DTIC FILE COPY

Working Paper No. 361

AD-A204 651

Contract N00014-85-K-0143

JOB-FILL INVENTORIES FOR SEQUENCES OF JOBS

by

John W. Mamer

and

Stephen A. Smith

August 1988

DTIC
ELECTE
FEB 2 7 1989
S D
D

WESTERN MANAGEMENT SCIENCE INSTITUTE
University of California, Los Angeles

89 2 21 068

Job-Fill Inventories For Sequences of Jobs

John W. Mamer

UCLA Graduate School of Management

Los Angeles, CA 90024

Stephen A. Smith

Dept. of Decision and Information Sciences

Leavey School of Business, Santa Clara University

Santa Clara, CA 95053

DTIC COPY INSPECTED 1

August 18, 1988

ABSTRACT

We consider the problem of determining optimal stock levels for a multi-item inventory system in which demands for items occur in randomly selected groups, causing in interdependent demand processes. This structure introduces dependencies that significantly affect the selection of item stock levels in systems whose performance is measured in terms of number of jobs completed or time until first stockout. Performance bounds and stock level selection techniques for several such systems are developed and analyzed. (SDW)

Accesion For
NTIS CRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification

By per HP
Distribution/

Availability Codes

Dist | Avail and / or Special
A-1

-A-

1 Introduction

We consider a multi-item inventory system of the following sort. Demands arrive randomly in the form of requests for subsets of the multiple items. A demand is met only when its entire subset of items is available. The key difference between this system and many typical inventory models is that it is the number of demands met (or not met) that is important, rather than the number of items supplied (or short).

A classical example of such an inventory system arises in the repair of machines in the field, e.g., the repair of computer terminals, photocopiers or other equipment at customer sites. Service personnel are equipped with repair kits containing a certain collection of parts and tools that are used for field repairs. Each repair job requires some particular set of parts and tools for completion. The kit is typically adequate only for some subset of the possible repair jobs. The perceived performance of the kit will be the expected fraction of the repair jobs for which it is adequate or the expected number of jobs which can be completed before stockout. This type of inventory system is often called a "job completion" or "job fill" inventory system. Other job completion inventory systems arise in stocking spare parts to assure the completion of a specific mission. Examples are voyages at sea, space flights or military missions. The performance criterion of the inventory is the probability that the mission will be completed before a required part is not available or the expected length of time to stockout. Another application arises in filling orders for merchandise. If orders occur in the form of sets of multiple items, the fraction of orders that can be filled completely is an important measure of the system's performance. An analogous problem arises in "build to order" production systems in which final products have common components.

Job completion based inventory systems have been considered in the management science literature in a variety of contexts. Sherbrooke (1971) considered an inventory system in which a failed line replaceable unit (LRU) could be replaced, i.e., the repair job completed, only if all failed subassemblies were available as spares. Silver (1972) developed a dynamic programming solution for this problem. These formulations assumed that subassemblies failed independently of one another, according to Poisson processes. Smith, Chambers and Shlifer (1980) obtained a closed form optimal stocking condition for this type of problem in the one period case, if the item stock levels are restricted to be 0, 1. Graves (1982) reformulated this as a 0,1 knapsack problem, which allows for the inclusion of linear constraints, such as inventory space requirements. Hausman (1982) developed an exchange curve analysis for this inventory stocking problem and Schaefer (1983) integrated the independent demand job completion formulation into an overall inventory management system.

A key assumption in the models discussed above is that individual item demands are independent of one another, so that the probability that any particular combination of items is required on a job is determined by the marginal demand rates of the individual items. This property does not hold in many of the examples cited earlier. In machine failures, for example, certain part failures may contribute to the failures of other parts, while in order filling systems, demands may be correlated (either positively or negatively) for a variety of reasons. Mamer and Smith (1982) developed an optimization algorithm, using the maximum flow/minimum cut theory, to solve the 0,1 stock level, multi-item case, when individual item demands are arbitrarily correlated and extended this approach to include the use of spare replacement machines [Mamer and Smith (1985).] March and Scudder (1984) developed an algorithm for determining optimal 0,1 stocking policies, subject to budget constraints. Recently, multi-item inventories are also being studied in the context of production processes with parts commonality among products. [Baker (1985), Baker, Magazine and Nuttle (1986).] As the examples analyzed in these papers indicate, determination of optimal stocking policies in such systems is difficult, even for products with only one common part. Given the large number of part types used in many production applications, the ability to analyze larger problems appears important in this application area as well. In a recent paper (1988) the authors develop a lower bound heuristic for a model which is quite similar to the one considered here, but in which the job arrival process is a Poisson process. We obtain bounds on system performance by exploiting the association property of part demand processes which flows from the Poisson arrivals assumption.

This paper considers a formulation that is more general than those considered previously for large problems. Item demands are dependent in an arbitrary manner, as in the more recent models discussed above. Stock levels are not restricted to 0,1 and multiple items of a given type are permitted on a particular job. Several alternative performance criteria are considered, including the probability of completing a fixed number of jobs, the expected number of jobs completed before first stockout, the expected time until stockout and the probability of serving all jobs arriving within a fixed period of time. The resupply process for this analysis, as in previous models, is fairly simple. It is assumed that once a request is unfilled (a job cannot be completed), all items are restocked to their originally specified levels. This neglects the effect of lead times and the possibility of stockouts at higher inventory echelons. This assumption appears to be quite reasonable for the field repair systems application discussed above, since repair personnel typically restock completely whenever they are required to return to the parts supply center. In one period inventory systems that support the completion of a single mission, there is no restocking process, so this model captures the relevant features of these systems. For the performance criteria involving a fixed time interval, this assumption is equivalent to a periodic resupply in which order levels can be readjusted up to the time of delivery.

This paper derives and compares a variety of upper and lower bounds for system performance. Since the tightness of the bounds varies by application, a set of upper and lower bounds is advantageous in that the maximum lower bound and minimum upper bound can then be chosen in any specific example.

Because of the complexity of the models considered in this paper, closed form solutions for optimal inventory policies cannot be obtained. In fact, the system performance for a given inventory level can be evaluated exactly only for problems having a very small number of different jobs and parts. System performance for a specific stocking policy can always be determined by Monte Carlo methods, however. When job arrivals form a renewal process upper and lower bounds on system performance can be obtained, which are expressed as simple functions of the item stock levels. The simple forms of the bounds allow optimization, subject to budget constraints and other linear constraints, such as inventory space. Optimization using bounds is important for this inventory system because it is observed that objective functions involving the exact performance criteria are neither convex or concave in the stock levels. Several example calculations are included in the paper to illustrate the relationships between the bounds and more precise calculations.

