How Much Cash Should A Bank Maintain? – An Optimization Approach

by

Meng-Huai Chen

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Master of Science in Operations Research

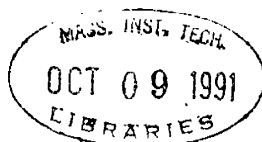at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1991

Author ....
Sloan School of Management
August 19, 1991

Certified by ....
Dimitris Bertsimas
Assistant Professor
Thesis Supervisor

Accepted by ....
Thomas L. Magnanti
Co-Director, Operations Research Center

How Much Cash Should A Bank Maintain? — An Optimization Approach

by

Meng-Huai Chen

Abstract

We study two aspects of cash management for a bank. First, we formulate a constrained dynamic programming problem that takes into account a fixed lead time for the delivery of cash, a known holding and ordering cost and a limit on the extent of shortage that is tolerable. We implement an algorithm for the optimal ordering policy and present numerical results. Secondly, we study a non-stationary inventory model for developing the optimal stocking policy for an automatic teller machine. Demand is assumed to follow a cyclic pattern during the week. An efficient algorithm is implemented for exponentially distributed demand taking into account no delivery lead times or set-up costs. Finally, we propose two heuristics for the generalization of this model to include set-up costs and present numerical results supporting their validity and their relative effectiveness.

Acknowledgments

I would like express my gratitude to my advisor, Prof. Dimitris Bertsimas, whose support, advice and constant encouragement made this thesis possible.

Contents

Chapter 1

Introduction

Large banks have a large network of local branches. These branches perform a number of activities involving cash. As a result, they need to maintain large amounts of cash. In each branch, the main characteristic of the daily demand for cash is that it is stochastic and time varying.

Maintaining excessive amount of cash is uneconomical simply because one could invest it. On the other hand, maintaining small amounts of cash can be disasterous in the event that a certain branch runs out of cash. As a result, it is intuitively clear that there is an optimal amount of cash that a branch should maintain. On a daily basis, given the stochastic and dynamic variability of the demand for cash, the optimal level will vary from day to day or even within the same day. Hence, the optimal level will be time varying and one could continuously control the level of cash and adaptively decide when to increase the level of cash maintained.

The problem this thesis addresses is to decide how much cash to order at regular time intervals so as to meet the depositor's demand for cash every day. This is similar to most of the mathematical inventory models, in which the two fundamental issues are: when should a replenishment order be placed, and how much should the order quantity be. The complexity of the model depends on the assumptions that one makes about the demand, the cost structure and the physical characteristics of the system.

In Chapter 2, we present and solve the problem of optimal cash allocation of a

branch of the bank. This problem can be viewed in terms of inventory theory as an infinite time horizon problem with independent and identically distributed demands each week, known holding and ordering costs, fixed non zero delivery time lags and a constraint on the expected level (or probability) of shortage that may arise if demand exceeds available supply. We review the related literature and consider how existing models may be modified so as to be applied to the current problem. We present an efficient algorithm for computing a close approximation to the optimal policy and present the computational results from the implementation of the algorithm.

In Chapter 3, we analyze a special demand pattern. Specifically, we study the impact of non-stationary demand patterns on the optimal stocking policy of automatic teller machines. Clearly, during the ordering cycles (days) the demands for ATMs are certainly not identically distributed. These demands may be assumed to be independently distributed. The cost structure analyzed in this chapter is similar to that studied in the earlier chapter. We assume, however, that there is no time required for the delivery of cash when ordered. We review the related literature that treats such problems. We provide an efficient algorithm for computing the optimal policy for the ATM problem. We implemented the algorithm and present computational results for computing the optimal stocking policy.

In Chapter 4, we study the following generalization of the model presented in Chapter 3: we assume that there is a set-up cost that is incurred every time an order is placed. This assumption would be necessary when designing the optimal stocking policy for an automatic teller machine that is remotely located and would therefore involve an additional fixed expense every time the machine has to be restocked. We propose two heuristic approaches and present computational results for their performance.

7

Chapter 2

The I. I. D. Demand Model

2.1 The Problem

The problem we discuss here is as follows: how to optimally order quantities of cash at the fixed time intervals, say, each Friday of the week, to meet the weekly stochastic cash demand, so as to minimize the total cost of this particular operation of a branch of the bank.

The assumptions we make are:

(1) The demands which are not satisfied during each period (i.e., week) are lost. So the on hand inventory at the end of each period is always non-negative (as opposed to the case of backlogging of excess demand).

(2) There is a time lag τ for cash delivery from the central bank to an individual branch. For example, if the branch orders the cash this Friday, it will arrive on Wednesday the week after (i.e., in 12 days). We have considered $\tau = 0, 1, 2$.

(3) The model can be over a finite or an infinite time horizon. We have focused primarily on the infinite time horizon, since we assume that the bank is operating under stable conditions.

For tractability purposes, our model does not include set up costs. Although there are cases in which this is an important limitation, one could assume that the costs incurred by the bank in distributing cash to its various branches is largely unaffected by the specific ordering patterns of an individual branch. This would suggest that the

8

fixed set up cost may be negligible, but this is an area that needs further exploration.

We use a dynamic programming model, where the stage is the period, the state is the cash level on hand at the beginning of the period, the control variables are the amount of cash that needs to be ordered at the beginning of every period. We assume the holding cost of excess cash in the inventory is linear. There is no set up cost, but there is linear ordering cost.

2.2   The Model Formulation

We assume that orders for cash that are placed at the beginning of period t will arrive at the beginning of period t+2. This deviates somewhat from the real situation described above. A survey of available inventory models reveals that, in general, it is assumed that orders are placed and deliveries are received at the beginning (or end of) fixed periods. We use the following notation:

y_t: inventory on hand at the end of period t.

x_t: amount ordered at the beginning of period $t-2$, this amount will be delivered at the beginning of period t.

d_t: demand in period t, which is a bounded random variable such that $0 \leq d_t \leq D$ and has a known p.d.f of $f_d(d)$ for all t.

α: safety factor which is assumed to be given; this is an upper bound on the expected level of shortage during any period.

h: cost per dollar held in the inventory. This represents the interest that can be attained if the cash in the inventory were invested.

c: cost per dollar ordered.

β: discount factor used in computing present values of future costs incurred.

n: number of periods in the time horizon.

The cost incurred during each period t consists of two components: (1) ordering cost cx_t and (2) the holding cost for excess inventory ny_t. So the total expected cost over n periods is: $E(\sum_{t=0}^{n-1} \beta^t (hy_t + cx_t))$; our objective is to minimize this cost by proper choice of the orders $x_0, x_1, ...x_{n-1}$ subject to the natural constraint $x_t \geq 0$,

9

$t = 0, 1, 2, ..., n - 1$ and the constraint that the expected cash shortage in period t should be less than a given safety factor α. The reason we include this as a constraint is due to the difficulty of assigning penalty cost to the shortage of cash. Finally, y_{-1}, the opening inventory at the beginning of period 0, is assumed to be given. The model can be formulated as follows:

Problem P

$$Min \quad Z = E(\textstyle\sum_{i=0}^{n-1} \beta^i(hy_i + cx_i)) \tag{2.1}$$

$$Subject\ to \quad y_i = \max(0, y_{i-1} + x_i - d_i) \quad for\ i = 0, ...n - 1 \tag{2.2}$$

$$E(\max(0, d_i - y_{i-1} - x_i)) \le \alpha \quad for\ i = 0, ...n - 1 \tag{2.3}$$

$$x_i \ge 0 \qquad\qquad for \quad i = 0, ...n - 1 \tag{2.4}$$

There are several things we should address in the formulation. First, we have used an approximation for computing the expected holding cost. We use the on hand inventory at the end of the period instead of using precisely the average on hand inventory level during the period. Secondly, we assume that the ordering cost will be incurred when the amount is ordered. Thirdly, in contrast with most inventory models, instead of having a service level constraint, we have a constraint which requires that the expected inventory shortage be less than α. The major question here is what will be the proper choice of α. An intuitive relation between the value of α and the service level can be stated as follows: the smaller the value of α is, the better the service should be. We have to admit that this constraint is weaker than constraints on the probability of shortage. Clearly, the constraint in the above form is considerably harder to conceptualize as opposed to the constraint on the probability of a shortage. But formulating it in this fashion proves to be advantageous in later analysis. Morever, in Section 2.7 we will include results from the case in which we impose constraints on the probability of shortage. In the following section, we introduce the idea of the penalty method for solving the problem **P** .

2.3 The Relaxation Problem

Problem **P** is a stochastic and dynamic problem with nonlinear constraints (2.3). Our approach is to relax constraints (2.3) multiplying each constraint by $\lambda_i = \lambda\beta^i$. We thus obtain:

The Problem P_λ :

$$Min \quad Z_\lambda = \sum_{i=0}^{n-1} \beta^i \{E(h \times max(0, y_{i-1} + x_i - d_i) + cx_i)$$

$$+\lambda E(max(0, d_i - x_i - y_{i-1}) - \alpha)\} \tag{2.5}$$

$$Subject \ to \quad y_i = max(0, y_{i-1} + x_i - d_i) \quad for \ i = 0, ...n - 1 \tag{2.6}$$

$$x_i \geq 0, \quad i = 0, ..n - 1 \tag{2.7}$$

The parameter λ has a natural interpretation as a penalty of shortage of cash. As we will see in the later section, λ can be seen as a parameter to control the tradeoff among the cost of holding, ordering and the shortage. The constraints (2.2) in the original problem **P** are not convex over the variables x_i. As a result, the solution to the relaxation problem P_λ will only provide a lower bound to the solution of the original problem, i.e. $Z_\lambda \leq Z$, for all $\lambda \geq 0$.

2.4 The DP Algorithm to solve P_λ :

Let s_k denote the total available cash at the beginning of period k. Then $s_k = y_{k-1} + x_k$. Let $L(s_k)$ denote the one period expected holding and shortage cost. Then

$$L(s_k) = E\{h \times max(0, s_k - d_k) + \lambda max(0, d_k - s_k)\}$$

We define:

(s_k, x_{k+1}): the state variable at the beginning of period k.

x_{k+2}: the decision variable at the beginning of period k.

k: the stage of the system.

$J_k(s_k, x_{k+1})$: the minimum total discounted cost incurred in period k through $n - 1$ given that (s_k, x_{k+1}) is the state variable at the beginning of the period k.

11

Observe that x_n and x_{n-1} is 0 in the case of the finite time horizon problem with n periods and 2 periods time lag. In other words, there is no decision to be made at the beginning of the last two periods.

The **DP** equation is:

$$J_k(s_k, x_{k+1}) = min_{x_{k+2} \geq 0}\{L(s_k) + cx_{k+2} + \beta E(J_{k+1}(s_{k+1}, x_{k+2}))\} \qquad (2.8)$$

Notice that whenever demand d_k exceeds s_k in period k, x_{k+1} represents the total cash available at the start of the following period, otherwise, $s_k + x_{k+1} - d_k$ is the available stock size for the next period. Then:

$$s_{k+1} = \begin{cases} x_{k+1} & \text{if } s_k \leq d_k \\ s_k + x_{k+1} - d_k & \text{if } s_k > d_k \end{cases}$$

Hence, the dynamic recursive equation for $J_k(s_k, x_{k+1})$ satisfied is:

$$J_k(s_k, x_{k+1}) = min_{x_{k+2} \geq 0}\{cx_{k+2} + L(s_k) + \beta J_{k+1}(x_{k+1}, x_{k+2}) \int_{s_k}^{\infty} f(\xi)d\xi$$

$$+ \beta \int_0^{s_k} J_{k+1}(s_k + x_{k+1} - \xi, x_{k+2})f(\xi)d\xi\} \quad k = 0, 1, ..., n-2 \qquad (2.9)$$

$$J_{n-1}(s_{n-1}, x_n) = L(s_{n-1}). \qquad (2.10)$$

2.5 The Existence and The Structure of The Optimal Solution of P_λ

Now we fix λ and solve problem P_λ. In Karlin[3] the existence and the structure of the optimal solution of (2.9) for the one period time lag has been well studied. In this section, we will state that result and Morton's[4] extension.

The dynamic recursive equationwhen the time lag is one period is given by:

$$J_k(s_k) = min_{x_{k+1} \geq 0}\{cx_{k+1} + L(s_k) + \beta J_{k+1}(x_{k+1}) \int_{s_k}^{\infty} f(\xi)d\xi$$

$$+ \beta \int_0^{s_k} J_{k+1}(s_k + x_{k+1} - \xi)f(\xi)d\xi\} \quad k = 0, 1, ..., n-2 \qquad (2.11)$$

$$J_{n-1}(s_{n-1}) = L(s_{n-1}). \qquad (2.12)$$

12

To perform the minimization operation, we differentiate (2.11) with respect to x_{k+1}. This gives:

$$c + \beta J'_{k+1}(x_{k+1}) \int_{s_k}^{\infty} f(\xi)d\xi + \beta \int_0^{s_k} J'_{k+1}(s_k + x_{k+1} - \xi)f(\xi)d\xi = G_k(s_k, x_{k+1})$$

Let $x^*_{k+1}(s_k)$ denote the solution of $G_k(s_k, x_{k+1}) = 0$.

Theorem 2.1 *(Karlin[3]) If there exist s^*_k such that $G_k(s^*_k, x^*_{k+1}) = 0$, then the optimal policy $x^*_{k+1}(s_k)$ has the property that $x^*_{k+1}(s_k)$ is continuous in s_k and of the form:*

$$x^*_{k+1}(s_k) = \begin{cases} > 0 & \text{if } s_k < s^*_k \\ = 0 & \text{if } s_k \geq s^*_k \end{cases}$$

Moreover, $x^*_{k+1}(s_k)$ is strictly decreasing for $s_k < s^*_k$ and $|\frac{\partial x^*_{k+1}(s_k)}{\partial s_k}| < 1$.

