Combining make to order and make to stock

Eindhoven University
of Technology

Department of Mathematics and Computing Science

Memorandum COSOR 96-31

Combining make to order and make to stock

I.J.B.F. Adan
J. van der Wal

Eindhoven, October 1996
The Netherlands

COMBINING MAKE TO ORDER AND MAKE TO STOCK

IVO ADAN AND JAN VAN DER WAL*

Abstract. In inventory control and production planning one is tempted to use one of two strategies: produce all demand to stock or produce all demand to order. The disadvantages are well-known. In the 'make everything to order' case (MTO) the response times may become quite long if the load is high, in the 'make everything to stock' case (MTS) one gets an enormous inventory if the number of different products is large. In this paper we study two simple models which combine MTO and MTS, and investigate the effect of combining MTO and MTS on the production lead times.

Key words. inventory control, production planning, Markov process, queueing model

AMS subject classifications. 60K25, 90B22

1. Introduction. Lesser and lesser production systems are fully Make to Stock (MTS) organized. In MTS systems the clients receive their products from a warehouse near the client. A lot of research concerns the performance and control of these systems (multi-echelon inventory control). Some other systems are completely Make to order (MTO). No product is made without a client. The analysis of these systems calls for queueing models. An important problem in MTO systems is the relation between the lead time and the utiization of the capacity. It is well-known that the lead time grows with the utilization as $1/(1-\rho)$. Utilizations of 80% are often considered to be far to low. On the other hand a utilization of 95% may give lead times which are not acceptable any more.

If one wants high utilizations and short lead times then one has to be able to deal with the natural day to day fluctuations in the demand (we are not thinking of seasonal variations). One possibility is flexible capacity. Another option studied in this paper is the combination of MTO with some MTS.

We will consider two models in which this mixture is studied. In the first one we assume that there is one product which is standard and can be made to stock. If the machine is idle it will produce a limited number of these standard products to stock.

In the second model we think of products being produced in two phases. The first phase is standard and common for all products. The second phase is customer specific. If the machine is idle it can continue to produce a limited number of 'first phases' as a capacity inventory.

Assuming exponential production and interarrival times the models can be described by Markov processes on a two-dimensional grid. One dimension indicates the number of products on stock, and thus is finite. We show that the equilibrium distributions of the Markov processes can be determined analytically. From this distribution it is straightforward to compute the distribution and moments of the production lead time. Numerical results show that combining MTO and MTS effectively reduces production lead times.

The literature on these hybrid systems is growing fast. The papers [7, 2, 6, 3] study models where several items are produced by a common facility with limited capacity. Some of these items are produced to stock, others are produced to order. The MTO items have priority over MTS items. For these systems a variety of questions

* Eindhoven University of Technology, Department of Mathematics and Computing Science, P.O.Box 513, 5600 MB - Eindhoven, The Netherlands.

is addressed, such as: 'Which items should be stocked or made to order?' and 'How must production capacity allocated among MTO and MTS items?'

In section 2 we first consider some simple queueing models for inventory systems related to the produce-to-stock systems in [1], and investigate the relation between inventory and the production lead time. In particular tails of the production lead time distribution tend to be rather thick. Section 3 deals with the two type system, standard products to stock, specific products to order. In section 4 we consider the production in two phases, the first one being standard, the second one customer specific.

2. Inventories and production lead times. Let us first consider the following model. Jobs arrive according to a Poisson process with rate $\rho(<1)$, the production times are exponential with mean 1. In this system the production lead time is exponential with mean $1/(1-\rho)$. We now assume that some of the idle time is used to produce items on stock, say at most M items. As state space for this Markov process we take the set $\{-M, -M+1, \ldots, 0, 1, 2, \ldots\}$ where the negative numbers refer to a number of items on stock and the positive numbers to jobs waiting to be processed. The equilibrium probability p_n for the system to be in state n clearly equals

$$p_n = (1-\rho)\rho^{n+M}, \quad n = -M, -M+1, \ldots.$$

So the production lead time $S(M)$ is equal to 0 with probability $1-\rho^M$ and exponential with mean $1/(1-\rho)$ with probability ρ^M. Hence, for ρ close to 1, the maximum stock M has to be fairly large to get a significant effect. Table 1 shows the mean production lead time and the probability of a production lead time greater than 10 times the mean production time as a function of ρ and M.

TABLE 1

Performance characteristics for exponential production times.

		$E[S(M)]$						$P[S(M)>10]$					
M		0	8	16	24	32	40	0	8	16	24	32	40
	.8	5	.839	.141	.024	.004	.001	.135	.023	.004	.001	.000	.000
ρ	.9	10	4.30	1.85	.798	.343	.148	.368	.158	.068	.029	.013	.005
	.95	20	13.3	8.80	5.84	3.87	2.57	.607	.402	.267	.177	.118	.078

In case the production times have an Erlang distribution with r phases, each phase with mean $1/r$ (so the mean production time is 1), and the maximum stock is M items the analysis proceeds in an almost identical way. Now we take as state space the set $\{-Mr, -Mr+1, \ldots, 0, 1, 2, \ldots\}$ where the negative numbers refer to a number of phases completed in advance (so in state $-r$ there is one item on stock) and the positive numbers to the number of uncompleted phases in the system. The equilibrium probabilities p_n are for $n = -Mr, -Mr+1, \ldots$ given by (see e.g. [4])

$$p_n = \sum_{i=1}^{r} A_i x_i^{n+Mr},$$

where the x_i's are the roots of the equation

$$x^r - \rho(x^{r-1} + x^{r-2} + \ldots + 1)/r = 0,$$

and the coefficients A_i satisfy

$$A_i = (1-\rho)/\prod_{j\neq i}(1 - x_j/x_i).$$

The production lead time $S(M)$ is equal to 0 with probability $\sum_{-Mr}^{-r} p_n$ and it is the sum of $n+r$ exponentials with mean $1/r$ with probability p_n for $n = -r+1, -r+2, \ldots$ So for $t \geq 0$ we obtain that

$$\begin{aligned} P[S(M) > t] &= \sum_{n=-r+1}^{\infty} p_n \sum_{l=0}^{n+r-1} e^{-rt}\frac{(rt)^l}{l!} \\ &= \sum_{i=1}^{r} \frac{A_i}{1-x_i} x_i^{-r+1+Mr} e^{-r(1-x_i)t}, \end{aligned}$$

and thus also

$$E[S(M)] = \sum_{i=1}^{r} \frac{A_i}{r(1-x_i)^2} x_i^{-r+1+Mr}.$$

Table 2 shows how the production lead times are reduced by the production to stock of the standard items.