2 The Model

2.1 Specifications

The key demand data for problems of the type we are considering are contained in the "job matrix" J, which is defined as follows:

$$\begin{aligned} J_{ij} &= \text{the number of items of type } i \text{ required} \\ &\quad \text{to satisfy a request (complete a job) of type } j. \end{aligned}$$

[In production applications, the concept of a repair job is replaced by a final assembled product that contains some collection of items or subassemblies.]

Assumption 1:

Jobs or requests for items are assumed to have random independent interarrival times with mean $1/\lambda$. The marginal probabilities for job types are defined as follows:

$$p_j = P\{\text{an arriving job or request is of type } j\}.$$

The type of each job is assumed to be independent of the types of past and future jobs. The marginal demand rates for individual items λ_i $i = 1, \ldots, n$ are then determined by:

$$\lambda_i = \lambda \sum_j J_{ij} p_j. \tag{1}$$

where λ_i = average demand rate per unit time for item i, $i = 1, \ldots, n$.

For any fixed time t, we define the random variables

$N_j(t)$ = number of requests (jobs) of type j arriving by time t, $j = 1, \ldots, m$

$X_i(t)$ = number of items of type i demanded by time t, $i = 1, \ldots, n$.

Letting

$$\begin{aligned} X(t) &= (X_1(t), \ldots, X_n(t)) \\ N(t) &= (N_1(t), \ldots, N_m(t)), \end{aligned}$$

we clearly have the matrix relationship

$$X(t) = JN(t). \tag{2}$$

Let $M(t),\ \ t \geq 0$ denote the counting process of job arrivals. That is, $M(t)$ is the number of jobs which have arrived by time t; i.e.,

$$M(t) = \sum_{j=1}^{M} N_j(t).$$

Throughout this paper we assume $M(t)$ is a renewal process, which restarts at each restocking point. The J matrix introduces dependence among the components of $X(t)$. Even when each job uses a single distinct part, the components of $X(t)$ may be correlated. We summarize the dependence among the components of $N(t)$ and $X(t)$:

For any $t \geq 0$

$$Cov(N_i(t), N_j(t)) = p_i p_j (Var(M(t)) - E(M(t))).$$

If $Var(M(t)) - E(M(t)) \geq 0$ then for all i, j

$$Cov(X_i(t), X_j(t)) \geq 0.$$

2.2 Defining Performance Measures

In this subsection, we derive expressions for performance measures as a function of the initial stock level. These are defined in terms of two random variables:

τ = time of the first stockout

σ = number of jobs completed (demands met) before stockout

The performance measures are:

- probability first stockout is after time t $\quad P\{\tau > t\}$
- expected time to first stockout $E(\tau)$
- probability that at least k jobs completed before stockout $P\{\sigma > k\}$
- expected number of jobs completed before stockout $E(\sigma)$

The most appropriate measure depends upon the application. For mission completion problems, $P\{\tau > t\}$ would be appropriate for a mission of given duration t, or $E(\tau)$ if the duration is uncertain. For field repair applications $E(\sigma)$ or $P\{\sigma > k\}$

would be appropriate, because these determine the probability that a repairman can complete a given number of calls without stocking out. In the field repair setting, the occurrence of a stockout for a single part forces the repairman to return to the warehouse for the needed part and to restock his kit at the same time. For the problem of repairing ships at sea, a stockout in a single critical part entails a return to port or some form of emergency delivery. For specified item holding costs h_i, these performance criteria lead to corresponding stock level optimization problems. Examples include minimizing the inventory cost of achieving a specified performance level, or maximizing the performance that can be achieved with an inventory budget constraint and/or space constraints.

For a given vector $s = s_1, \ldots, s_n$ of stock levels, where s_i = initial stock level for item type i, τ the time to first stockout is defined by:

$$\tau = \inf\{t : X_i(t) \geq s_i + 1 \text{ for some } i,\ i = 1, \ldots, n\}.$$

In words, τ is the first time that the demand for some part or parts strictly exceeds supply. Thus

$$P\{\text{All requests filled up to time } t\} = P\{X(t) \leq s\} = P\{\tau > t\}. \tag{3}$$

When the time horizon is not known with certainty, a reasonable performance criterion is the expected time to first stockout, $E(\tau)$,

$$E(\tau) = \int_0^\infty P\{\tau > t\} dt.$$

Wald's equation provides a relationship between $E(\tau)$ and $E(\sigma)$. Define σ by

$$\sigma = \min\{k : X_i(k) \geq s_i + 1 \text{ some } 1 \leq i \leq n\}$$

Let Y denote the (random) time between arrivals of jobs, where $E(Y) = 1/\lambda$. Wald's equation yields the identity

$$E(\tau) = E(Y)E(\sigma) = E(\sigma)/\lambda. \tag{4}$$

Let

$$N_j(k) = \begin{array}{l}\text{number of jobs (demands) of type } j \text{ in the} \\ \text{first } k \text{ arriving jobs since restocking.}\end{array}$$

The marginal probability distribution for $N_j(k)$ is binomial

$$P\{N_j(k) = u\} = \begin{pmatrix} k \\ u \end{pmatrix} p_j^u (1 - p_j)^{k-u}.$$

Let T_k denote the time of arrival of the k^{th} job. We denote the functions of the discrete index k by $N_j(k) = N_j(T_k)$ and $X_i(k) = X_i(T_k)$.

For fixed k, the distribution of jobs by type is a multinomial, since the type of an arriving job is independent of both its interarrival time and its neighboring job types. Thus,

$$P\{\sigma > k\} = \sum_{z \in I_k \cap H} \frac{k!}{z_1! \ldots z_m!} p_1^{z_1} \cdots p_m^{z_m}, \tag{5}$$

where $I_k = \{z | \sum_i z_i = k, z_i \geq 0\}$ and $H = \{z | z \geq 0, Jz \leq s\}$. Now,

$$E(\sigma) = \sum_{k=0}^{\infty} \sum_{z \in H \cap I_k} \frac{n!}{z_1! \cdots z_m!} p_1^{z_1} \cdots p_m^{z_m} = \sum_{z \in H} \frac{(\sum_{i=1}^m z_i)!}{z_1! \ldots z_m!} p_1^{z_1} \ldots p_m^{z_m}. \tag{6}$$

Equation (5) can be used to obtain $P\{\tau > t\}$ via the identity,

$$\begin{aligned} P\{\tau > t\} &= \textstyle\sum_{\ell=0}^{\infty} P\{\sigma > \ell | M(t) = \ell\} P\{M(t) = \ell\} \\ &= \textstyle\sum_{\ell=0}^{\infty} P\{\sigma > \ell\} P\{M(t) = \ell\} \end{aligned}$$

where the last equality follows since the job types are independent of the job arrival process.