The previous theorem provides a characterization of the optimal policy for the case in which the time lag is $\tau = 1$ and there is a finite time horizon n. Morton[4] has proven also the analog of theorem 2.1 for $\tau = 2$. Unfortunately, these theorems are not useful computationally.

In order to propose useful implementable policies, we now consider infinite time horizon problems and use the approach of Morton[4] in order to provide bounds for the optimal policy. In the infinite horizon case, Morton[4] proves the following theorem:

Theorem 2.2 *Let $F^{(i)}(x) = \int_0^x F^{(i-1)}(x - \xi)f(\xi)d\xi$, $i = 1, 2, 3$, with $F^{(0)}(x) = 1$, i.e. $F^{(i)}(x)$ is the cdf of the demand after i days. Let $P = \frac{\beta^2 \lambda - c}{\beta^2(\lambda + h) - \beta c}$. Let $\xi^i = (F^{(i)})^{-1}(P)$, $i = 1, 2, 3$. For $\tau = 2$, if the state of the sytem is (s_k, x_{k+1}), then the optimal order amount x^*_{k+2} satisfies*

$$0 \leq x^*_{k+2} \leq max\{0, min(\xi^1, \xi^2 - x_{k+1}, \xi^3 - x_{k+2} - s_k)\}$$

This theorem provides an adaptive algoruthm to calcualte approximately the optimal policy (for fixed λ), where we have already order x_{k+1} and the on hand cash is s_k.

Algorithm A_2 ($\tau = 2$, infinite time horizon)

1. Calculate ξ^1, ξ^2, ξ^3.

13

2. Find $\xi^* = max\{0, min(\xi^1, \xi^2 - x_{k+1}, \xi^3 - x_{k+2} - s_k)\}$.

3. Order $x_{k+2}^* = \xi^*$.

For $\tau = 1$, the algorithm becomes:

Algorithm A_1 ($\tau = 1$, infinite time horizon)

1. Calculate ξ^1, ξ^2.

2. Find $\xi^* = max\{0, min(\xi^1, \xi^2 - s_k)\}$.

3. Order $x_{k+1}^* = \xi^*$.

For $\tau = 0$, the algorithm becomes:

Algorithm A_0 ($\tau = 0$, infinite time horizon)

1. Calculate ξ^1.

2. Find $\xi^* = max\{0, \xi^1\}$.

3. Order $s_k^* = \xi^*$.

The above algorithms, for $\tau = 0, 1, 2$ do not find the optimal cash order exactly. They rather overestimated the optimal order, trying to be more conservative. This means that the probability of stockout under this approximate policy will be lower than under the optimal policy.

We have also explored the possibility of improving the previous policy. Since under the proposed policy, the optimal ordering amount is overestimated, we experimented with decreasing the ordering amount. In particular we used

$$x_{k+2}^* = F \times max\{0, min(\xi^1, \xi^2 - x_{k+1}, \xi^3 - x_{k+2} - s_k)\} \qquad (2.13)$$

with $0 \leq F \leq 1$. For $\tau = 0$, algorithm A_0 is _exact_, since in this case, the problem becomes the newsboy's problem.

Another important issue is that we do not know the penalty λ of stockout. In the following section, we used simulation and a search technique to provide a tradeoff curve between cost and probability of stckout.

14

2.6 Computational Results

The underlying assumption in the previous section is that the penalty cost λ is a known parameter. Unfortunatly, the penalty cost associated with the shortage of funds is not easily computable. Instead, we have used a constraint on the expected shortage of funds in each period in the formulation of the problem. By using the relaxation approach, we have transformed the constrains of the allowable shortage to a penalty λ. We first notice that the larger we set λ, the lower the expected shortage will be; however, extremely large values of λ would result in an excessively conservative policy. The basic idea is to find the smallest value of the penalty cost λ that ensures that the constraint on expected shortage cost is satisfied. In order to do that, we establish a range within which we expect the best value of λ to lie.

We achieve this as follows: Lower bounds on λ are established by using the fact that both the numerator and denominator of the probability $\frac{\beta^\tau \lambda - c}{\beta^\tau (h+\lambda) - \beta c}$ for $\tau = 0, 1, 2$ should be greater than 0. Notice that when $\tau = 0$, the above probability becomes $\frac{\lambda - c}{(h+\lambda) - \beta c}$, which is the probability of no stockout in the newsboy's problem. Hence, the upper bound for λ is computed by choosing a value such that $\frac{\beta^\tau \lambda - c}{\beta^\tau (h+\lambda) - \beta c} > 1 - \alpha$ for $\tau = 0, 1, 2$

Having established a range for λ, we narrow the range iteratively by computing the optimal policy for the chosen value of λ, simulating the system for a fixed but long period of time, and checking to see if the constraint on expected shortage is satisfied. Several replications of each simulation are performed to get an accurate estimate of the expected shortage per period. Eliminating ranges of values to be searched, the range for the optimal λ can be narrowed to a desirable of accuracy.

Our computational results were obtained under the assumption that the demand during each week was gamma distributed with a mean of 500 and a standard deviation of 33 (all figures are expressed in thousands of dollars). The gamma distribution is clearly more acceptable than the normal for such an application, since it is a non-negative random variable.

We assume that the annual internal rate of return is 15%. This is converted to a

weekly interest rate by using $(1 + i_{weekly})^{52} = (1 + i_{annual})$. In order to convert to the infinite stream of costs into a present cost, the appropriate discount factor is $\frac{1}{1+i_{weekly}}$. The holding cost is the same as the weekly interest rate since the commodity for which we are computing the holding cost is itself cash. The ordering cost was assumed to be linear with a rate of $0.001 per thousand dollars.

We repeat the above procedure for different values of lead time τ. Specifically, we compare the results of no lead time ($\tau = 0$) with those of $\tau = 1$ and $\tau = 2$ periods.

The results of our numerical computations are presented in the form of six graphs.

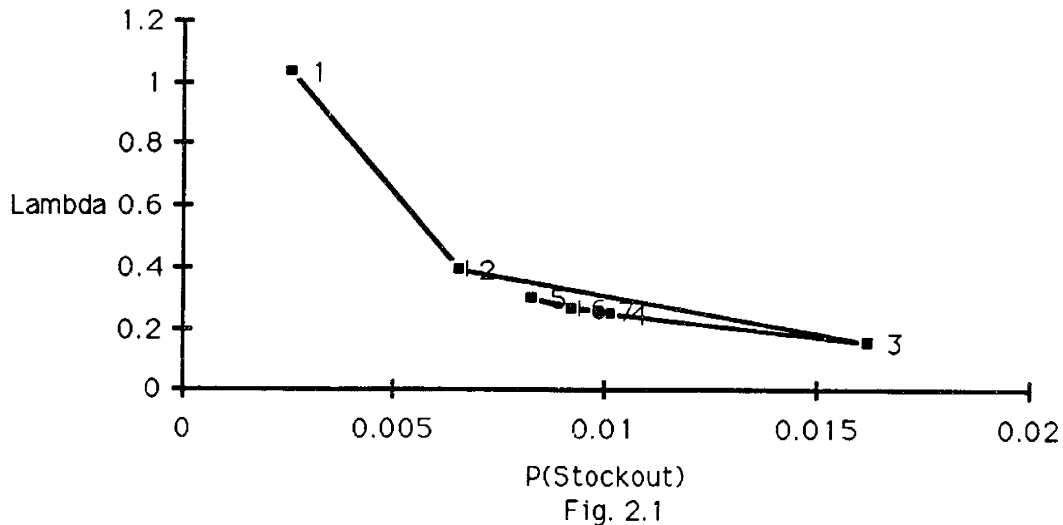## Path of Lambda during the search v/s P(Stockout) : Tau=1

Fig. 2.1

Figure 2.1 indicates the search path of λ to solve a typical problem. At the conclusion of the simulation for a particular value of λ, the estimate of the probability of stock out (the fraction of periods in which a stock out is observed in the course of the simulation) is compared with the prespecified limit of 0.01. If the constraint is violated we increase λ, while if it is satisfied we decrease λ. In this particular case, this procedure takes seven steps before the range of search is narrowed beyond a prespecified precision. Notice that the check for constraint satisfaction involves a certain amount of error, since we can only compare an *estimate* of the probability of stockout or the expected stockout. This often results in a termination of the

16

procedure at a point where the constraint is slightly violated or at a point which appears to correspond to a more conservative policy than strictly required by the constraint.

Total Cost v/s E(Stockout)

Fig 2.2
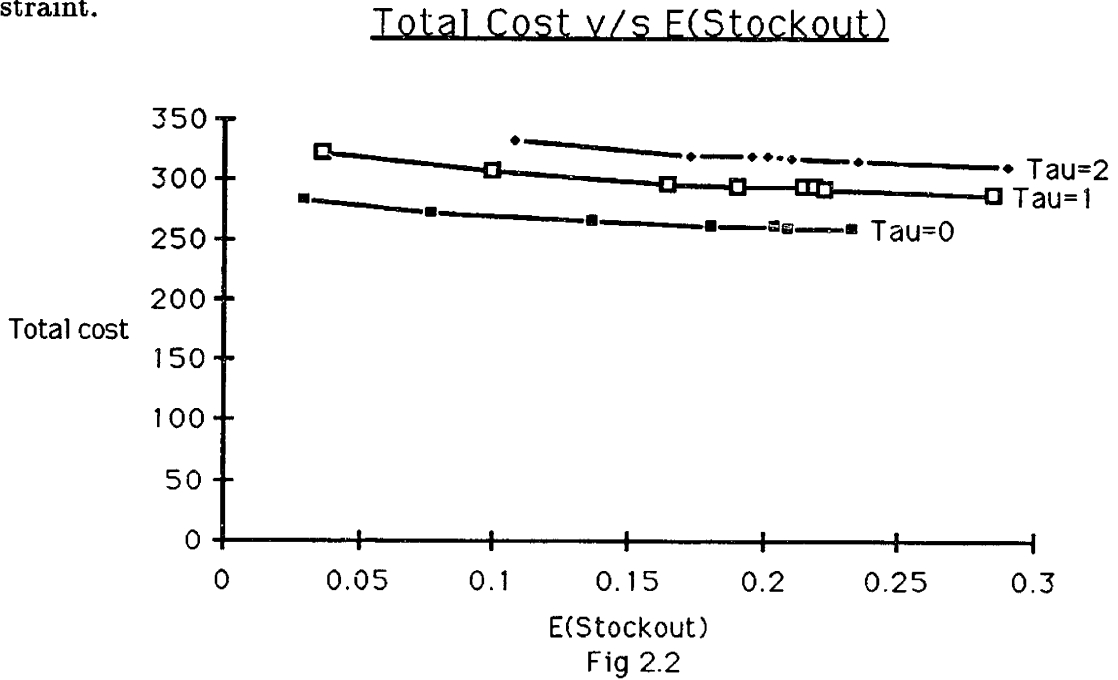
We next study in Figure 2.2 the tradeoff between the total discounted cost (ordering+holding cost) and the expected amount of stockout in a period. The two main observations that one can make follow our intuition; first, for a given delivery lead time, the total cost decreases as we permit larger values of expected stock out; secondly, for a given level of permissible stock out, there is a significant increase in the total cost for longer delivery times. The first observation is intuitive, since a higher level of permissible stockout would permit less conservative ordering policies. The second observation is also intuitive, since a longer lead time is equivalent to planning for a later time in the future. This would imply that the decisions are subject to a greater number of unknown factors, namely, the demand in the delivery period. This would require a more conservative policy for longer delivery lead times in order to stay within the same bounds of expected stock out.

