BUFFERING AGAINST DEMAND UNCERTAINTY IN MATERIAL REQUIREMENTS PLANNING SYSTEMS--SYSTEMS WITH NO EMERGENCY SETUPS

Robert C. Carlson
Department of Industrial Engineering
and Engineering Management
Stanford University

Candace Arai Yano
Department of Industrial & Operations Engineering
University of Michigan
Ann Arbor, Michigan

Technical Report 84-13

April 1981 - Revised March 1985

BUFFERING AGAINST DEMAND UNCERTAINTY IN MATERIAL REQUIREMENTS PLANNING SYSTEMS-- SYSTEMS WITH NO EMERGENCY SETUPS

Candace Arai Yano
Department of Industrial and
Operations Engineering
University of Michigan
Ann Arbor, Michigan

Robert C. Carlson
Department of Industrial Engineering
and Engineering Management
Stanford University

Revised March 1985

ABSTRACT

Material Requirements Planning (MRP) originally was designed to aid in production planning given known, deterministic demand for finished product. However, MRP systems often operate under conditions of stochastic demand, and safety stock is used to maintain customer service levels. An algorithm is developed to determine cost-effective safety stock levels for a single item with a two-level product structure. Emergency production setups are not permitted. Application of the algorithm over a wide range of parameter values indicates that in most instances, only finished product safety stock should be held. These results are confirmed with simulation studies. The general results also may be extended to other single-item arborescent assembly structures.

BUFFERING AGAINST DEMAND UNCERTAINTY IN MATERIAL REQUIREMENTS PLANNING SYSTEMS--SYSTEMS WITH NO EMERGENCY SETUPS

1. INTRODUCTION

Material Requirement Planning (MRP) was designed to operate with the assumption of known demand for the finished product, but MRP systems often operate in conditions of demand uncertainty. This not only creates uncertainty about the quantity required, but may also create uncertainty about the timing of a requirement. The latter may arise because the lot-sizing algorithm employed shifts the timing of a planned production run, or because an emergency production setup is required to avert shortages. A buffering mechanism is required if acceptable customer service levels and stable production schedules are to be maintained.

Several researchers have proposed operational policies to buffer against demand uncertainty, while others report results from simulation studies. Researchers who have proposed operational policies include Eichert (1974), Meal (1979), Miller (1979), Moore (1973), New (1975), Orlicky (1973), and Welch (1973). Eichert suggests that standard statistical inventory methods be applied to unplanned demand only. Meal proposes a method for measuring forecast errors for each item in the MRP system and suggests that these measurements be incorporated into standard statistical inventory models. Miller advocates a technique which takes into account decreasing forecast error variance as the period in question approaches, thereby maintaining a larger proportion of safety stock at points in the product structure where inventory is less expensive. Moore states that safety stock should be determined for finished goods based on the cumulative leadtime.

New points out that a practical rule of thumb is to set safety stock equal to the "maximum demand likely to occur in a single period" (p.8). Orlicky states that "safety stock is properly applied only to inventory items subject to independent demand" (p. 79), but also indicates that one exception to this rule is safety stock to compensate for uncertainty of supply. Welch advocates the use of an order point equal to the maximum planned demand during a "planning period."

Simulation studies in the literature are numerous. One of these studies was reported by Whybark and Williams (1976). Their study indicates that safety stock is better than safety time for buffering against demand uncertainty, while safety time is better than safety stock for buffering against timing uncertainty.

Lambrecht, Muckstadt, and Luyten (1984) show that the form of the optimal policy for a multi-stage periodic review production system with stochastic demand, setup costs, and shortage costs (per unit per period) only for the final product, is of the (s,S) type. They also propose a heuristic solution procedure which is shown to be effective and efficient.

Nahmias and Schmidt (1983) characterize the form of the optimal policy for a two-stage assembly system with no setup costs and a shortage cost per unit per period for the finished product. They show that the optimal policies are of the "order-up-to" type insofar as permitted by component availability.

Many production operations prefer (and use) a cyclic schedule when possible (see, for example, Caie and Maxwell (1981)) because of ease of implementation, because it simplifies capacity allocation, and because an (s,S) policy cannot guarantee a cyclic schedule. Furthermore, shortage costs that are required for an (s,S) type of policy are difficult to ascertain in real applications, while desired fill-rates are somewhat easier to ascertain. We therefore chose to address the problem of determining placement of safety stock to achieve a desired fill-rate at minimum cost when cyclic schedules are used. The reader is referred to Carlson and Yano (1984) for analyses of a system in which both timing and quantity of batches are allowed to vary over time.

Our objective in this paper is to develop a fundamental understanding of the economics related to "location" and quantity of safety stock and the impact of mating of parts in assembly for relatively high volume products. Because this work represents the first analytically-based research on safety stocks for assembly systems using a fill-rate criterion, we cannot hope to solve the problem for the most general case. Therefore, we examine a simple case first to develop some fundamental insights into how more general problems might be approached.

In section 2, we describe the problem. Section 3 outlines our approach to the problem, and section 4 then details the development of the algorithm. Section 5 discusses the implementation of the algorithm and experimental results. Extensions of the results and a summary are presented in Section 6.

2. DESCRIPTION OF THE PROBLEM

We address the problem of determining cost-effective safety stock levels for each component in the product structure of a single finished product under conditions of stochastic demand. We assume that the item is a homogeneous, medium to high volume product produced to inventory (i.e., not produced to order), and that there is complete backordering. We examine a simple assembly structure in which two components are purchased or produced and assembled into the finished product. We assume that these components are not common to any other finished products. Less restrictive assumptions would make the analyses much more complex and is beyond the scope of this initial investigation.

We assume that there are no production or storage capacity constraints and that no other uncertainty exists in the system. Production leadtimes, yields, and both timing and quantity of supply are assumed to be deterministic. While these are strong assumptions, our intent is to focus on the effects of demand uncertainty, which is a critical problem in make-to-stock situations.

Demand is assumed to be stationary, which is a reasonable first-cut approximation to reality in many make-to-stock environments. To be consistent with this assumption and to incorporate unbiased forecasts, the forecast of demand for the finished product is equal to mean demand. In this paper, we assume that forecast errors are normally distributed with mean zero. However, our algorithm may be adapted to any infinitely divisible distribution of forecast errors. We assume that production leadtimes are positive since safety stock is unnecessary when leadtimes are zero.

To incorporate the dynamic nature of MRP systems, we use rolling schedules (Baker (1977)). A rolling production schedule is a plan which is composed of the first period plan of each of a series of finite horizon plans. Operationally, one would establish a plan based on known demand requirements or forecasts for a finite horizon and implement only the production plan for the first period in that horizon. One would then observe actual demand in the current period, add a new requirement or forecast to the end of the previous horizon, update other requirements/forecasts as necessary, devise a new finite horizon production plan, and implement the first period of that plan. The process continues in this manner.

Although we are using rolling schedules, which by their nature are dynamic, we examine a situation in which the timing of production setups is fixed far in advance of the setup. Therefore, no "emergency" setups or expediting may occur. However, the planned production quantity may change, and is not fixed until the production run is begun.

The objective is to determine fixed (time independent) safety stock quantities (and hence order-up-to levels) which minimize total setup and inventory holding costs subject to a service level constraint, where only the service level of the finished product (end-item) is considered critical. Service level is measured in terms of percent of demand filled immediately from stock, often referred to as "fill-rate," which is a commonly used criterion (in industry) for medium to high volume products.

We assume that the desired fill-rate is sufficiently high that shortages of the finished product and components occur only in the last period with positive demand prior to an order arrival. In most realistic situations and at most commonly used target service levels, this is a reasonable assumption. It would be possible to incorporate less restrictive assumptions in portions of the analyses. However, in other portions (as will be apparent later), the complexity of the analysis prohibits this. We will discuss this point in more detail in Section 4.

In the simulation studies which will be discussed later, we use the Wagner-Whitin (1958) dynamic lot-sizing algorithm independently for each item in the product structure, beginning with the finished product and continuing toward the raw materials. The use of this particular lot-sizing algorithm is not critical to the development of our algorithm. What is important is that the Wagner-Whitin algorithm produces regular, cyclic production schedules when demand (or the demand forecast) is relatively constant from period to period. The setup cost can be set to obtain any desired cycle length. Given a cyclic schedule, the form of the optimal replenishment policy is of the (R,T) type (order up to R every T periods) when demand is stationary. Therefore, the results presented here are applicable to all such situations, although the exact lot-sizing policy used may differ. For instance, the Silver-Meal (1973) or Period Order Quantity (POQ) lot sizing heuristics with appropriate provision for safety stock can be used equally well in this framework.

The system operates as follows. The Wagner-Whitin algorithm is used to determine the "optimal" time between production runs. Since the quantity to be produced is based on forecasts (since demand is stochastic), one produces up to a target inventory level if requisite components are available. Otherwise one produces as much as available components permit. The order-up-to quantity is the sum of mean demand during a cycle plus leadtime, and safety stock. Additional details of the simulation model can be found in Appendix A.

