1967

Production scheduling with fixed lot sizes and constrained total capacity

David G. Jackson
Lehigh University

PRODUCTION SCHEDULING WITH FIXED LOT SIZES AND CONSTRAINED TOTAL CAPACITY

by

David G. Jackson

A Thesis

Presented to the Graduate Faculty

of Lehigh University

in Candidacy for the Degree of

Master of Science

Lehigh University

1967

CERTIFICATE OF APPROVAL

This thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

May 11, 1967
Date

Professor in Charge

Head of the Department of
Industrial Engineering

TABLE OF CONTENTS

ABSTRACT

A heuristic model is developed for obtaining the monthly production schedule in a multiproduct shop with a constraint on aggregate capacity.

The effects of certain approximations employed in the model development are investigated through computer simulation, the criterion for evaluation being the achieved service level. The results indicate that while the service level obtained satisfactorily follows the specified service level in the aggregate, the model presented favors the products with the higher mean monthly demands.

CHAPTER I

INTRODUCTION

Background

The topic of "production smoothing" has been given considerable attention in the literature during the past several years. The concept generally involves balancing the costs associated with carrying inventory on the one hand against the costs of changing the level of production on the other, the objective being to minimize the overall cost.

Production smoothing problems may be roughly divided into two categories:

1. Smoothing production for a single product over n periods.
2. Smoothing production from period to period when several products are competing for facilities.

Methods of handling the single product category, which will not be discussed further, vary from exact mathematical solution for simpler problems to approximate solutions for more complex situations, which, of course, are dependent on the particular assumptions made in each case (11,14,15,17,19).

The second category, production smoothing when more than one product is involved, has also been handled in a variety of ways. The manner in which a problem is formulated depends upon many factors, not all of which are under the control of the planner. (Some, for example, may be a matter of "policy".) A number of these considerations are presented below, not as a complete list, but only to indicate the variety of factors that influence, and complicate, the method of solution.

<u>Planning Horizon</u>: Is there a finite period of time over which demands are assumed known, or is it a continuing, "infinite horizon" situation? Is forecasting to be used?

<u>Nature of Demands</u>: Are demands deterministic, or probabilistic? Is the market static, or dynamic?

<u>Type of Production</u>: Is production to be in batches? Are lot sizes fixed, or will they vary from period to period? Is sequence of production important?

<u>Data-Processing Availability</u>: Are procedures to be handled manually, or by machine? Is updating done continuously, or periodically?

<u>"Degree of Smoothing"</u>: Is a cost associated with changing the level or production? Or is it <u>fixed</u> in some range by physical limitations or management decree? (The latter might be termed "absolute smoothing".)

<u>Service Considerations</u>: Is demand backordered when stock isn't available, or are the sales lost? Are backorders to be controlled by assigning a cost to stockouts, or by specifying a service criterion?

Techniques employed in formulating production smoothing problems in this second category include: linear decision rules applied after using quadratic approximation of cost functions (13), linear programming (5), dynamic programming in combination with linear programming (18), and Lagrangian multipliers (21). Each technique, when applied to a specific problem, must work within its own peculiar set of assumptions, but all the methods have one thing in common: an attempt to solve a problem of real-world size without resorting to some degree of approximation usually results in excessive computation. As a result, in order to provide a realistic method of solution, i.e., one that can be applied in an operating environment, optimality is often sacrificed for simplicity.

Problem Definition

Consider a shop engaged in the manufacture of a variety of products, all of which require similar processing steps, i.e., there are no specialized production lines. It is desired to manufacture on a batch basis, since changing production from one product to another requires adjustment and resetting of some facilities, thus incurring a set-up cost. Raw materials and component parts are always available from an adjacent storeroom, i.e., the quantity of any product scheduled for manufacture is not limited by raw material availability. Production output is made to meet inventory requirements, and if demands occur when stocks are depleted, the units are backordered and the backorder is filled the next time that particular product is manufactured. Because of the cost associated with continuously monitoring inventory levels of the many products, and because data processing facilities are available only on a preplanned basis, inventory is reviewed periodically and at the same time production is scheduled for the next period. The demands for all products are stable, and this situation is expected to continue in the future. For this reason, it is not desirable to change the size of the workforce. However, some fluctuation of the workload is acceptable, since a limited amount of shifting of operators between adjacent shops can be done when required. Hence, capacity is constrained within a range.

In addition to the above, the following assumptions are stated:

1. Demand for each product is normally distributed, with known mean and standard deviation.
2. A product scheduled for production at review time will be available for inventory at the end of the period.
3. Set-up times are neglected in arriving at the production schedule for the period.
4. No attempt is made to set the sequence of product manufacture during the period.

Objective

The objective of this study is to apply the Lagrange multiplier technique to the problem outlined above, and evaluate the effect of certain approximations employed. The criterion for evaluation will be the service level achieved while operating under a doctrine designed to minimize the sum of ordering costs, inventory holding costs, and backorder costs.

CHAPTER II

HEURISTIC MODEL FOR PRODUCTION SCHEDULING

In the previous chapter, the method of Lagrange multipliers was mentioned as one means of minimizing inventory costs in the presence of a constraint on capacity. In this chapter, a Lagrange formulation will be developed in detail, and then certain approximations will be introduced to make the method more computationally attractive (21).

First, for a single product, let the expected annual cost include:

$$C(R,Q) = \text{order cost} + \text{inventory holding cost} + \text{stockout cost}$$

$$= C_o \frac{12\overline{S}}{Q} + CuI(\tfrac{1}{2}Q+R-\overline{S}_L) + C_z \frac{12\overline{S}}{Q} \bullet \int_R^{\infty}(S_L-R)f(S_L)dS_L$$

where

C_o = order or set-up cost

C_u = unit cost

C_z = stockout cost per unit

I = inventory holding cost rate, \$/unit/year

$\overline{S}$ = expected monthly demand

$\overline{S}_L$ = expected lead time demand

$f(S_L)$ = probability density function of demand during lead time

Q = lot size

R = reorder point, or "trigger level"

Since the shop being considered manufactures n products, the total cost is the sum of n similar equations, with the addition of the Lagrangian term to handle the constraint on capacity:

$$TC = \sum_i C(R_i, Q_i) + \lambda\left(\sum_i h_i Q_i - P_T\right) \qquad i = 1, 2, \ldots, n$$

where h_i is a factor to convert units of product to hours, P_T is the total available production time, and λ is a Lagrangian multiplier.

Since decisions are to be made once each month and it is not known in advance which products are expected to reach the reorder point, all products will be initially included. Differentiating the cost function with respect to R_i and Q_i, the first order conditions for a minimum are obtained:

$$\text{(1)} \qquad \frac{\partial TC}{\partial R_i} = F(R_i, Q_i) = 0 \qquad i = 1, 2, \ldots, n$$

$$\frac{\partial TC}{\partial Q_i} = G(R_i, Q_i) + \lambda h_i = 0$$

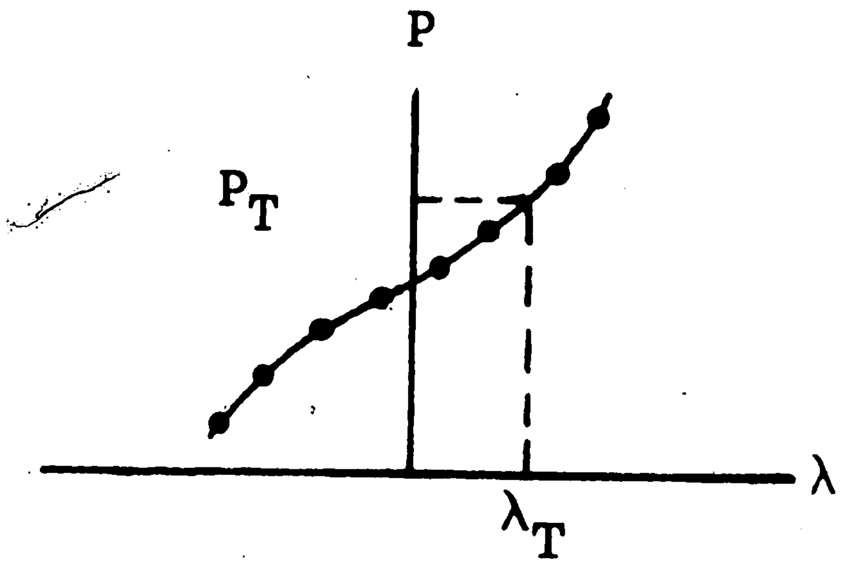
To solve each of these n pairs of equations, a λ is selected, and using the standard lot size (or Wilson lot size) as an initial Q_i, R_i may be obtained from the second equation. Then R_i is substituted in the first equation, thus obtaining a second estimate of Q_i. Iteration in this manner eventually yields satisfactory values of R_i and Q_i for the selected value of λ.