The Figure 2.3 results from performing the procedure outlined above except that the choice of λ was governed by constraints on the probability of stockout instead of expected stockout. We continue to apply the rule for optimal ordering policy from
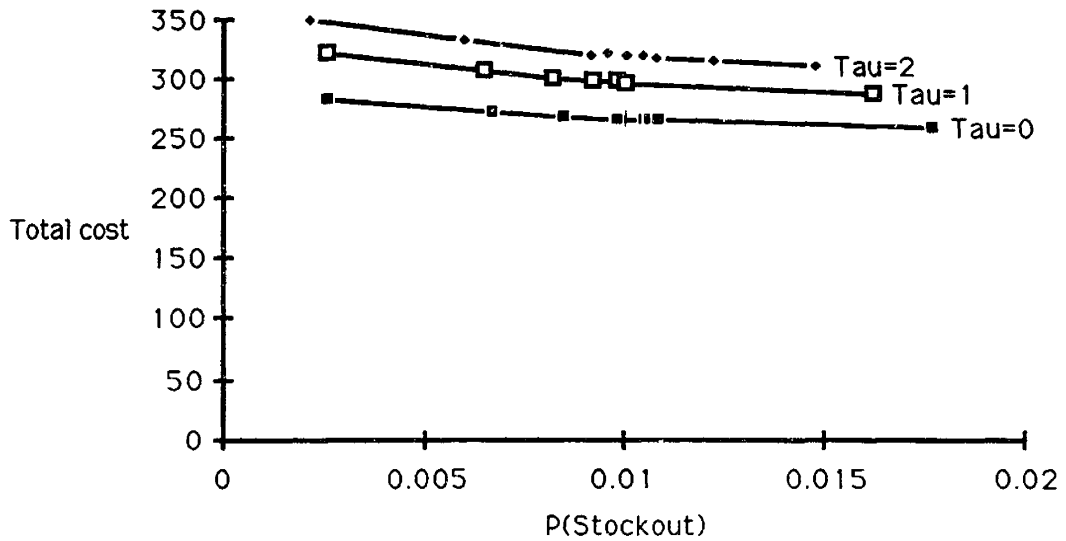
Total Cost v/s P(Stockout)

Fig. 2.3

section 2.5 for different values of the penalty cost; at each step we check to see if the estimate of the probability of stock out is within the permissible limit of 0.01. As expected, the results are very similar to those obtained in the case of the constraint on expected stock out.
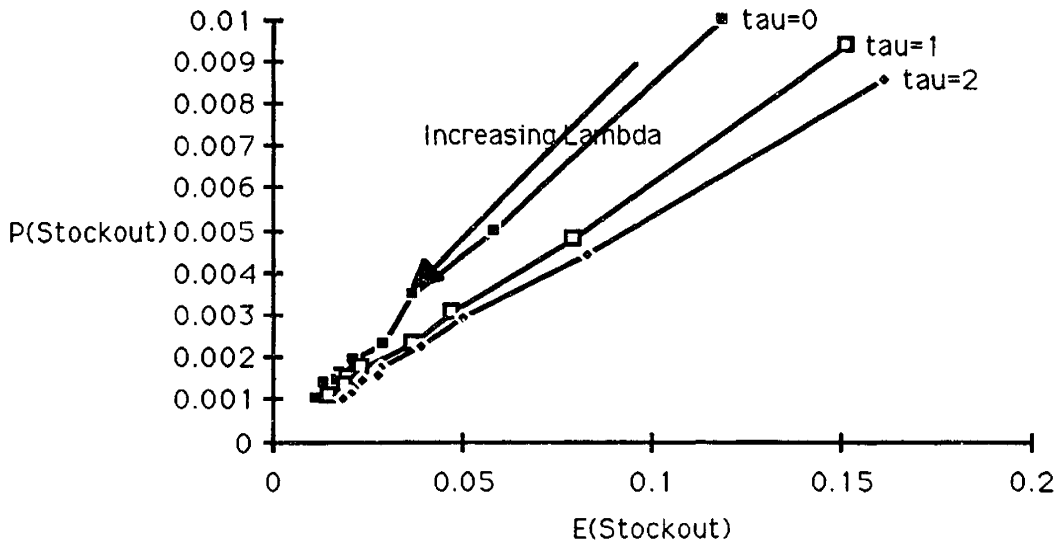
P(Stockout) v/s E(Stockout)

Fig. 2.4

We next studied in Figure 2.4 the correlation of of the probability of stockout and the expected stockout under the proposed policy. It can be seen that increasing the

18

penalty cost produces decreasing effects on both the probability and the expectation of stock out. It also reveals a highly linear relationship between the probability and the expectation. This would imply that at least under our assumptions one can translate a probability constraint into a constraint on the expectation without introducing any substantial error. Another interesting observation is that, for a given expected amount of stockout, the probability of stockout is highest for the case with the shortest lead time. This leads one to conclude that there is a larger number of smaller shortages when the lead time is shorter; this again indicates that under shorter lead time conditions, one has greater control over the operations including shortages.
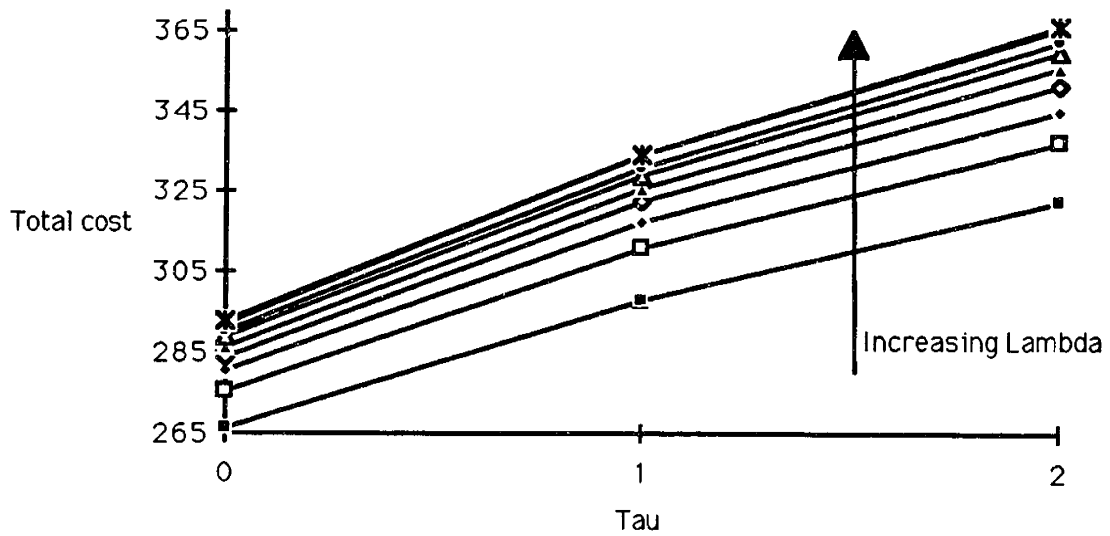

Fig. 2.5

We investigate in Figure 2.5 the dependence of the total cost on different lead times for various values of the penalty cost. We conclude that higher penalty costs lead to diminishing changes in the total cost. A similar observation can be made from Figure 2.6 where we investigate the dependence of the expected amount of stockout on different lead times.

Finally, we investigated whether we can further improve the proposed policy by using a factor $F < 1$, (see equation (2.13)). In Figs. 2.7 and 2.8 we see that the total cost function does decrease very slightly when the reduction factor is reducted to 0.99

19

E(Stockout) v/s Tau (varying lambda)
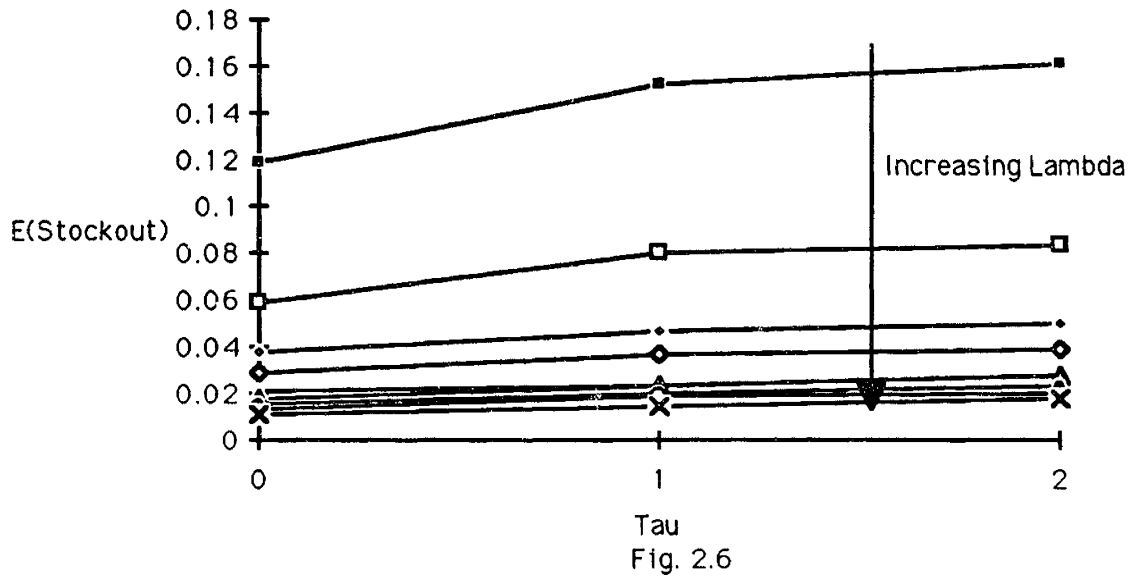
Fig. 2.6

Plot of total cost v/s Reduction factor used on Morton's bounds on optimal policy (tau=1)

Fig. 2.7
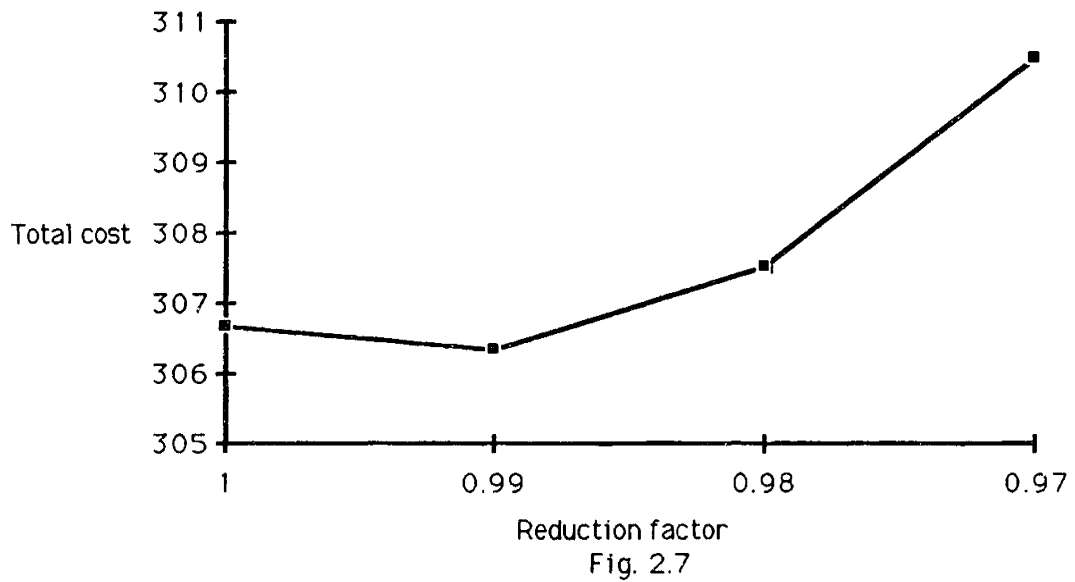
Plot of total cost v/s Reduction factor used on Morton's bounds
on optimal policy (tau=2)
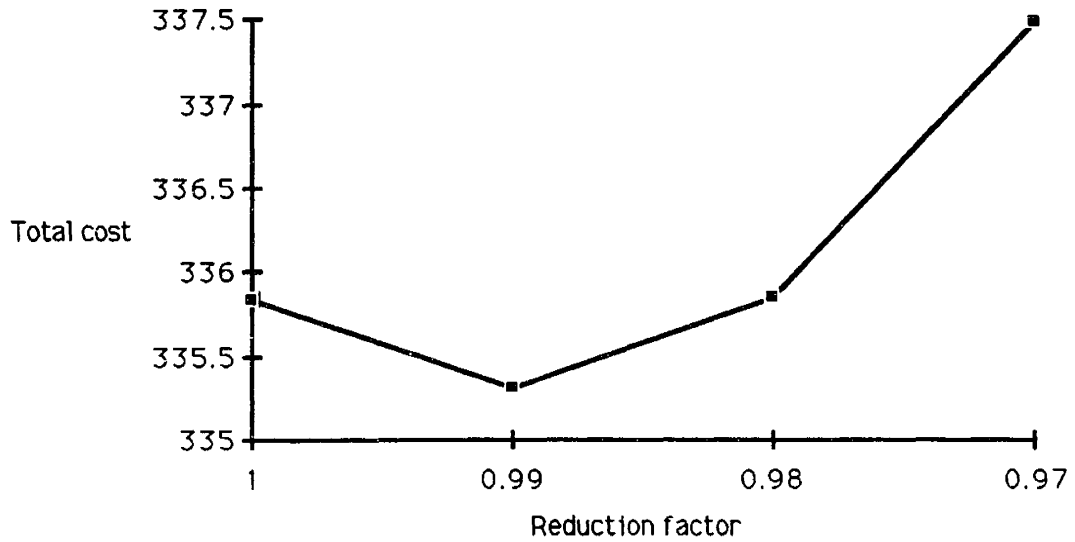

Fig. 2.8

Plot of E(Stockout) v/s Reduction factor used on Morton's
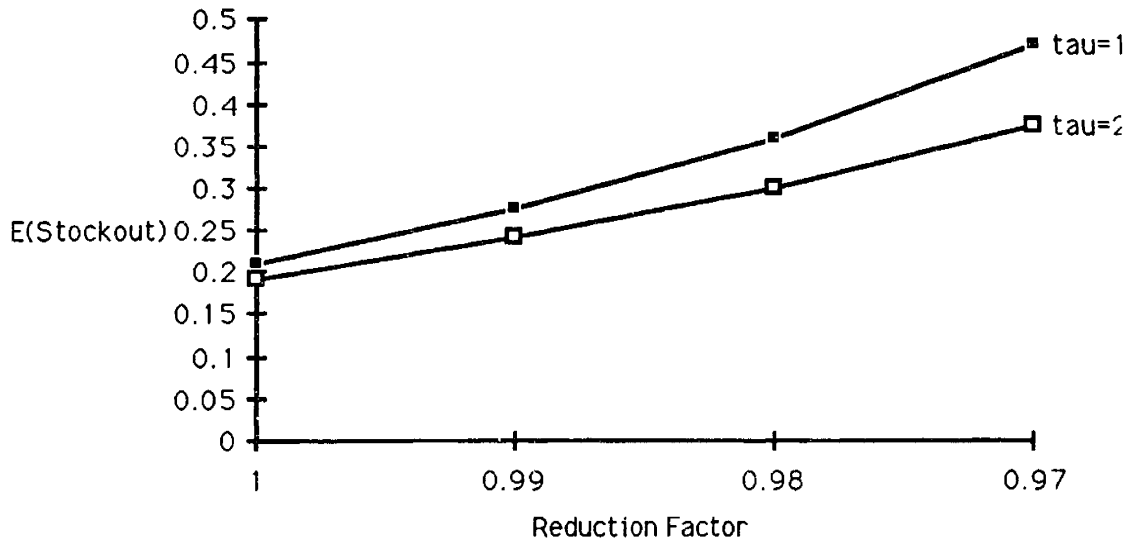bounds on optimal policy

Fig. 2.9

but rises rapidly if we decrease F in both cases. This suggests that the optimal order quantity is very close to the proposed policy. Moreover, from Figure 2.9 it is evident that when we use a reduction factor that is less than 1.0, the expected stockout rises rapidly. This implies that in order to achieve the same level of expected shortage we would need to use a larger penalty cost. In general, this sensitivity analysis strongly suggests that the proposed algorithm is quite robust as well as being near optimal.