3. APPROACH TO THE PROBLEM

The problem can be formulated as in an intractable non-linear integer optimization problem with stochastic elements. This formulation appears in Appendix B. While it is possible to represent some of the terms in mathematical notation, a few cannot be represented in closed form in terms of known parameters. In particular, some of the terms related to expected holding costs and expected backorders cannot be represented in closed form.

Hence, we cannot even <u>evaluate</u> the objective function or the constraint, much less attempt to optimize using traditional approaches. We therefore developed an alternative approach to modeling the problem mathematically in an attempt to find a simple method for determining approximate solutions.

We began our study by simulating MRP systems for two- and three-level product structures, varying the following factors: (1) holding cost rates for components, (2) natural cycle lengths of the end-item and components, (3) leadtimes, (4) variability of demand, and (5) safety stock levels. The purpose of these studies was to gain insight into the impact of these factors on system performance measured in terms of total cost and service level. The term "natural cycle" refers to the average number of periods between adjacent production setups.

The primary finding of consequence here is that the most important tradeoff in determining cost-effective safety stock levels is the tradeoff between end-item and second-level component safety stock. Predecessors of the finished product are referred to as second-level components. Simulation studies showed that in most situations, component safety stock is not cost-effective, as illustrated in Figure 1. Under special conditions, however, some positive levels of component safety stock are cost-effective. A typical cost versus service level diagram for such conditions is depicted in Figure 2. By examining the characteristics of these situations, we determined that the elements which are most important in this tradeoff are: (1) holding cost of the component relative to that of the end-item, (2) proportion of the value of the end-item added at the last assembly/manufacturing stage, (3) frequency of setups of the second-level component relative to that of the end-item, and (4) availability of "partner" components with which a particular component must be mated.

<u>FIGURES 1 AND 2</u>

4. DEVELOPMENT OF THE ALGORITHM

We desire a measure of the relative cost-effectiveness of safety stock of the end-item and of second-level components which is possible to model mathematically. Insight into the principal factors in the tradeoff discussed in the last section led to the development of such a measure. Since the timing of setups is fixed (with frequency approximately equal to the natural cycle), changing safety stock quantities can affect inventory holding costs and fill-rate but not setup costs. Therefore, a marginal analysis approach using inventory costs alone is possible if one accepts the premise of decreasing marginal returns as safety stock increases.

An intuitively appealing approach would be to consider the fill-rate changes and costs resulting from one additional unit of safety stock for each item in the product structure. This approach, unfortunately, has one of the same drawbacks as the original formulation: the fill-rate cannot be determined (and likewise changes in the fill rate).

To circumvent this problem, we would like a measure which indicates the cost, on average, resulting from adding enough safety stock of an item, say i, to avert one additional (end-item) shortage during an end-item natural cycle. Thus, we would have a well-defined unit of change in the fill rate (i.e., 1 unit divided by average demand during an end-item natural cycle), and the associated costs for the various safety stock alternatives. The problem of determining how many units to add still remains. Furthermore, determining their associated costs is difficult because each additional unit results in different (increasing) marginal costs.

An alternative approach, which is the one we chose, is based on an approximation to this approach. We can determine expected costs resulting from an additional unit of safety stock for a given item. We can also determine the probability that the additional unit of safety stock will be instrumental in reducing end-item shortages during the course of an end-item cycle (i.e., a shortage of at least one unit would have occurred if the additional unit were not available). The ratio of the expected cost and the probability is a value which approximates, on the margin, "the expected cost of averting an incremental shortage per end-item cycle." We refer to this relative cost measure as C_i, where the index i refers to the item for which safety stock is added. In general terms, this relative cost can be expressed as

$$\left(\begin{array}{l}\text{expected holding cost per}\\ \text{cycle from an incremental}\\ \text{unit of safety stock}\end{array}\right) \Bigg/ \left(\begin{array}{l}\text{probability that the}\\ \text{incremental unit averts}\\ \text{an end-item shortage}\\ \text{during a typical end-}\\ \text{item cycle}\end{array}\right) \tag{1}$$

The algorithm developed in this paper uses these relative cost measures in conjunction with an approximation to the shortage cost implied by the fill-rate to achieve (approximately) the desired fill-rate.

In section 4.1 we develop this relative cost measure for end-item safety stock. Section 4.2 discusses a framework for modeling the relative cost measure for component safety stock.

Throughout the paper we use the following notation:

T_i = natural cycle of item i

h_i = holding cost of item i

L_i = leadtime for item i

k_i = safety stock multiplier for item i

D = mean end-item demand during a period

σ = standard deviation of demand during one period

$\Phi(\cdot)$ = cumulative standard normal distribution

$\phi(\cdot)$ = standard normal density

We will represent safety stock quantities as a multiple of the standard deviation of demand during an order cycle plus leadtime. This convention is not critical to the analysis; it simply permits us to let k_i represent the usual safety stock multiplier. If the safety stock must buffer against demand variability over a shorter or longer time horizon, this fact is incorporated explicitly into the analyses.

4.1 Expected Cost of Averting an Incremental Shortage Per End-Item Cycle Using End-Item Safety Stock

We first must determine the expected holding cost term in the numerator of the relative cost given by (1). It can be expressed as the holding cost per unit time, h_1 multiplied by the expected period of time the unit will be held.

Recall that we made an assumption in section 2 that shortages occur only in the last period with positive demand prior to an order arrival. Since demand for the finished product is positive in each period, shortages are assumed to occur only in the period immediately prior to an order arrival. Observe that the probability that a shortage occurs two periods prior to an order arrival is

$$1- \Phi[(D + k_i\sqrt{T_i + L_i}\,\sigma)/(\sqrt{T_i - 1 + L_i}\,\sigma)]$$

which is negligible in most instances. Therefore, an incremental unit of safety stock will be held for T_1 periods (the entire cycle) if it is not needed to satisfy an end-item demand, and $T_1 - 1$ periods if it is needed. Thus, that incremental unit will be held for T_1 periods with probability

$$\Phi(k_1)$$

and for $T_1 - 1$ period with probability

$$1 - \Phi(k_1)$$

The expected length of time it will be held is

$$\bar{T}_1 = T_1 \Phi(k_1) + (T_1 - 1) [1 - \Phi(k_1)]$$

We next determine the probability term in the denominator. An incremental unit of end-item safety stock will avert an end-item shortage only if demand exceeds current cycle stock plus safety stock. This is equivalent to the event that demand during the leadtime plus natural cycle is greater than k_1 standard deviations from its mean. This event occurs with probability

$$1 - \Phi(k_1)$$

Notice that since the probability is applicable to each end-item cycle, it is applicable to the "typical" end-item cycle as well.

The expected cost of averting an incremental end-item shortage using end-item safety stock is thus

$$C_1 = \frac{h_1 \bar{T}_1}{(1 - \Phi(k_1))}$$

4.2 Expected Cost of Averting an Incremental End-Item Shortage Using Safety Stock of Second-Level Components

An incremental unit of safety stock for a second level component will avert an end-item shortage only if

(a) the component is susceptible to a shortage,

(b) that incremental unit of safety stock averts a shortage of that component (i.e., level 1 is able to obtain more than would have been available otherwise), given that the component is susceptible to a shortage,

(c) "partners" for that second level component are available, and

(d) an end-item shortage is averted as a result.

Let us first discuss (a). If a component is produced in a lot size which is approximately a multiple n (greater than 1) of the end-item lot size, then for the first n - 1 withdrawals of components, shortages are extremely unlikely, particularly if desired fill-rates are high. The actual multiple, n, in this problem is T_i/T_1. In only the last of the T_i/T_1 withdrawals is a shortage of components likely to occur. Therefore, the fraction of the time that a withdrawal is susceptible to shortage is T_1/T_i for second level component i. If an item is susceptible to a shortage, a shortage situation may or may not occur depending upon actual demand. However, we assume that shortages <u>do not</u> occur if an item is not susceptible to shortage. In the event that $T_1 = T_i$, the possibility that there is insufficient stock of component i always exists (i.e., it occurs a fraction of the time equal to $T_1/T_i = 1$).

Given that item i is susceptible to a shortage, we can diagram the possible sequence of events in an event tree as illustrated in Figure 3.

FIGURE 3

This is not an event tree in a decision analytic sense; however, it is useful for exposition purposes. Therefore, we will continue to use this construct throughout the paper. Also indicated for each path are the resultant expected holding costs which are developed in Section 4.2.2.

4.2.1 Calculation of Path Probabilities

The analysis of this system would be simple if the probability associated with each of the branches were independent of predecessor and successor branches. Unfortunately, this is not the case.