Using the reorder point or trigger level thus obtained, form the ratio:

$$\text{(2)} \qquad \frac{R_i(\lambda)}{I_i - \bar{S}_{Li}}$$

where I_i is the inventory at review time. If this ratio is greater than or equal to 1, the product is expected to trigger during the month, and the $Q_i(\lambda)$ is credited to used capacity.

When this process is repeated for all products (using the same value of λ), it will yield one point on a λ, P graph (P = hours of production). See accompanying figure.



Now a different λ is selected, and the whole process repeated again to get another point on the graph. When enough points have been obtained, the multiplier λ_T, which gives the desired total production P_T may be read off. This value of λ_T is then used in each pair or equations (1) to solve for R_i and Q_i.

This procedure, while workable, has several disadvantages. First, even when the calculations are programmed, the computational time is extremely long, and it must be repeated each month for n pairs of non-linear equations. Furthermore, the solution obtained is still an approximation in that it uses inventory rules for independent products and then makes aggregate production decisions based on the results.

So, due primarily to the effort involved, a further approximation in the solution is made. The number of calculations is cut down considerably by constraining only one of the variables, either R_i or Q_i, and letting the other take on its unconstrained value. Studies have shown that constraining only the trigger level (reorder point) and using the unconstrained lot size causes a smaller increase in cost above the "optimal" than does using the constrained lot size and the unconstrained trigger level (21). However, when using the unconstrained (constant) lot size, there is only one decision variable remaining, R_i, and the derivatives with respect to R_i do not contain λ (see equations (1)); hence the solution appears to have degenerated.

Introduction of an additional approximation will, however, permit solution. This is to use the unconstrained R_i, obtained by solving equations (1) with $\lambda = 0$, and compute priorities from the ratio shown in (2). Then the products are ranked in order of descending priority, and the hours required to produce the unconstrained lot sizes are accumulated until the capacity constraint is reached. This then gives the products that should be manufactured during the month.

The procedure, at review time, for obtaining the list of products to be produced during the month, may be summarized as follows:

1. Obtain the unconstrained reorder point, R_i, from equation (1), with $\lambda = 0$.
2. Obtain the priority ratio shown in equation (2), and rank the products in order of descending priority.
3. Beginning with the highest priority, accumulate the hours required to manufacture the unconstrained lot size Q_i until the capacity constraint P_T is reached.

The products selected will be manufactured during the next month.

In the next chapter, the method of determining the unconstrained lot sizes and reorder points for the specific production inventory system described by the assumptions in Chapter I will be developed.

CHAPTER III

APPLICATION OF THE SCHEDULING METHOD

Determination of Unconstrained Lot Size

From Chapter II, the expected annual cost for a single product is:

$$C(R,Q) = C_o \frac{S_Y}{Q} + C_u I(\tfrac{1}{2}Q+R-\bar{S}_L) + C_z \frac{S_Y}{Q} \bullet \int_R^{\infty} (S_L-R) f(S_L) dS_L$$

where S_Y is the expected yearly demand.

The last term states that the annual backorder cost is obtained by multiplying together the backorder cost per unit (C_z), the expected number of order periods per year (S_Y/Q), and the expected number of backorders per period given that the reorder point is R. Let the expected number of backordered units per cycle be given by:

$$N(S_L,R) = \begin{cases} 0 \text{ if } S_L-R \leq 0 \\ S_L-R \text{ if } S_L-R > 0 \end{cases}$$

(product subscript suppressed)

or

$$N = \int_R^{\infty} (S_L-R) f(S_L) dS_L$$

Differentiating the cost function with respect to the two decision variables, as in equation (1), Chapter II:

$$\frac{\partial C}{\partial Q} = -\frac{C_o S_Y}{Q^2} + \frac{C_u I}{2} - \frac{C_z S_Y}{Q^2} N = 0$$

$$\frac{\partial C}{\partial R} = C_u I + C_z \frac{S_Y}{Q} \bullet \frac{\partial N}{\partial R}$$

If the expected number of backorders per year is negligible:

$$\frac{\partial C}{\partial Q} = -\frac{C_o S_Y}{Q^2} + \frac{C_u I}{2} = 0$$

and solving for Q:

$$(1) \qquad Q = \sqrt{\frac{2\ C_o S_Y}{C_u I}}$$

Thus, if backorders are temporarily neglected, the unconstrained Q is found to be the standard, or Wilson lot size. This is the lot size that will be assumed throughout the remainder of the development.

Determination of Unconstrained Reorder Point

The unconstrained reorder point may be obtained for each product through the use of the cost equations described in Chapter II or by the specification of a service level criterion. The latter alternative will be used here. There are two reasons for this choice. First, the stockout cost, or backorder cost is difficult to evaluate, and even though specification of a service level imputes a cost to stockouts, it is often more desirable from a management standpoint to specify a desired service level. Secondly, when demands are considered normally distributed, the determination of the reorder point using equation (1) of Chapter II is extremely tedious (1, page 352).

There are numerous ways of defining service level (1, page 333). The one used here will be:

$$Z = \frac{\text{number of units shipped without delay}}{\text{number of units demanded}}$$

This has the advantage of permitting the service levels for all products to be specified identically over a given period of time, regardless of the number of order periods encountered, or the length of the lead time period. This is in contrast with the normal method of using a service level criterion to establish the reorder point, which neglects the length of the lead time and the number of order periods per year. In other words, the normal method of computing the reorder points from a specified service level establishes the same level of protection for all products during their respective lead time periods, but a different level of protection when viewed over a common period of time, e.g., a year.

Since it is more logical to specify reorder points that establish the same over-all service level for all products, it is necessary to modify the usual method of determining the reorder points to account for different order periods and lead time periods. This entails obtaining the relation between the over-all service level, designated Z_O, and the lead time service level, designated Z_R.

According to the definition of service level given above, the expected over-all service level for a period of T months may be expressed as:

$$Z_o = \frac{\bar{S} \cdot T - N_T}{\bar{S} \cdot T} = 1 - \frac{N_T}{\bar{S} \cdot T}$$

where

$\bar{S}$ = mean monthly demand

N_T = expected number of backorders in T months

The expected number of backorders in period T may be expressed as the expected number of backorders in a lead time period, N_R, times the number of lead time periods in T months.