Equations (5) and (6) can, in principle, be used to calculate the performance criteria for a given stock level. The calculations, however, are quite time consuming even for fairly small numbers of items and stock levels. The difficulty is compounded when we seek optimal stock levels, a process which may require evaluating a large number of policies. Furthermore, computational experience has demonstrated that $E(\sigma)$ and $P\{\sigma > \ell\}$ are not concave functions of the stock level s. Thus global optimality cannot be guaranteed, even when local optima are located.

2.3 Bounds

In this section, we develop a series of upper and lower bounds for the performance measures. These bounds are important for two reasons: (1) the exact performance measures are tedious to evaluate and for large problems the computation becomes infeasible (2) the exact performance measures do not produce concave or convex optimization problems, respectively, when maximizing performance subject to a budget constraint or minimizing cost subject to a performance constraint. The bounds we develop, however, do lead to concave or convex problems, as a function of the stock levels. Thus the bounds can play the role of surrogate objective functions that allow approximately optimal stock levels to be easily determined. Furthermore, when the bounds are known to lie below the true performance level, the optimized bound can be set so that a performance target is guaranteed to be met. When several such lower bounds can be computed, as shown in later examples, the best (greatest) lower bound can be selected as the best estimate of performance.

The power of the bounds is increased in stock level optimization problems by combining them with more accurate (and time consuming) Monte Carlo evaluations of the exact performance measures. That is, we use the bounds in defining well behaved objective functions for determining "good" stock levels. We then use Monte Carol to evaluate the exact criterion $P\{\sigma > t\}$, $E(\sigma)$ etc. at that stock level. Based on this evaluation, the constraints in the surrogate optimization could then be adjusted and a new improved stock determined that more closely matches the exact performance level target.

We begin with upper and lower bounds on expected time to stockout.

Lemma 1 *With only Assumption 1, we have the following upper bound on expected time to stockout*

$$E(\tau) \leq (1/\lambda)U \tag{7}$$

where

$$U = \min_i\{\frac{s_i + 1 + \max_j\{J_{ij}\}}{\sum_j J_{ij}p_j}\}.$$

<u>Proof</u>: Define

$$\sigma_i = \inf\{k : X_i(k) \geq s_i + 1\}.$$

Clearly,

$$E\sigma = E\min\{\sigma_1, \ldots, \sigma_n\}.$$

It follows from Jensen's inequality that

$$E\sigma \leq \min\{E\sigma_1, \ldots, E\sigma_n\}.$$

The process $X_i(k)$ experiences an increment of J_{ij} with probability p_j. If we let $Z_1, \ldots$ denote the sequence of increments then we have

$$s_i + 1 \leq \sum_{i=1}^{\sigma_i} Z_i \leq s_i + 1 + \max_j \{J_{ij}\},$$

where the term on the far right is the maximum amount by which $X_i(\sigma)$ will exceed $s_i + 1$. Taking expectations on both sides and applying Wald's equation again yields

$$\frac{s_i + 1}{E(Z_i)} \leq E(\sigma_i) \leq \frac{s_i + 1 + \max_j(J_{ij})}{E(Z_i)}.$$

and $E(Z_i) = \sum_{j=1}^{m} p_j J_{ij}$. This yields

$$E(\sigma) \leq \min_i \left\{ \frac{s_i + \max_j\{J_{ij}\} + 1}{\sum_j J_{ij} p_j} \right\} = U \tag{8}$$

Q.E.D.

The bound in (8) is simple to compute. Thus the right hand side of (7) is easily determined.

It is often more important to obtain a lower bound on time to stockout. The lower bound provides a "pessimistic" estimate of performance. This is useful because, by purchasing sufficient inventory, the pessimistic bound can usually be brought to any specified level of system performance. This assures that acceptable performance will be achieved, although some excess inventory investment is likely because of the approximate nature of the bound.

Lemma 2 *With probability 1, σ the number of jobs until stockout satisfies*

$$\sigma_* \leq \sigma,$$

where

$$\sigma_* = \inf\{k | \sum_{j=1}^{m} \frac{N_j(k)}{R_j} \geq 1\}$$

and R_j denotes the maximum number of consecutive jobs of type j which can be served before stockout,

$$R_j = \min_{\{i | J_{ij} > 0\}} \{\frac{s_i + 1}{J_{ij}}\}.$$

Proof: It is sufficient to show,

$$\{\omega : JN(k) < s+1\} \supseteq \{\omega : \sum_{j=1}^{m} \frac{N_j(k)}{R_j} < 1\}.$$

Suppose that for some ω, $\sum_{j=1}^{m} \frac{N_j(k)}{R_j} < 1$, the definition of R_j then gives

$$\sum_{j=1}^{m} \frac{1}{\min_{\{i|J_{ij}>0\}}\{\frac{s_i+1}{J_{ij}}\}} N_j(k) < 1$$

Now

$$(\min_{\{i|J_{ij}>0\}}\{\frac{s_i+1}{J_{ij}}\})^{-1} = \max_i\{\frac{J_{ij}}{s_i+1}\}$$

Thus

$$\sum_{j=1}^{n} \max\{\frac{J_{ij}}{s_i+1}\} N_j(k) < 1$$

or

$$\sum_{j=1}^{n} J_{ij} N_j(k) < s_i + 1. \qquad i = 1, \ldots, n,$$

Q.E.D.

Lemma 2 produces a major simplification because the vector relationship $JN(k) < s+1$ is reduced to a one dimensional relationship expressed as a weighted sum of binomial random variables. Lemma 2 obtains a pessimistic bound for the probability distribution of the number jobs that are completed before stockout. In some applications, the probability distribution of the time until stockout is a more appropriate performance measure. It can be noted in the proof of Lemma 2, that all the implications are equally valid if $N_j(k)$ is replaced with $N_j(t)$, the number of jobs of type j arriving by time t. Thus, using arguments analogous to those in Lemma 2, we can establish the following lower bound on the probability distribution of the time to stockout.