Since we have an "order-up-to" type of system, each order quantity is equal to the demand for that item since the last such order was placed. The demand for second-level components is generated by level 1 orders for them, which, in turn, are equivalent to end-item demand over some period of time. Therefore, the occurrence of each of the events in the event tree depends upon end-item demand in some time frame. Consider a particular (but typical) item i (component) lot for which an order is placed in some period t. That lot becomes available in period $t + L_i$ and is utilized to satisfy assembly requirements in periods $t + L_i$, $t + L_i + 1, \ldots, t + L_i + T_i - T_1$. Notice that the <u>next</u> item i lot will satisfy assembly requirements starting in period $t + L_i + T_i$. Since assembly is done only every T_1 periods, so there are no component withdrawals between $t + L_i + T_i - T_1$ and $t + L_i + T_i$. Since an "order-up-to" system is used, the assembly requirement in period $t + L_i + T_i - T_1$ depends upon end-item demand through period $t + L_i + T_i - T_1 - 1$. The item i lot in question was ordered in period t but must buffer against demand fluctuations through period $t + L_i + T_i - T_1 - 1$. Therefore, whether or not an additional unit of safety stock averts a shortage of item i depends upon demand in periods

$$t, \ldots, t + L_i + T_i - T_1 - 1$$

Therefore, if the sum of demands during these periods exceeds mean demand during a cycle plus the previous level of safety stock, an additional unit of safety stock will avert an item i shortage. Otherwise, the extra unit remains as component inventory.

Let us now examine the availability of partner components. For each possible pair of second level components there will be simultaneous production runs or deliveries at intervals equal to the least common multiple of the natural cycle lengths of that pair. For instance, if $T_i = 4$ and $T_j = 6$, the production runs for items i and j will be completed simultaneously every 12 periods. In most realisitic situations, shortages occur only in the last period with positive demand prior to a production run, just as shortages generally occur only in the last period prior to a production run of the end-item. Therefore, an incremental unit of safety stock of a particular second-level component will almost certainly have partners available the fraction of the time that the partner's order arrivals do <u>not</u> coincide with that of the component in question. The remainder of the time, the partner may be lacking.

Let item j be the partner component in a situation with two components on the second level. Items i and j will have simultaneous order arrivals at intervals equal to the least common multiple of the two natural cycles. Let T_{LCM} = least common multiple of T_i and T_j. Then there is a potential item j shortage situation a fraction T_i/T_{LCM} of the time, and item j will be assumed to be available a fraction of the time equal to $(1 - T_i/T_{LCM})$.

Using this knowledge, we perform an analysis of the dynamics of the system similar to that done before. We note that the event that item j is available, given that the next order arrives simultaneously with an item i order, depends upon demand realizations in periods

$$t + T_i + L_i - T_j - L_j, \ldots, t + L_i + T_i - T_l - 1$$

In theory, it is possible to determine the probability that all partners of a particular component are available. However, the complexity of the analysis limits us to the case of two second-level components here.

We turn now to the final branches in the tree. An additional unit of item i safety stock (eventually) averts an end-item shortage only if end-item demand exceeds available stock from all sources. Available end-item stock is constrained by available component stock at the time of production plus end-item safety stock. Recall that an end-item shortage may occur even when there is sufficient second-level stock if the end-item safety stock level is not large enough to buffer against the demand variability which occurs. However, we are concerned here with those shortages which are a direct result of shortages of second-level stock.

An analysis of the system dynamics provides the result that the event that an end-item shortage occurs because of an item i shortage depends upon demand in periods

$$t + L_i, \ldots, t + L_i + T_i + L_1 - 1$$

In short, we are concerned with the following:

(1) Fraction of item i order arrivals which coincide with item j order arrivals

$$\frac{T_i}{T_{LCM}}$$

(2) The event that an incremental unit of item i averts an item i shortage, which depends upon demand in periods

(*) $\quad t, \ldots, t + L_i + T_i - T_1 - 1$

(3) The event that item j is available, given the next order arrives simultaneously with an item i order, which depends on demand in periods

(***) $\quad t + T_i + L_i - T_j - L_j, \ldots, t + L_i + T_i + T_1 - 1$

(4) The event that an item 1 (end-item) shortage occurs because of an item i shortage, which depends upon demand in periods

(**) $\quad t + L_i, \ldots, t + L_i + T_i + L_1 - 1$

The event that an incremental unit of item i averts an item i shortage will occur if demand in the periods indicated in (2) above exceeds

$$\text{mean demand} + k_i\sqrt{T_i + L_i}\,\sigma$$

The event that item j is available given the next order arrives simultaneously with an item i order occurs if demand in the periods indicated in (3) above is less than

$$\text{mean demand} + k_j\sqrt{T_j + L_j}\,\sigma$$

The event of an item 1 shortage due to an item i shortage will occur if demand in the periods indicated in (4) above exceeds

$$\text{mean demand} + k_1\sqrt{T_1 + L_1}\,\sigma + \min\ \{k_i\sqrt{T_i + L_i}\,\sigma\ ,\ k_j\sqrt{T_j + L_j}\,\sigma\}$$

if the next order arrivals of items i and j are simultaneous, or

$$\text{mean demand} + k_1\sqrt{T_1 + L_1}\,\sigma + k_i\sqrt{T_i + L_i}\,\sigma$$

if they are not.

Our algorithm incorporates the demand characteristics in each of these (usually overlapping) time spans directly as discussed below.

The duration spanned by the intervals in (*), (**), and (***) can be broken into independent subintervals to simplify the computation of the required joint probabilities. Since demand in each subinterval is independent of demand in other subintervals, each of the joint probabilities can be expressed as a multiple integral in which only the limits of the integrals are related. However, even in their simplest forms, the joint probabilities can be expressed only as indeterminate integrals. This difficulty arises, in part, because of our use of normally distributed forecast errors. But the problem would exist for any commonly used, continuous, infinitely divisible distribution of forecast errors. Probability mass functions provide little help in solving this difficulty. In order to circumvent this problem, we use discrete approximations for these indeterminate forms, as they can be computed quickly by a computer.

Notice that when $T_i = T_j$ and $L_i = L_j$, it is essential that the safety stocks of item i and j be equal. There are other situations such as when $T_i = T_j$ with L_i and L_j unequal, in which an even more detailed coordination is required. In such situations, purchases of the component with the shorter procurement leadtime should be directly tied to the <u>known</u> future availability of the component with the longer leadtime. (See Nahmias and Schmidt (1983) for further discussion). Observe, however, that the vast majority of MRP systems do <u>not</u> have the sophisticated intra-level coordination required to accomplish this, and only fixed order-up-to points can be used. In all other cases, the coordination required between component safety stock levels is taken into account explicitly through the time spans indicated in (*), (**), and (***).

It is important to note here that to model the general case with arbitrary leadtimes and arbitrary but integer lot timing multiples, twenty-eight (28) different relationships arise among the subintervals mentioned above. In other words, the correct procedure to use to calculate the joint probabilities depends upon the leadtime and natural cycle values, and it must be chosen from among 28 such procedures. Because of the complexity associated with this aspect of modeling the probabilities, and because it is generally not possible to implement detailed intra-level coordination within existing MRP systems, we do not address detailed coordination in the few situations in which it may be desirable.

4.2.2 Calculation of Expected Holding Cost Terms

The expected cost of holding an incremental unit of item i safety stock is determined as follows. The incremental unit of item i will be held for T_i periods at a cost of h_i per period if it is not assembled into a unit of item 1. There are two ways in which this may occur. First, this unit of item i may not be required for the final assembly stage because demand realizations are less than the forecast or because end-item safety stock is expected to be sufficient to cover deviations of demand from the forecast. Second, there may not be a unit of item j available with which to mate it. These are paths 1 and 4, respectively, in Figure 3.

If the incremental unit of item i is assembled into a unit of item 1, it will be held at a cost of h_i per period for $T_i - T_1$ periods. This is the duration between the item i order arrival and the last level 1 order (for components to be assembled) which is drawn from that item i lot. After it is assembled, the holding cost associated with the value-added in the assembly stage must be incurred as well. The holding cost associated with the value-added is

$$h_1 - h_i - h_j$$

Therefore, the holding cost for the incremental unit of item i plus the value-added is

$$h_i + (h_1 - h_i - h_j) = h_1 - h_j$$

Notice that the unit of item j already existed and any holding cost associated with it is a sunk cost. At this point this unit of what is now item 1 behaves as a unit of item 1 safety stock. We must incur the holding cost $h_1 - h_j$ for T_1 periods if the unit does not avert a shortage (i.e., is not used to satisfy a customer demand) and for $T_1 - 1$ periods if it is used to avert a shortage.

We can now identify the costs associated with each path in the tree in Figure 3. The expected holding cost associated with an incremental unit of item i safety stock is:

$$\bar{h}_i(\underset{\sim}{k}) = h_i T_i (p_1 + p_4) + h_i(T_i - T_1)(p_2 + p_3) + (h_1 - h_j)(T_1 - 1)p_2$$
$$+ (h_1 - h_j)T_1 p_3$$

where p_ℓ is the probability associated with the ℓth path in the tree. The value of h_i depends upon $\underset{\sim}{k}$, the safety stock vector, because the values of p_ℓ are a function of $\underset{\sim}{k}$. Using this information and the facts that

(1) item i must be susceptible to a shortage in order for an incremental unit of safety stock to have an impact on customer service, and

(2) given that item i is susceptible to a shortage, an incremental unit of item i safety stock will affect customer service with the probability associated with path 2 in the tree,

we can represent our relative cost measure, C_i, as

$$C_i = \frac{\bar{h}_i(\underset{\sim}{k})}{(T_1/T_i)\ p_2} \qquad \text{i on second level}$$

in the general case and

$$C_i = \frac{\bar{h}_i(\underset{\sim}{k})}{p_2} \qquad \text{i on second level}$$

in the special case where all second level components have natural cycles equal to that of the end-item.