The number of lead time periods in period T is:

$$\frac{T}{Q/\bar{S}} = \frac{T \cdot \bar{S}}{Q}$$

Hence the expected number of backorders in period T is:

$$N_T = N_R \left[\frac{T \cdot \bar{S}}{Q} \right]$$

This is substituted in the expression for the over-all service level in a period of T months:

$$Z_o = 1 - \frac{N_T}{\bar{S} \cdot T} = 1 - \frac{N_R}{Q}$$

Now the definition of service level is applied to the lead time period in order to obtain the expected service level when the inventory reaches the reorder point R:

$$Z_R = \frac{\bar{S}_L - N_R}{\bar{S}_L} = 1 - \frac{N_R}{\bar{S}_L}$$

where

$\bar{S}_L$ = expected lead time demand

Solving for N_R:

$$N_R = \bar{S}_L(1-Z_R)$$

This is substituted in the expression for over-all service level:

$$Z_o = 1 - \frac{N_R}{Q} = 1 - \frac{\bar{S}_L}{Q}(1-Z_R)$$

yielding the desired relation between the lead time service level Z_R and the over-all service level Z_o:

$$(2) \qquad Z_R = 1 - \left(\frac{Q}{\bar{S}_L}\right)(1-Z_o)$$

Thus, using equation (2), it is possible to begin with a desired service level specified by management, Z_o, and obtain the service level necessary during the lead time period, Z_R, in order to support Z_o. The reorder point R must now be determined from the necessary lead time service level Z_R. This development follows:

In deriving equation (2) above the definition of service level was applied to the lead time, i.e.:

$$Z_R = 1 - \frac{N_R}{\bar{S}_L}$$

Buchan (1) expresses this service level in terms of a "shortage factor" and the coefficient of variation (ratio of standard deviation of lead time demand to the mean lead time demand):

$$(3) \qquad Z_R = 1 - \mu F_R$$

where

F_R = shortage factor when inventory is at level R

$\mu = \sigma_L / \bar{S}_L$

The shortage factor F_R is dependent on the distribution of demand. For the normal distribution this is:

(4) $$F_R = \phi(t_R) - t_R\, \psi(t_R)$$

where

(5) $$t_R = \frac{R - \bar{S}_L}{\sigma_L}$$

$$\phi(t_R) = \frac{1}{\sqrt{2\pi}} \exp\left[-\tfrac{1}{2}\, t_R^{\,2}\right]$$

$$\psi(t_R) = \int_{t_R}^{\infty} \phi(x)\,dx$$

(The derivation of the shortage factor F_R, consistent with the service level as defined here, is given in Appendix A.)

Once the lead time service level Z_R is computed from equation (2), the shortage factor F_R may be obtained from equation (3). Then the normal deviate t_R may be computed from equation (4), using trial and error solution and any set of normal tables containing $\phi(x)$ and $\psi(x)$. Finally, the reorder point R is obtained from t_R using equation (5) above.

The second part of Appendix A contains an example illustrating the calculation of a reorder point by the procedure described above, and compares the result with that obtained using the "standard" method.

Determination of Priority

An obvious difficulty presents itself in the ratio used to establish priorities:

$$\frac{R}{I - \overline{S}_L} \text{ (product subscript suppressed)}$$

If production were not constrained, an order would be placed when this ratio reached one. However, in the presence of the constraint, a ranking procedure is used, with the positive ratios greater than one being scheduled first, as described in Chapter II. Herein lies the difficulty. Should the demand between review periods be great enough so that the inventory at review time, I, is less than the expected lead time demand, $\overline{S}_L$, the ratio shown above will be negative (assuming that the reorder point, R, is positive). A negative priority ratio would be ranked so low that the product would not be scheduled. (This, of course, would not be the case if the inventory system was humanly monitored, but could happen in a computerized system). As a result, once the inventory I was reduced to a level below the expected lead time demand without an order being triggered, the inventory could never recover. (Note that this predicament is due to the periodic review scheme, and could not occur under a continuous review inventory system).

Further complication is caused by the fact that it is possible for reorder points to be negative (see discussion under "Sample Calculation," Appendix A). This can cause incorrect ranking if the ratio above is used without modification.

These difficulties can be resolved by relating the reorder points (R) and the expected end-of-month inventories ($I-\overline{S}_L$) to some fictitious reference point, analogous to "absolute zero." This is accomplished by adding the same positive constant to all reorder points, and to all expected end-of month inventories before computing priority. The magnitude of the positive constant is such that the denominator of the priority ratio is positive for every product. In this manner, the two difficulties described above are eliminated, and the ratio continues to indicate which products are expected to reach the reorder point during the month. (The effect of this procedure is a "compressing" of the ratios about one, but the relative order is still maintained.)

In this chapter, the method of determining the unconstrained lot size, unconstrained reorder point, and priority ratio for use in the scheduling algorithm of Chapter II has been presented. The procedure may be summarized as follows:

1. Compute the unconstrained lot size:

(1) $$Q = \sqrt{\frac{2C_o S_Y}{C_u I}}$$

2. Given the over-all service level to be maintained, Z_0, compute Z_R, the required service level during lead time for reorder point R:

(2) $$Z_R = 1 - \frac{Q}{S_L}(1 - Z_0)$$

3. Determine the shortage factor F_R from:

(3) $$Z_R = 1 - \mu F_R$$

4. Using trial and error, along with standard normal tables, solve for t_R from:

(4) $F_R = \phi(t_R) - t_R \psi(t_R)$

5. Compute the reorder point:

(5) $R = \bar{S}_L + t_R \sigma_L$

6. Compute the priority ratio:

$$\frac{R_i + C}{\hat{I}_i - \bar{S}_{L_i} + C}$$

Where C is a positive constant such that the denominator is positive for all products.

7. Beginning with the highest priority, accumulate the hours required to manufacture the unconstrained lot size Q_i until the capacity constraint P_T is reached.

CHAPTER IV

TESTING THE MODEL

The method of determining the production schedule, described in Chapter II, contained some approximations and "compromises" incorporated to reduce a problem of considerable complexity and size to one that could be more readily handled with a realistic amount of computation. Then in Chapter III some further simplifying assumptions were introduced in order to obtain the unconstrained lot sizes and reorder points, and to arrive at a usable priority determination.

Each one of these simplifying steps would naturally force the solution further away from the optimal. However, as indicated in Chapter I, even the more rigorous methods of solution to the type of problem considered here eventually involve some degree of approximation. The main interest is therefore in the effect of some of the more gross approximations resorted to in arriving at the three-step procedure for determining the products to be manufactured during the month (Chapter II), and in obtaining the lot sizes and reorder points used in that procedure (Chapter III).

Probably the most significant approximation employed was the introduction of the unconstrained lot size in equation (1), Chapter II. This rendered inoperative the Lagrangian term, forcing the determination of the production schedule to be made on a priority basis, with no guarantee of a product being reordered when the priority reached a value of 1, i.e., when the reorder point was reached. Due to the fact that production is in batches, more products might reach the reorder point (priority = 1) than can be accommodated by the constrained production time. When this happens, the manufacture of some of the products is postponed until the following month, and since the reorder point has already been penetrated but no order placed, the safety stock may not be adequate to satisfy demands occurring before replenishment. This chain of events, should it occur, will culminate in a reduced level of service. In order to investigate the significance of the approximation then, the service level may be monitored while allowing the system to operate under the priority rule for a period of time. Two aspects are of interest here:

(1) The service levels of the individual products.

(2) The aggregate service level, i.e., for all products combined.

Both aspects should be investigated since the reorder points and lot sizes are computed for the products independently, but the production schedule is limited by an aggregate constraint.

Several additional factors are of interest either because of the specific approximations employed or because of their influence on, or being influenced by, the priority method of scheduling. These factors of interest are discussed below:

(1) Another approximation introduced was to neglect the backorders and use the Wilson lot size as the unconstrained lot size. It can be shown that if backorders are not neglected, the lot sizes for the individual products

would be larger than the Wilson lot size (6, p. 17). However, there is another force at work in this case. This is the constraint on total production, which would tend to shrink lot sizes in order to produce all products that reach the reorder point in a given month. These two divergent interests would appear to make the lot size a critical factor in the service level that is realized over a period of time.