Lemma 3

$$P\{X(t) \le s\} = P\{\tau > t\} \ge P\{\tau_* > t\}$$

where

$$\tau_* = \inf\{t | \sum_{j=1}^{m} N_j(t)/R_j \ge 1\}$$

and

$$R_j = \min_{\{i|J_{ij}>0\}}\{(s_i+1)/J_{ij}\}.$$

The situation in Lemmas 2 and 3 is depicted graphically in Figure 1. Suppose there are 2 possible jobs and 2 parts. The first type of job involves part 1 only and the second parts 1 and 2. Suppose we stock 3 of part 1 and 2 of part 2. Plotted along each axis are the number of jobs of type 1 and 2 observed. The solid lines depict the constraints $Jx \leq s+1$. Thus the lattice points within this region indicate combinations of jobs which can be repaired from stock on hand. The dashed line is the set of points such that $\sum_{i=1}^{m} \frac{x_i}{R_i} = 1$. The process $N(k)$ moves from lattice point to lattice point in a random fashion, but always to the northeast. (A typical realization is given by the line marked with arrows). The lemma states that the $N(t)$ process always crosses the dashed lines before crossing the solid boundary. Lemma 2 thus implies that the complementary cumulative distribution of σ_* lies to the left of that of σ in the sense that $P\{\sigma_* > k\} \leq P\{\sigma > k\}$ all k. Thus the distribution of σ_* provides a bound on the distribution of σ.

2.4 Computing the Bounds

Probability of Stockout

The distribution of σ_* is tedious to calculate exactly, but it is considerably easier to compute than the distribution of σ. Define

$$H_k = \{(\ell \in Z_+^m) : \sum_{j=1}^{m} \ell_j = k, \sum_{j=1}^{m} \frac{\ell_j}{R_j} \leq 1\}.$$

Then

$$P_r\{\sigma_* > k\} = \sum_{\ell \in H_k} \cdots \sum \frac{k!}{\ell_1! \ldots \ell_m!} p_1^{\ell_1} \ldots p_m^{\ell_m}. \tag{9}$$

For large values of k, $(P\{\sigma_* > k\}$ can be approximated by the tail of a Normal distribution. Since the $N_j(k)$ $j = 1, \ldots, m$ are multinomially distributed, we let

$$\mu(k) = E(\sum_{j=1}^{m} \frac{N_j(k)}{R_j}) = \sum_{j=1}^{m} \frac{kp_j}{R_j}$$

$$\sigma^2(k) = Var(\sum_{j=1}^{m} \frac{N_j(k)}{R_j}) = k\sum_{j=1}^{m}(1/R_j)^2 p_j(1-p_j) - 2k\sum_{i=1}^{n} \sum_{j=i+1}^{m} (1/R_i)(1/R_j)p_i p_j.$$

Approximating the multinomial by a multivariate normal with the same mean and covariance matrix yields

$$P\{\sigma_* > k\} = P\{\sum_{j=1}^{m} \frac{N_j(k)}{R_j} \leq 1\} \approx \Phi(1|\mu(k), \sigma^2(k)),$$

where $\Phi(\cdot|a,b)$ represents a univariate normal distribution with mean a and variance b.

The Normal approximation can also be used to compute $P\{\tau_* > t\} = P\{\sum_{i=1}^m N_i(t)/R_i \leq 1\}$ for Lemma 3. Here we simply replace $N_j(k)$ above with $N_j(t)$ and determine the revised mean and variance $\mu(t)$ and $\sigma^2(t)$. We then have by the Central Limit Theorem that

$$P\{\tau_* > t\} \approx \Phi(1/\mu(t), \sigma^2(t)).$$

This approximation is best when t is such that the $N_j(t)$ are fairly large. This in turn implies that the s_i should be large enough to serve a substantial number of jobs. Thus, the Normal approximations are appropriate when both the stock levels and the number of jobs to be completed before stockout are large.

Expected Time to Stockout

The expected time to stockout $E(\tau)$ can be bounded below by using the probability distribution $P\{\tau_* > t\}$. That is,

$$E(\tau_*) = \int_0^\infty P\{\tau_* > t\}dt < \int_0^\infty P\{\tau > t\}dt = E(\tau). \tag{10}$$

Thus $E(\tau_*)$ provides a lower bound on $E(\tau)$. Given the complexity of the analytical expression for the distribution of σ_* (and hence τ_*), the integral in (10) may be difficult to calculate. We can, however, develop a simple, precise approximation to $E(\sigma_*)$ (and hence to $E(\tau_*)$).

Lemma 4 *The pessimistic bound $E(\sigma_*)$ on number of jobs until stockout satisfies*

$$1/(\sum_{j=1}^m p_j/R_j) \leq E(\sigma_*) \leq [1 + \max_j\{1/R_j\}]/(\sum_{j=1}^m p_j/R_j).$$

Proof: Define W_k by

$$W_k = N(k) - pk$$

where $p = \{p_1, \ldots, p_m\}$; W_k is a discrete time, R^m valued martingale. To see this, note

$$E(W_{k+1} \mid W_k) = E\{W_k + w_{k+1} - p \mid \mathcal{W}_k\},$$

where $P(w_{k+1} = e_i) = p_i$, e_i representing the i^{th} unit vector in R^m. It follows then that

$$\begin{aligned} E(W_{k+1} \mid W_k) &= E(W_k \mid W_k) + E(w_{k+1} - p) \\ &= W_k. \end{aligned}$$

For each n, $\| W_k \| \leq n$ establishing the martingale property. Note that since $N(k)$ moves only to the northeast (increases according to the usual vector partial order) this implies that $P(\sigma_* < \infty) = T$; hence, $\sigma_* \leq \sum_{i=1}^{n}(s_i + 1)$. The optional sampling theorem applied to W_k yields

$$0 = E(W_{\sigma_*}) = E(N(\sigma_*)) - pE(\sigma_*). \tag{11}$$

Thus $E(N(\sigma_*)) = pE(\sigma_*)$. It follows from the definition of σ_*, however, that with probability 1

$$1 \leq \sum_{j=1}^{m} \frac{1}{R_j} N_j(\sigma_*) \leq 1 + \max_j(\frac{1}{R_j}) \tag{12}$$

Taking inner products on both sides of the vector equation (12) with respect to the vector $(1/R_1, \ldots, 1/R_m)$ and applying (13) yields

$$1 \leq \sum_{j=1}^{m} \frac{p_j}{R_j} E(\sigma_*) \leq 1 + \max(1/R_j). \tag{13}$$

Where the last term of (15) bounds the amount by which $N(k)$ may overshoot the boundary of the set $\sum_{j=1}^{n} \frac{p_j}{R_j} \leq 1$. Solving (14) yields a lower bound L

$$L = (\sum_{j=1}^{m} \frac{p_j}{R_j})^{-1} \leq E(\sigma_*) \leq (\sum_{j=1}^{m} \frac{p_j}{R_j})^{-1}(1 + \max_j(1/R_j)). \tag{14}$$

Q.E.D.

Note that in the extreme case in which each job uses every part, the bound L is tight. In view of equation (7), L/λ is a lower bound for $E(\tau)$. Summarizing, we have obtained the following bounds and approximations for the performance measures.