2.7 Conclusions

Our analysis demonstrated:

1. Our method gives interesting tradeoffs between cost and service (i.e. probability of stockout or expected stockout).

2. Shorter delivery times are preferable under this model; they produce lower cost and lower probability of stockout.

3. There is a roughly linear relationship between the probability of stockout and the expected stockout under the proposed policy. This is useful since it allows us to use either of the two measures to control the service level of the system. If the decision maker prefers to state his or her preferences in terms of a maximum permissibile average shortage or in terms of a maximum permissible frequency of shortage, then we would be able to derive an approximately optimal policy by converting it to a corresponding implicity penalty cost and then applying our techniques.

4. Although the proposed method is approximate, it is very close to the optimal policy as shown in our sensitivity analysis.

It has to be noted that the problem in real life would involve some additional factors that have not been included in the present model. First, in all the literature related to inventory control under delivery lead time assumptions that was reviewed, orders are placed and received only at the beginning of periods. In the real situation, the bank places its orders on Fridays and receives the ordered amount on the Wednesday of the following week. Such a situation has not been studied due to the resulting complications.

Chapter 3

The Cyclic Demand Model

Another aspect of the bank cash management problem is the control of the amount of cash in the Automatic Teller Machines (ATM). At the beginning of each day the bank reviews the amount of cash in an ATM. Depending on the distribution of the projected demand for cash during that day, the implicit cost of shortage of cash during that day, the cost of holding a certain amount of cash and the cost of ordering cash, the decision maker has to determine the amount of cash to be ordered. It is clear that the amounts of cash demanded on different days at an ATM are not necessarily identically distributed. But it is a reasonalbe assumption that the demand for cash on different days is independent. Further, from our knowledge of demand for cash, we could assume that the demand for cash on specific days (say, Mondays) is identically distributed. Of course, there are other factors that influence demand for cash (like the incidence of long weekends, holidays etc.), but the demand pattern may be seen primarily as a weekly pattern. Since we formulate the problem for the ATM on a weekly basis, the demand pattern for the ATM is cyclic each week, hence it is suitable that we assume the demand for cash seeing by the ATM is cyclic. The model we apply here is again an inventory model with infinite planning time horizon, periodic review, stochastic cyclic daily demand, no delivery lead time and lost sales.

The crucial difference with the model of the previous chapter is that we assumed the demand during each period for cash is cyclic. For example, during the weekend, one would expect a heavier demand for cash than during a week day. Hence we assume

the demand for each day is independent but not necessary identically distributed. On the other hand, we assume that the demand for each day-type (i.e., Monday, Tuesday, etc.) is identically distributed. In other words, the demand pattern is cyclic with a period of one week.

3.1 The Cyclic Demand Model

The model we use is similar to that of Iglehart and Karlin [1]. In their model, they consider the situation where the demand is not i.i.d. but correlated. They have assumed that there is a finite number of states of demand; further, the transitions between states of demand obey a Markov transition law. Our problem is a special case of the problem they studied in the sense that the transition matrix is a simple permutation matrix. The notation we use in this model is as follows:

d_n: demand on day n, $n = 1, 2, ...7$.

$f_n(d_n)$: the demand density function for day n.

y_n: cash available at the beginning of day n.

x_n: cash level at the end of day $n - 1$.

β: discount factor.

The decision we need to make at the beginning of day n is how much to order, given the the on hand inventory, x_{n-1}. There are three main costs that influence the ordering decisions. There is a linear ordering cost c per unit ordered; holding cost is charged at the rate of h for the stock of inventories on hand and the shortage cost associated with the failure to meet the demands is proportional to the extent of shortage with the rate of λ. We assume that the holding and shortage cost are charged at the end of each period. The expected holding and shortage cost as $L_i(y_i)$ is then:

$$L_i(y_i) = h \int_0^{y_i} (y_i - \xi) f_i(\xi) d\xi + \lambda \int_{y_i}^{\infty} (\xi - y_i) f_i(\xi) d\xi$$

Let $J_i(x)$ denote the minimum discounted expected cost incurred during an infinite sequence of time periods if x is the cash available and an optimal ordering rule is used

at each ordering opportunity. Assuming that excess demands are lost and there is no time lag in delivery, the system of the **DP** equations is:

$$J_i(x) = min_{y_i \geq x}\{c(y_i - x) + L_i(y_i) + \beta J_{i+1}(0) \int_{y_i}^{\infty} f_i(\xi)d\xi +$$
$$\beta \int_0^{y_i} J_{i+1}(y_i - \xi)f_i(\xi)d\xi\} \qquad i = 1, ..., 6 \qquad (3.1)$$

$$J_7(x) = min_{y_7 \geq x}\{c(y_7 - x) + L_7(y_7) + \beta J_1(0) \int_{y_7}^{\infty} f_7(\xi)d\xi +$$
$$\beta \int_0^{y_7} J_1(y_7 - \xi)f_7(\xi)d\xi\} \qquad (3.2)$$

These equations can be explained as follows: $J_i(x)$ is the optimal total discounted cost given that we have x at the beginning of a day of type i. The inventory can be raised to y_i by the immediate delivery of an order of size $y_i - x$. Clearly $y_i \geq x$. The expected cost incurred on the first day is then the cost of ordering $y_i - x$ and the expected holding and penalty cost resulting when the starting inventory is y_i. If the demand ξ exceeds y_i, then the closing inventory is going to be 0 (since we are assuming that demand is not backlogged); on the other hand, if the demand ξ is less than the closing inventory is $y_i - \xi$. The subsequent cost is then the optimal total discounted cost given that we have these quantities at the beginning of a day of type $i + 1$ (or 1, if i=7) discounted by the appropriate discounting factor.

3.2   The Algorithm

Notice that $L_i'(y_i) = (h + \lambda)F_i(y_i) - \lambda$ and $L_i''(y_i) = (h + \lambda)f_i(y_i) > 0$, hence the function $L_i(y_i)$ is always convex. We impose the following additional technical conditions on the model:

$$h + \lambda - \beta c > 0 \qquad (3.3)$$

$$L_i'(0) + c < 0, \quad (i = 1, 2, ...7) \qquad (3.4)$$

Notice the second condition is intuitive since $L(0)$ is the holding plus penalty cost given the current openning inventory is zero. Therefor if we order 1 unit and raise the inventory from 0 to 1, the increases in the cost is $c + L'(0)$ and if this quantity is greater than zero, there is no point even ordering at the first place.

For $i = 1, 2, ...7$, let

$$G_i(y_i) \; = cy_i + L_i(y_i) + \beta J_{i+1}(0) \int_{y_i}^{\infty} f_i(\xi)d\xi + \beta \int_0^{y_i} J_{i+1}(y_i - \xi)f_i(\xi)d\xi \qquad (3.5)$$

(Note that the above equation needs to be modified for the case of i=7; as in the earlier section, J_8 should be replaced by J_1 in this case. Similar modifications should be introduced in the arguments that follow wherever necessary.)

Then

$$J_i(x) = min_{y_i \geq x}\{-cx + G_i(y_i)\} = -cx + min_{y_i \geq x}G_i(y_i) \qquad (3.6)$$

Taking the derivative of function $G_i(y_i)$ with respect to y_i, we get:

$$G_i'(y_i) = c + L_i'(y_i) + \beta \int_0^{y_i} J_{i+1}'(y_i - \xi)f_i(\xi)d\xi \quad i = 1, ..., 7 \qquad (3.7)$$

Define y_i^* such that $G_i'(y_i^*) = 0$. We can actually solve the following equations to get y_i^*:

$$c + L_i'(y_i) + \beta \int_0^{y_i} J_{i+1}'(y_i - \xi)f_i(\xi)d\xi = 0 \qquad i = 1, 2, ...7$$

We will show in section 3.3 that these roots are indeed the optimizers of the functions $G_i(y_i)$. Therefore, the optimal policy calls for ordering of cash to the level of y_i^*, when $x \leq y_i^*$ and not ordering when $x > y_i^*$.

The Algorithm

We now state the algorithm that can be used to calculate these critical numbers. For ease of exposition we assume that the y_i^*, $i = 1...7$ are distinct. We denote the smallest critical number by $y_{[1]}$, second smallest by $y_{[2]}$ and so on.

Step 1 Calculate the smallest number $y_{[1]}$

Consider the equations

$$H_i^1(y) = c + L_i'(y) + \beta \int_0^y T_{i+1}^0(y - \xi)f_i(\xi)d\xi = 0 \qquad i = 1....7$$

26

where $T_i^0(y) = -c$, for $i = 1...7$. Differentiating each function $H_i^1(y)$, we get:

$$\frac{\partial}{\partial y} H_i^1(y) = L_i''(y) - \beta c f_i(y)$$
$$= (h + \lambda - \beta c) f_i(y)$$

By assumption (3.3), it is easy to verify that each $H_i^1(y)$ is strictly increasing. Using assumption (3.4) it follows that $H_i^1(0) < 0$ for each i. Moreover, we have $\liminf_{y \to \infty} H_i^1(y) \geq c(1 - \beta) > 0$ for $i = 1...7$. Hence each $H_i^1(y) = 0$ has a unique root. Now determine the root of each $H_i^1(y) = 0$ and denote these roots by $y_1^1, y_2^1, ...y_7^1$ and let :

$$y_{[1]} = min_{1 \leq j \leq 7}(y_j^1)$$

The subscript [1] simply stands for the index value equal to that j for which y_j^1 is minimum.

Step 2 Calculate the second smallest number $y_{[2]}$

To calculate the second smallest critical number we construct the functions

$$H_i^2(y) = c + L_i'(y) + \beta \int_0^y T_{i+1}^1(y - \xi) f_i(\xi) d\xi \qquad i = 1,.....7$$

where

$$T_j^1(y) = -c \quad (j \neq [1])$$

$$T_{[1]}^1(y) = \begin{cases} -c & \text{if } 0 \leq y \leq y_{[1]}, \\ g_{[1]}^1(y) & \text{if } y_{[1]} < y, \end{cases}$$

and the function $g_{[1]}^1(y)$ is determined as the solution of the integral equation

$$g_{[1]}^1(y) = L_{[1]}'(y) + \beta \int_{J_0}^y T_{[1]+1}^1(y - \xi) f_{[1]}(\xi) d\xi \qquad (y > y_{[1]})$$

It is clear that by using assumption (3.3) and the fact that $L_i(y)$ is convex, $g_{[1]}^1(y)$ is nondecreasing, and in particular that $T_{[1]}^1(y) \geq -c$ for $y \geq 0$. Moreover, $T_{[1]}^{(1)'}(y) \geq 0$ for all nonnegative y with the possible exception of the point $y = y_{[1]}$, where only left and right-hand bounded deritives may exist. The above facts ensure that

$$\liminf_{y \to \infty} H_i^2(y) \geq c(1 - \beta) > 0 \qquad for \quad i = 1,.....7$$

27

and each $H_i^2(y)$ is nondecreasing. Also we know by condition (3.4) that $H_i^2(0) < 0$ for each i. Hence each equation $H_i^2(y) = 0$ has a unique root. Denote these roots by $y_1^2, y_2^2, ... y_7^2$. Observe that $y_i^2 \leq y_i^1$, since $H_i^2(y) \geq H_i^1(y)$ for all y. Moreover, the minimum of the y_i^2 is exactly $y_{[1]}$, since $H_i^1(y) = H_i^2(y)$ for $y \leq y_{[1]}$. Now define

$$y_{[2]} = min_{(1 \leq j \leq k \ and \ j \neq [1])}(y_j^2)$$

Step r+1 Calculate the $(r + 1)^{th}$ smallest number $y_{[r+1]}$

Suppose that the r smallest critical numbers are determined: $y_{[1]} < y_{[2]} < .... < y_{[r]}$. Construct the functions

$$H_i^{r+1}(y) = c + L_i'(y) + \beta \int_0^y T_{i+1}^r(y - \xi) f_i(\xi) d\xi \qquad i = 1...7$$

where

$$T_j^r(y) = -c \quad (j \neq [1][2]...[r])$$

$$T_{[j]}^r(y) = \begin{cases} -c & \text{if } 0 \leq y \leq y_{[j]}, \\ g_{[j]}^1(y) & \text{if } y_{[j]} < y \leq y_{[j+1]}, \\ g_{[j]}^2(y) & \text{if } y_{[j+1]} < y \leq y_{[j+2]}, \\ g_{[j]}^{r+1-j}(y) & \text{if } y_{[r]} < y, \end{cases}$$

Here the $g_{[j]}^i(y)$ for $i = 1, 2, ... r - j$ and $j = 1, 2, ... r - 1$ were determined during the analysis of the previous steps of the algorithm. Now we solve for $g_{[j]}^{r+1-j}(y)$, where $j = 1, 2, ... r$ as the solutions of the system of equations:

$$g_{[j]}^{r+1-j}(y) = L_{[j]}'(y) + \beta \int_0^y T_{[j]+1}^1(y - \xi) f_{[j]}(\xi) d\xi \quad (y > y_{[r]} \ and \ j = 1, 2, ... r) \quad (3.8)$$

Determine the root of each $H_i^{r+1}(y) = 0$ and denote these roots by $y_1^{r+1}, y_2^{r+1}, ... y_7^{r+1}$ and let :

$$y_{[r+1]} = min_{(1 \leq j \leq 7 \ and \ j \neq [1], [2], ... [r])}(y_j^{r+1}).$$

We get the $r + 1$th smallest critical number.