(2) The coefficient of variation, μ, plays an important role in the determination of the unconstrained reorder point (see equations (3), (4), and (5), Chapter III). From equation (3), Chapter III, it is seen that a larger μ will result in a lower lead time service level, Z_R, which in turn adversely affects the over-all service level Z_O. Or, stated in another way, the safety stock required to maintain a given service level increases as the coefficient of variation increases. It is of interest to know whether this fact in conjunction with the possible postponing of production (described above) will make the service level realized over a given period of time a function of the coefficient of variation of the product demand.

(3) Obviously an important factor having an effect on the resulting service level is the specification of the production constraint, P_T. Should the postponing of the manufacture of certain products occur, it would appear

that the situation could be alleviated by increasing P_T, permitting more products to be manufactured. However, it can be reasoned that allowing production in excess of the amount required to meet the mean aggregate demand will, over the long run, permit inventories to build up, since demands are stationary. On the other hand, a constraint set lower than that required to meet mean aggregate demand will permit depletion of safety stocks over the long run.

(4) Another factor of interest is in the service level criteria originally specified by management. It is seen from equation (1), Chapter III that this specification of Z_0 is the first step toward arriving at the unconstrained reorder point. However, as indicated above, there is no guarantee that a product will be reordered when this point is reached. Does the specification of the desired over-all service level, Z_0, actually affect the level achieved over a period of time, or does, in fact, the system set its own level of service due to the constraint on production?

Initial Investigation

The four areas of interest mentioned above were investigated through simulation of the inventory system, governed by the methods, rules, and assumptions of Chapters I, II, and III. These four areas indicate four factors, all of which may occur at different levels, which might affect the service levels of the individual products, or of the aggregate collection of products, as described. This suggests a factorial experiment, the factors being lot size Q, coefficient of variation μ, production constraint P_T, and the management specified service level Z_0. Initially the primary interest is in the direction and amount of the main effect of a factor (if indeed there is an effect) and not in the shape of the response curve. For this reason the initial experiment was designed as a four factor experiment, with each factor at two levels. The specific levels of the factors for the initial experiment were chosen as follows:

Lot Size: As described in Chapter III, the unconstrained lot size is computed using the standard or Wilson lot size formula. This is accomplished by neglecting backorders, which, if not neglected would tend to increase the lot size. Hence, it is reasoned that products should be manufactured in lots no smaller than the standard lot size, and when runs are made with this factor at the "low" level; lot sizes will be computed by the standard lot size formula. When the "high" level of the lot size factor is required, lot sizes will be determined by increasing the standard lot size by 50%. This method was chosen in order to provide a systematic means of varying the lot sizes for all products, and it was determined through preliminary runs that more than a 50% increase caused severe problems in obtaining a lot size solution within the allowed capacity range, thus invalidating the procedure for obtaining a production schedule. The two levels are therefore representative of the extreme values for the range of interest.

<u>Coefficient of Variation</u>: Here also some systematic means of specifying the levels of the factor over all products was required. It was decided that any effect of the coefficient of variation on the service level was of interest only for the products in the aggregate. That is, the same coefficient of variation was used for all products, and conclusions are limited to determining the effect on service level when the demands for the aggregate collection of products are more volatile in one case (coefficient of variation, μ, at the high level) than in the other (μ at low level). The two levels were again chosen at the extremes of the range of interest. Since a coefficient of variation too close to zero indicates almost deterministic demands, the lower level of the coefficient of variation factor was chosen at $\mu = .2$. Choice of the high level was guided by the fact that the "normal" distribution is assumed for demands, and a high μ requires "throwing away" too much of the distribution (demand is not allowed to be negative). A $\mu = .6$ means that only about 5% of the lower tail of the demand distribution is eliminated. It is believed that this much can be safely eliminated without introducing too much error into the experiment, since the theoretical normal distribution is used in computing the reorder points (Chapter III). In addition, this value provides a standard deviation equal to 60% of the mean of the distribution, which represents a fairly volatile demand situation.

Production Constraint: As discussed previously, if the constraint on hours of production is set lower than that required to fulfill average demand over the long run, depletion of inventories and degradation of service level will result with certainty. On the other hand, too much excess capacity over that required to meet average demand will permit inventories to build up unnecessarily. With this in mind, the low and high levels of the capacity constraint factor will be set at the number of hours to meet mean demand, and 10% over this value, this being a reasonable range, at least for the initial investigation.

Specified Service Level: The levels of this factor were chosen to represent the values of service level commonly used in practice. The low and high levels selected were 95% and 98% respectively for initial investigation. If it is found that these levels are too close together to provide any useful information about the effect of the specified service level on the actual performance of the system, other values will be chosen the later experimentation.

Determining the Type of Experiment

It was previously stated that the initial experiment should be a four-factor experiment, with each factor at two levels. Since it was suspected that the three factor interactions would be small compared to the main effects, it was decided to use a half replicate of the factorial experiment, in which the three factor interactions are aliased with the main effects, thus permitting a smaller experiment for the initial investigation. The design chosen also aliases the two factor interactions, and if the results show that this procedure causes the two factor interactions to appear significant, additional investigation will be necessary in order to determine which interaction is responsible.

Because of the number of products involved, it was decided to make the first analysis on the results of the aggregate service level, rather than be concerned with the effect of the factors on the individual products. This latter effect could be investigated in later experiments.

A more complete description of the $\frac{1}{2} \times 2^4$ experiment is given in Appendix B.

Determining the Amount of Simulation Required

The length of each run could be determined in an intuitive manner. Since the process being monitored is an inventory system, the length of any particular run may be set by determining the period of time for which the results would be of interest. In a production-inventory system like the one under investigation, over an extended period of time the product mix would probably change, demand patterns would vary, and costs would, more than likely, not remain constant. Thus, to allow one time series to cover a period of ten years, for example, would be meaningless, since conditions in a production-inventory system are not likely to remain static for that length of time. It would be more meaningful from a practical standpoint to learn what happens under a given set of conditions within a period of two or three years. With this in mind, it was decided that each replication of the experiment would include 36 months of activity.

For the initial investigation, it was decided to use a single time series of demands, randomly chosen, for all treatment combinations selected. This would provide an indication of the effects of the factors (discussed above) for the particular sequence of demands encountered. Then in order to determine how much of the variability in response might be expected if different demands were experienced, a separate one-way experiment could be run for any factor that appears to have a large effect on the service level. This second smaller experiment would involve both levels of the factor in question with the other three factors held constant at a randomly selected combination of levels. Replicating this experiment with different sets of random demands will provide an indication of how much of the response is due to the change in level of the factor as compared to the amount that can be attributed solely to the fact that different demands were encountered.

The results of the initial investigation follow.

Results of the Initial Investigation

Figure IV-1 shows the data obtained from the $\frac{1}{2} \times 2^4$ experiment, a description of which appears in Appendix B.

For this experiment, the factors are defined as:

A = coefficient of variation

B = capacity constraint

C = lot size

D = specified service level

For convenience, the observations recorded were "fraction backordered" rather than service level. However, as can be seen from the definition of service level in Chapter III:

Service Level = 1 - fraction backordered

		A_1		A_2	
		B_1	B_2	B_1	B_2
C_1	D_1		.0002	.0186	
	D_2	.0092			.0003
C_2	D_1	.0163			.0012
	D_2		.0002	.0169	

$\frac{1}{2} \times 2^4$, Aggregate Fraction Backordered

FIG. IV-1

Since the response measured is a proportion, a transformation is made on the observed values (4, p. 45):

$$x = \text{arc sin} \sqrt{b}$$

where b is the observed fraction backordered. Performing the analysis on the transformed variables yields the following table:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square
A(=BCD)	1.82882	1	1.82882
B(=ACD)	69.31941	1	69.31941
AB(=CD)	.15429	1	.15429
C(=ABD)	.75337	1	.75337
AC(=BD)	.18210	1	.18210
BC(=AD)	.02749	1	.02749
ABC(=D)	1.27440	1	1.27440
Total	73.53987	7	

FIG. IV-2

Factor B (capacity constraint) is seen to have a very large effect on the service level, as was surmised. All other main effects (which are aliased with 3-factor interactions) are larger than the 2-factor interactions. Note also that factor D (specified service level) is seen to have little effect on the resulting aggregate service level.