TABLE 2
Performance characteristics for Erlang-10 production times.

		$E[S(M)]$						$P[S(M) > 10]$					
M		0	8	16	24	32	40	0	8	16	24	32	40
	.8	3.2	.140	.006	.000	.000	.000	.027	.001	.000	.000	.000	.000
ρ	.9	5.95	1.32	.291	.064	.014	.003	.172	.038	.008	.002	.000	.000
	.95	11.5	5.47	2.61	1.24	.592	.282	.420	.200	.096	.046	.022	.010

By comparing the results in table 1 and table 2 we see that in case of Erlang production times the reductions in the production lead time are much larger than in case of exponential production times. This is also illustrated in figure 1.

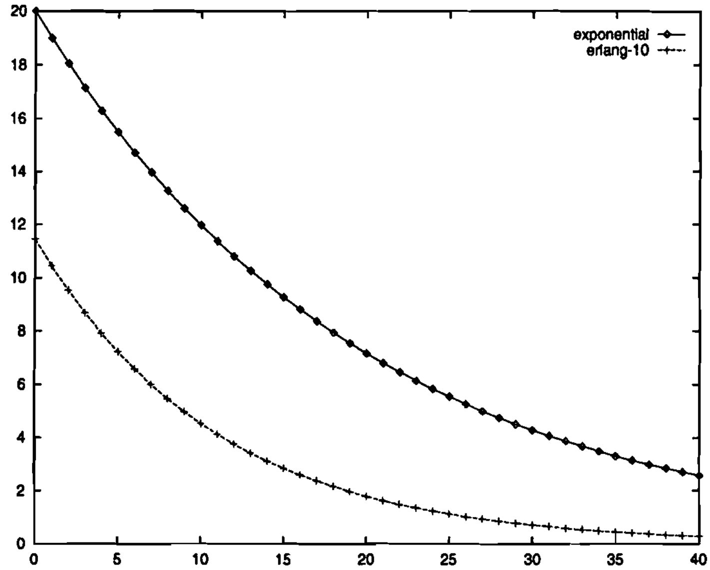

FIG. 1. *The mean production lead time as a function of the maximum stock M for $\rho = 0.95$ in case of exponential and Erlang-10 production times, resp.*

We finally note that for general phase-type production time distributions the analysis is equally simple.

3. Two types of jobs, standard (MTS) and customer specific (MTO). We now turn to a model which combines make to order and make to stock. There are two types of products, standard and non-standard products. When the machine is idle (i.e. there are no orders) it is used to produce standard products until the stock reaches a a certain maximum level, M say. A customer asking for a standard product receives it directly (if possible) from stock. If the stock is empty the customer order joins the queue. Non-standard, customer specific products are never delivered from stock, but always produced to order. Demand for standard products arrives according to a Poisson process with rate λ_1, the customer specific demand is Poisson as well with rate λ_2. Further we set $\lambda = \lambda_1 + \lambda_2$. The production times for both types of products are exponential with the same mean $1/\mu$. Production to stock is preempted by production to order. For stability we assume that $\rho = \lambda/\mu < 1$. Below we will first present an exact analysis of this system and then investigate the effect of producing to stock on the production lead times of standard and non-standard orders.

The system can be described by a Markov process with states (n, m) where n is the number of orders in the production queue and m the number of items on stock. If $m > 0$ all n jobs in the queue are type 2, but if $m = 0$ the n jobs are a mixture of type 1 and type 2 jobs. The flow diagram of this Markov process is depicted in figure 2.

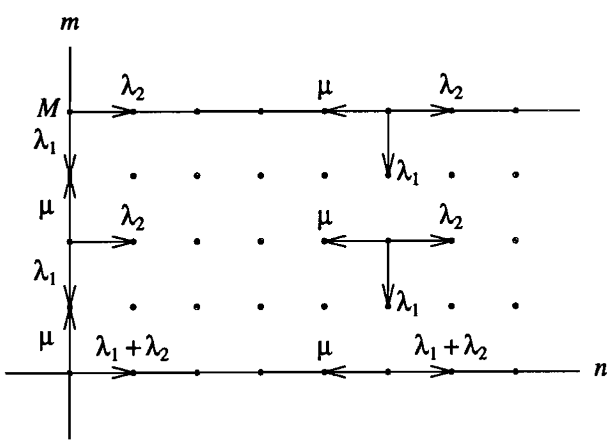

FIG. 2. *Flow diagram for the model with standard and customer specific jobs*

By balancing in each state (n, m) the flow out of and into that state we get the following set of equations:

$$
\begin{aligned}
&(1) \quad & p(n,M)(\lambda+\mu) &= p(n-1,M)\lambda_2 + p(n+1,M)\mu, \quad n > 0, \\
&(2) \quad & p(n,m)(\lambda+\mu) &= p(n,m+1)\lambda_1 + p(n-1,m)\lambda_2 + p(n+1,m)\mu, \\
& & & \qquad n > 0, 0 < m < M, \\
&(3) \quad & p(n,0)(\lambda+\mu) &= p(n,1)\lambda_1 + p(n-1,0)\lambda + p(n+1,0)\mu, \quad n > 0, \\
&(4) \quad & p(0,M)\lambda &= p(0,M-1)\mu + p(1,M)\mu, \\
&(5) \quad & p(0,m)(\lambda+\mu) &= p(0,m+1)\lambda_1 + p(0,m-1)\mu + p(1,m)\mu, \\
& & & \qquad 0 < m < M, \\
&(6) \quad & p(0,0)(\lambda+\mu) &= p(0,1)\lambda_1 + p(1,0)\mu.
\end{aligned}
$$