28

3.3 Proof of The Correctness of The Algorithm

In this section we show that Iglehart and Karlin's algorithm of section 3.2 correctly calculates the optimal y_i^*. We first show that (3.8) indeed has a solution.

Proposition 3.1 *There exists a unique solution to the system of equations:*

$$g_{[j]}^{r+1-j}(y) = L'_{[j]}(y) + \beta \int_0^y T_{[j]+1}^1(y - \xi)f_{[1]}(\xi)d\xi \qquad (y > y_{[j]} \quad and \quad j = 1, 2, ...r)$$

Proof:

Observe that the construction of $g_{[j]}^{r+1-j}(y)$ compels $T_{[j]}^{r+1-j}(y)$ to be continuous at $y_{[j]}^{r+1-j}$, since $H_{[j]}^r(y_{[j]}) = 0$. Now introduce the function $h_{[j]}^{r+1-j}(y)$, which is the translation of $g_{[j]}^{r+1-j}$ given by:

$$h_{[j]}^{r+1-j}(y) = g_{[j]}^{r+1-j}(y + y_{[r]}) \qquad for \quad y > 0$$

We reduce $g_{[j]}^{r+1-j}(y)$ to a system of renewal equations as follows:

$$h_{[j]}^{r+1-j}(y) = L'_{[j]}(y + y_{[r]}) + \beta \int_0^{y+y_{[r]}} T_{[j]+1}^{r+1-j}(y + y_{[r]} - \xi)f_{[j]}(\xi)d\xi \quad y > 0$$

$$h_{[j]}^{r+1-j}(y) = \begin{cases} L'_{[j]}(y + y_{[r]}) - \beta c F_{[j]}(y + y_{[r]}) & \text{if } not\ \exists m : [m] = [j] + 1 \\ & \text{for } m = 1, 2, ...r \\ \\ L'_{[j]}(y + y_{[r]}) & \\ + \beta \int_0^{y+y_{[r]}} T_{[m]}^r(y + y_{[r]} - \xi)f_{[j]}(\xi)d\xi & \text{if } \exists m \\ & [m] = l = [j] + 1 \end{cases}$$

Now consider the second case of $h_{[j]}^{r+1-j}(y)$:

Substituting $u = y + y_{[r]} - \xi$, we have:

$$\int_0^{y+y_{[r]}} T_{[m]}^r(y + y_{[r]} - \xi)f_{[j]}(\xi)d\xi = -c \int_0^{y_{[m]}} f_{[j]}(y + y_{[r]} - u)du$$

$$+ \sum_{n=m}^{r-1} \int_{y_{[n]}}^{y_{[n+1]}} g_{[m]}^{n+1-m}(u)f_{[j]}(y + y_{[r]} - u)du$$

$$+ \int_{y_{[r]}}^{y+y_{[r]}} g_{[m]}^{r+1-m}(u)f_{[j]}(y + y_{[r]} - u)du$$

29

Substituting back $\xi = y + y_{[r]} - u$, we have:

$$\int_0^{y+y_{[r]}} T_{[m]}^r(y + y_{[r]} - \xi) f_{[j]}(\xi) d\xi = -c \int_{y+y_{[r]}-y_{[m]}}^{y+y_{[r]}} f_{[j]}(\xi) d\xi$$

$$+ \sum_{n=m}^{r-1} \int_{y+y_{[r]}-y_{[n+1]}}^{y+y_{[r]}-y_{[n]}} h_{[m]}^{n+1-m}(y - \xi) f_{[j]}(\xi) d\xi$$

$$+ \int_0^y h_{[m]}^{r+1-m}(y - \xi) f_{[j]}(\xi) d\xi$$

So,

$$h_{[j]}^{r+1-j}(y) = a_{[j]}(y) + \beta \times \int_0^y h_{[m]}^{r+1-m}(y - \xi) f_{[j]}(\xi) d\xi$$

where

$$a_{[j]} = L_{[j]}'(y + y_{[r]}) - \beta c \int_{y+y_{[r]}-y_{[m]}}^{y+y_{[r]}} f_{[j]}(\xi) d\xi$$

$$+ \beta \sum_{n=m}^{r-1} \int_{y+y_{[r]}-y_{[n+1]}}^{y+y_{[r]}-y_{[n]}} h_{[m]}^{n+1-m}(y - \xi) f_{[j]}(\xi) d\xi$$

Differentiate $h_{[j]}^{r+1-j}(y)$ with respect to y, we have:

$$\frac{d}{dy} h_{[j]}^{r+1-j}(y) = \begin{cases} L_{[j]}''(y + y_{[r]}) - \beta c f_{[j]}(y + y_{[r]}) & \text{if not } \exists \, [m] = [j] + 1 \\ & \text{for } m = 1, ..., r \\ a_{[j]}'(y) + \beta \int_0^y \frac{d}{dy} h_{[m]}^{r+1-m}(y - \xi) f_{[j]}(\xi) d\xi & \\ + \beta h_{[m]}^{r+1-m}(0) f_{[j]}(y) & \text{otherwise} \end{cases} \tag{3.9}$$

Where

$$a_{[j]}'(y) = L_{[j]}''(y + y_{[r]}) - \beta c f_{[j]}(y + y_{[r]}) - \beta h_{[m]}^{r+1-m}(0) f_{[j]}(y)$$

$$+ \beta \sum_{n=m}^{r-1} (h_{[m]}^{n+1-m}(y_{[n]} - y_{[r]}) f_{[j]}(y + y_{[r]} - y_{[n]})$$

$$- h_{[m]}^{n+1-m}(y_{[n+1]} - y_{[r]}) f_{[j]}(y + y_{[r]} - y_{[n+1]}))$$

$$+ \beta \sum_{n=m}^{r-1} \int_{y+y_{[r]}-y_{[n+1]}}^{y+y_{[r]}-y_{[n]}} \frac{d}{dy} h_{[m]}^{n+1-m}(y - \xi) f_{[j]}(\xi) d\xi \tag{3.10}$$

By using the continuity of $T^r_{[m]}(y)$ at the critical points, the assumption that $h + \lambda - \beta c > 0$ and observing that $h^{r+1-m}_{[m]}(0)$ is $-c$, $a'_{[j]}(y)$ can be simplified as:

$$
\begin{aligned}
a'_{[j]}(y) &= L''_{[j]}(y + y_{[r]}) - \beta c f_{[j]}(y + y_{[r]}) - \beta h^{r+1-m}_{[m]}(0) f_{[j]}(y) \\
&\quad + \beta \sum_{n=m}^{r-1} \int_{y+y_{[r]}-y_{[n+1]}}^{y+y_{[r]}-y_{[n]}} \frac{d}{dy} h^{n+1-m}_{[m]}(y - \xi) f_{[j]}(\xi) d\xi \\
&= (h + \lambda - \beta c) f_{[j]}(y + y_{[r]}) + \beta c f_{[j]}(y) \\
&\quad + \beta \sum_{n=m}^{r-1} \int_{y+y_{[r]}-y_{[n+1]}}^{y+y_{[r]}-y_{[n]}} \frac{d}{dy} h^{n+1-m}_{[m]}(y - \xi) f_{[j]}(\xi) d\xi
\end{aligned}
$$

Now by induction, we have that $\frac{d}{dy} h^{n+1-m}_{[m]}$ is positive in the range of integration; hence we have $a'_{[j]}(y) > 0$.

In order to show equation (3.9) has a solution, we will use the following theorem of Feller[5]

Theorem 3.1 *(Feller[5]) Let f_i denote density functions and suppose that the functions $a_i(x)$ are continuous on the positive real axis. Then the system of renewal equations*

$$
u_i(x) = a_i(x) + \sum_{j=1}^{k} \int_0^x u_j(x - t) f_i(t) dt \quad x \geq 0 \quad i = 1, ..., k
$$

possesses a unique set of solutions. Moreover, if $a_i(x) \geq 0$ for $x \geq 0$ and $i = 1, ..., k$, then $u_i(x) \geq 0$ for $i = 1, ..., k$.

Now applying the above theorem to (3.9), we conclude that $\frac{d}{dy} h^{r+1-j}_{[j]}(y) > 0$, i.e., $h^{r+1-j}_{[j]}(y)$ is a non-decreasing function of y and that $\frac{d}{dy} h^{r+1-j}_{[j]}(y)$ are unique. Since h is a translated version of g, we can conclude that g, hence T, is a non-decreasing unique function of y. □

Figure 3.1 shows that T^r_j *for $j = 1, ..., 7$, $r = 1, 2, ...7$* are indeed continuous and non-decreasing.

Notice that there is a significant difference among T's. Ignoring these differences by averaging the demand over the days or any other such technique which leads one to not take the difference between days into account in constructing an policy will be misleading.

Under the conditions of section 3.2, we now establish the following proposition.

Plot of T functions for all days of the week

Fig. 3.1
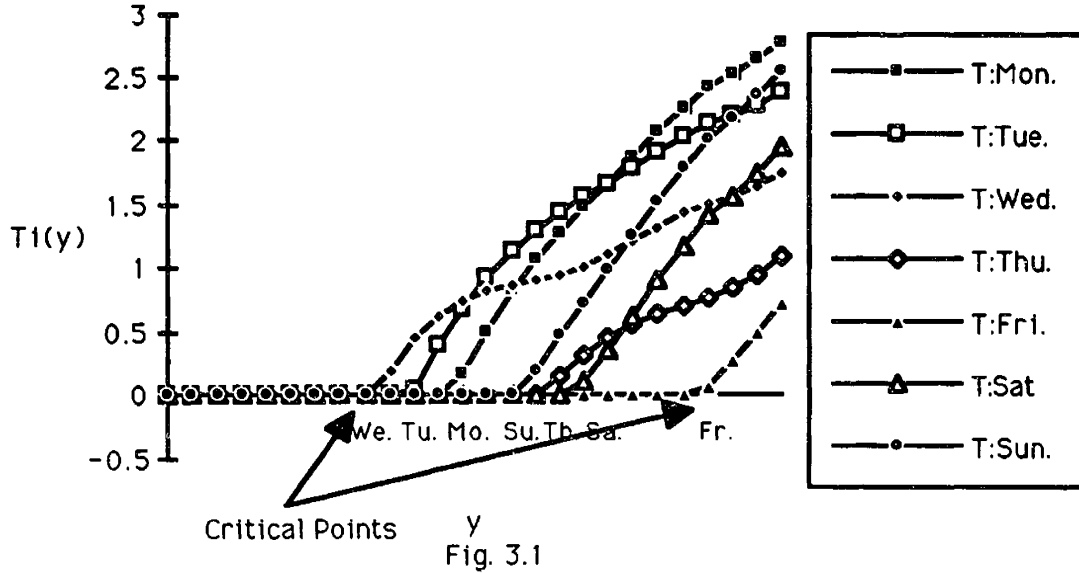
Proposition 3.2 *Let $J_i(x)$ denote the minimum discounted expected cost if the initial stock level is x. If y_i^* is the optimal critical number for the first period, then $J_i(x)$ is convex and has a continuous derivative and that $J_i'(x) = -c$ for $(x \leq y_i^*)$.*

Proof:

See Karlin[3] (pp.142-49).   □

Proposition 3.3 *If the functions $L_i(y) - \beta c \int_y^\infty (\xi - y) f_i(\xi) d\xi$ $(i = 1, ...7)$ are convex in y, the optimal policy when we do not backlog excess demand is characterized by a set of 7 critical numbers $y_1^*, y_2^*, ...y_7^*$.*

Proof:

For details, see Karlin[3] (pp149-53, 162-69).   □

Proposition 3.4 *The function $G_i(y_i)$ is convex, and $G_i'(y_i) = 0$ has a unique solution.*

Proof:

Recall

$$G_i(y_i) = cy_i + L_i(y_i) + \beta J_{i+1}(0) \int_{y_i}^\infty f_i(\xi) d\xi + \beta \int_0^{y_i} J_{i+1}(y_i - \xi) f_i(\xi) d\xi \quad i = 1,.,7$$

Since the function $L_i(y_i)$ is convex, by proposition (3.2) $J_i(y_i)$ is convex, hence $G_i(y_i)$ is convex. Moreover,

$$G_i'(y_i) = c + L_i'(y_i) + \beta \int_0^{y_i} J_{i+1}'(y_i - \xi) f_i(\xi) d\xi \quad i = 1, ..., 7$$

hence, $G_i'(y_i)$ are monotone-increasing and by assumption (3.4) , we have $G_i'(0) < 0$. Hence the equations $G_i'(y_i) = 0$, $i = 1, ..., 7$ have unique roots. $\qquad\qquad\square$