Without doing a formal analysis, it is seen that factor B (capacity constraint) is the only factor that appears to have a significant effect on the service level, and as stated previously, a one-way experiment is called for, using additional random demand patterns (replications) with B at its high and low levels, and A, C, and D held constant. The results are shown in Figure IV-3. An analysis of variance table is shown in Figure IV-4.

	B_1	B_2
Replication 1	.0004	.0000
Replication 2	.0041	.0004
Replication 3	.0062	.0000
Replication 4	.0068	.0001

One-Way Classification (Factor B), Fraction Backordered

FIG. IV-3

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
B	25.6458	1	25.6458	65.944
Error	2.3332	6	.3889	
Total	27.9790	7		

FIG. IV-4

Factor B (capacity constraint) is seen to have a highly significant effect on the resulting service level even when different demand patterns are encountered.

From the results of the $\frac{1}{2} \times 2^4$ experiment (Figure IV-1), it is seen that for the treatment combinations used, the resulting aggregate service level was greater than that specified in all cases, even when the lower capacity constraint (factor B at low level) was used. Then in the smaller experiment, investigating factor B for different demand patterns, it is seen that the effect of factor B is still significant. This is verification of the fact that setting the capacity constraint higher than the amount to meet mean aggregate demand will only permit inventories to build up unnecessarily. Thus, for further experimentation, the constraint on capacity will be set at the amount required to manufacture for the mean aggregate demand, and the other factors can be investigated in more detail, while still permitting an experiment of reasonable size.

Further Investigation

The initial investigation indicated that the aggregate service level could be maintained satisfactorily with the capacity constraint set at the number of hours required to manufacture for the mean aggregate demand. Thus, the high level of the capacity constraint factor was eliminated from any further experimentation, as setting the capacity higher than that required to meet mean aggregate demand would permit unnecessary inventories to be carried.

Results of the initial experiment also seemed to indicate that, even though the aggregate service level was higher than the specified service level, some products (those with the lower mean monthly demands) were receiving much less than the desired degree of protection. This, if actually true, would indicate that the model did not adequately distribute the inventory among the products.

With this in mind, the second phase of experimentation concentrated on determining if the service levels achieved were fairly uniform over all products. Because of the large number of products, it is possible that any real difference in achieved service level due to differences in demand might be masked by noise, if the products are considered individually. Hence it would be more meaningful if the products were divided into groups according to the mean monthly demands. The separating of products into groups was some-what arbitrary. The specific demand levels defining the groups are not particularly important. The emphasis was on keeping the number of groups small enough to be effectively handled in an experiment, while still keeping the demand range within a group narrow enough so that each group is fairly homogeneous, i.e., so that no single product is vastly different from the rest of the group, which would bias the group results too heavily. This arbitrary grouping resulted in the following three classifications:

Group 1: Products with mean monthly demand between 100 and 600 units.

Group 2: Products with mean monthly demand between 70 and 100 units.

Group 3: Products with mean monthly demand less than 70 units.

These classifications encompassed all but one of the products. The excluded product had a mean monthly demand so much greater than the rest of the products that it would heavily weight the aggregate service level of Group 1, and therefore exaggerate any real difference in achieved service level. The fact still remains that there will be a certain amount of variability of service levels of the individual products within each group, but much less than for the whole population of products. At any rate, the grouping should provide a suitable estimate of the degree of difference in attained service levels that is caused by differences in demand.

The other factors of interest in this experiment are the same as in the initial investigation (with the exception of the capacity constraint, previously discussed). Two levels of the coefficient of variation are used (μ = .2, and μ = .6). The effect of lot size on the resulting service levels is investigated through the use of the standard lot size and 1.5 times the standard lot size. The fourth factor, initially specified service level, Z_0, is examined at the two levels used during the initial investigation (Z_0 = .95 and Z_0 = .98) plus a third level (Z_0 = .92) added to provide information on the effect of the specified service level over a wider range.

This phase of investigation thus utilized a complete 2 x 3 x 3 x 2 factorial experiment. It was decided to make two replications (series of random demands) in order to obtain an estimate of experimental error, i.e., the variability that can be attributed solely to the difference in demand patterns encountered.

The results of the experiment, and the appropriate analysis, are given below.

Results of the Final Experiment

The data shown in Figure IV-5 were obtained from the 2 x 3 x 3 x 2 factorial experiment, with the factor levels as defined above. The factors are:

A = coefficient of variation

B = specified service level

C = demand group

D = lot size

		A_1			A_2		
		B_1	B_2	B_3	B_1	B_2	B_3
D_1	C_1	.0043 .0012	.0043 .0012	.0026 .0004	.0329 .0107	.0241 .0053	.0147 .0034
	C_2	.0406 .0295	.0360 .0483	.0268 .0087	.0659 .0940	.0548 .0708	.0241 .0425
	C_3	.0920 .0777	.0836 .1002	.0678 .0340	.1195 .1089	.0991 .0736	.0353 .0401
D_2	C_1	.0056 .0094	.0044 .0040	.0000 .0018	.0276 .0361	.0225 .0262	.0091 .0068
	C_2	.0581 .0817	.0589 .0632	.0143 .0284	.1139 .1698	.0740 .1028	.0385 .0287
	C_3	.1141 .1336	.1072 .0967	.0394 .0643	.1624 .2023	.1266 .1400	.0807 .0527

2 x 3x 3 x 2 Factorial, Fraction Backordered

FIG. IV-5

As in the initial experiment, since the response measured is a proportion, a transformation is made on the observed values:

$$x = \arc \sin \sqrt{b}$$

where b is the observed fraction backordered. Performing the analysis on the transformed variables yields the following table:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
A	197.8427	1	197.8427	66.1173
B	452.8098	2	226.4049	75.6625
C	1869.6780	2	934.8390	312.2415
D	95.1422	1	95.1422	31.7957
AB	29.6445	2	14.8222	4.9534
AC	26.1979	2	13.0989	4.3775
AD	7.5888	1	7.5888	2.5361
BC	51.6666	4	12.9166	4.3166
BD	31.4286	2	15.7143	5.2516
CD	13.5921	2	6.7960	2.2712
ABC	3.9556	4	.9888	--
ABD	.4625	2	.2312	--
ACD	6.0076	2	3.0038	1.0038
BCD	4.1377	4	1.0344	--
Error	119.6927	40	2.9923	
Total	2909.8473	71		

$F_{1,40,.99} = 7.31,\ F_{2,40,.99} = 5.18,\ F_{4,40,.99} = 3.83$

FIG. IV-6

The last column shows that while some of the first order interactions are statistically significant at the 99% confidence level, the F-ratios for the 4 main effects are several orders of magnitude greater. By far the most significant effect is factor C, the demand group, indicating that there is indeed a difference in the achieved service level for products of different mean demands. The accompanying Two-Way Table of Averages for Factors B and C (Figure IV-8) shows that the higher demand group (Group 1) consistently achieved a higher service level.

Factors A (coefficient of variation) and B (initially specified service level) are next in level of significance. This indicates that the coefficient of variation for the product demands affects the service levels achieved to about the same extent as does the initially specified service level. The relative effects of the various levels of the factors on the outcome can be seen from the accompanying Two-Way Table of Averages for Factors A and B (Figure IV-7).