Note that in the equations (2) there are no probabilities $p(l,k)$ with $k < m$, since it is only possible to consume and not produce stock items in states (l,k) with $l > 0$. Hence the balance equations can be solved as follows. We start to solve the equations for $m = M$ and then work our way down from $m = M - 1$ to $m = 0$. Equation (1) is a difference equation in n only. Its solution is given by

$$p(n, M) = A_{0,0}x^n,$$

where $A_{0,0}$ is (for the time being) free and x is a root of

$$x(\lambda + \mu) = \lambda_2 + x^2\mu. \tag{7}$$

One of the two roots of (7) is larger than 1, and thus not useful. The other one equals

$$x = \frac{\lambda + \mu - \sqrt{(\lambda+\mu)^2 - 4\lambda_2\mu}}{2\mu}.$$

This root is positive and less than 1. Substituting $p(n, M) = A_{0,0}x^n$ into equation (2) for $m = M - 1$ we obtain an inhomogeneous recurrence relation for the probabilities $p(n, M - 1)$, with the $p(n, M)$'s as the inhomogeneous terms. The solution of this recurrence relation is of the form

$$p(n, M-1) = A_1 x^n + A_2 n x^n .$$

The first term is the solution of the homogeneous equation, the second one is a particular solution of the inhomogeneous equation. It is convenient to rewrite the expression for $p(n, M - 1)$ in the form

$$p(n, M-1) = \sum_{j=0}^{1} A_{1,j} \binom{1+n}{j} x^n . \tag{8}$$

Next we substitute this expression into (2) for $m = M - 2$ and solve the probabilities $p(n, M - 2)$. Repeating this procedure we can work our way down from $M - 1$ to 1. This leads to the solution

$$p(n, M-k) = \sum_{j=0}^{k} A_{k,j} \binom{k+n}{j} x^n, \quad k = 0, \ldots, M-1, \ n = 0, 1, \ldots, \tag{9}$$

where for $k = 1, \ldots, M - 1$ the constants $A_{k,j}$ satisfy

$$A_{k,k} = \frac{\lambda_1 A_{k-1,k-1}}{\lambda + \mu - 2\mu x},$$

$$A_{k,j} = \frac{\mu x A_{k,j+1} + \lambda_1 A_{k-1,j-1}}{\lambda + \mu - 2\mu x}, \quad j = 1, \ldots, k-1.$$

The constants $A_{0,0}, A_{1,0}, \ldots, A_{M-1,0}$ are still free. Equation (3) is a difference equation in $p(n, 0)$ with the $p(n, 1)$'s given by (9) as inhomogeneous terms. The solution of this equation is given by

$$p(n, 0) = B\left(\frac{\lambda}{\mu}\right)^n + \sum_{j=0}^{M-1} A_{M,j} \binom{M+n}{j} x^n. \tag{10}$$

The first term with B free is the solution of the homogeneous equation. The sum is a particular solution of the inhomogeneous equation. The constants $A_{M,j}$ in this sum satisfy

$$A_{M,M-1} = -A_{M-1,M-1}x,$$

and

$$A_{M,j} = \lambda_1^{-1}\left[\{(\lambda+\mu)x - 2\mu x^2\}A_{M,j+1} - \mu x^2 A_{M,j+2} - \lambda_1 x A_{M-1,j}\right],$$

for $j = 0, \ldots, M-2$, where $A_{M,M} = 0$ by definition. We can conclude that the solution of the equations (1)-(3) is given by (9) and (10) in which the constants $A_{0,0}, \ldots, A_{M-1,0}$ and B are free. These constants can finally be determined from the boundary equations (4)-(5) together with the normalization condition.

Once the probabilities $p(n,m)$ are known, performance characteristics, such as e.g. the mean number of orders in the production queue and the mean production lead time for the two types of orders can be easily obtained. For instance, let $L(M)$ be the number of orders in the production queue. Then, by inserting the expressions for the probabilities $p(n,m)$ we find that

$$E[L(M)] = \sum_{n,m} np(n,m) = B\frac{\lambda/\mu}{1-\lambda/\mu} + \sum_{k=0}^{M}\sum_{j=0}^{k} A_{k,j}\sum_{n=0}^{\infty}\binom{k+n}{j}nx^n,$$

where the infinite sum can be reduced to a finite one by using

$$\sum_{n=0}^{\infty}\binom{k+n}{j}nx^n = \frac{j-k+(k+1)x}{x^{k-j}(1-x)^{j+2}} - \sum_{n=j-k}^{-1}\binom{k+n}{j}nx^n.$$

The mean overall production lead time $E[S(M)]$ (i.e. for type 1 and 2 orders together) can be obtained from $E[L(M)]$ by application of Little's law, i.e.,

$$E[S(M)] = E[L(M)]/\lambda.$$

Denote by $S_1(M)$ and $S_2(M)$ the production lead time for a standard and non-standard order, respectively. From the PASTA property (see [8]) we obtain

$$E[S_2(M)] = (E[L(M)]+1)/\mu.$$

Further, it holds that

$$E[S(M)] = \frac{\lambda_1}{\lambda}E[S_1(M)] + \frac{\lambda_2}{\lambda}E[S_2(M)],$$

from which $E[S_1(M)]$ directly follows.