Table 2.1. Summary of Bounds Obtained

$E(\sigma)$	expected no of jobs completed
LOWER	$(1+\max_j\{1/R_j\})/\sum_{j=1}^{m} p_j/R_j$
UPPER	$U = \min_i \left\{\frac{s_i+1+\max_j\{J_{ij}\}}{\sum_j J_{ij}p_j}\right\}$
$P\{X(t) \le S\}\| = P\{\text{no stockout by } t\}\| = P\{\tau > t\}$	
LOWER	$P\{\sum_{j=1}^{m} N_j(t)/R_j \ge 1\}$
$P\{\sigma > k\}\| = P\{\text{at least } k \text{ jobs served}\}\|$	
LOWER	$P\{\sum_{j=1}^{m} N_j(k)/R_j \le 1\}$
	$\doteq \Phi(1/\mu(k)_j, \sigma^2(k))$ for large s, k

where:

$$R_j = \min_{\{i/J_{ij}>0\}}\{(s_i+1)/J_{ij}\},$$
$$\mu(k) = \sum_{j=1}^{m} \frac{kp_j}{R_j}$$
$$\sigma^2(k) = k\sum_{j=1}^{m}(1/R_j)^2 p_j(1-p_j) - 2k\sum_{i=1}^{n}\sum_{j=i+1}^{m} 1/R_i 1/R_j p_i p_j$$

3 Finding Approximately Optimal Stocking Policies

The goal of our inventory analysis is to provide "good" policies for stocking. In order to do this, we need to evaluate the performance of a large number of policies. This is a difficult task in view of the complexity of equations (5) and (6).

This section uses the heuristic of replacing the actual performance criterion by our bounds on the performance criterion. The bounds thereby provide surrogate optimization problems whose solutions lead to "good" inventory policies. The bounds specified above offer particularly easy optimization. In addition, optimizations using the lower bounds offer guarantees of minimum performance. In our example application, we concentrate on the criterion of the expected time until stockout.

Consider the problem of stocking a mobile repairman who operates in the following way. Starting with an initial kit of parts $(s_1, \ldots, s_n)$, he answers repair calls one at a time until encountering a job which cannot be repaired from the remaining contents of his kit. At this point he returns to the parts depot to obtain the parts for that job as well as to restock the kit. The trip to the parts depot results in lost work time and customer good will. We take as our objective the maximization of the expected number of repair jobs that can be completed before restocking is required. Space is often tightly constrained in such service systems since the repairman's part supply, as well as tools, must fit inside his car. In addition, budgetary constraints are an important consideration. Our objective will be to maximize the expected number of repair jobs completed between restockings, subject to the budget and space constraints. Let the cost of part i for inventory purposes be h_i and the space required by r_i. Then the problem to be solved can be stated as

$$\max_{(s_1,\ldots,s_n)} E(\tau) \tag{15}$$

s.t.

$$\sum_{i=1}^{n} h_i s_i \leq V$$
$$\sum_{i=1}^{n} r_i s_i \leq R.$$

This example, as well as similar practical problems, can be written as

$$\max_{(s_1,\ldots,s_n)} E(\tau) \tag{16}$$

s.t.

$$As \leq b$$
$$s \geq 0,$$

where A is a matrix containing the coefficients of the constraints on the stock levels.

In (15), $E(\tau)$ is calculated from $E(\tau) = E(\sigma)/\lambda$. The previously discussed difficulty in calculating $E(\sigma)$ is compounded in the problem of finding an optimal stock level vector; any approach to optimization by enumeration requires the evaluation of $E(\sigma)$ for a large number of stock level vectors. The possible number of such vectors grows roughly in proportion to $\prod_{i=1}^{n} s_i$, where $(s_1, \ldots, s_n)$ are the maximal feasible stock levels. It is interesting to note that the problem of maximizing the expected time to stockout is difficult even when the jobs consist of distinct groups of parts (i.e., when part demands are independent). This is so because the time to stockout is the minimum time to stockout of each of the parts.

Because of the difficulty encountered in repeatedly evaluating $E(\sigma)$, we propose two heuristic approaches to optimization. Our approaches involve optimizing the upper and lower bounds for $E(\sigma)$. Let $\hat{V}(\hat{s}_1, \ldots, \hat{s}_n)$ denote the value of the problem (16) at an optimal policy. We define

$$\hat{V}^*(\hat{s}_1^*, \ldots, \hat{s}_n^*) = \max_{s_1, \ldots, s_n} U(s_1, \ldots, s_n)(1/\lambda) \tag{17}$$

s.t.

$$\begin{aligned} As &\le b \\ s &\ge 0 \end{aligned}$$

where U is given in (8) (the dependence on $s_1, \ldots, s_n$ is made explicit in (17) for clarity). Rewriting (17) yields

$$\hat{V}^* = \max_{(s_1, \ldots, s_n)} \{\min_i \{ \frac{s_i + \max_{ij}\{J_{ij}\} + 1}{\sum_{j=1}^{m} J_{ij} p_j} \}\}(1/\lambda)$$

s.t.

$$As \ge b,$$

$$s_i \ge 0.$$

This expression for $\hat{V}^*$ can be rewritten by introducing additional inequalities:

$$\max_{(s_1, \ldots, s_n, t)} t/\lambda \tag{18}$$

s.t.

$$\begin{aligned} \sum_{j=1}^{m} J_{ij} p_j t - s_i &\le \max_{ij}\{J_{ij}\} + 1 \\ As &\le b \\ s_i &\ge 0, i = 1, \ldots, n. \end{aligned}$$

If, as seems prudent given the approximate nature of (18), we ignore the integer restrictions on the s_i, (18) is a standard linear program. A moment's reflection reveals that this is exactly the linear program one would obtain by ignoring the random nature of the breakdowns and solving a deterministic stocking problem in which each breakdown required $\sum_{j=1}^{m} J_{ij} p_j$ of part i.

On the other hand, maximizing our lower bound (equation (14)), gives a conservative stocking policy and yields the optimization

$$\hat{V}_* = \max_{(s_1,\ldots,s_n,R_1,\ldots,R_n)} (\sum_{j=1}^{m} \frac{p_j}{R_j})^{-1}(1/\lambda) \tag{19}$$

$$\text{s.t.} \quad As \leq b$$
$$R_j \leq (s_i + 1)/J_{ij} \text{ each } i \text{ such that } J_{ij} > 0, \text{ all } j$$
$$s_i \geq 0, i = 1, \ldots, n$$

which is equivalent to

$$1/\lambda[\min_{(s_1,\ldots,s_n;R_1,\ldots,R_m)} \sum_{j=1}^{m} p_j 1/R_j]^{-1} \tag{20}$$

$$\text{s.t. } R_j \leq (s_i + 1)/J_{ij} \text{ each } i \text{ such that } J_{ij} > 0 \text{ all } j.$$
$$As \leq b$$
$$s_i \geq 0, i = 1, \ldots, n.$$

Problem (20) can be solved as with linear constraints and a separable convex objective.