Theorem 3.2 *(Iglehart and Karlin [1]) The (r+1)th smallest critical number is $y_{[r+1]}$, where $y_{[r+1]}$ is the (r+1)th smallest root of the equations $H_i^{r+1}(y) = 0$, for $i = 1, 2, ...7$.*

Proof:

The optimal critical number y_i^* should occur at the absolute minimum of $G_i(y)$, since $G_i'(y_i) = c + L_i'(y_i) + \beta \int_0^{y_i} J_{i+1}'(y_i - \xi) f_i(\xi) d\xi \quad i = 1, ..., 7$

and $G_i'(0) \leq 0$, we infer that y_i^* is a root of the equation :

$$G_i'(y_i) = c + L_i'(y_i) + \beta \int_0^{y_i} J_{i+1}'(y_i - \xi) f_i(\xi) d\xi = 0 \quad i = 1, ..., 7$$

We now show that y_i^* can be evaluated by the algorithm given above. Since $J_i(x) = min_{y \geq x}\{-cx + G_i(y)\}$, $J_i'(x) = -c$, for $x \leq min_{1 \leq i \leq k}\{y_i^*\}$. Hence the smallest root of the equations $H_i^1(y) = 0$ is $y_{[1]} = min_{1 \leq i \leq k}\{y_i^*\}$ Now when y ranges on the interval from $y_{[1]}$ to $y_{[2]}$, the second smallest critical number, we have $J_i'(y) = -c$ for $y \leq y_{[2]}$, $i \neq [1]$, and

$$J_{[1]}'(y) = L_{[1]}'(y) + \beta \int_o^y J_{[1]+1}'(y - \xi) f_{[1]}(\xi) d\xi \qquad for\ y \geq y_{[1]}$$

Observe that $g_{[1]}^1(y) = J_{[1]}'(y)$ on the range $y_{[1]} \leq y \leq y_{[2]}$, hence $G_i'(y) = H_i^2(y)$ for $y \leq y_{[2]}$. Thus the second smallest root of $H_i^2(y) = 0$ is $y_{[2]}$. At step three, we note that $J_i'(y) = -c$ for $y \leq y_{[3]}$ and $i \neq [1], [2]$. From step 2, we already have $J_{[1]}'(y)$ for $y \geq y_{[1]}$ and for $y \geq y_{[2]}$

$$J_{[2]}'(y) = L_{[2]}'(y) + \beta \int_0^y J_{[2]+1}'(y - \xi) f_{[2]}(\xi) d\xi$$

Observe that $J_{[1]}'(y)$ with $y \geq y_{[1]}$ and $J_{[2]}(y)$ with $y \geq y_{[2]}$ can be now written as:

33

$$J'_{[1]}(y) = \begin{cases} -c & \text{if } y \le y_{[1]}, \\ g^1_{[1]}(y) & \text{if } y_{[1]} < y \le y_{[2]}, \\ g^2_{[1]}(y) & \text{if } y_{[2]} < y \le y_{[3]}, \end{cases}$$

and

$$J'_{[2]}(y) = \begin{cases} -c & \text{if } y \le y_{[2]}, \\ g^1_{[2]}(y) & \text{if } y_{[2]} < y \le y_{[3]}, \end{cases}$$

Hence $G'_i(y) = H^3_i(y)$ for $y \le y_{[3]}$, and the third smallest root of $H^3_i(y) = 0$ is $y_{[3]}$. Repeating the above argument, we prove the theorem. $\qquad \square$

3.4 The Exponential Cyclic Demand Case

The algorithm outlined in the section 3.2 is quite complex primarily because of the need to solve the integral equation (3.9). In order to get insight about the behavior of the algorithm, we will further assume that the demand is exponentially distributed. In this case, we will show that the system of simultaneous renewal equations (3.9) reduces to a system of simultaneous first order ordinary differential equations with constant coefficients, which we solve numerically.

Proposition 3.5 *Let* $f_i(y) = \lambda_i e^{-\lambda_i y}$, *then equation* (3.9) *reduce to*

$$\frac{d}{dy} h^{k+1-j}_{[j]}(y) = \beta \lambda_{[j]} h^{k+1-m}_{[m]}(y) - \lambda_{[j]} h^{k+1-j}_{[j]}(y) + a'_{[j]}(y) + \lambda_{[j]} a_j(y)$$
$$\text{for} \quad k = 1, 2, \dots r$$

Proof:

It is obviously true when

$$\frac{d}{dy} h^{r+1-j}_{[j]}(y) = L''_{[j]}(y + y_{[r]}) - \beta c f_{[j]}(y + y_{[r]})$$

For the second case of $\frac{d}{dy} h^{r+1-j}_{[j]}(y)$, it is:

$$\frac{d}{dy} h^{r+1-j}_{[j]}(y) = a'_{[j]}(y) + \beta \int_0^y \frac{d}{dy} h^{r+1-m}_{[m]}(y - \xi) \lambda_{[j]} e^{-\lambda_{[j]} \xi} d\xi + \beta h^{r+1-m}_{[m]}(0) \lambda_{[j]} e^{-\lambda_{[j]} y}$$

34

Integrating by parts, we have

$$\beta \int_0^y \frac{d}{dy} h_{[m]}^{r+1-j}(y-\xi)\lambda_{[j]} e^{-\lambda_{[j]}\xi} d\xi = -\beta \int_0^y \lambda_{[j]} e^{-\lambda_{[j]}\xi} d(h_{[m]}^{r+1-j}(y-\xi))$$

$$= -\beta\lambda_{[j]} e^{-\lambda_{[j]}y} h_{[m]}^{r+1-j}(0) + \beta\lambda_{[j]} h_{[m]}^{r+1-j}(y)$$

$$+ \lambda_{[j]}(h_{[j]}^{r+1-j}(y) - a_{[j]}(y))$$

Hence the differential equation is given by:

$$\frac{d}{dy} h_{[j]}^{k+1-j}(y) = \beta\lambda_{[j]} h_{[m]}^{k+1-m}(y) - \lambda_{[j]} h_{[j]}^{k+1-j}(y) + a_{[j]}'(y) + \lambda_{[j]} a_{j]}(y) \quad (3.11)$$

$$for \quad k = 1, 2, ... r$$

□

Therefore the algorithm in section 3.2 in the case of exponential demand reduces to:

Initialization (i.e., set demand parameter values, costs etc.);

$t = \frac{\sqrt{5}-1}{2}$;

Set λ range to $(\lambda_{min}, \lambda_{max})$;

While $(\lambda_{max} - \lambda_{min} > \epsilon)$

{

   $\lambda = \lambda_{min} + (1 - t) * (\lambda_{max} - \lambda_{min})$;

   Initialize $T_j = -c$ for $j = 1, ..., 7$;

   Iteration_count = 1;

   Solved_set = ϕ;

   While$(r \leq 7)$

   {

   Solve successively the equations $H_j = 0$ for $j = 1, ..., 7$;

   for (j not in *Solved_set*)

       Denote the rth smallest root as the rth critical number;

   If(the solution of equation j produced the rth smallest root)

   then include j in *Solved_set*;

   Update T_j for $x > r^{th}$ critical point and j in *Solved_set* by solving

the system of simultaneous differential equations for h_j;

Increment *Iteration_count* by 1;

}

Simulate the system with current iterate of critical numbers computing an

estimate of the expected amount of stock out by cumulating the amount of

stock out on each day (if any) and dividing by the

total length of the simulations;

If (the estimate of E(Stock out) exceeds prespecified bound) then

$\lambda_{min} = \lambda;$

Else

$\lambda_{max} = \lambda;$

}

3.5   Computational Results

The algorithm described in the earlier sections was implemented in FORTRAN due to the need to access a library routine for the solution of simultaneous differential equations. The package chosen was Numerical Algorithms (NAG). The implementation of this algorithm required the use of numerical techniques to compute solutions to transcendental equations, computing and storing solutions to differential equations by constructing a suitably chosen grid of points that included the range of interest.

The approach to choose the penalty used here was similar to that used in chapter 2 (see section (2.6). The proposed algorithm is clearly valid when the value of the penalty cost is known. In our case, the desire for no shortage (or small stockout instances) is expressed in terms of a constraint on the probability of stockout. We use an identical procedure of searching for the value of the penalty that results in an acceptable probability of stockout at the lowest possible cost.

As we mentioned we have assumed that daily demand was exponentially distributed with different means on different days of the week. The exact values of the

parameters chosen were selected to reflect an increased demand on Friday and Saturday with Tuesday and Wednesday representing the slowest of the days. For the specific values of these parameters, see the following table. (In this case we used the following values of the input parameters : the discounting factor is the daily equivalent of a 15% annual interest rate, the ordering cost is 0.001 per unit of cash ordered, the holding cost is computed using the daily interest rate, and the maximum permissible value for the probability of shortage is 0.01.)

Day	MeanDemand($\times 10^3$)	CriticalNumber($\times 10^3$)
Mon	0.35	1.53
Tue	0.30	1.32
Wed	0.25	1.13
Thu	0.45	2.04
Fri	0.70	2.82
Sat	0.50	2.17
Sun	0.45	1.89

In the above table, the critical numbers are those which are optimal for a case in which the penalty cost was 90.3; this corresponds to a probability of stock out of approximately 1%. It is clear that the critical numbers are significantly larger than the mean demand for that day; this can be explained by the fact that the exponential distribution has a long tail and our constraint requires that the probability of stock out be very small.

Some observations can be made from the following figures.

Figure 3.2 plots critical numbers for each day as the function of the penalty cost. Notice that the critical numbers do not change a lot as the penalty changes.

In figure 3.3 and 3.4 we investigated the depence of the total cost and the probability of stockout as a function of the penalty cost. The total cost was computed from the simulation experiment. As expected the total cost is increasing while the probability of stockout is decreasing with λ.
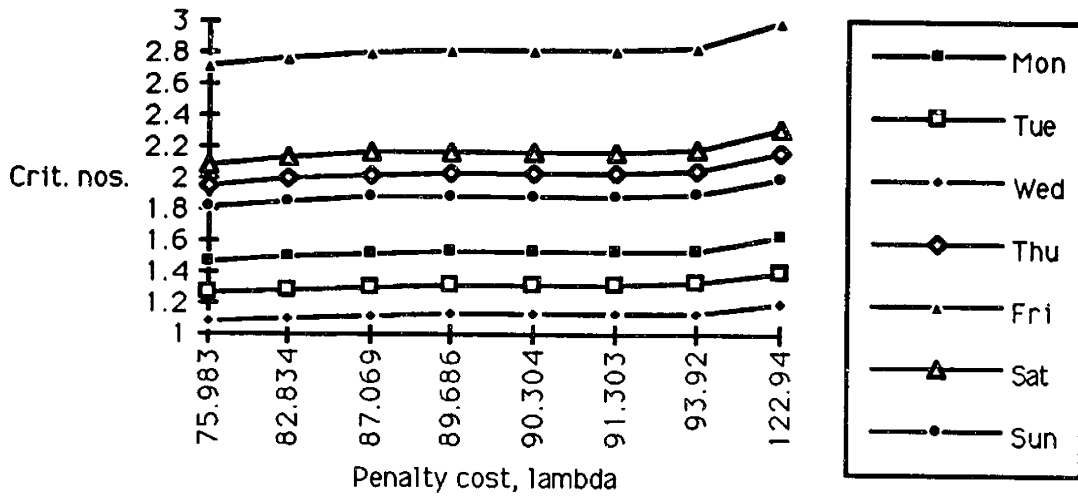
Critical numbers for each day v/s penalty cost

Crit. nos.