The figures 3 and 4 illustrate for $\rho = 0.9$ and $\mu = 1$ how the mean production lead times for standard and non-standard orders decrease with M for the ratios λ_1/λ_2 equal to 3, 1 and 1/3, respectively. Note that for $M = 0$ and $M = \infty$ we have

$$E[S_1(0)] = E[S_2(0)] = \frac{1/\mu}{1-\rho_1-\rho_2},$$

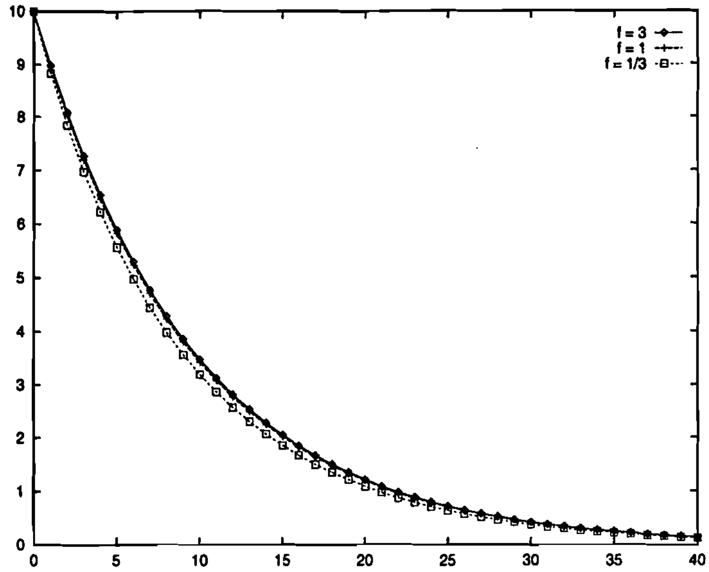

FIG. 3. *The mean production lead time for standard orders as a function of M for $\mu = 1$ and $\rho = 0.9$ where $f = \lambda_1/\lambda_2$ is varied as* 3, 1 *and* 1/3, *resp.*

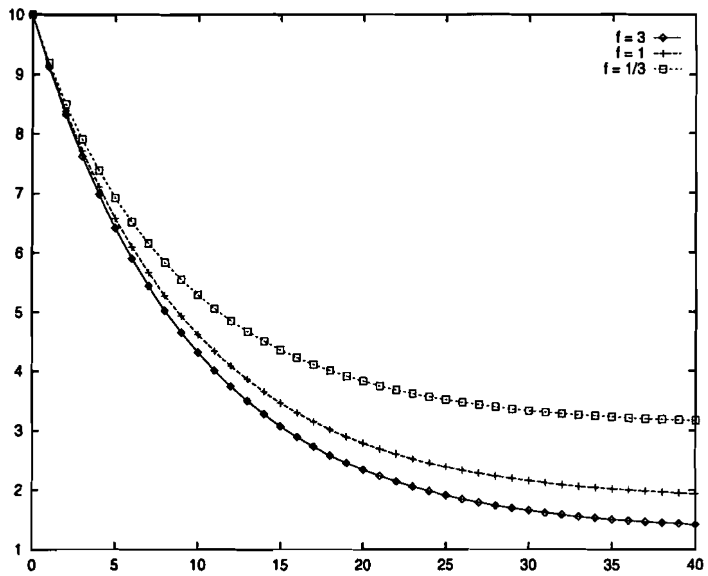

FIG. 4. *The mean production lead time for non-standard orders as a function of M for $\mu = 1$ and $\rho = 0.9$ where $f = \lambda_1/\lambda_2$ is varied as* 3, 1 *and* 1/3, *resp.*

$$E[S_1(\infty)] = 0\,, \quad E[S_2(\infty)] = \frac{1/\mu}{1-\rho_2}\,,$$

where $\rho_1 = \lambda_1/\mu$ and $\rho_2 = \lambda_2/\mu$.

We see that significant reductions in the lead time are possible, and that, apparently, the reductions for the standard orders are quite insensitive to the ratio λ_1/λ_2. Most of the reduction, 80% say, is already achieved for moderate M. This is also shown in table 3 where we list the minimal M for which $\alpha\%$ of the possible reduction in the lead time is realized, i.e.

$$E[S_1(0)] - E[S_1(M)] \geq \frac{\alpha}{100} E[S_1(0)] \, ,$$

$$E[S_2(0)] - E[S_2(M)] \geq \frac{\alpha}{100} (E[S_1(0)] - E[S_1(\infty)]) \, .$$

Again observe that the minimal M is quite insensitive to the ratio λ_1/λ_2.

TABLE 3

Minimal M to achieve α% of the possible reduction in the lead time

			M			
$\lambda_1/\lambda_2 = 3$		α	20	50	80	90
	.8		2	5	9	12
ρ	.9		4	8	17	23
	.95		6	15	33	46
$\lambda_1/\lambda_2 = 1$						
	.8		2	4	8	11
ρ	.9		4	8	17	23
	.95		6	15	33	46
$\lambda_1/\lambda_2 = 1/3$						
	.8		2	4	7	10
ρ	.9		3	7	16	22
	.95		6	15	32	46

The figures 3 and 4 and table 3 illustrate the reductions in the lead time achieved for given *maximum* stock M. The maximum stock size is important for the allocation of space needed for storage, the average stock is relevant for the calculation of holding costs. The average stock can be determined as follows. There is a simple relation between the probability $p(0, m-1)$ and the probability $q(m)$ that there are m items on stock. Namely, balancing the flow between the sets $\{(0, m-1), (1, m-1), \ldots\}$ and $\{(0, m), (1, m), \ldots\}$ yields

$$p(0, m-1)\mu = q(m)\lambda_1 \, , \quad m = 1, \ldots, M \, .$$

From this relation we can compute the probabilities $q(m)$ and thus also the average stock. In figure 5 we show for $\rho = 0.9$ and $\mu = 1$ the average stock as a function of M for the ratios λ_1/λ_2 equal to 3, 1 and 1/3, respectively. Observe that from $M = 15$ say, the lines are nearly straight with slope 1, so that from there on it is not worth to stock another item.