The least significant main effect is lot size. The Two-Way Table of Averages, Factors C and D (Figure IV-9) shows that the standard or Wilson lot size (D_1) gave better results for all demand groups than did the larger lot size used in the study.

		Specified Service Level		
		B_1 (92%)	B_2 (95%)	B_3 (98%)
Coefficient of Variation	A_1	95.75	96.02	98.31
	A_2	91.49	93.90	97.24

Two-Way Table of Averages, Factors A and B

(Values Shown are Average Observed Service Level, %)

FIG. IV-7

		Demand Group			
		C_1	C_2	C_3	Mean
Specified Service Level	B_1 (92%)	98.69	92.31	87.50	92.83
	B_2 (95%)	99.07	93.73	89.72	94.17
	B_3 (98%)	99.65	97.47	94.94	97.69

Two-Way Table of Averages, Factors B and C

(Values Shown are Average Observed Service Level, %)

FIG. IV-8

		Lot Size	
		D_1	D_2
Demand Group	C_1	99.36	99.01
	C_2	95.75	93.64
	C_3	92.39	89.40

Two-Way Table of Averages, Factors C and D

(Values Shown are Average Observed Service Level, %)

FIG. IV-9

Discussion of the Rationale for Determining the Amount of Simulation Required

Two significant problems in designing a simulation experiment are to determine (1) the length of time the simulation should cover, and (2) the number of replications that should be made in order to draw meaningful conclusions. These aspects are particularly important in monitoring stochastic processes in which the results are autocorrelated, as is the case in this simulation where the measurement of backorders is involved. The autocorrelation of the backorders in the time series generated by the experiment prevents the utilization of "standard" statistical techniques, which apply to independent observations, in determining the number of months of activity that must be observed in order to draw valid conclusions.

The number of months activity that a single replication should include was governed by the period of time for which the performance of the system was of interest, and as discussed above, this was specified as 36 months. The number of replications required, however, cannot be determined in such a straight forward manner, as it must be concluded whether any change in response is due to the occurrence of a particular sequence of demands, or to the experimental conditions (coefficient of variation, capacity constraint, lot size determination, and desired service level) used. The total number of observations required to draw conclusions in a statistical sense was the subject of two papers (7,8), a brief discussion of which follows.

Geisler's approach (8) was to run a preliminary experiment for a number of periods that could be considered to provide an "infinite sample size" and from this determine an estimate of the true value of backorders, an estimate of the standard deviation of the number of backorders, and an estimate of the correlation coefficient for the number of backorders over an empirically determined correlation time. From this, the minimum sample size (number of accumulated backorders) could be determined for a given confidence level and interval, and each experiment could then be run until this number of backorders was accumulated.

The primary disadvantage in Geisler's approach is in the amount of computer time required to run the preliminary experiment for a sufficient period of time to constitute an "infinite sample size," which he arbitrarily assumed to be 500 periods. For a single product, this would not be prohibitive, provided that the number of combinations of experimental parameters was not too large. The estimates of the mean, standard deviation, and correlation coefficient are applicable only for the one set of experimental conditions; a 500 period run would have to be made for each set of parameters. Then since the production-inventory system under consideration involves multiple products, the length of the final experimental runs would be determined by the largest of the minimal sample sizes for the different products, under the set of parameters (coefficient of variation, capacity constraint, lot size, and desired service level) that required the longest individual runs to draw the desired conclusions. For the inventory system under consideration, it was estimated that this procedure would require better than 100 hours of computer time for the preliminary runs in order to determine how many months of data would be required for the final experimental runs.

The approach taken by Fishman (7) was to determine the number of observations required on the autocorrelated time series which would be equivalent to the number of independent observations necessary to obtain the desired degree of confidence in the results. To accomplish this, some heuristic arguments were employed to relate spectral analysis methods to the analysis of autocorrelated time series. This approach, like Geisler's, requires extensive simulation in order to determine the correlation time, and through this, the estimate of the large sample variance.

Both approaches require initial arbitrary decisions for determining the length of preliminary runs, and more seriously, both require a prohibitive amount of computer time in order to determine the sample size required for the final experiments. If the results were sufficiently general in nature, or if the published findings of either method could be applied directly to an inventory system such as the one under study, either method would appear to give a fairly rigorous means of determining sample size (converted to number of months of activity) necessary to interpret results with a given degree of confidence. However, due primarily to the amount of simulation time required to determine the adequate sample size, the use of either method for the particular problem at hand was not feasible.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER STUDY

The conclusions from this study are:

1. The use of the priority method of scheduling the fixed production quantities did not cause a degradation of service from the specified level when the products are considered in aggregate. This is borne out by the results of the initial investigation, presented in Figure IV-1, which shows that the achieved service level for the products taken in aggregate was higher than the level specified, for all combinations of conditions investigated.
2. The use of the priority method of scheduling did permit significant degradation of service for some products (and hence a higher than specified level for others). This is evidenced by the results of the final investigation presented in Figure IV-5, which indicated that the level of service achieved was dependent to a great extent on the mean monthly demand for the product. The Two-Way Table of Averages, Factors C and D (Figure IV-9) shows that for both lot size determinations, the higher demand products achieved higher service levels. Furthermore, Figure IV-8 shows that the product group with the lowest mean monthly demand range consistently received less than the specified degree of protection against stockouts.

What is the reason for the apparent failure of the system to preserve the service levels of some of the products? It was stated in Chapter IV that while the reorder points and lot sizes are computed for the products independently (an outgrowth of some of the approximations employed in Chapter II), the production schedule is ultimately limited by an aggregate constraint. Thus, somewhere in the scheduling algorithm there must be a mechanism which gives proper weighting to the individual products so that the dominant objective of the system is satisfied. This mechanism is the priority. The problem arises because in the scheduling procedure summarized in Chapter III, it is not clear which objective is dominant, the objective of meeting the over-all (aggregate) service level, or the objective of meeting the service levels of the individual products. The procedure of relating the reorder point and expected end-of-month inventory position to some fictitious "absolute zero" inventory level (Step 6 in the scheduling procedure, Chapter III), while introduced to eliminate the ambiguity in the meaning of negative priorities, does in fact establish the dominance of the aggregate service level, since the inventories and reorder points of all products are then linked to a common reference point.

Once the dominance of the aggregate service level is established, the system will naturally favor the higher demand products, because these products have a greater influence in the aggregate service level computation (see definition of service level, Chapter III).

3. The approximation which neglected backorders in determining the unconstrained lot size did not cause a degradation of service level. As indicated in Chapter IV, if backorders were not neglected, an unconstrained lot size larger than the Wilson lot size would result. Thus the model was tested with the standard or Wilson lot size as the smallest value used. The ANOVA Table, Figure IV-6, shows that for the range considered (see Chapter IV) the lot size used does indeed have a significant effect on the achieved service level. Furthermore, it can be seen from Figure IV-9 that the Wilson lot size provided the better service level. This was true for all demand groups.

4. The service level achieved is a function of the coefficient of variation of demand (ratio of standard deviation of demand to mean monthly demand). This aspect was investigated on the basis of the coefficient of variation for the whole population of products. In other words, the model was tested with what might be considered as two sets of products which were identical in all respects except that one set of products experienced more volatile demands than the other set. The case in which product demands were more volatile (higher coefficient of variation) consistently showed lower resultant service levels (See Figure IV-7). Since the reorder point computation takes into account the coefficient of variation of demand in setting the safety stock, the poorer performance in the presence of

higher demand volatility is probably due to the inability of the system to respond quickly in two respects:

(a) The reorder point may be reached shortly after a review, but can't be recognized until the next review time.

(b) Once the "ideal" reorder point is reached (priority = 1), the product may not be scheduled due to competition for production time.