	Mon
	Tue
	Wed
	Thu
	Fri
	Sat
	Sun

Penalty cost, lambda

Fig. 3.2

Total Cost v/s penalty cost

Total Cost

Penalty cost, lambda

Fig. 3.3

38
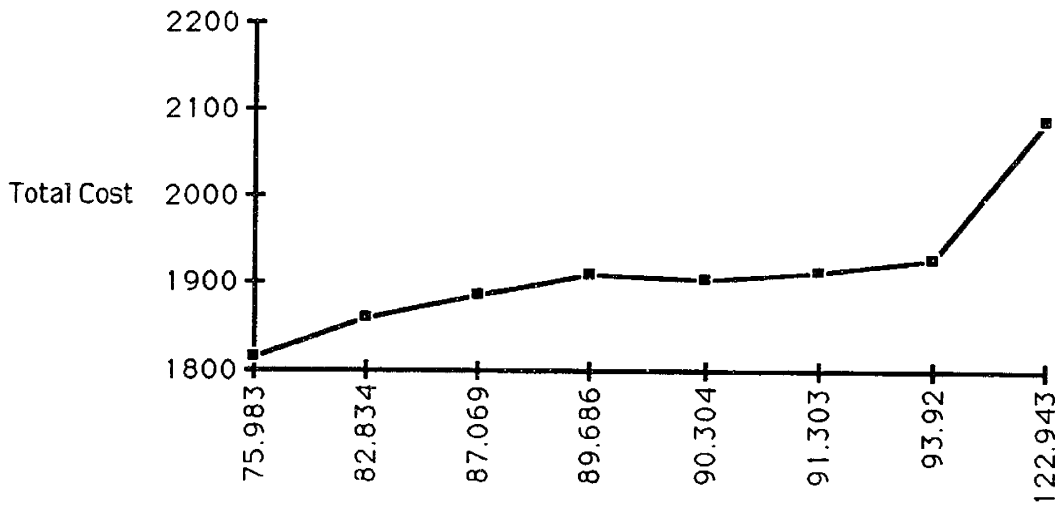
Probability of Stockout v/s penalty cost

Penalty cost, lambda
Fig. 3.4
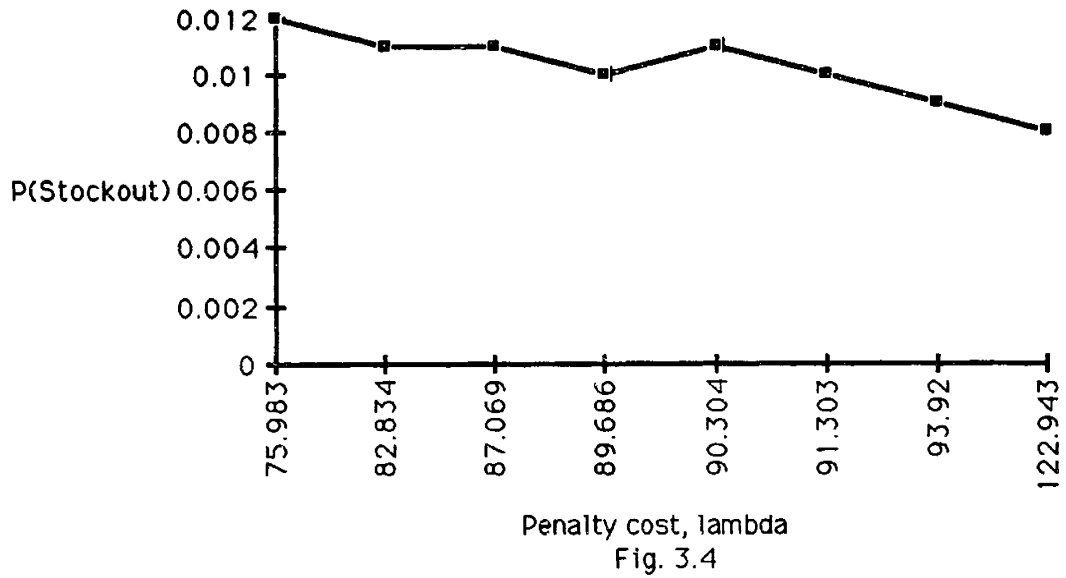
Total discounted holding and ordering cost v/s probability of stockout
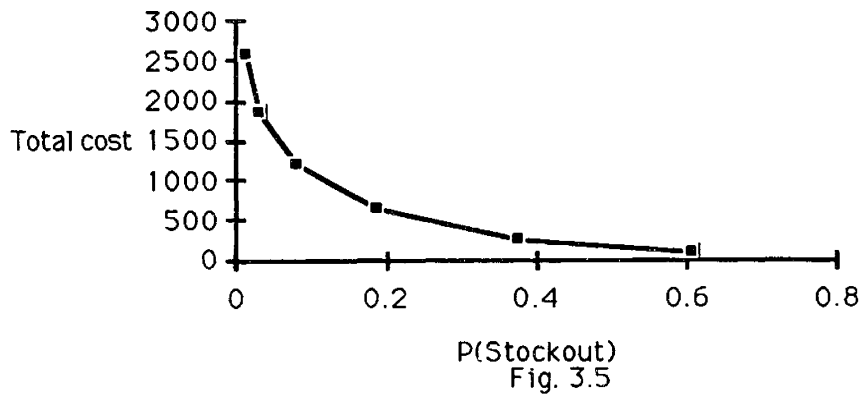
P(Stockout)
Fig. 3.5

In order to further understand the behavior of the system, we have ploted in Figure 3.5 the total cost as a function of the probability of stockout. This curve could be useful in operating the ATM system.

3.6 Summary

We believe that our study reached the following conclusions:

1. The algorithm we used is an exact algorithm which provides the optimal solution to the cyclic demand model (infinite time horizon, no delivery time)

2. We have implemented the algorithm with the exponential demand case.

3. We have provided the tradeoff curves between the total cost and the probability of stockout, which is useful for the decision maker to operate the ATM system.

4. In its current form, our implementation may not seem to have much practical utility since the exponential assumption is highly unrealistic but we believe that the procedure can be extended to include all distributions of the gamma family.

A limitation of this model is that it ignores set-up costs. This assumption may be considered to be tolerable only in the case in which ATM machines which are located at or near a source of cash so that re-stocking can be done at small or no cost. It would not be very applicable for remotely located machines which would require substantial cost and effort to refill.

Hence in the next chapter, we develop this model to include the set up cost

Chapter 4

The Cyclic Demand Model with Set-up Cost

In the preceding chapter, we considered the problem of control of the amount of cash in the Automatic Teller Machine (ATM). As we stated there, we have assumed that at the beginning of each day the supervisor of the bank reviews the amount of cash in an ATM. In this chapter we attempt to study the following variation: once the decision maker places an order, there will be a fixed cost. The relevant model is an inventory model with infinite planning time horizon, periodic review, stochastic daily demand, no delivery lead time, lost sales and **set-up costs**.

As we did in the previous chapter, we assume that the demand for cash during each period is cyclic. We also assume that the demand for each day-type (i.e., Monday, Tuesday, etc.) is independent but not identically distributed.

4.1 The Formulation of Cyclic Demand Model with Set-up Cost

We will use the following natation:

d_n: demand on day n, $n = 1, 2, ...7$.

$f_n(d_n)$: the demand density function for day n.

y_n: cash available at the beginning of day n.

x_n: cash level at the end of day $n - 1$.

β: discount factor.

K: Cost to be incurred for each instance of re-ordering.

The decision we need to make at the beginning of day n is how much to order, given the on hand inventory is x_{n-1}. There are four main costs that influence the ordering decisions. There is a linear ordering cost c per unit ordered; holding cost is charged at the rate of h for the stock of inventories on hand at the end of the period; the shortage cost associated with the failure to meet the demands is proportional to the extent of shortage with the rate of λ; and a fixed amount of K that is charge everytime an order is placed. We assume that the holding and shortage cost are charged at the end of each period. Now denote the one-day expected holding and shortage cost as $L_i(y_i)$, the demand density on that day is $f_i(d_i)$ and the excess demand is lost. Then:

$$L_i(y_i) = h \int_0^{y_i} (y_i - \xi) f_i(\xi) d\xi + \lambda \int_{y_i}^{\infty} (\xi - y_i) f_i(\xi) d\xi$$

Let $J_i(x)$ denote the minimum discounted expected cost incurred during an infinite sequence of time periods if x is the initial level of stock, f_i is the demand density in the first period, and an optimal ordering rule is used at each ordering opportunity. Assuming that excess demands are lost and there is no time lag in delivery, the system of functional equations are:

$$J_i(x) = min_{y_i \geq x}\{K\delta(y_i - x) + c(y_i - x) + L_i(y_i) + \beta J_{i+1}(0) \int_{y_i}^{\infty} f_i(\xi)d\xi +$$
$$\beta \int_0^{y_i} J_{i+1}(y_i - \xi)f_i(\xi)d\xi\} \qquad i = 1, ..., 6$$

$$J_7(x) = min_{y_7 \geq x}\{K\delta(y_7 - x) + c(y_7 - x) + L_7(y_7) + \beta J_1(0) \int_{y_7}^{\infty} f_7(\xi)d\xi +$$
$$\beta \int_0^{y_7} J_1(y_7 - \xi)f_7(\xi)d\xi\}$$
$$where \quad \delta(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{otherwise} \end{cases}$$

Unlike in the previous case, these functions do not possess the convexity property even when the functions $L_i(y_i)$ are convex. Specifically, the purchasing or ordering cost, if one should include the set-up cost in the ordering cost, is no longer convex.

If we should replace the infinite time horizon assumption with a finite time horizon assumption, this problem can be re-formulated as a finite time horizon dynamic programming problem. Clearly, without further work, it would be impossible to establish even the form of the optimal policy for the cyclic model with set-up cost. The literature has not addressed this problem. Given the cyclic nature of the demand pattern, it also appears likely that this approach would be computationally very complex. This suggests that we should explore heuristic solutions to our problem.

4.2   Two Proposed Heuristic Approaches

The only known variation of this problem whose solution appears to have been reported in the literature is the infinite time horizon case with i.i.d. demand distribution. Under these assumptions, it has been shown that the optimal policy can be characterized by two numbers s and S; when the closing inventory at the end of a period is less than s, we order upto S, else no order is placed. Specifically, when the demand is exponentially distributed, Scarf [7] shows that

$$Q = \sqrt{\frac{2Km}{h}} \tag{4.1}$$

$$e^{-\frac{s}{m}} = \frac{h}{p}(1 + \frac{Q}{m}) \tag{4.2}$$

where K=set-up cost, $Q = S - s$, m=mean daily demand, h=holding cost per day, p=penalty cost for shortage.

We propose to use (4.1) and (4.2) as guidelines for our proposed heuristic approaches. We calculate the values of s and S for each day, ignoring the fact that we have a cyclic demand pattern (and denote them s_i and S_i for day i). We now argue intuitively that for any policy of this kind, the values of s_i would remain the same as those computed using (4.1) and (4.2).

Consider the value of s_i, according to the policy, if at the beginning of the day of type i, the inventory is below s_i, we place an order to raise the stock level upto S_i. Therefore, this number can be viewed as the inventory level at which the tradeoff between placing an order (and incurring the resulting costs) and the expected shortage cost that is saved is balanced. Therefore, s_i could be viewed as being affected only by the demand for that day. Hence, the value of s_i would remain unchanged even if the demand pattern for the succeeding day is different from that of this day.

Heuristic I:

We adopted (4.1) and (4.2) for our cyclic demand case. For each day, the mean demand m_i is known. Hence we calculate Q_i, the order amount for day of type i and s_i by:

$$Q_i = \sqrt{\frac{2Km_i}{h}} \tag{4.3}$$

$$e^{-\frac{s_i}{m_i}} = \frac{h}{\lambda}(1 + \frac{Q_i}{m_i}) \tag{4.4}$$

The limitation of the heuristic I is that we have applied (4.1) and (4.2) to the cyclic demand case. This policy would be close to the optimal policy, if the variation in the demand between the periods is not very large, since (4.1) and (4.2) is derived for the i.i.d demand case. The quality of this solution will degenerate when the differences become large.

Consider the value of S_i used in heuristic I. This number was derived from the optimal policy for the i.i.d. demand case and it reflects a certain probability of needing to re-order on each of the days following the day on which one places an order. For instance, if one places an order at the beginning of day of type i, the probability that one will not need to re-order at the beginning of next is equal to the probability that the demand for today is less than $S_i - s_i$; similarly, the probability that one will not need to re-order in two days is equal to the probability that the demand for two days is less than $S_i - s_i$, and so on. In general, we denote: $F_i^{(k)}(S_i - s_i)$, the probability of no re-order after k days of an order on day of type i given i.i.d. demand, where $F_i^{(k)}$ is the k-fold convolution of the demand distribution function of day type i.

44

Denote by $\overline{S}_i$, the modified S_i. We attempt to alter the order upto points S_i in such a way that overcome the shortcoming of heuristic I. An intuitive approach would be to ensure that there is an increase in the values of the order upto points for days that precede days of heavier demands.

Heuristic II:

1. Compute (s_i, S_i) from heuristic I;
2. $k = 1; \overline{S}_i = 0;$
3. Compute $(F_i * ... * F_{i+k-1})(\overline{S}_i - s_{i+k});$
4. for $i = 1$ *to* 7;

if $(F_i * ... * F_{i+k-1})(\overline{S}_i - s_{i+k}) < F_i^{(k)}(S_i - s_i);$

increase $\overline{S}_i;$

go to step 3;

5. Find $\overline{S}_i^*;$

Heuristic II will ensure that the probabilities of stock out on each of the days succeeding the placing of an order is at least as large as the corresponding probabilities in the case of i.i.d. demands. The Motivation is that if one should take into account the future in making our current ordering decisions, we are more likely to reduce the number of orders thus reducing the number of times that the set-up cost is incurred.

The limitation of this heuristic is that it ignores the holding costs. Notice that by ordering larger quantities in anticipation of days with larger demands, we risk increasing the holding cost to a point that the reduction in set-up cost can no longer justify this approach. The other limitation is that heuristic II works well only for a certain range of K, (see Figure 4.1). For larger K, we need to look few more days ahead instead of just consider the next day.

4.3 Computational Results

Since there is no available literature that describes an optimal policy for an inventory problem of this kind, our approach was to compare the two proposed heuristics. We

ran simulations similar to those described in the previous chapter. The value of the penalty cost, λ, is set at the same value that was found to be optimal for the cyclic demand model studied in the previous chapter, i.e., we set $\lambda = 90.3$. The holding cost $h = \beta$ and the ordering cost $c = 0.001$ per unit.

The systems were simulated for a period of 10000 days and the total discounted cost was computed by summing the discounted costs of ordering and holding. Also, for each simulation, we estimate the expected stockout per day by computing the total amount of shortage experienced during the simulation and dividing by the length of the simulation.

In order to compare the policies for the case of no set-up costs and also the validation of these two heuristics, we set $K = 0$. In this case, heuristic I and heuristic II are indentical, since there is no need to modify the S_i. Notice that when $K = 0$, $S_i = s_i$ and we expect that S_i is roughly y_i^* which we get from chapter 3.

Heuristic critical numbers S_i^*	Optimal critical numbers y_i^*
1.5762	1.5300
1.3511	1.3200
1.1259	1.1300
2.0266	2.0400
3.1525	2.8200
2.2518	2.1700
2.0266	1.8900

Clearly the policies are somewhat different but the similarity would suggest that the heuristic policies are reasonably good approximations and can be used as a basis to extend our results to the case of the problem with set-up costs.