It is also possible to compute the distribution of the production lead time from the equilibrium probabilities $p(n, m)$. The production lead time $S_1(M)$ for a standard order is the sum of $n+1$ exponentials with mean $1/\mu$ with probability $p(n, 0)$ and it is equal to 0 otherwise. For a non-standard order the production lead time $S_2(M)$ is equal to the sum of $n+1$ exponentials with probability $p(n, m)$. Hence, for $t \geq 0$ we obtain that

$$P[S_1(M) > t] = \sum_{n=0}^{\infty} p(n, 0) \sum_{l=0}^{n} e^{-\mu t} \frac{(\mu t)^l}{l!} \, ,$$

$$P[S_2(M) > t] = \sum_{n=0}^{\infty} \sum_{m=0}^{M} p(n, m) \sum_{l=0}^{n} e^{-\mu t} \frac{(\mu t)^l}{l!} \, .$$

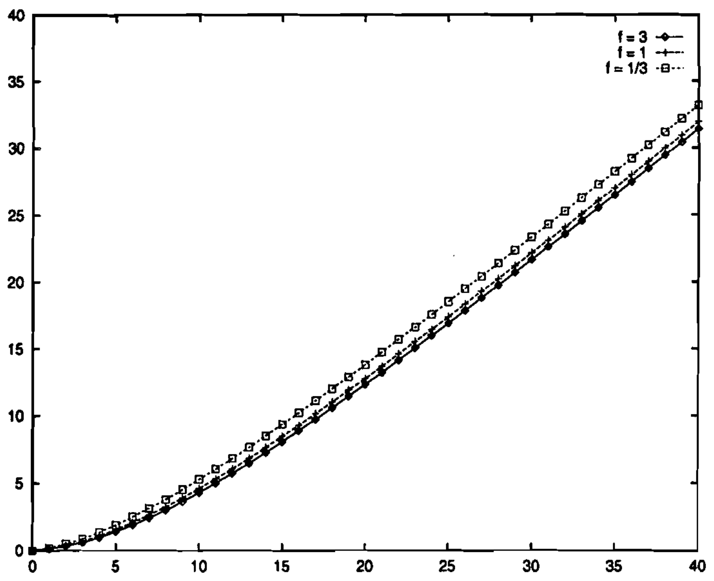

FIG. 5. *The average stock as a function of M for $\mu = 1$ and $\rho = 0.9$ where $f = \lambda_1/\lambda_2$ is varied as 3, 1 and 1/3, resp.*

Table 4 lists the percentages of the standard and non-standard orders that will get a production lead time more than 10 times the mean production time.

TABLE 4

Percentages of the standard and non-standard orders with a lead time longer than 10 times the mean production time

		$P[S_1(M) > 10]$					$P[S_2(M) > 10]$					
$\lambda_1/\lambda_2 = 3$	M	0	8	16	24	32	0	8	16	24	32	∞
ρ .8		14	2.3	.38	.06	.01	14	2.3	.41	.10	.04	.03
.9		37	16	6.8	2.9	1.3	37	16	6.8	3.0	1.3	.04
.95		61	40	27	18	12	61	40	27	18	12	.05
$\lambda_1/\lambda_2 = 1$												
ρ .8		14	2.3	.38	.06	.01	14	2.4	.61	.31	.26	.25
.9		37	16	6.8	2.9	1.3	36	16	7.1	3.3	1.6	.41
.95		61	40	27	18	12	61	40	27	18	12	.52
$\lambda_1/\lambda_2 = 1/3$												
ρ .8		14	2.1	.32	.05	.01	14	3.4	2.1	1.9	1.8	1.8
.9		37	15	6.5	2.8	1.2	37	17	9.4	6.2	4.9	3.9
.95		61	40	26	17	12	61	41	29	21	16	5.6

4. Half-finished products to stock.

In this section we consider products that are produced in two phases. The first phase is standard and identical for all products. The second phase is customer specific. We will analyze how the production lead time decreases if we store some, at most M, half-finished products, that is products with phase 1 completed. The first phase of a product takes an exponential time with mean $1/\mu_1$, the second phase is exponential with mean $1/\mu_2$. Orders arrive for one item at a time according to a Poisson process with rate λ. For stability we assume that

$$\frac{\lambda}{\mu_1} + \frac{\lambda}{\mu_2} < 1. \tag{11}$$

The system can be described by a Markov process with states (n, m) with n the number of orders in the system and m the number of half-finished products in stock or in use. So state $(1, 2)$ denotes the situation with 1 job in the system for which the machine is processing phase 2, and 1 half-finished product on the shelf. If phase 2 is completed the state changes to $(0, 1)$ and the machine continues with the production of phase 1 products until the limit M is reached, or a new customer order enters. So, a newly arriving order preempts a stock production. In state $(n, 0)$ with n positive the machine is working on phase 1. If phase 1 is completed the state changes to $(n, 1)$ and the machine continues with phase 2. The flow diagram of this Markov process is depicted in figure 6.