Intuitively, $\hat{V}^*$ is the result of optimizing the "certainty equivalent" problem in which the random part demands are replaced by deterministic demands equal to the expected demand per job. On the other hand $\hat{V}_*$ represents a weighted max-min criterion; the number of each job that can be served is determined by the part that runs out of stock soonest. The time to stockout is estimated by the weighted average of these stockout times. It seems reasonable that the policy given by the optimization of the lower bound will be good when there is a high level of part commonality between jobs. Indeed, in the extreme case, in which each part is used in every job, $\hat{V}_*$ gives the optimal stocking policy. For the other performance criteria in Table 2.1, conservative stocking policies can be obtained in a similar manner by using the lower bounds to form surrogate objective functions.

4 Examples

We now turn to specific illustrative numerical examples of the applications described in section 4. The nature of the solution, as well as the efficacy of our approaches to it, depend critically on the nature of the job matrix J. Our specific examples are designed to illustrate typical situations which may occur in practice, rather than the "average" performance of our techniques applied to "randomly generated" problems. This approach will give some qualitative insights into the effects of various types of part demand dependencies.

Our numerical examples concentrate on the problem of maximizing the expected time to stockout, subject to a budget constraint. We consider 3 policies, the policies resulting from optimizing the upper and lower bounds and the policy that results from a part-fill type calculation, e.g., in which the dependence between parts is ignored and stock levels are set so that the minimal expected time to stockout is maximized for each of the marginal demand processes (subject to a budget constraint). This is obtained from the optimization

$$
\begin{aligned}
\max_{t, s_1, \ldots, s_n} \quad & t/\lambda \\
\text{s.t.} \quad & t \le \frac{s_i}{\sum_{j=1}^{m} J_{ij} p_j} \quad i = 1, \ldots, n \\
& As \le b \\
& t \ge 0, s_i \ge 0 \quad i = 1, \ldots, n
\end{aligned}
\tag{21}
$$

The problem (21) is similar to the upper bound optimization (18) (the righthand sides are changed by a constant). The solution to problem (21) has the effect of maximizing the minimum expected time to achieve a 0 stock level for each of the marginal demand processes considered separately. As it concentrates only on the marginal part demands we call this the "part-fill" heuristic.

In order to minimize the integer effects caused by part usages which occur in batches, each of the examples the part usage matrix J was chosen to be a 0-1 matrix. For simplicity, we assume that the matrix A consists of a single budget constraint of $\sum_i h_i s_i \le b$ where h_i are the unit holding costs.

Each example has 20 parts and 10 jobs, the job probabilities were assumed equal and part costs for examples are given in table 1. The job matrices are given in Tables 2 – 7. In problem 1 each job consists of two distinct parts. Problem 2 entails a low level of commonality between the jobs and problem 3 a high level of commonality. Problem 4 involves a cascading part usage structure, each job requiring a subset of the parts used in previous (lower numbered) jobs. Problem 5 consists of a mixture of the overlapping and non-overlapping cases. Parts 1 – 6 are common to many jobs; 7 – 20 are used in different jobs. Problem 6 has the same part usage matrix as problem 5, but the part costs are changed so that $h_1 = \ldots = h_6 = 5$ and $h_7 = h_8 = \ldots = h_{20} = 0.01$. In each case we take $b = 500$ and $\lambda = 1$. Clearly, our choices of structures for J do not exhaust all possible commonality types. They are meant to represent in a general way the structure of part commonalities experienced in practice. Such structures can be characterized as "separate" (problem 1), "minimal overlap" (problem 2), "significant overlap" (problem 3), "cascading" (problem 4), and a mix of "overlapped" and "separate" (problem 5 and 6).

Table 7 compares the mean and variance of the time until stockout obtained via Monte Carlo simulation. In each case we took the time between jobs to be exactly 1 [$\lambda = 1$]. The policy obtained using the upper bound as a surrogate objective function was generally dominated by the policy obtained from the "part-fill" optimization. The policy obtained from optimizing the lower bound performed relatively well in those cases in which there was a high level of commonality of parts between the jobs, and relatively poorly when there was little commonality. While increasing the level of part commonality between jobs improved the relative performance of the lower bound heuristic, part commonality alone is not enough to cause the lower bound heuristic to perform much better than the part fill heuristic.

Problem 6 demonstrates the conditions under which the lower bound heuristic performs better than either the upper bound or the part-fill heuristic. In problem 6, those parts common to several jobs (1 – 6) are expensive and those parts used in a single job (7 – 20) are cheap. No job can be performed without at least one part specific to that job alone. By concentrating on the marginal demand rates, the upper bound and part-fill heuristics over invest in those parts common to several jobs and under invest in the unique parts. In contrast, the lower bound heuristic equalizes the number of each type of job to be done and thus equalizes its parts expenditure. For each of the problems tested the lower bound heuristic resulted in nearly identical part stocking levels across parts, and the upper bound and part-fill heuristics resulted in widely varying stock levels. This is due in part to the assumed equal job arrival rates and in part to the "max-min" nature of the lower bound heuristic.

To further test the hypothesis that the quality of the solution obtained from the lower bound improves as the problem becomes more dependent, we generated 2 sets of 10 problems with randomly constructed part usage matrices. The first set had an average of 2 parts per job, the second, 12 parts per job. Table 8 summarizes the results of the experiment. Once again, the solution based on using the lower bound as a surrogate objective fared better when there was a high level of part commonality between jobs. We also see that the variance of the time to stockout was smaller for the lower bound heuristic policy them for the part-fill policy. The reason for this stems from the structure of the objective function in (21). With a high level of part commonality, problem (21) tended to produce policies for which the number of each type of job which could be completed from the kit is similar. For a policy in which the number of each kind of job completable from the kit is exactly the same, the variance in time to stockout is 0. Finally, it should be noted that, as expected, the quality of the lower bound as an estimate of system performance improved dramatically with higher levels of part commonality.

5 Conclusion

A key purpose of this paper has been to call attention to the importance of the dependence among item demands in a multi-item inventory system. The type of dependence considered in this paper arises naturally in many multi-item settings, as discussed in the introduction.

When part or item demands show significant dependence, the potential error arising from ignoring this dependence can be large. While ignoring the dependence between item demands may lead to serious errors in assessing system performance, exact stock level optimization with item demand dependence seems fraught with difficulty. Evaluation of the exact formulas for system performance for a single stocking policy is computationally demanding, even for systems with very modest size. Monte Carlo is effective for evaluation of individual stocking policies, but is not well suited for optimization because the objective functions are not concave or convex.