5. ALGORITHM

We now have all the components necessary to create an algorithm. Recall that the policy will be to add the unit of safety stock with the lowest value of C_i. However, because we cannot compute the fill rate as a function of the safety stock multipliers k_i, we cannot simply add the "cheapest" safety stock until we achieve the desired service level.

When only end-item safety stock is held, there is a strong linear relationship between fill-rates actually achieved and fill-rates computed as if the system had a single stage with a leadtime equal to the cumulative leadtime of the true system. (The reader is referred to Yano (1981) for details). Therefore, it is possible to determine the requisite value for k_1 to achieve any desired fill-rate. The value of k_1, in turn, implies a shortage cost per unit which is approximated by C_1. Given a value of C_1, we can use the formulas for C_i developed earlier to determine an approximately optimal quantities of component safety stock. The "most economical" type of component safety stock is added until the cost of averting an incremental shortage, C_i, equals the "shortage cost per unit," C_1.

ALGORITHM

Step 1. Set k_1 so that the achieved service level will be equal to desired service level.

Step 2. Set $k_i = 0$ for i on the second level.

Step 3. Choose a step size, Δk, which is the incremental step size for the second-level safety stock.

Step 4. Calculate

$$C_1 = \frac{h_1 \bar{T}_1}{1 - \Phi(k_1)}$$

Step 5. Calculate

$$C_i = \frac{\bar{h}_i(\underset{\sim}{k})}{(T_1/T_i)\ p_2} \quad \text{for all i on the second level}$$

Step 6. Find $z = \min_i C_i$, i on second level

If $z > C_1$, stop.

Otherwise:

Set $k_i = k_i + \Delta k$, where i = minimand C_i.

Return to Step 5.

As an alternative, one can begin with $\underset{\sim}{k} = (0,0,0)$ and proceed with a greedy approach in which an increment of the most "cost-effective" safety stock is added at each iteration. In order to ensure that the fill-rate achieved is close to the target value, it is necessary to choose a threshold shortage cost using the formula for C_1 in conjunction with an appropriate value of k_1.

We also note that if $T_i = T_j$ <u>and</u> $L_i = L_j$, we can collapse the product structure into a serial system by using an "aggregate" component with holding cost equal to $h_i + h_j$. This forces equal safety stock quantities for the two components.

5.1 Results

Results from the algorithm for a large number of problems indicate that for nearly all realistic situations, the algorithm determines that safety stock level for second-level components should be zero. Simulation results for a large number of problems confirm this fact. Initially we designed a complete factorial experiment. The parameter values for these experiments are shown in Table 1. We ran the algorithm for most of these problems using various values of k_1 from 0 to 1.2, which would be expected to yield fill-rates over 95% in most cases and over 90% in almost all others. In nearly all of these cases, these algorithm provided solutions with no component safety stock. Furthermore, when the algorithm was initialized at $\underset{\sim}{k} = (0,0,0)$ and was permitted to increment k_1 as well, finished product safety stock remained the preferred buffering technique.

Two examples are selected for illustration. Figures 4 and 5 show average cost versus average service level relationships as end-item safety stock is increased and as second-level safety stock (for one component, or both, as applicable) is increased. These figures are for two situations in which the algorithm indicates that no second-level safety stock should be used. The plotted values represent averages obtained from 50 simulation runs, each with a 24 period horizon. The same set of stochastic conditions (same problem set) was used for each set of safety stock multipliers, so as to reduce the variance of the differences among the results. The results are not amenable to statistical analyses because of the bivariate nature of the data, but with the large number of simulation runs and variance reduction techniques, we can place a reasonble degree of confidence in the results.

Note that Figure 4 represents a case in which parameter values were chosen so as to afford advantage to second-level safety stock relative to end-item safety stock. First, $T_i = T_1$, so that the T_1/T_i factor in the denominator of C_i is at a maximum. Second, the second-level components have very low holding costs so that the h_i term is low. Finally, the joint probability term is large relative to other situations because the partner component is always available.

The case illustrated in Figure 5 is more typical. Only 20% of the value of the product is added at the last stage (versus 80% in Figure 4). In addition $T_2 = T_3 > T_1$, so that the T_1/T_i factor is not at a maximum. In this case, second-level safety stock is extremely costly and provides little increase in the service level.

Through the C_i values, the algorithm accurately predicts how second-level safety stock will perform relative to end-item safety stock. For instance, the values of C_i at $\underset{\sim}{k} = (0,0,0)$ for the case illustrated in Figure 4 are

$$C_1 = 3.00; \; C_2 + C_3 = 3.40$$

while the values of C for the case illustrated in Figure 5 are

$$C_1 = 3.00; \; C_2 + C_3 = 288.$$

Recall that items 2 and 3 have equal natural cycle lengths and leadtimes so that safety stock must be added simultaneously for both items. The large value of $C_2 + C_3$ for the second case shows that it is extremely expensive to utilize second-level safety stock in order to increase the service level. The fact that the value of $C_2 + C_3$ is only slightly larger than C_1 for the first case indicates that second-level safety stock is only slighty more costly than end-item safety stock to attain small increases in the service level.

Since it is preferable to carry no second-level safety stock in nearly all situations, the algorithm does not need to be modified from its present form for the purpose of adjusting k_1 in most cases. Therefore, the value of k_1 can be set so as to achieve the desired service level.

Additional results from the algorithm and from simulation studies indicate that several factors must be present simultaneously in order for some second-level safety stock to be cost-effective. They are:

(1) The holding cost of the component must be very small relative to that of the end-item (i.e., 10% or less).

(2) The partner components must have high availability brought about by long natural cycles.

(3) The desired service level must be extremely high (i.e., 98% or more) leading to a large value for C_1.

If these conditions exist, then the savings from using second-level safety stock depends upon the ratio of the holding cost associated with the value-added in assembly (echelon holding cost for finished product) to the holding cost of the end-item. As this ratio decreases, the benefit from using component safety stock increases. These findings provide strong confirmation of the logic on which the algorithm is based. The reader is referred to Yano (1981) for additional details.

After observing these results, we decided to move parameters systematically in directions that would appear to favor component safety stock. In this way, we were able to determine (approximately) the conditions necessary for positive component safety stock to be cost-effective. An example of such a situation is:

$$\underset{\sim}{T} = (4,4,12)$$

$$\underset{\sim}{h} = (1.0,\ 0.1,\ 0.1)$$

$$\underset{\sim}{L} = (1,1,1)$$

The algorithm gives a solution of $k_2^* = 0.25$ and $k_3^* = 0$ for an initial value of $k_1 = 0$. Simulation results representing the average of 50 problems, each with a 24 period horizon are shown in Figure 6. The same stochastic conditions are used for each of the safety stock levels indicated in the figure. It appears that our algorithm has determined (approximately) the point at which it is no longer economical to add component safety stock (i.e., additional finished product safety stock should be added before component safety stock is increased).

We have done additional studies of situations in which the desired fill-rate is high (greater than 98%), necessitating fairly large quantities of finished product safety stock. In these situations, the cost-effectiveness from additional finished product safety stock declines rapidly as the quantity is increased. Suppose that one chooses an initial value of k_1 and uses the algorithm to determine k_2 and k_3. It appears at first that the algorithm severely understates the "optimal" values of component safety stock as determined by a grid search. An example of such a situation is illustrated in Figure 7. Observe, however, that in order to increase the fill-rate further one should then add a small increment of finished product of safety stock and recompute C_1. We would expect the recomputed value of C_1 to be significantly higher than the previous value (as is apparent in the declining slope of the cost - fill-rate curve). It is extremely likely that considerably more component safety stock would now be cost-effective and should be added if additional increases in the fill-rate are desired. The only unsolved problem is one of determining the fill-rate obtained from a given safety stock vector in which component safety stock is positive. This remains a problem for further research.

6. EXTENSIONS AND SUMMARY

The results obtained here may be applicable to other product structures by simple induction. If there are more than two components on the second level, the likelihood that all partner components are available decreases. Therefore, the expected cost of averting an incremental shortage using safety stock of any particular component increases. This in turn causes a decrease in the optimal safety stock quantity and the potential savings to be gained from using component stock.

An evaluation of the relative cost measure for a component on the third level can be done in the same way as we have done here for two levels, although we do not claim that it is easy to do so. However, intuition would indicate that as one moves deeper into an arborescent product structure, a much larger number of events must occur simultaneously in order for component safety stock to have a beneficial effect on the end-item service level. The joint probability that all the advantageous events occur simultaneously decreases approximately geometrically. The expected holding costs arising from an incremental unit of safety stock tends to decline at a slower rate, since inevitably some of the additional third-level safety stock will be incorporated not only into units of the (second-level) parent item but also into units of the end-item which would not have existed otherwise. Therefore, the expected cost of averting an incremental end-item shortage tends to increase as one moves deeper into the product structure.