Item (a) above is a matter of stated policy, so any degradation of service due to periodic review cannot be attributed to the approximations employed in developing the algorithm. It can therefore be reasoned that only a part of the variability shown in the ANOVA Table (Figure IV-6) as being due to the coefficient of variation is caused by the approximations and/or simplifications utilized; the remainder would be present even if the exact mathematical solution were used.

5. The tightness of the capacity constraint, i.e., the number of hours of allowed production in relation to the number of hours required to meet mean aggregate demand, probably has the greatest effect of any factor considered on the service level, as evidenced by the results of the initial investigation (ANOVA Table, Figure IV-2). However, for the case in which demands are considered stationary with time, the importance is diminished by the fact that a constraint other than the number of hours required to manufacture for mean aggregate demand will, over

the long run, result in either a depletion or an excess of inventory, neither of which is desirable. Since the results of the initial investigation indicated that the desired service level could be met, at least in the aggregate, with capacity sufficient to satisfy mean demand, the problem becomes one of inventory balance rather than sufficient capacity.

The range in which the month-to-month capacity is allowed to vary is however of some importance. In this study, this "tolerance" was arbitrarily set at 10%, which was considered reasonable for the problem framework set forth in Chapter I. However, from a rigid computational standpoint, it can be seen that this range must be sufficient to accommodate the time required to manufacture the largest fixed lot size (in terms of manufacturing hours) of the population of products considered. Otherwise there would exist the possibility of not being able to reach a fixed lot size solution within the capacity range. This is, however, a trivial point, as in reality if this predicament arose some manual adjustment could easily be made to avoid the impasse. Furthermore, the likelihood of a fixed lot size solution not being attainable diminishes as the number of products increases, since any given lot size is then a smaller percentage of the capacity necessary to meet mean aggregate demand.

6. The initially specified service level does have a significant effect on the service level actually achieved. From the Two-Way Table of Averages, Factors B and C (Figure IV-8), it can be seen that the specified service levels and achieved service levels compare as follows:

Z_0 (specified)	Z (achieved)
92%	92.83%
95%	94.17%
98%	97.69%

Hence, on the average, the system is seen to "track" fairly well, for the range of service levels investigated. However, it must be remembered that the production capacity is rigidly constrained within a range, so that for very low (say under 90%) or very high (over 99%) specified service levels, the ability of the system to track will be encumbered.

The degree of influence of the specified service level is not the same for all demand groups however. It may also be observed from Figure IV-8 that for the group of products with the higher mean monthly demands (Group 1), the achieved service level varied over a range of only 1 percentage point (98.69% to 99.65%) while the specified service level had a range of 6 percentage points (92% to 98%). On the other hand, the achieved service level for the group of products with the lower mean monthly demands varied over a range of 7 percentage points (87.50% to 94.94%), while the range of the specified service level was the same as indicated above. This further illustrates the apparent bias of the system to service the higher demand products.

Recommendations for Further Study

Probably the most severe limitation on the scheduling model presented in this paper is that it is applicable only in the static demand situation. However, the system could be extended to handle the dynamic demand problem by incorporating one of the available forecasting techniques to obtain an estimated demand for each product, and hence the estimated aggregate production capacity. Scheduling priority would then be based on forecasted demands instead of mean monthly demands, as was the case in this paper.

If an n-period planning horizon is employed, one of the more sophisticated smoothing and scheduling techniques such as that presented in (18) is justified.

APPENDIX A

Derivation of Shortage Factor

The shortage factor, F_R, is developed below for demand occurring from a normal distribution. (Notation is that used in Chapters II and III).

For any continuous distribution of demand, the expected number of backorders that will be incurred during the lead time period, given that the inventory is I at the beginning of the period is:

$$N_I = \int_I^{\infty} (S_L - I) f(S_L)\, dS_L$$

For a normally distributed S_L, this becomes:

$$N_I = \int_I^{\infty} (S_L - I) \frac{1}{\sigma_L \sqrt{2\pi}} \exp\left[-\tfrac{1}{2}\left(\frac{S_L - \bar{S}_L}{\sigma_L}\right)^2\right] dS_L$$

Making the change of variable:

$$t = \frac{S_L - \bar{S}_L}{\sigma_L} \text{ and } t_I = \frac{I - \bar{S}_L}{\sigma_L}$$

$$N_I = \int_{t=t_I}^{\infty} (\sigma_L t + \bar{S}_L - \sigma_L t_I - \bar{S}_L) \frac{1}{\sqrt{2\pi}} \exp\left[-\tfrac{1}{2}t^2\right] dt$$

$$= \sigma_L \int_{t=t_I}^{\infty} t \cdot \frac{1}{\sqrt{2\pi}} \exp\left[-\tfrac{1}{2}t^2\right] dt - t_I \sigma_L \int_{t=t_I}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\tfrac{1}{2}t^2\right] dt$$

but

$$\frac{1}{\sqrt{2\pi}} \exp\left[-\tfrac{1}{2}t^2\right] = \phi(t)$$

and

$$\int_{t_I}^{\infty} \phi(t)\ dt = \psi(t_I)$$

Hence

$$N_I = \sigma_L \cdot \int_{t_I}^{\infty} t \cdot \phi(t)dt - t_I \sigma_L\, \psi(t_I)$$

and since

$$\int_{t_I}^{\infty} t \cdot \phi(t)dt = \phi(t_I)$$

$$N_I = \sigma_L\, \phi(t_I) - \sigma_L t_I\, \psi(t_I) = \sigma_L \left[\phi(t_I) - t_I\, \psi(t_I)\right]$$

This is the expected number of backorders that will be incurred during the lead time period, given that the inventory is I at the beginning of the period.

Assume that the inventory I was the reorder point R, i.e., $N_I = N_R$. Let the expected number of backorders be expressed as the product of the standard deviation of lead time demand and a term called "shortage factor", F_R.

$$N_R = \sigma_L F_R$$

By the expression for the expected number of backorders derived above:

$$N_R = \sigma_L \left[\phi(t_R) - t_R\, \psi(t_R)\right]$$

or

$$F_R = \phi(t_R) - t_R\, \psi(t_R)$$

Sample Calculation

The following example illustrates the procedure for calculating the reorder point, as presented in Chapter III.

Let a hypothetical product have the following characteristics:

$$\overline{S}_L = 133$$

$$\sigma_L = 30$$

Suppose the lot size has previously been determined:

$$Q = 897$$

If the desired service level is .95, then by equation (2) Chapter III:

$$Z_R = 1 - \frac{897}{133}(.05) = .663$$

By equation (3):

$$F_R = \frac{1}{\mu}(1-Z_R) = \frac{\overline{S}_L}{\sigma_L}(1-Z_R) = \frac{133}{30}(.337) = 1.49$$

By equation (4):

$$F_R = \phi(t_R) - t_R\,\psi(t_R)$$

Using trial and error solution, let

$$t_R = -1.48$$

$$F_R = \phi(-1.48) - (-1.48)\,\psi(-1.48)$$

$$= .133 + 1.38 = 1.513 \qquad (>1.49)$$

As a second trial, let

$$t_R = -1.46$$

$$F_R = .137 + 1.355 = 1.492$$

Therefore $t_R = -1.46$ yields approximately $F_R = 1.49$, and by equation (4):

$$R = \overline{S}_L + t_R\sigma_L = 133 - 1.46\ (30) = 90$$

Observe that the reorder point obtained by taking the order period and lead time into account is much less than the expected lead time demand. As a matter of fact, using the method presented, it is possible to obtain a negative reorder point, i.e., back-orders will be incurred with certainty. The coefficient of variation μ, and/or the number of orders per year could be so low that backorders are deliberately incurred while still maintaining the desired overall service level Z_0.

It is interesting to compare the reorder point obtained above with that obtained by the "standard" procedure.