Our solution to this problem is to develop easily computable convex or concave bounds on system performance. In Section 4, these bounds were used as surrogate objective functions to find policies which maximized the performance of the bound, subject to a budget and other constraints. The simplicity of the bounds gives rise to relatively simple surrogate optimization problems. Monte Carlo is then used to test the actual performance of the stock levels determined by true optimization.

The results of our limited experience with the bounds are encouraging. The part-fill heuristic gave reasonable performance when there was a low level of part demand dependence and the lower bound heuristic gives very good performance for high levels of item demand dependence. In each case, the bounds themselves may give rather large over- or underestimates of system performance, but the policies resulting from the optimization of the bound performed well in the Monte Carlo experiment.

Clearly, many interesting and unsolved problems remain in analyzing stocking policies for interdependent multi-item inventory systems. Given the increased interest in multi-item inventories in production, as well as in repair kit inventory problems, we hope that the results obtained in this paper will stimulate further progress in this research area.

References

Baker, K.R., M.J. Magazine and H.L. Nuttle, "The Effect of Commonality on Safety Stock in a Simple Inventory System," *Management Science*, **32**, No. 8 (August 1986), 982 – 988.

Baker, K.R., "Safety Stocks and Component Commonality," *Journal of Operations Management*, **6**, No. 1 (November 1985), 13 – 21.

Graves, S. C., "A Multiple-item Inventory Model with a Job Completion Criterion," *Management Science*, **28**, 11 (November 1982), 1334 – 1337.

Hausman, W. H., "On Optimal Repair Kits Under a Job Completion Criterion," *Management Science*, **28**, 11 (November 1982), 1350 – 1351.

Mamer, J. W. and S. A. Smith, "Optimal Field Repair Kits Based on Job Completion Rate," *Management Science*, **28**, 11 (1982), 1325 – 1333.

Mamer, J. W. and S. A. Smith, "Job Completion Based Inventory Systems: Optimal Policies for Repair Kits and Space Machines," *Management Science*, **31**, 6, (June 1985), 703 – 718.

Mamer, J. and S. A. Smith, "A Lower Bound Heuristic for Selecting Multi-Item Inventories for a General Job Completion Objective," Manuscript, Santa Clara University, 1988.

March, S. T. and G. D. Scudder, "On Optimizing Field Repair Kits Based on Job Completion Rate", it Management Science, 30, No. 8, (August 1984), 1025 – 1028.

Schaefer, M. K., "A Multiitem Maintenance Center Model for Low Demand Repairable Items," *Management Science*, **29**, No. 7, (Sept. 1983), 1062 – 1068.

Sherbrooke, C. C., "An Evaluator for the Number of Operationally Ready Aircraft in a Multi Level System," *Operations Research*, **19**, 3, (May-June 1971), 618 – 635.

Silver, E. A., "Inventory Allocation Among an Assembly and Its Repairable Sub-Assemblies," *Naval Research Logistics Quarterly*, **19**, 2 (June 1972), 261 – 280.

Smith, S. A., J. Chambers, and E. Shlifer, "Optimal Inventories Based on Job Completion Rate for Repairs Requiring Multiple Items," *Management Science*, **26**, 8 (August 1980), 849 – 852.

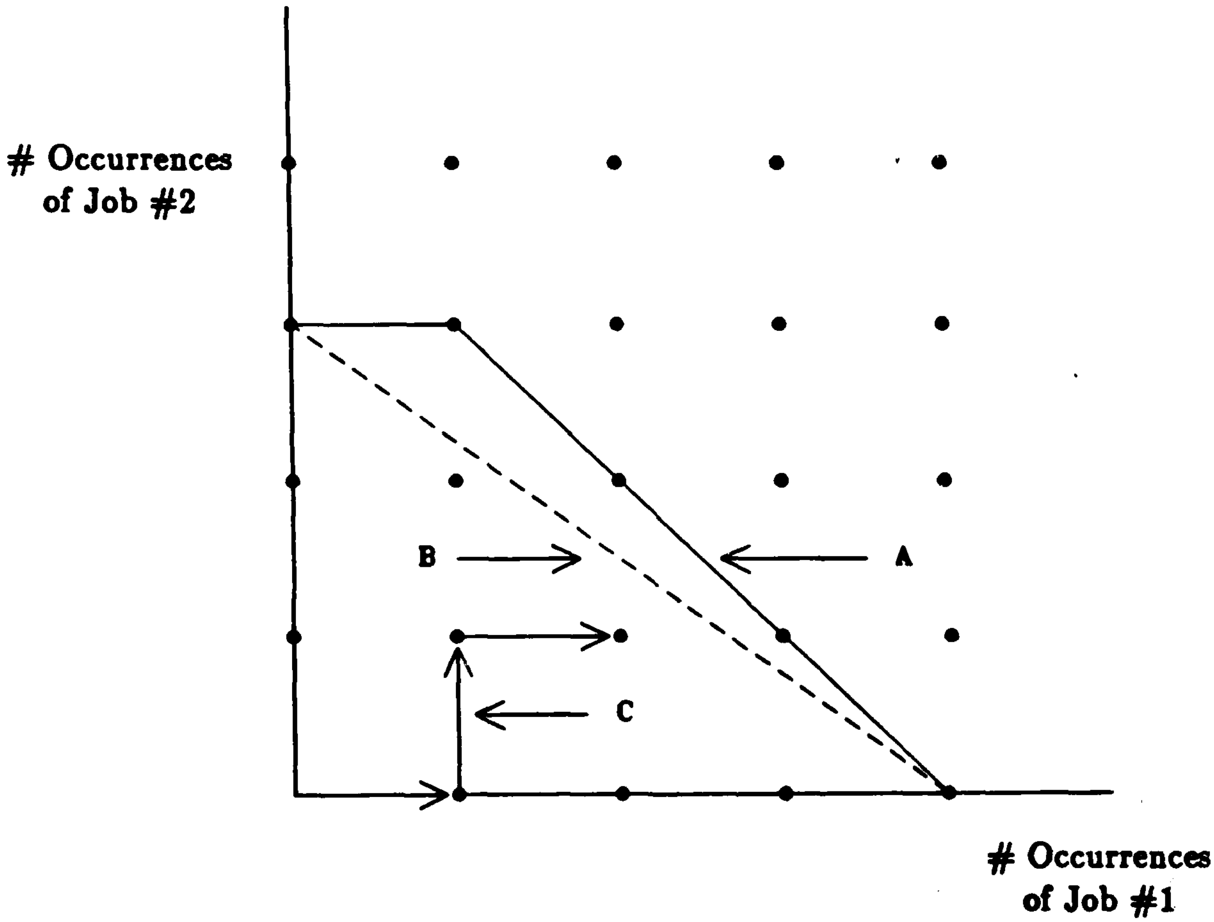

FIGURE 1

A: Boundary of $\{x | Jx \leq s + 1\}$.