One must use caution when attempting to compare the results obtained here with related research on multi-stage systems with stochastic demand. Other research (see Nahmias and Schmidt (1983) and Lambrecht, Muckstadt, and Luyten (1984)) has used a shortage cost per unit per period while we have used a fill-rate criterion which is essentially equivalent to a shortage cost per unit. (Observe, however, that the relationship between the shortage cost and the fill-rate is not as clearly defined in a two-stage system as in a single-stage system). The use of a shortage cost per unit per period will provide greater incentive to hold component safety stock than will the use of a shortage cost per unit. Observe that if a shortage cost per unit per period is imposed, within a period of time no longer than the assembly natural cycle plus leadtime, component safety stock can be transformed into usable finished product (provided a partner component is available). This finished product can then reduce the shortage costs which otherwise would be incurred from that point on. This is not true when a "one-time-only" shortage cost per unit is used. While the actual results depend upon the problem data, component safety stock may play a more important role when time-weighted shortage costs are used. Shortage costs imposed at intermediate stages in the production process may also increase the desirability of some component safety stock.

We have provided an approach to the MRP safety stock problem which, while not optimal, has provided a method for determining when component safety stock is cost-effective in a simple product structure when emergency setups are not permitted. A variation of the algorithm can aid in determining "where" additional safety stock should be added to improve the fill-rate most economically.

Managers usually determine timing of production runs and safety stock quantity sequentially. The problem of obtaining truly optimal solutions for large systems is simply too complex and too costly, particularly in the presence of setup costs. However our relatively simple approach to the problem provides important managerial guidelines. We have provided intuition regarding the factors which favor component safety stock when a fill-rate criterion is used. Further research is needed to solve more complex problems involving multiple finished products with common parts and materials.

TABLE 1

Initial Problem Parameters

Parameter	Values
T_1	2,4
T_2, T_3	T_1, $2T_1$, $3T_1$
h_1	1.0
h_2, h_3	.10, .25, .40
D	200/period
σ	10, 30, 50 per period
L_1, L_2, L_3	1,5

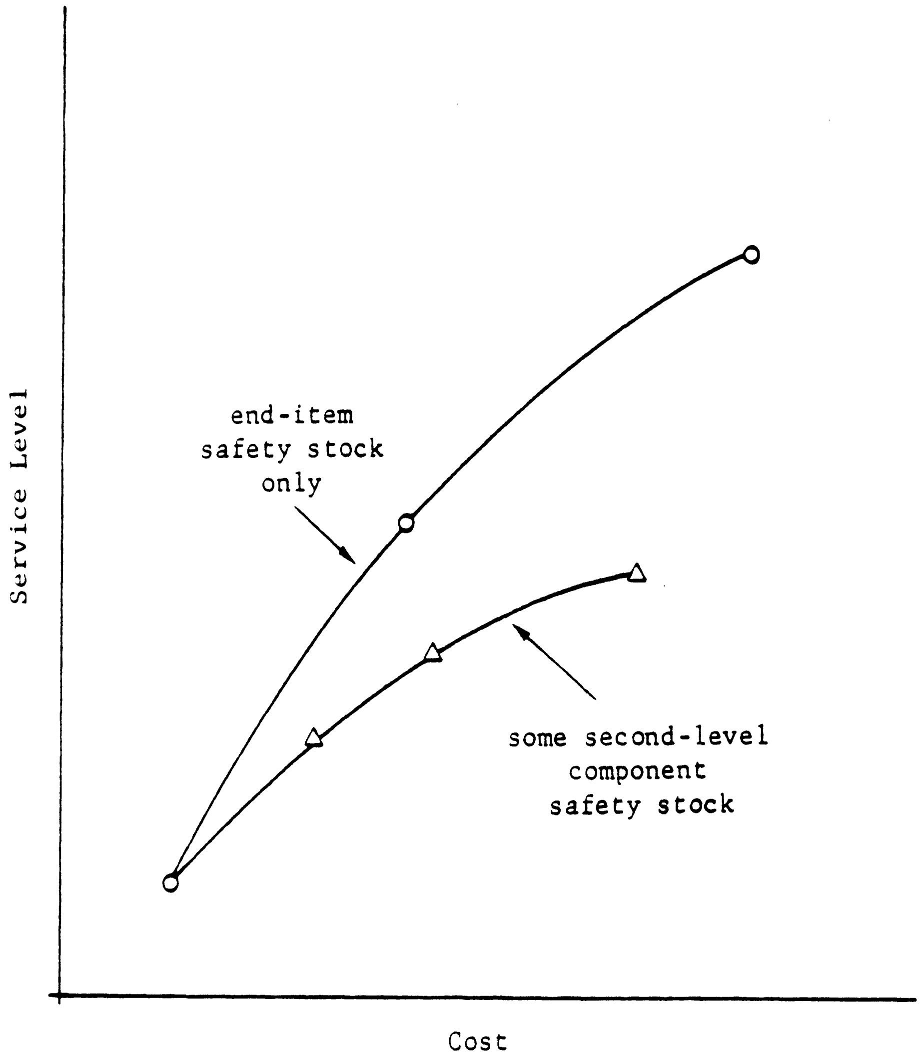


Figure 1

Case Where Zero Second-Level Component Safety Stock
is Cost-Effective

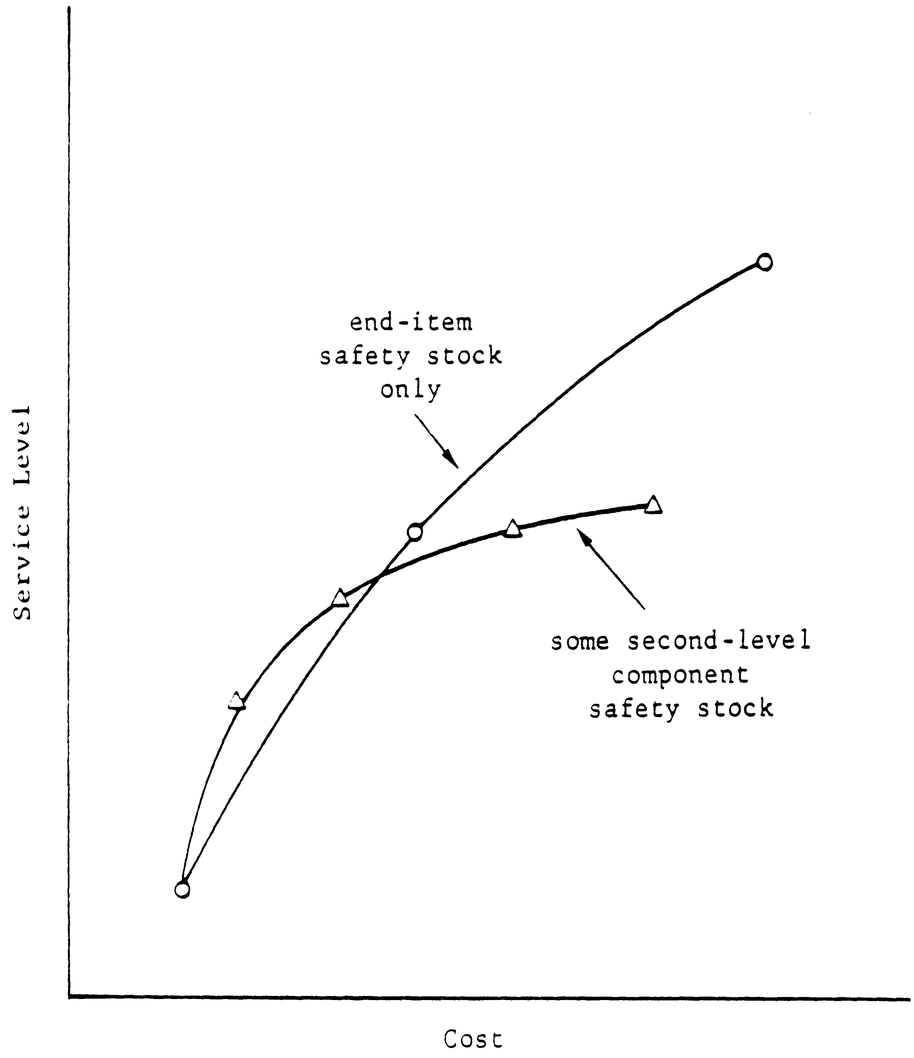
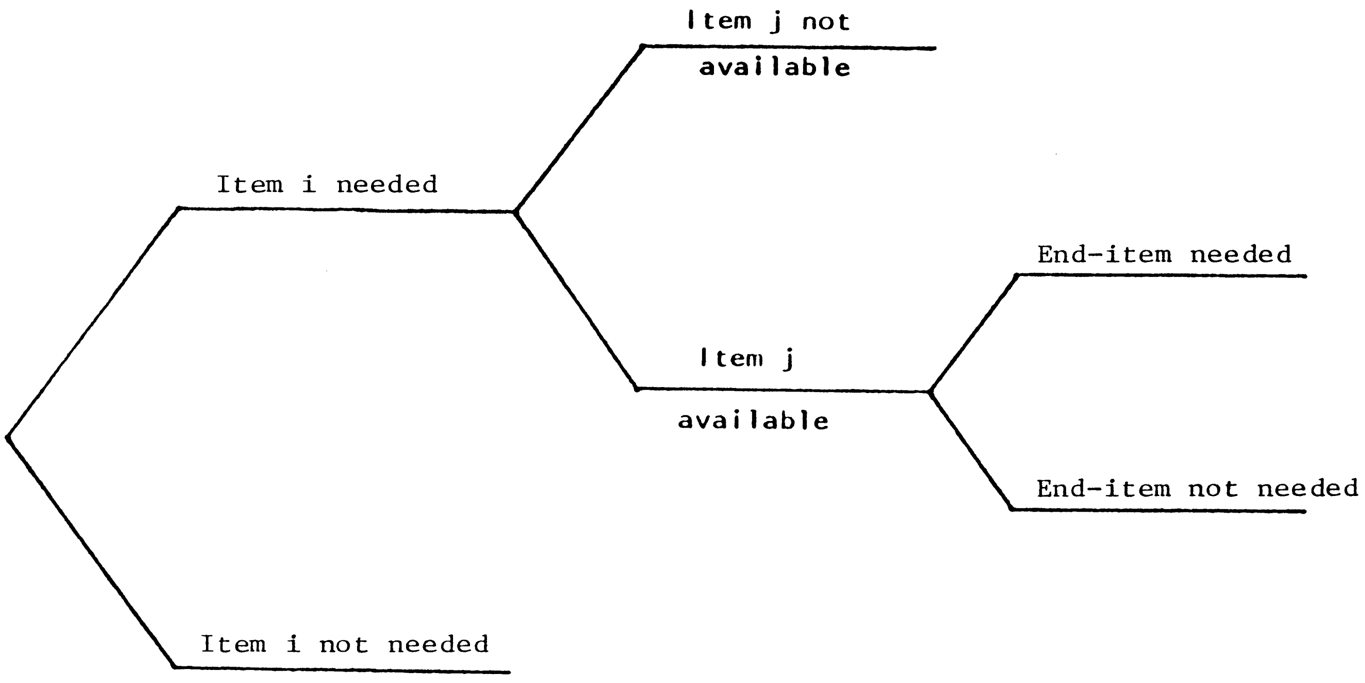