The normal method of using a service level criterion to establish the reorder point is:

$$R = \overline{S}_L + K\sigma_L$$

where

$\overline{S}_L$ = expected lead time demand

σ_L = Standard deviation of lead time demand

K = normal deviate corresponding to the desired service level.

Suppose it is desired to maintain an overall service level $Z_0 = .95$. Using the standard method, this requires a safety stock of:

$$K\sigma_L = 1.65\sigma_L$$

or

$$R = \bar{S}_L + 1.65\,\sigma_L$$

For the hypothetical product discussed above:

$$R = 133 + 1.65\ (30) = 183$$

Thus, the reorder point is seen to be much higher than that obtained when the order period and lead time are considered. However, if the product is reordered when the inventory reaches this reorder point, there is only a 5% chance that a stockout will be incurred. But the remainder of the time, i.e., before the reorder point is reached, the probability of a stockout is 0. Thus, the over-all service level is greater than the 95% specified.

Suppose, in the example being considered, $\bar{S}_L = \bar{S}$, i.e., the lead time is one month. Then the average number of times the product is ordered per year is:

$$\frac{12\ \bar{S}}{Q} = \frac{12 \cdot 133}{897} \approx 1.78$$

Since the lead time is one month, this product is subject to stockout only $\frac{1.78 \times 1}{12}$ of a year. The other $\frac{12 - 1.78}{12}$ of the year there is no risk of stocking out and the over-all probability of incurring a backorder is:

$$P = (\frac{12 - 1.78}{12})(0) + (\frac{1.78}{12})(.05) = .0074$$

or the actual over-all service level is:

$$Z_0 = 1 - P = 99.26\%$$

Thus, it is seen that by specifying the reorder point in the standard manner using a service level criterion, a higher level of service is achieved, but at the expense of carrying additional safety stock. Furthermore, a different level of protection may be afforded each product, depending upon the number of times the product reaches the reorder point during the year (the number of order cycles), and the length of the lead time.

APPENDIX B

Design of the Initial Experiment

As discussed in Chapter IV, the initial experiment proposed was a half replicate of a 2^4 factorial. The defining contrast is:

$$I = ABCD$$

The aliased effects are then:

$$A = BCD$$
$$B = ACD$$
$$C = ABD$$
$$D = ABC$$
$$AB = CD$$
$$BC = AD$$
$$AC = BD$$

indicating that the three-factor interactions are the aliases of the main effects, which can be tolerated as long as the three factor interactions are not large. Note also that the two factor interactions have other two factor interactions as aliases, which is tolerable if interactions are not too large, and as long as no attempt is made to extract conclusions about specific interactions from the results.

When the defining contrast is ABCD, the two half-replicates are:

a	(1)
b	ab
abc	bc
c	ac
bcd	abcd
acd	cd
d	ad
abd	bd

Selecting the first half replicate, the eight treatment combinations necessary are:

	Factors Levels
a	A_2 B_1 C_1 D_1
b	A_1 B_2 C_1 D_1
abc	A_2 B_2 C_2 D_1
c	A_1 B_1 C_2 D_1
bcd	A_1 B_2 C_2 D_2
acd	A_2 B_1 C_2 D_2
d	A_1 B_1 C_1 D_2
abd	A_2 B_2 C_1 D_2

where the factors represent:

A = coefficient of variation
B = capacity constraint
C = lot size
D = specified service level

and the levels of the factors are as specified in Chapter IV.

BIBLIOGRAPHY

1. Buchan, J., and E. Koenigsberg, Scientific Inventory Management, Englewood Cliffs, N. J.: Prentice-Hall, 1963.

2. Churchman, C. W., R. L. Ackoff, and E. L. Arnoff, Introduction to Operations Research, New York, N. Y.: John Wiley & Sons, Inc., 1957.

3. Cox, D. R., Planning of Experiments, New York, N. Y.: John Wiley & Sons, Inc., 1958.

4. Davies, O. L., ed., The Design and Analysis of Industrial Experiments, New York: Hafner Publishing Company, 1963.

5. Elmaghraby, S. E., and R. T. Cole, "On the Control of Production in Small Job Shops," The Journal of Industrial Engineering, Vol. XIV, No. 4, July-August, 1963, pp. 186-196.

6. Fetter, R. B., and W. C. Dalleck, Decision Models for Inventory Management, Homewood, Illinois: Irwin, 1961.

7. Fishman, G. S., "Problems in the Statistical Analysis of Simulation Experiments: The Comparison of Means and the Length of Sample Records," Memorandum RM-4880-PR, The RAND Corporation, February, 1966.

8. Geisler, M. A., "The Sizes of Simulation Samples Required to Compute Certain Inventory Characteristics with Stated Precision and Confidence," Management Science, January, 1964, pp. 261-286.

9. Hadley, G., and T. M. Whitin, Analysis of Inventory Systems, Englewood Cliffs, N. J.: Prentice-Hall, 1963.

10. Hicks, C. R., Fundamental Concepts in the Design of Experiments, New York: Holt, Rinehart and Winston, 1964.

11. Hoffman, A. J., and W. Jacobs, "Smooth Patterns of Production," Management Science, October, 1954, pp. 86-91.

12. Holt, C. C., F. Modigliani, J. F. Muth, and H. A. Simon, Planning Production, Inventories, and Work Force, Englewood Cliffs, New Jersey: Prentice-Hall, 1962.

13. Holt, C. C., F. Modigliani, and H. A. Simon, "A Linear Decision Rule for Production and Employment Scheduling," Management Science, October, 1955, pp. 1-30.

BIBLIOGRAPHY (cont'd)

14. Manne, A. S., "A Note on the Modigliani-Hohn Production Smoothing Model," Management Science, Vol. 3, No. 4, 1957, pp. 371-379.

15. Mikes, D. R., "A Stochastic Programming Model to Perform a Production Smoothing Function," Master's Thesis in Industrial Engineering, Bethlehem, Pa.: Lehigh University, June, 1965.

16. Miller, I., and J. E. Fruend, Probability and Statistics for Engineers, Englewood Cliffs, New Jersey: Prentice-Hall, 1965.

17. Modigliani, F., and F. E. Hohn, "Production Planning Over Time and the Nature of the Expectation and Planning Horizon," Econometrica, January, 1955, pp. 46-66.

18. O'Malley, R. L., S. E. Elmaghraby, and J. W. Jeske, Jr., "An Operational System for Smoothing Batch-Type Production," Management Science, June, 1966, pp. 433-449.

19. Wagner, H. M., and T. M. Whitin, "Dynamic Version of the Economic Lot Size Model," Management Science, October, 1958, pp. 89-96.

20. Whitin, T. M., Theory of Inventory Management, Princeton, New Jersey: Princeton University Press, 1957.

21. Winters, Peter R., "Constrained Inventory Rules for Production Smoothing," Management Science, July, 1962, pp. 470-481.

VITA

PERSONAL HISTORY

Name:	David Glenn Jackson
Date of Birth:	January 24, 1936
Place of Birth:	St. Louis, Missouri
Parents:	Merle M. and Jean L. Jackson
Wife:	Sally W. Jackson
Children:	Nancy Elizabeth

EDUCATIONAL BACKGROUND

Webster Groves High School	Graduated 1954
University of Missouri Bachelor of Science in Electrical Engineering	Graduated 1961
Lehigh University Candidate for Master of Science in Industrial Engineering	1967

HONORS

Eta Kappa Nu

PROFESSIONAL EXPERIENCE

IBM Corporation
Owego, New York
Engineer - Radar Systems Development

Western Electric Company, Inc.
Columbus, Ohio
Engineer/Development Engineer - Electronic Switching Systems

Western Electric Company, Inc.
Princeton, New Jersey
Research Engineer - Lehigh Master's Program

PROFESSIONAL ORGANIZATIONS

Institute of Electrical and Electronic Engineers