B: $\{x | \sum_{i=1}^{m} \frac{x_i}{R_i} = 1\}$.

C: Path of $N(k)$.

Table 1.

Part No.	Cost (h_i)
1	1.5
2	3.8
3	9.8
4	8.2
5	1.6
6	2.4
7	4.5
8	2.3
9	5.3
10	1.6
11	8.6
12	1.9
13	2.8
14	7.7
15	1.4
16	7.2
17	5.5
18	9.8
19	8.4
20	5.1

Table 2.
Problem 1.

parts	jobs 1	2	3	4	5	6	7	8	9	10
1	1	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	0	0	0
4	0	1	0	0	0	0	0	0	0	0
5	0	0	1	0	0	0	0	0	0	0
6	0	0	0	1	0	0	0	0	0	0
7	0	0	1	0	0	0	0	0	0	0
8	0	0	0	1	0	0	0	0	0	0
9	0	0	0	0	1	0	0	0	0	0
10	0	0	0	0	0	1	0	0	0	0
11	0	0	0	0	1	0	0	0	0	0
12	0	0	0	0	0	1	0	0	0	0
13	0	0	0	0	0	0	1	0	0	0
14	0	0	0	0	0	0	0	1	0	0
15	0	0	0	0	0	0	1	0	0	0
16	0	0	0	0	0	0	0	1	0	0
17	0	0	0	0	0	0	0	0	1	0
18	0	0	0	0	0	0	0	0	0	1
19	0	0	0	0	0	0	0	0	1	0
20	0	0	0	0	0	0	0	0	0	1

Table 3.
Problem 2.

parts	jobs 1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1
4	0	1	0	0	0	0	0	0	0	0
5	0	0	1	0	0	0	0	0	0	0
6	0	0	0	1	0	0	0	0	0	0
7	0	0	1	0	0	0	0	0	0	0
8	0	0	0	1	0	0	0	0	0	0
9	0	0	0	0	1	0	0	0	0	0
10	0	0	0	0	0	1	0	0	0	0
11	0	0	0	0	1	0	0	0	0	0
12	0	0	0	0	0	1	0	0	0	0
13	0	0	0	0	0	0	1	0	0	0
14	0	0	0	0	0	0	0	1	0	0
15	0	0	0	0	0	0	1	0	0	0
16	0	0	0	0	0	0	0	1	0	0
17	0	0	0	0	0	0	0	0	1	0
18	0	0	0	0	0	0	0	0	0	1
19	0	0	0	0	0	0	0	0	1	0
20	0	0	0	0	0	0	0	0	0	1

Table 4.
Problem 3.

parts	jobs 1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1	1	1
10	0	0	0	0	0	1	0	0	0	0
11	0	0	0	0	1	0	0	0	0	0
12	0	0	0	0	0	1	0	0	0	0
13	0	0	0	0	0	0	1	0	0	0
14	0	0	0	0	0	0	0	1	0	0
15	0	0	0	0	0	0	1	0	0	0
16	0	0	0	0	0	0	0	1	0	0
17	0	0	0	0	0	0	0	0	1	0
18	0	0	0	0	0	0	0	0	0	1
19	0	0	0	0	0	0	0	0	1	0
20	0	0	0	0	0	0	0	0	0	1

Table 5.
Problem 4.

parts	jobs 1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	0
4	1	1	1	1	1	1	1	1	0	0
5	1	1	1	1	1	1	1	0	0	0
6	1	1	1	1	1	1	0	0	0	0
7	1	1	1	1	1	0	0	0	0	0
8	1	1	1	1	0	0	0	0	0	0
9	1	1	0	0	0	0	0	0	0	0
10	1	1	1	0	0	0	0	0	0	0
11	1	1	0	0	0	0	0	0	0	0
12	1	1	1	1	1	1	1	0	0	0
13	1	1	0	0	0	0	0	0	0	0
14	1	1	1	1	1	0	0	0	0	0
15	1	1	0	0	0	0	0	0	0	0
16	1	0	0	0	0	0	0	0	0	0
17	1	1	1	1	0	0	0	0	0	0
18	1	0	0	0	0	0	0	0	0	0
19	1	1	0	0	0	0	0	0	0	0
20	1	0	0	0	0	0	0	0	0	0

Table 6
Problems 5 & 6.

parts	jobs 1	2	3	4	5	6	7	8	9	10
1	1	1	1	0	1	0	1	0	1	1
2	0	1	0	1	0	1	0	1	0	1
3	1	0	1	0	1	0	1	0	1	0
4	0	1	0	1	0	1	0	1	0	1
5	1	0	1	0	1	0	1	0	1	1
6	0	1	0	1	0	1	0	1	0	0
7	0	0	0	0	0	0	1	0	0	0
8	0	0	0	0	0	0	0	1	0	0
9	0	0	0	0	0	0	0	0	1	0
10	0	0	0	0	0	0	0	0	0	1
11	0	0	0	0	1	0	0	0	0	0
12	0	0	0	1	0	0	0	0	0	0
13	1	0	0	0	0	0	0	0	0	0
14	0	1	0	0	0	0	0	0	0	0
15	0	0	0	0	1	0	0	0	0	0
16	0	0	1	0	0	0	0	0	0	0
17	0	0	0	1	0	0	0	0	0	0
18	0	0	0	0	0	1	0	0	0	0
19	0	0	0	0	0	0	1	0	0	0
20	0	0	0	0	0	0	0	0	0	1

Table 7.
Montecarlo Results Sample Mean and Variance
of Time to Stockout

Example	Lower Bound Mean	Heuristic Variance	Upper Bound Mean	Heuristic Variance	Fill-Rate Mean	Heuristic Variance
1	27.26	55.12	28.61	42.99	28.36	37.90
2	7.00	0.001	4.95	3.78	10.13	11.66
3	7.00	0.00	1.68	1.07	6.31	6.78
4	7.94	0.10	6.69	18.64	9.70	8.72
5	7.90	1.44	4.71	2.94	9.73	9.04
6	25.14	8.06	4.63	3.03	15.49	16.84

Table 8.
Montecarlo Results Sample Mean and Variance
of Time to Stockout - Randomly Generated Problems.

	Lower Bound Mean	Heuristic Variance	Fill-Rate Mean	Heuristic Variance
2 parts per job (average).	24.08	22.94	32.19	44.78
12 parts per job (average).	6.28	0.21	6.44	1.11