Figure 2

Case Where Some Second-Level Component Safety Stock is Cost-Effective



Holding Costs/Cycle	Path
$h_i T_i$	1
$h_i(T_i - T_1) + (h_1 - h_j)(T_1 - 1)$	2
$h_i(T_i - T_1) + (h_1 - h_j)T_1$	3
$h_i T_i$	4

FIGURE 3

Event Tree with Associated Costs

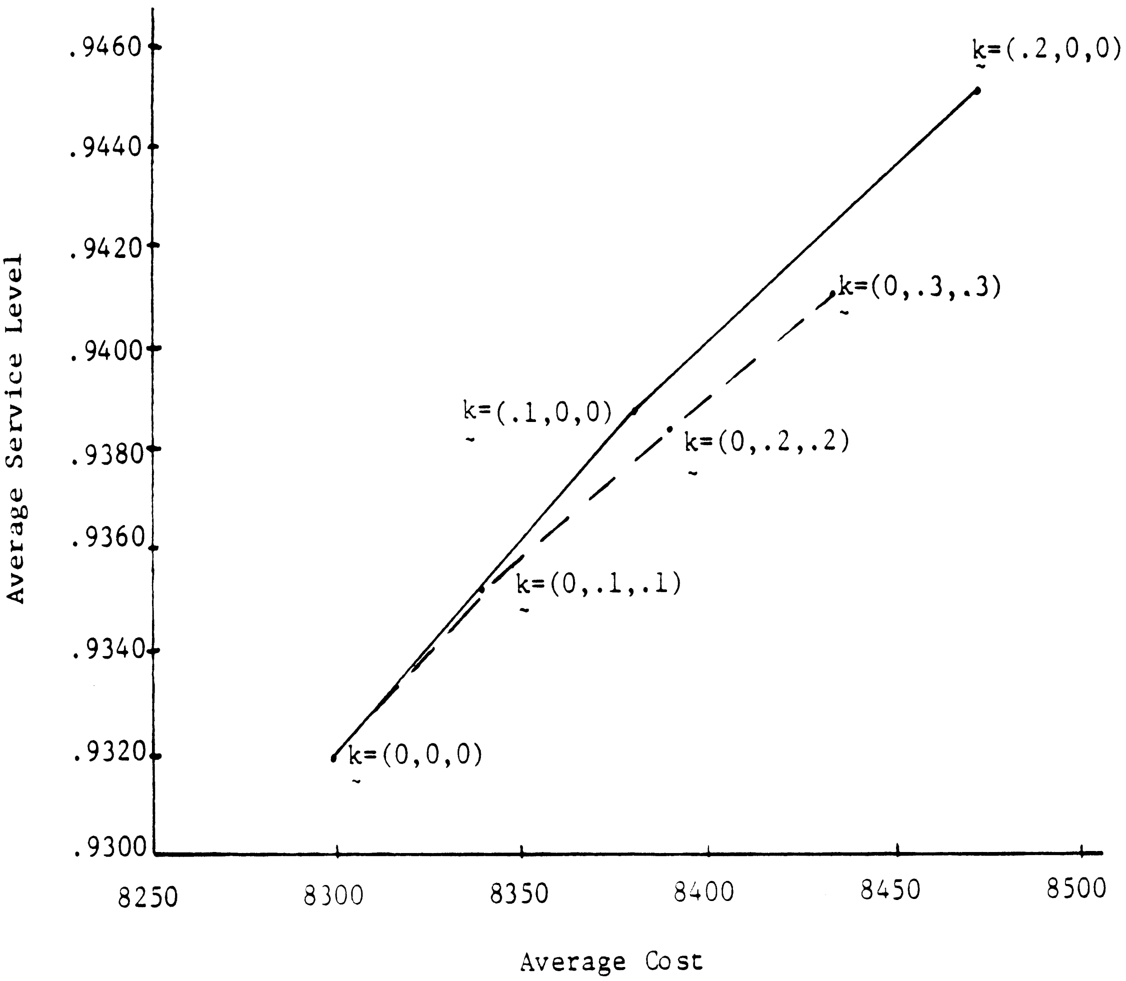


Figure 4

Results for $\underset{\sim}{T} = (2,2,2)$, $\underset{\sim}{h} = (1,0.1,0.1)$,
$\underset{\sim}{L} = (1,1,1)$, $r=30$

