. The following hypotheses are tested using simulations.

Hypothesis 1  The optimal planned lead time of an item is longer if the utilization rate of the resource it is produced on is higher.

Hypothesis 2  The optimal planned lead time of an item is longer if the variation in the demand is higher.

In the approximation of the waiting time, equation (2), we see that higher utilization rates and more variation result in longer waiting times. Thus, these two hypotheses are consistent with the findings in section 2. Hence, we expect that the optimal planned lead time increases if the utilization rate and/or the demand variation increase.

Hypothesis 3  The optimal planned lead time of an item is longer if the difference in holding costs between this item and the end item is larger.

Since longer planned lead times may reduce the safety stocks and thus the involved costs, longer planned lead times also create more work-in-process costs. Hence, the relation between the savings of the safety stocks reduction and the extra costs of increased work-in-process inventory influences the optimal planned lead time. So, we expect that the optimal lead time is longer if the difference between holding an item as safety stock and the work-in-process costs is large.

4.3 Experimental parameters

In the simulations different input parameters are used to verify the hypotheses. These parameters are, in accordance with the hypotheses, the utilization rates, the squared coefficient of variation, and the holding costs.

In the simulation experiments, for each of the structures, a full factorial experimental design is used. In Table 1, the different utilization rates are given. Since there are two capacitated resources in structures 2 and 3, we look at two situations. In one of the situation both capacitated resources have equal utilization rates and in the other situation one of the resources, the bottleneck resource, has
a higher utilization rate. For each structure we consider two squared coefficients of variation, namely $c_d^2 = 0.25$ and $c_d^2 = 2$. The holding costs are determined by the added value when producing an item. We use a small added value of 1 on all but one echelon, on the remaining echelon the added value is equal to 4. For structures 1 and 2, this means three different holding costs scenarios, indicated by A, B, and C. Structure 3 consists of 4 echelons, so four different holding costs scenarios, A to D, are considered. In table 2, the holding costs are given. We see that the difference in holding costs between the end item and the item produced on the capacitated resource is largest in holding costs scenario A.

Table 1: The utilization rates used in the experiments.

<table>
<thead>
<tr>
<th>Structure 1 Item 2</th>
<th>Structure 2 Item 2</th>
<th>Structure 2 Item 3</th>
<th>Structure 3 Item 2</th>
<th>Structure 3 Item 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_2 = 0.85$</td>
<td>$\rho_2 = 0.85$</td>
<td>$\rho_2 = 0.85$</td>
<td>$\rho_2 = 0.85$</td>
<td></td>
</tr>
<tr>
<td>$\rho_2 = 0.88$</td>
<td>$\rho_2 = 0.88$</td>
<td>$\rho_3 = 0.85$</td>
<td>$\rho_3 = 0.85$</td>
<td></td>
</tr>
<tr>
<td>$\rho_2 = 0.91$</td>
<td>$\rho_2 = 0.88$</td>
<td>$\rho_3 = 0.88$</td>
<td>$\rho_3 = 0.88$</td>
<td></td>
</tr>
<tr>
<td>$\rho_2 = 0.94$</td>
<td>$\rho_2 = 0.88$</td>
<td>$\rho_3 = 0.88$</td>
<td>$\rho_3 = 0.88$</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: The holding costs for the different items used in the experiments.

We use a full factorial design, so in total, we have 24 different situations for structure 1, 18 different situations for structure 2, and 32 different situations for structure 3.

4.4 Results

In the experiments we vary the utilization rate, the squared coefficient of variation in the demand distribution, and the holding costs of the different type of items. To find the optimal planned lead times, we execute the experiments for different
planned lead times, and compare the inventory plus work-in-process costs. The planned lead time which results in the lowest costs is optimal.

<table>
<thead>
<tr>
<th>str.</th>
<th>item</th>
<th>$\rho$</th>
<th>$c_d^2 = 0.25$</th>
<th>$c_d^2 = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.85</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.88</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.91</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.94</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3-2</td>
<td>0.85 - 0.85</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3-2</td>
<td>0.85 - 0.88</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3-2</td>
<td>0.88 - 0.88</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3-2</td>
<td>0.85 - 0.85</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3-2</td>
<td>0.85 - 0.88</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3-2</td>
<td>0.88 - 0.85</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3-2</td>
<td>0.88 - 0.88</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3: Optimal planned lead times.

The results of these experiments are given in table 3. When comparing the results by the squared coefficient of variation, and by the difference in holding costs, we see increasing optimal planned lead times if one or more of these parameters increase. Thus we can conclude that hypotheses 2 and 3 hold. When we look at supply chain structure 2, with $c_d^2 = 0.25$ and holding costs scenario B, we see that the optimal planned lead time is equal to 2 for item 2 and equal to 1 for item 3 if $\rho_2 = 0.88$ and $\rho_3 = 0.85$. If the utilization rate of the resource on which item 3 is produced increases to $\rho_3 = 0.88$, we see a decrease in the optimal planned lead time of item 2. This is in contradiction with hypothesis 1, so this hypothesis does not hold. The cause of this contradiction can be found in the holding costs structure. Supply chain structure 2 is symmetric, so in the case with equal utilization rates we expect equal optimal planned lead times for items 2 and 3. Knowing the optimal planned lead times for the situation for the situation $\rho_2 = 0.88$ and $\rho_3 = 0.85$, we expect optimal planned lead times equal to 2 or larger. If we compare the safety stock for the situation with $\tau_2 = \tau_3 = 1$ (327 items) with the situation with $\tau_2 = \tau_3 = 2$ (259 items), we see smaller safety stock for the longer planned lead times. Unfortunately the decrease in safety stocks, and thus in the safety stock costs is less than the extra work-in-process cost of the increased planned lead times. Thus the difference in holding costs between the item produced on the capacitated resource and the end item is not large enough to cancel out the extra work-in-process costs by the decreasing safety stock costs. So, we conclude that hypothesis 1 only holds if the difference in holding costs between the item on the capacitated resource and the end item is large enough.
Table 4: Relative difference in inventory plus work-in-process costs between the optimal planned lead times and planned lead times equal to 1 period.