In order to compare the two heuristics, we consider a certain range of set-up costs. Because the ordering cost and the holding cost are very small, both less than 1.0, we

46

initially considered the range $0 \leq K \leq 5$ with increments of 0.5. The result are summarized in the figures 4.1 through 4.3.

The overall picture we get from plotting the results for heuristic I and II is that for a certain range of K, heuristic II behaves better than heuristic I in terms of total cost. Hence, we narrowed the range to $0 \leq K \leq 2$ with increment 0.01 to get a better idea of the behavior over the range of interest.
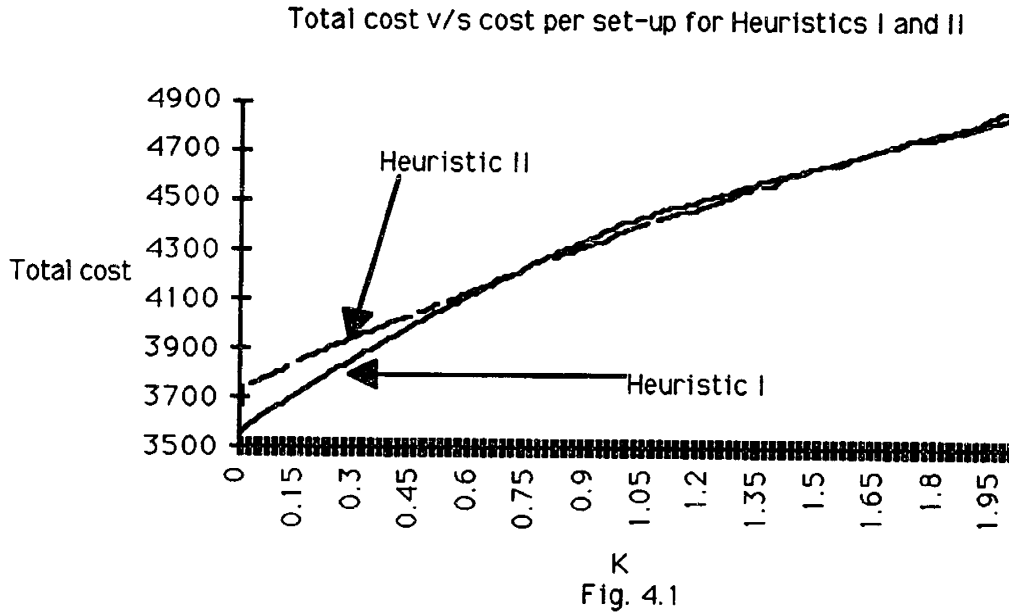
Total cost v/s cost per set-up for Heuristics I and II

Fig. 4.1

Figure 4.1 is the plot of total cost (described above) as the function of the cost per set-up, K. Clearly, for both heuristic I and heuristic II, as the set up cost increases, the total cost increases. Comparing heuristics I and II for the range of $0.75 \leq K \leq 1.5$ we notice that heuristic II results in a lower total cost. An intuitive explanation of this behavior is as follows: When the set-up cost is small in comparison to the other costs, frequent ordering is not very expensive. So any attempt to reduce ordering frequency is not economical due to the larger increase in holding costs. On the other hand, when set-up costs are large the levels to which inventory is raised on re-order are so high that the probability of demand exceeding the difference between S_i and s_{i+1} is so small that there is almost no difference between S_i and $\overline{S}_i$. Hence heuristic II is not expected to make any improvement on the total cost. As a result it is only in the intermediate values of K that heuristic II improves upon heuristic I.
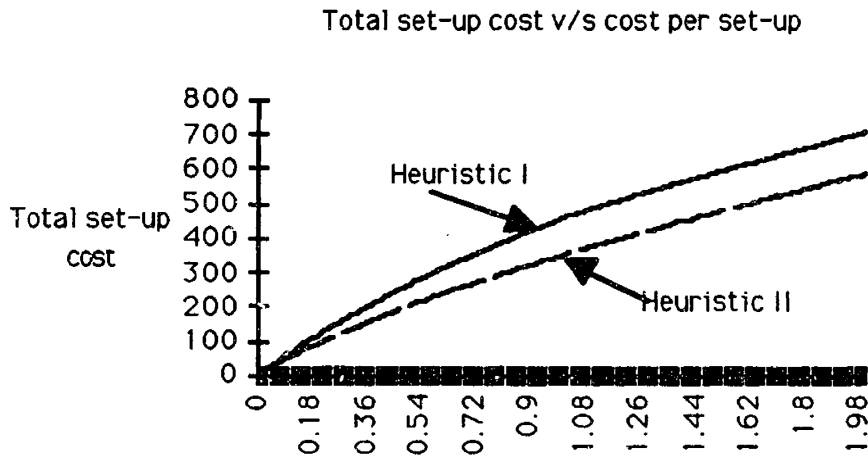
Total set-up cost v/s cost per set-up

Fig. 4.2

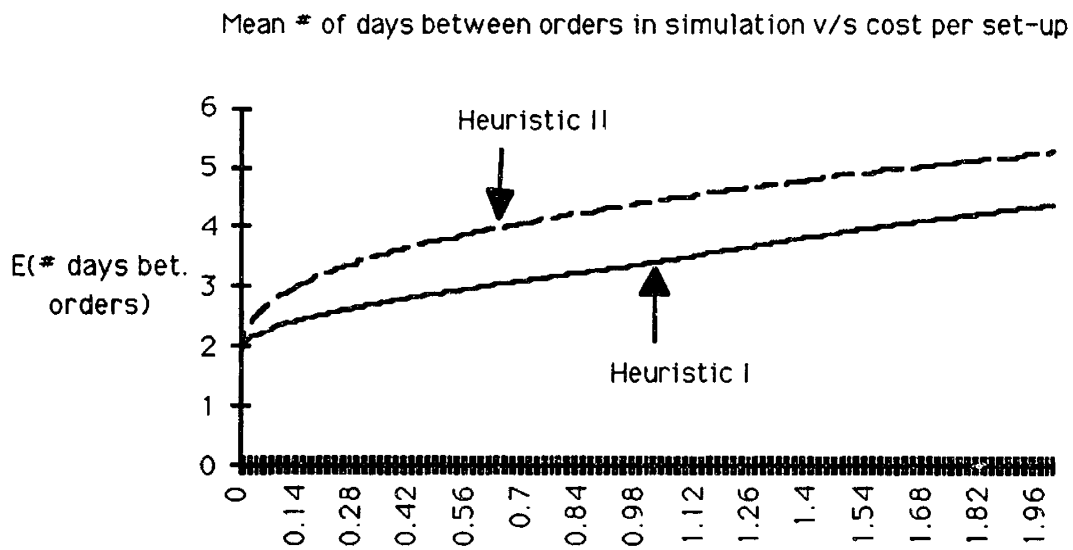Mean # of days between orders in simulation v/s cost per set-up

Fig. 4.3

Figure 4.2 investigates the behavior of the total set-up cost as a function of K. It is very clear that heuristic II always has lower total set-up cost than heuristic I has. The reason is because that by modifying S_i to $\overline{S}_i$, heuristic II's policy tends to order less frequently than heuristic I's policy does.

Figure 4.3 investigates the dependence of the expected number of days between orders on K for both heuristics I and II. We conclude that the policy provided by heuristic II results in less frequent orders than Heuristic I, since the number of expected days between orders is always larger for heuristic II than heuristic I.

Figure 4.4 investigates the tradeoff between total cost and expected stockout.
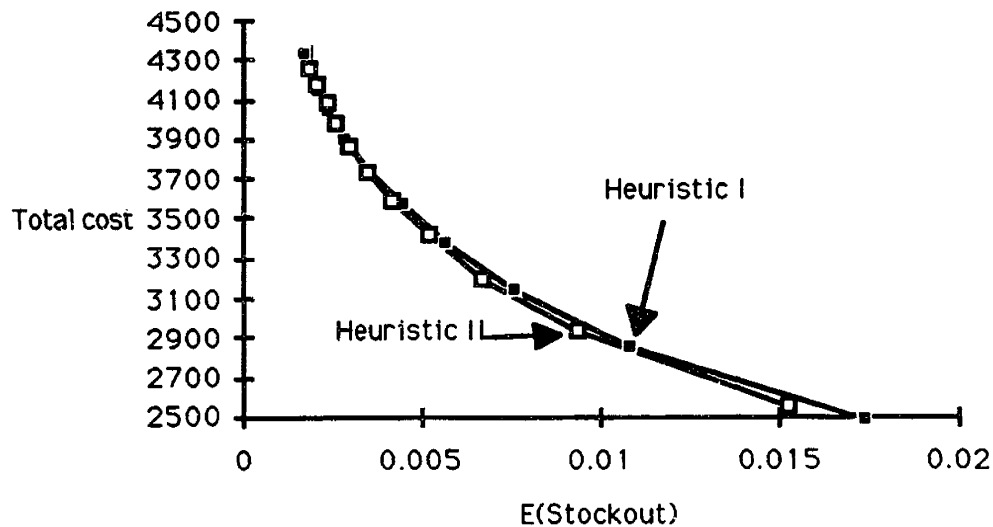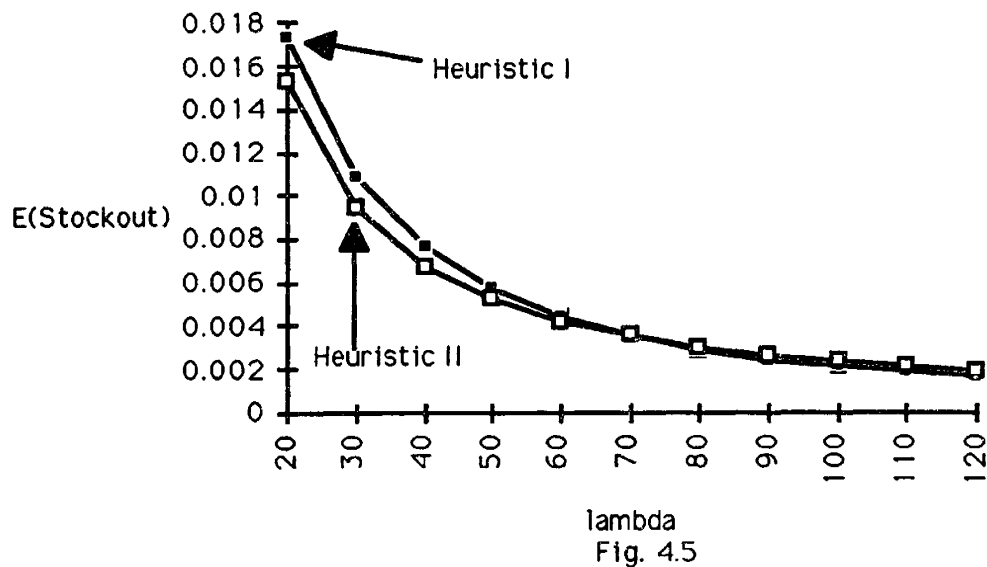
48

E(Stockout)
Fig. 4.4
Estimate of E(Stockout) v/s lambda for heuristics I & II

lambda
Fig. 4.5

Corresponding to each value of λ that is chosen to calculate the parameters of each heuristic policy, there is a total discounted cost and an expected shortage. Choosing an extremely large value of the λ would result in a highly conservative ordering policy, that would result in a relatively small expected shortage but high costs. This graph, which has been plotted with $K = 1.0$, shows the tradeoff between cost and shortages. It appears that the heuristic II dominates the heuristic I.

In figure 4.5 we investigated the dependence of the expected stockout on λ.

4.4   Further Generalizations

Additional features can be included such as:

1. Nonstationary demands with set up cost and delivery times.

2. Existence of exceptional days (periods in which the demand violates the regular pattern).

3. Non-integer lead times. Although our current methods provide good approximate solution we believe that some theoretical work is needed if we include these features in our model.

Bibliography

[1] Iglehart, D. L., and S. Karlin *Optimal Policy for Dynamic Inventory Process with Nonstationary Stochastic Demands.* Studies in Applied Probability and management Science, Stanford, Calif.,University Press, 1962.

[2] Karlin, S. and A. Fabens. *A Stationary Inventory Model with Markovian Demand.* Chapter 11 in Arrow, K. J., S. Karlin, and P. Suppes (eds.), Mathematical Methods in the Social Science, 1959, Stanford Univ. Press, 1960.

[3] Karlin, Samuel and Herbert Scarf *In Studies in the Mathematical Theory of Inventory and Production.* Stanford University Press, 1958.

[4] Morton, Thomas E. *Bounds On the Solution of the Lagged Optimal Inventory Equation With No Demand Backlogging and Proportional Costs.* SIAM Review, Vol. 11, No. 4, Oct. 1969.

[5] Feller, W. *On the Integral Equation of Renewal Theory.* Ann. Math. Stat., 1941, 12, 243-66.

[6] Nahmias, Steven *Simple Approximations for a Variety of Dynamic Leadtime Lost-Sale Inventory Models.* Operations Research, Vol 27, No. 5 ,Sept.

[7] Scarf, Herbert E. *A Survey of Analytic Techniques in Inventory Theory.* Multistage Inventory Models and Techiques,Stanford, Calif. Univ. Press, 1963.

[8] Whitin, T. M. *The Theory of Inventory Management.* 2d ed., Princeton, N.J., Princeton Univ. Press, 1957.