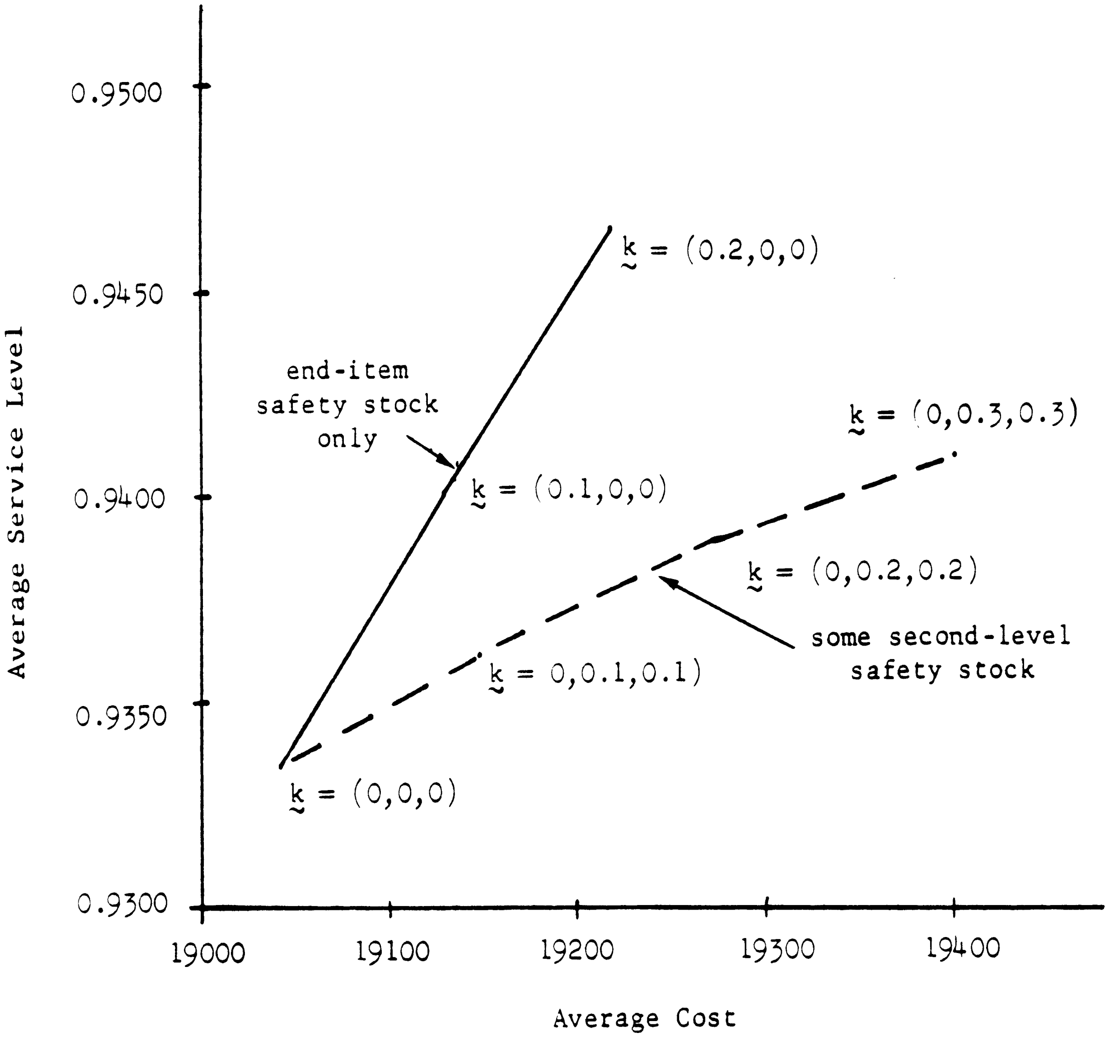


Figure 5

Comparison of Cost-Effectiveness of End-Item and Second-Level Safety Stock – $\underset{\sim}{T} = (2,4,4)$, $\underset{\sim}{h} = (1,0.4,0.4)$, $\underset{\sim}{L} = (1,1,1)$

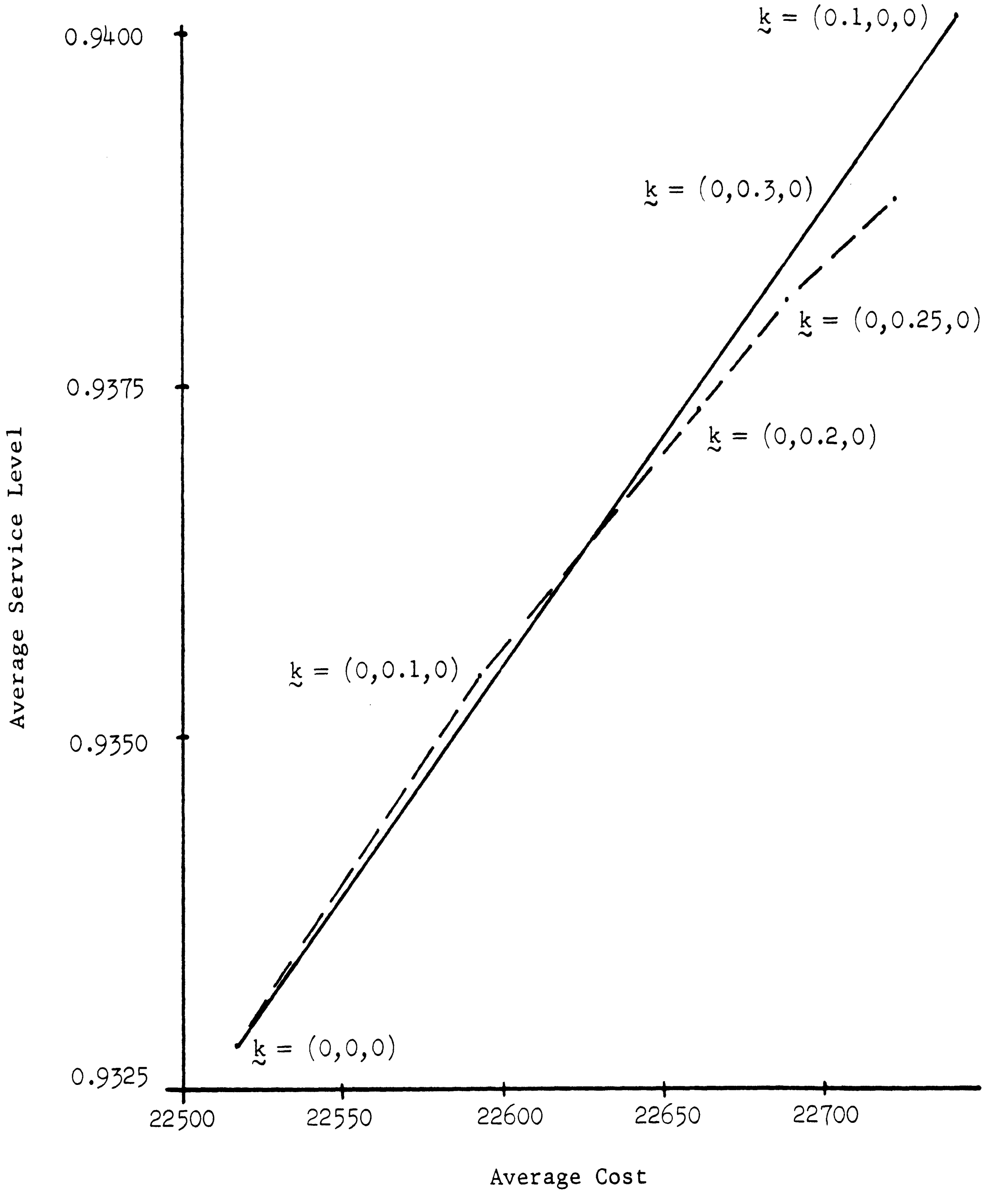


Figure 6

Results for $\underset{\sim}{T} = (4,4,12)$, $\underset{\sim}{h} = (1,0.1,0.1)$, $\underset{\sim}{L} = (1,1,1)$
Under Fixed Scheduling

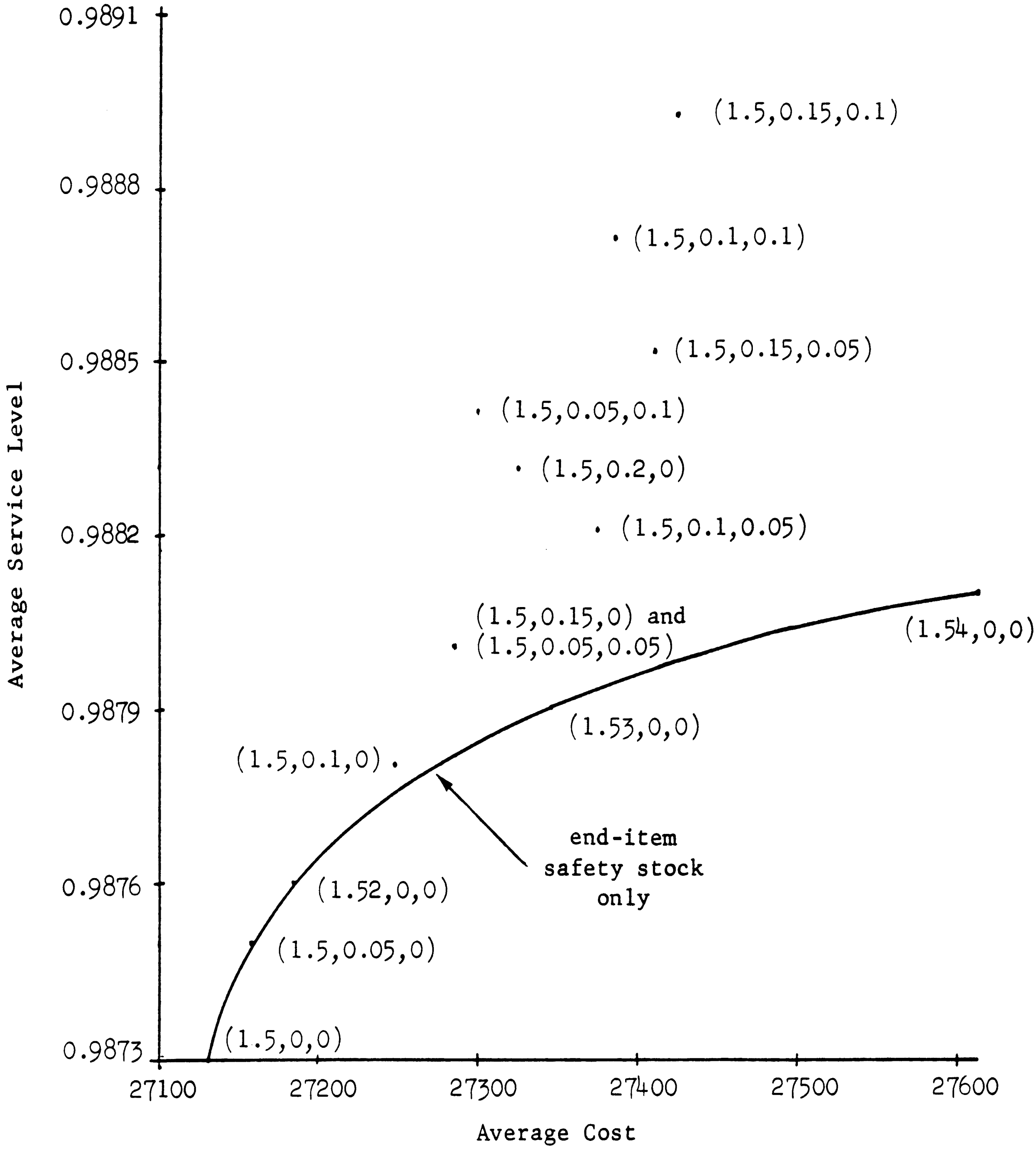


Figure 7

Results for $\underset{\sim}{T} = (4,4,12)$, $\underset{\sim}{h} = (1,0.1,0.1)$, $\underset{\sim}{L} = (1,1,1)$
With Fixed Scheduling at High Service Levels

(Safety Stock Vectors Indicated in Parentheses)

REFERENCES

Baker, K. R., "An Experimental Study of the Effectiveness of Rolling Schedules in Production Planning." Decision Sciences, 8, (1977), 17-27.