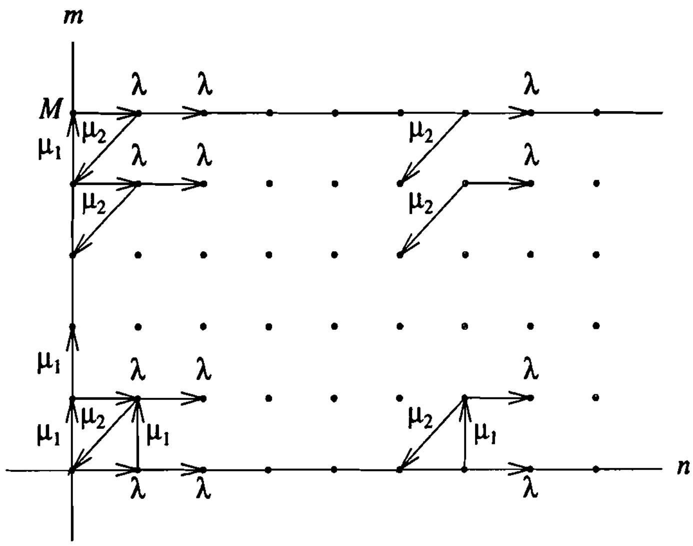

FIG. 6. *Flow diagram for the model with half-finished products*

By equating in each state (n, m) the flow out of and into that state we get:

$$
\begin{array}{llcl}
(12) & p(n,M)(\lambda+\mu_2) & = & p(n-1,M)\lambda, \quad n>0, \\
(13) & p(n,m)(\lambda+\mu_2) & = & p(n-1,m)\lambda + p(n+1,m+1)\mu_2, \\
 & & & \quad n>0, 1<m<M, \\
(14) & p(n,1)(\lambda+\mu_2) & = & p(n-1,1)\lambda + p(n,0)\mu_1 + p(n+1,2)\mu_2, \quad n>0, \\
(15) & p(n,0)(\lambda+\mu_1) & = & p(n-1,0)\lambda + p(n+1,1)\mu_2, \quad n>0, \\
(16) & p(0,M)\lambda & = & p(0,M-1)\mu_1, \\
(17) & p(0,m)(\lambda+\mu_1) & = & p(0,m-1)\mu_1 + p(1,m+1)\mu_2, \\
 & & & \quad 2<m<M, \\
(18) & p(0,2)(\lambda+\mu_1) & = & p(0,1)\mu_1 + p(1,3)\mu_2, \\
(19) & p(0,1)(\lambda+\mu_1) & = & p(0,0)\mu_1 + p(1,2)\mu_2, \\
(20) & p(0,0)(\lambda+\mu_1) & = & p(1,1)\mu_2.
\end{array}
$$

The approach to solve the equations (12)-(20) is similar to the one in the previous section. We start with the equations for $m = M$ and then subsequently solve the equations for $m = M - 1$, $m = M - 2$ and so on. Equation (12) is a difference equation in n only. Its solution is given by

$$p(n, M) = A_{0,0} x^n ,$$

where $A_{0,0}$ is free and

$$x = \frac{\lambda}{\lambda + \mu_2} .$$

By substituting $p(n, M) = A_{0,0}x^n$ into (13) for $m = M - 1$ we get an inhomogeneous recurrence relation for the probabilities $p(n, M-1)$, with the $p(n, M)$ as inhomogeneous terms. The solution of this recurrence relation can be written as (cf. (8))

$$p(n, M-1) = \sum_{j=0}^{1} A_{1,j} \binom{1+n}{j} x^n .$$

Next we insert this expression into (13) for $m = M-2$. This yields an inhomogeneous recurrence relation for the probabilities $p(n, M-2)$, which can readily be solved. By repeating this procedure for $m = M - 3$ down to $m = 2$ we obtain

$$p(n, M-k) = \sum_{j=0}^{k} A_{k,j} \binom{n+j}{j} x^n , \quad k = 0, \ldots, M-2, \; n = 0, 1, \ldots, \tag{21}$$

where for $k = 1, \ldots, M - 2$ the constants $A_{k,j}$ satisfy

$$A_{k,k} = \frac{x\mu_2}{\lambda + \mu_2} A_{k-1,k-1} ,$$

$$A_{k,j} = A_{k,j+1} + \frac{x\mu_2}{\lambda + \mu_2} A_{k-1j-1} , \quad j = k-1, \ldots, 1 .$$

This completes the solution of the equations (12)-(13). The constants $A_{0,0}, \ldots, A_{M-2,0}$ in this solution are still free. We now consider the equations (14)-(15). Note that these equations are difference equations in $p(n, 0)$ and $p(n, 1)$ with the $p(n, 2)$'s given by (21) as inhomogeneous terms. The solution is given by

$$p(n, 1) = B_1 y_1^n + B_2 y_2^n + \sum_{j=0}^{M-2} A_{M-1,j} \binom{n+j}{j} x^n , \tag{22}$$

$$\begin{aligned} p(n, 0) = {} & B_1 \frac{\lambda + \mu_2 - \lambda y_1^{-1}}{\mu_1} y_1^n + B_2 \frac{\lambda + \mu_2 - \lambda y_2^{-1}}{\mu_1} y_2^n \\ & + \sum_{j=0}^{M-2} A_{M,j} \binom{n+j}{j} x^n . \end{aligned} \tag{23}$$

The first two terms in (22)-(23) are the solution of the homogeneous equations, where B_1 and B_2 are free and y_1 and y_2 are the two roots of

$$\mu_1 \mu_2 y^2 - [\lambda^2 + \lambda(\mu_1 + \mu_2)] y + \lambda^2 = 0 .$$

Condition (11) guarantees that the roots y_1 and y_2 are both positive and less than 1. The two sums in (22)-(23) are a particular solution of the inhomogeneous equations.