In table 4, we calculated the relative difference in inventory plus work-in-process cost between using the optimal planned lead times and planned lead times equal to 1 period by

\[
\frac{\text{cost for optimal } \tau - \text{cost for } \tau \text{ equal to 1}}{\text{cost for optimal } \tau} \times 100\% \quad (20)
\]

We see that savings up to 33.7% can be reached. Furthermore we see that the holding costs again play an important role in determining the optimal planned lead times. While in supply chain structure 1, with \( c_d^2 = 2 \) and \( \rho_2 = 0.94 \), the optimal planned lead time of 31 periods gives savings of 33.7% for cost scenario A, in scenarios B and C the optimal planned lead time is equal to 1 period. So no savings are reached by longer planned lead times in these two scenarios. Longer planned lead times do decrease the safety stocks in these scenarios, but the cost savings of the decreasing safety stock are smaller than the increase in the work-in-process cost. So longer planned lead times are only beneficial is the added value between the item produced on the capacitated resource and the end item is large enough.

Table 5: Approximated waiting times using equation (2).

In section 2, we determined an approximation for the waiting times in a
single item, single server queue. With this approximation we determined factors influencing the optimal planned lead times. Although demand distributions of individual items are different from the end item demand distribution, we would like to use this distribution and equation (2) to get good approximations of the optimal planned lead times. In table 5, the approximated waiting times for different demand distributions and utilization rates are given.

<table>
<thead>
<tr>
<th>str.</th>
<th>item</th>
<th>$\rho$</th>
<th>$c_d^2 = 0.25$</th>
<th>$c_d^2 = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.85</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.88</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.91</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.94</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>3 - 2</td>
<td>0.85 - 0.85</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>3 - 2</td>
<td>0.85 - 0.88</td>
<td>0.2%</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>3 - 2</td>
<td>0.88 - 0.88</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>3 - 2</td>
<td>0.85 - 0.85</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>3 - 2</td>
<td>0.85 - 0.88</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>3 - 2</td>
<td>0.88 - 0.85</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>3 - 2</td>
<td>0.88 - 0.88</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 6: Relative difference in inventory plus work-in-process costs between the optimal planned lead times and approximated planned lead times.

In table 6, we calculated the relative difference in inventory plus work-in-process costs between the optimal planned lead times and the planned lead times derived by rounding up the approximated waiting times by using

$$\frac{\text{cost for optimal } \tau - \text{cost for approximated } \tau}{\text{cost for optimal } \tau} \times 100\% \quad (21)$$

We see that, since differences up to 40% are reached, the approximated waiting times are not suitable for determining good planned lead times. Extended research has to be carried out to find heuristics for determining good planned lead times. In these studies not only the variation in demand and the utilization rate of the resource, but also the holding cost structures must be taken into account. We have seen that holding cost structures play a key role in determining the optimal planned lead times. Longer planned lead times can decrease the safety stock costs, but this is only advantageous if the work-in-process costs do not increase faster.
5 Conclusions

We considered the Supply Chain Operations Planning model for multi-item, multi-echelon systems under demand uncertainty in this paper. We used the model introduced in Spitter et al. [2002]. Instead of a fixed lead time, see e.g. Billington et al. [1983], where capacity is allocated at the fixed time offset, this model uses a planned lead time, where capacity can be allocated at any point in time within the planned lead time interval. The advantage of using planned lead times is twofold. By using planned lead times, the queueing effects, arising when finite resource availability is confronted with stochastic resource requirements, are incorporated in the planned lead time. Using planned lead times also gives the possibility to produce earlier than strict necessary. In Spitter et al. [2003], it is shown that this can be beneficial in cases of high utilization rates of the resources.

In this paper we have shown that it can be advantageous to have longer planned lead times. This proves that models with planned lead times results, providing that the length of the planned lead times is chosen well, in lower inventory costs.

Furthermore we have shown, using three representative supply chain structures, that the length of the optimal planned lead time of an item, is dependent on the coefficient of variation of the demand, the utilization rate of the resource the item is produced on, and the difference in the holding costs between this item and the end item it is used in. For the variation in the demand and the difference in holding costs it holds that if the value of one or both parameters increases, the optimal planned lead time becomes longer. For the utilization rate this characteristic only holds if the difference in holding cost between the item produced on the capacitated resource and the end items is large enough. The holding cost structure plays the leading role in the determination of the optimal planned lead time. Safety stocks can decrease by longer planned lead times if the variation in demand and/or the utilization rate is high, but this is only advantageous if the work-in-process costs do not increase too fast by long planned lead times.

This research shows that longer planned lead times can be beneficial, and gave us insights in the factors influencing the optimal planned lead time. For finding the optimal planned lead time we compared the inventory costs of different planned lead times. Further research should be done to find a heuristic for finding good planned lead times.

References