Blackburn, J. D. and R. A. Millen, "Heuristic Lot-Sizing Performance in a Rolling-Schedule Environment." Decision Sciences, 11 (1980), 691-701.

Caie, J. P. and W. L. Maxwell, "Hierarchical Machine Load Planning." In Multi-Level Production/Inventory Control Systems: Theory and Practice, L.B. Schwarz (ed.), New York: North Holland Publishing Co., 1981.

Carlson, R. C., S. L. Beckman and D. H. Kropp, "The Effectiveness of Extending the Horizon in Rolling Production Scheduling." Decision Sciences, 13 (1982), 129-146.

Carlson, R. C. and C. A. Yano, "Buffering Against Demand Uncertainty in Material Requirements Planning Systems -- Systems with Emergency Setups for Some Components." Technical Report 84-14, Department of Industrial and Operations Engineering, University of Michigan, 1984.

Eichert, E. S., III, "Accounting for Unplanned Inventory Demands in Material Requirements Planning Systems." Production and Inventory Management, 15 (1974), 69-77.

Lambrecht, M. R., J. A. Muckstadt, and R. Luyten, "Protective Stocks in Multi-Stage Productions Systems." Int. J. Prod. Res. 22, (1984), 1001-1027.

Meal, H. C., "Safety Stocks in MRP Systems." Technical Report #166, MIT Operations Research Center, 1979.

Miller, J., "Hedging the Master Schedule." Chapter 15 in Disaggregation: Problems in Manufacturing and Service Operations. L. Ritzman et al. (eds.), Hingham, MA: Martinus Nijhoff Publishers, 1979.

Moore, M. D., "The Use of Safety Stocks with Material Requirements Planning," APICS International Conference Proceedings, Atlantic City, NJ (1973), 128-136.

Nahmias, S. and C. Schmidt, "Optimal Policy for a Two Stage Assembly System Under Random Demand," Working Paper, University of Santa Clara, 1983.

New, C. "Safety Stocks for Requirement Planning." Production and Inventory Management, 16 (1975), 1-18.

Orlicky, J. A. Material Requirements Planning. New York: McGraw-Hill, Inc., 1975.

Silver, E.A. and H.C. Meal, "A Heuristic for Selecting Lot Size Quantities for the Case of Deterministic Time Varying Demand Rate and Discrete Opportunities for Replenishment." Production and Inventory Management 14 (1973), 64-74.

Wagner, H.M. and T. Whitin, "Dynamic Version of the Economic Lot Size Model." Management Science, 5, (1958), 89-96.

Welch, W. E., "MRP--It's Time Has Come," APICS International Conference Proceedings, Atlantic City, NJ (1973), 103-127.
Whybark, D. C. and J. G. Williams, "Material Requirements Planning Under Uncertainty," Decision Science, 7 (1976), 595-606.

Yano, C. A. "Safety Stocks in Material Requirements Planning Systems." Ph.D. Dissertation, Department of Industrial Engineering and Engineering Management, Stanford University, 1981.

APPENDIX A

Description of Simulation Model

The simulation operates as follows. Since the demand process is stationary, forecasts are set equal to mean demand. Available stock, backorders, and on-order quantities, if any, are then used to calculate net requirements using standard MRP logic. The Wagner-Whitin algorithm is implemented using these net requirements. If a lot must be produced in the current period, safety stock in the amount of $k_i\sqrt{t_i + L_i}\,\sigma$ is added to the production quantity.

We use a planning window of 24 periods, which for all cases in this study, is three or more times the length of the natural cycle of the component with the largest natural cycle. The fixed schedule is achieved by fixing the timing (but not the quantity) of all orders for a period of time equal to the largest integer multiple of the natural cycle less than the length of the planning window.

There are several reasons for this approach. First, research by Baker (1977) Blackburn and Millen (1980) and Carlson, Beckman and Kropp (1982) indicates that using a horizon equal to an integral multiple of the natural cycle is better than a non-intergral multiple when the Wagner-Whitin algorithm is implemented in a rolling horizon environment. Second, fixing the timing but not the quantity of the orders provides some latitude for responding to demand fluctuations without changing the ordering interval. Third, this technique limits production schedule changes to the end of the horizon, thereby essentially eliminating "nervousness" in the system. For instance, an item with a natural cycle of 4 would have its production schedule fixed for (24/4 -1) x 4 = 20 periods when the length of the planning window is 24 periods. This technique also serves to avoid scheduling setups whose timing may not be optimal because of end-of-horizon effects. The production schedule for the latter periods becomes fixed as the horizon rolls forward.

When this fixed scheduling technique is used, the Wagner-Whitin algorithm causes an order to be placed precisely every T_i periods (i.e., a cyclic schedule results) when the demand forecast is constant.

In each period, a plan is determined for a finite horizon, current decisions are implemented, a realization of the random demand process occurs, and the finite horizon rolls forward. The process then repeats. This differs from most of the earlier rolling horizon literature because demand is stochastic (not simply time varying) and is not know in advance.

APPENDIX B

We can formulate the problem with fixed scheduling as follows:

$$\underset{\underset{\sim}{k}}{\text{Minimize}} \left\{ S_1 + h_1 \sum_{\substack{0<n<T_1 \\ n \text{ integer}}} \sqrt{n}\,\sigma \int_{-\infty}^{\frac{T_1\bar{D}-n\bar{D}}{\sqrt{n}\,\sigma}} \left(\frac{T_1\bar{D} - n\bar{D}}{\sqrt{n}\,\sigma} - z \right) \phi(z)dz \right\} \Big/ T_1$$

$$+ \sum_{i=2}^{3} \left\{ \left[S_i + h_i \sum_{\substack{0<n<T_i/T_1 \\ n \text{ integer}}} \sqrt{n\ T_1}\,\sigma \int_{-\infty}^{\frac{T_i\bar{D}-nT_1\bar{D}}{\sqrt{nT_1}\,\sigma}} \left(\frac{\bar{T}_i D - n\bar{T}_1 D}{\sqrt{nT_1}\,\sigma} - z \right) \phi(z)dz \right] \Big/ T_i \right.$$

$$+ h_1 \cdot \tau(k_1)/T_1$$

$$+ \sum_{i=2}^{3} \left[H(\underset{\sim}{k})_i / T_i \right]$$

subject to

$$\frac{\sigma\sqrt{T_1 + L_1}\ G(\underset{\sim}{k},\underset{\sim}{T},\underset{\sim}{L})}{T_1\bar{D}} \leq 1 - \alpha$$

where [.] = integer part

- S_i = setup cost for item i
- $\underline{T}_i$ = natural cycle of item i
- L_i = leadtime for item i
- h_i = holding cost for item i
- σ = standard deviation of demand
- $\bar{D}$ = expected demand per period

$G(\cdot)$ = standard normal loss function given $\underset{\sim}{k}, \underset{\sim}{T}, \underset{\sim}{L}$

$H(\underset{\sim}{k})_i$ = expected holding cost arising from using safety stock multiplier k_i rather than $k_i = 0$, given k_1 and k_j

$h_1 \cdot \tau(k_1)$ = expected holding cost arising from using safety stock multiplier k_i rather than $k_1 = 0$

$$H(\underset{\sim}{k})_i = \sum_{\ell=0}^{[k_i \sqrt{T_i+L_i}\,\sigma]} \bar{h}_i\left(k_1, \frac{\ell}{\sqrt{T_i+L_1}\,\sigma}, k_j\right)$$

where $\bar{h}_i(k_1, k_i, k_j)$ = expected holding cost arising from an incremental unit of item i safety stock, given k_1, k_i and k_j

$$\tau(k_i) = \sum_{\ell=0}^{[k_1 \sqrt{T_1+L_1}\,\sigma]} \bar{T}_1\left(\frac{\ell}{\sqrt{T_1+L_1}\,\sigma}\right)$$

where $\bar{T}_1(k_1)$ = expected length of time an incremental unit of item 1 safety stock is held, given k_1