Note that the particular solution has the same form as the inhomogeneous term $p(n,2)$ given by (21). For the constants $A_{M,j}$ and $A_{M-1,j}$ in the two sums the following relations can be derived. The constants $A_{M,j}$ in (23) satisfy

$$A_{M,M-2} = -\frac{\mu_2 x}{\mu_1} A_{M-2,M-2} ,$$

and

$$\begin{aligned} A_{M,j} &= \frac{\lambda x^{-1}}{\mu_1 \mu_2 x} \left[(\mu_1 - \mu_2) A_{M,j+1} + \lambda x^{-1} A_{M,j+2} \right] \\ &\quad - \sum_{l=j+1}^{M-2} A_{M,l} - \frac{\mu_2 y}{\mu_1} \sum_{l=j}^{M-2} (l+1-j) A_{M-2,l} , \end{aligned}$$

for $j = M-3, \ldots, 0$, where $A_{M,M-1} = 0$ by definition. The constants $A_{M-1,j}$ in expression (22) follow from the relations

$$A_{M-1,j} = \frac{1}{\lambda x^{-1}} \left[\mu_1 A_{M,j-1} + \mu_2 x \sum_{m=j-1}^{M-2} A_{M-2,m} \right]$$

for $j = 1, \ldots, M-2$ and

$$A_{M-1,0} = \frac{1}{\mu_2 x^{-1}} \left[(\lambda + \mu_1 - \lambda x^{-1}) A_{M,0} + \lambda x^{-1} A_{M,1} - \mu_2 x \sum_{l=1}^{M-2} A_{M-1,l} \right] .$$

We can now conclude that the solution of the equations (12)-(17) is given by (21)-(23) in which the constants $A_{0,0}, \ldots, A_{M-2,0}$ together with B_1 and B_2 are free. These constants can finally be solved from the boundary equations (16)-(20) and the normalization condition.

It is straightforward to determine performance characteristics, such as the mean number of orders in the system and the mean lead time from the equilibrium probabilities $p(n,m)$. For instance, for the mean number of orders in the system we obtain

$$\begin{aligned} E[L(M)] &= \left(1 + \frac{\lambda + \mu_2 - \lambda y_1^{-1}}{\mu_1}\right) \frac{B_1 y_1}{(1-y_1)^2} + \left(1 + \frac{\lambda + \mu_2 - \lambda y_2^{-1}}{\mu_1}\right) \frac{B_2 y_2}{(1-y_2)^2} \\ &\quad + \sum_{k=0}^{M-2} \sum_{j=0}^{k} A_{k,j} \frac{(1+j)x}{(1-x)^{2+j}} \\ &\quad + \sum_{j=0}^{M-2} A_{M-1,j} \frac{(1+j)x}{(1-x)^{2+j}} + \sum_{j=0}^{M-2} A_{M,j} \frac{(1+j)x}{(1-x)^{2+j}} , \end{aligned}$$

where we used that

$$\sum_{n=0}^{\infty} \binom{n+j}{j} n x^n = \frac{(1+j)x}{(1-x)^{2+j}} .$$

The mean production lead time follows from $E[L(M)]$ and Little's law, i.e.,

$$E[S(M)] = E[L(M)]/\lambda .$$

Note that for $M = 0$ and $M = \infty$ it holds that

$$E[S(0)] = \frac{1/\mu_1 + 1/\mu_2 - \rho_2/\mu_1}{1 - \rho_1 - \rho_2},$$

$$E[S(\infty)] = \frac{1/\mu_2}{1 - \rho_2},$$

where $\rho_1 = \lambda/\mu_1$ and $\rho_2 = \lambda/\mu_2$. Figure 7 illustrates how the mean production lead time decreases with M for the ratios μ_1/μ_2 equal to 3, 1 and 1/3, respectively. In each case the mean production time (i.e., $1/\mu_1 + 1/\mu_2$) is equal to 1.

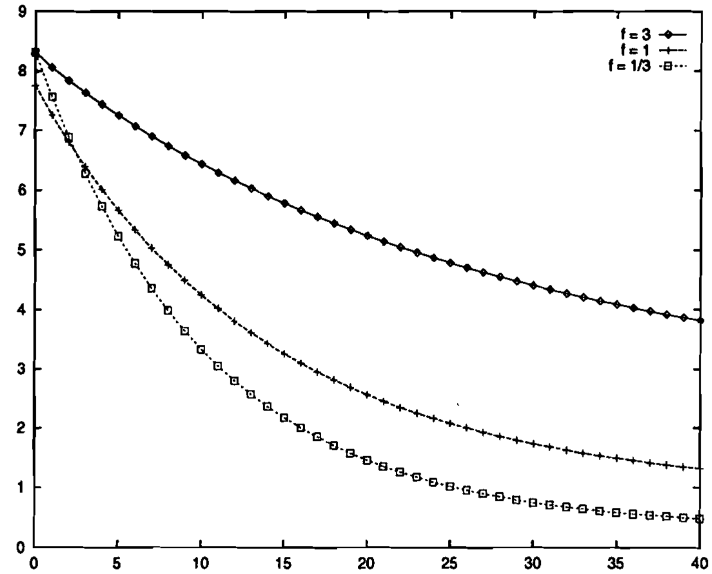

FIG. 7. *The mean production lead time as a function of M for mean production time* 1 *and* $\rho = 0.9$ *where* $f = \mu_1/\mu_2$ *is varied as* 3, 1 *and* 1/3, *resp.*

We see that significant reductions in the lead time are possible. The reduction is large if phase 1 constitutes a large part of the total production time, thus if μ_1/μ_2 is small. In table 5 we list the minimal M for which $\alpha\%$ of the possible reduction in the lead time is realized, i.e.

$$E[S(0)] - E[S(M)] \geq \frac{\alpha}{100}(E[S(0)] - E[S(\infty)].$$

Figure 7 and table 5 show the reductions in the lead time for given maximum stock level M. It is interesting to know the average stock in these cases. To compute the average stock we can use the following relation between the probability $q(m)$ that there are m half-finished products in stock and $p(0, m-1)$. By balancing the flow between the sets $\{(0, m-1), (1, m-1), \ldots\}$ and $\{(0, m), (1, m), \ldots\}$ it follows that

$$p(0, m-1)\mu_1 = (q(m) - p(0, m))\mu_2, \quad m = 2, \ldots, M.$$

Also note that $q(1) = p(0, 1)$. From this relation we can calculate the probabilities $q(m)$, and thus also the average stock. Figure 8 illustrates the average stock for the examples in figure 7. We see that for $\mu_1/\mu_2 = 1, 1/3$ the lines are straight with slope

TABLE 5

Minimal M to achieve $\alpha\%$ of the possible reduction in the lead time

		M			
$\mu_1/\mu_2 = 3$	α	20	50	80	90
	.8	3	8	20	29
ρ	.9	6	20	48	69
	.95	14	43	102	146
$\mu_1/\mu_2 = 1$					
	.8	2	5	11	16
ρ	.9	4	10	23	33
	.95	7	21	48	68
$\mu_1/\mu_2 = 1/3$					
	.8	2	4	8	12
ρ	.9	3	8	17	24
	.95	5	15	34	49

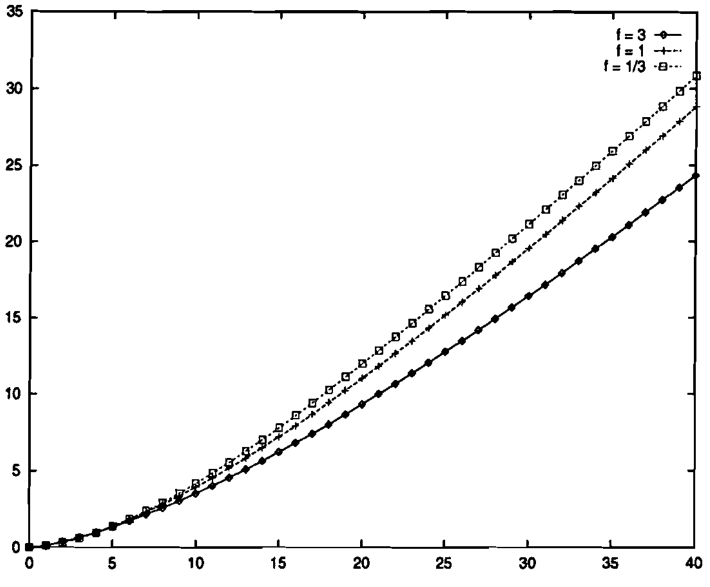

FIG. 8. *The average stock as a function of M for mean production time 1 and $\rho = 0.9$ where $f = \mu_1/\mu_2$ is varied as 3, 1 and 1/3, resp.*

1 from $M = 20$ say, so that from there on an extra half-finished product on stock will not leave the shelf. For $\mu_1/\mu_2 = 3$, however, it will help a little bit to produce extra half-finished products, even when there are already 40 items on stock.

It is also possible to compute the distribution of the production lead time from the equilibrium probabilities $p(n,m)$. Let $S(n,m)$ be the lead time if on arrival the system is in state (n,m). So, if $n+1 \leq m$, then $S(n,m)$ is the sum of $n+1$ phase 2 productions (and no phase 1 productions), and otherwise, $S(n,m)$ is the sum of $n+1-m$ phase 1 and $n+1$ phase 2 productions. Since an arriving order finds the system with probability $p(n,m)$ in state (n,m) we obtain for $t \geq 0$ that

$$P[S(M) > t] = \sum_{n,m} p(n,m) P[S(n,m) > t]\,.$$

From this relation it is straightforward to calculate the probability $P[S(M) > t]$ once the probabilities $p(n,m)$ are known. This is done for the examples in table 6 where the percentages of the orders that will get a production lead time more than 10 times the mean production time are listed.

TABLE 6

Percentages of the orders with a lead time longer than 10 times the mean production time

		$P[S(M) > 10]$					
$\mu_1/\mu_2 = 3$	M	0	8	16	24	32	∞
	.8	8.8	5.3	3.2	2.0	1.4	.48
ρ	.9	30	23	18	14	11	1.3
	.95	55	48	43	38	34	2.2
$\mu_1/\mu_2 = 1$							
	.8	7.2	2.2	.68	.21	.07	.00
ρ	.9	27	16	8.9	5.1	2.9	.00
	.95	52	40	30	23	18	.00
$\mu_1/\mu_2 = 1/3$							
	.8	8.8	1.7	.32	.06	.01	.00
ρ	.9	30	14	6.3	2.9	1.3	.00
	.95	55	37	26	18	12	.00

5. Concluding remarks. In this paper we studied two production-to-order systems and we investigated whether producing to stock in idle time can effectively reduce the production lead times. Indeed, it appeared that this simple combination of make to stock and make to order can significantly reduce the production lead time.

We explicitly solved the equilibrium equations of the models in sections 3 and 4 by combining products of powers. Another approach which is well suited to solve these equations is the matrix-geometric approach developed by Neuts [5]. In fact, for the models considered in this paper the corresponding rate matrices R can be solved exactly by a simple recursion.

REFERENCES

[1] J. A. BUZACOTT, J. G. SHANTHIKUMAR, *Stochastic models of manufacturing systems*, Prentice Hall, 1993.

[2] A. S. CARR, A. R. GÜLLÜ, P. L. JACKSON, J. A. MUCKSTADT, *An exact analysis of a production-inventory strategy for industrial suppliers*, Working paper, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, 1993.

[3] A. FEDERGRUEN, Z. KATALAN, *Make-to-stock or make-to-order: that is the question; novel answers to an ancient debate*, Working paper, Graduate School of Business, Columbia University, New York, 1994.

[4] L. KLEINROCK, *Queueing systems, Volume I: Theory*, J. Wiley & Sons, New York, 1975

[5] M. F. NEUTS, *Matrix-geometric solutions in stochastic models*, Johns Hopkins University Press, Baltimore, 1981.

[6] C. R. SOX, L. J. THOMAS, J. O. MCCLAIN, *Jellybeans: Sorting capacity to improve service*, Working paper, Department of Industrial Engineering, Auburn University, Alabama, 1994.

[7] T. M. WILLIAMS, *Special products and uncertainty in production/inventory systems*, EJOR, 15 (1984), pp. 46–54.

[8] R. W. WOLFF, *Poisson arrivals see time averages*, Opns. Res., 30 (1982), pp. 223–231.