WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

OPTIMAL SEQUENTIAL PRODUCTION UNDER UN-

CERTAINTY

by

Warren H. Hausman

February 1971 512-71

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

OPTIMAL SEQUENTIAL PRODUCTION UNDER UNCERTAINTY

by

Warren H. Hausman

February 1971 512-71

A B S T R A C T

Consider a product whose actual demand is stochastic but whose average demand rate follows a time-varying pattern with known shape over its life cycle. Initially, the level of the demand pattern may be quite uncertain; as experience is gained, however, the level of demand becomes somewhat more predictable. The general situation described here is particularly suitable for items such as replacement parts for a durable product which undergoes frequent model and/or substantive style changes. In this paper, it is assumed that revised forecasts of total demand over the entire product life cycle become available each period. Backorders for (only) one period are allowed, at a cost. Setup and variable production costs are non-decreasing over time. The manufacturer wishes to minimize the discounted expected cost over the item's life cycle. A dynamic programming formulation of the problem is presented, and computational aspects are discussed.

I. INTRODUCTION

Consider the following general problem: A manufacturer must select the timing and magnitude of production quantities for a product whose actual demand is stochastic but whose average demand rate has a time -varying pattern with known shape over the entire product life cycle. The actual level of the demand pattern is initially unknown, but as experience is gained, it becomes somewhat more predictable. Setup and variable production costs are non-decreasing over time; backorders are allowed for only one period, at a cost. It is assumed that some forecasting mechanism (either human or mechanical) exists which generates a revised forecast each period of total demand over the life of the product. The manufacturer's goal is to meet total product demand at minimum discounted expected cost. This paper describes some alternative forecasting methods which exhibit certain necessary statistical data-generating properties, and proceeds to integrate one statistical forecast revision process with a framework of decision choices and costs. The resulting system is formulated and solved as a dynamic programming problem.

Motivating Example

The methodology described here may be applicable in principle to any number of products with a relatively well-defined life cycle. However, the primary focus of the methodology is toward products such as spare parts for durable goods (consumer or industrial) in which the following characteristics are present:

(1) Wearout characteristics of the part and model obsolescence of the durable good create a situation of a time-varying average demand rate whose shape or pattern may be specified.

(2) Substantial uncertainty about the level of the demand pattern for the part is present initially.

(3) High setup costs indicate the desirability of large production quantities.

(4) Tracking of demands over time allows improved forecasts to be made as time passes.

Previous Research

The deterministic version of this problem has been formulated by Wagner and Whitin [18] as a dynamic programming problem. Moreover, under certain cost assumptions, they demonstrated a computationally important "planning horizon" result ensuring that in the optimal schedule one would receive a replenishment quantity only when on-hand inventory had fallen to zero. The stochastic version of the problem can be further partitioned into models assuming independent stochastic demands in each period, as opposed to models allowing for some type of sequential dependency. We will be concerned exclusively with ways of allowing for dependencies in demands.

Concerning stochastic models which allow for dependent demands, some previous work focuses primarily on the nature of the statistical data-generating process for sequential forecast revisions, while other work combines these concepts with an associated framework of decision alternatives and costs. Hertz and Shaffir [10] describe a straight-forward approach toward forecast revisions which is in wide practical use. Through analysis of historical data, some underlying pattern of average demand over the entire planning

horizon is obtained, and then actual cumulative demands to date are simply extrapolated to obtain a revised estimate of total demand over the horizon. In their case they were dealing with seasonal style goods, and the horizon was a few months; also, their data supported use of a cumulative Normal distribution as the underlying pattern of average demand over the short season.[1] Neither [10] or a similar study [11] integrated the forecast revision process with the related sequential decision and cost framework in an optimal manner, however. Murray and Silver [14] present a Bayesian model for forecast revision (based on an underlying binomial share-of-market model) which again presumes some underlying pattern of average demand over the entire horizon. In their case the initial forecast of total demand made at the outset continues to carry some weight as demands are accumulated, in contrast to the pure demand-to-date extrapolation described above. Chang and Fyffe [3] also describe a method of generating forecast revisions which, derived from work in optimal linear control by Kalman [12] and Shaw [16], continues to place appropriate (Bayesian) weight on the initial forecast. However, Chang and Fyffe explicitly omit consideration of "... the associated decision problems ..."[2] involved in their motivating example. In contrast, Murray and Silver <u>do</u> imbed their forecast revision process in an optimizing framework of sequential decisions and costs; they are concerned with ordering and reordering style goods for retail sale. In principle, the Murray and

---

[1] A similar approach was used by Ichihara, Neidell, Oliver and William-Powlett [11] as summarized in Hausman [6].

[2] Reference [3], p. 3.

Silver formulation could be used directly for the problem under consideration here. However, if their use of a binomial share-of-market approach to the demand data-generating process is not a natural representation of the situation, then it is not clear how one should "scale" their formulation.[3]

In related research, Brown, Lu and Wolfson [2] present a general framework in which a number of distinct demand data-generating processes are considered, and in which Markovian transitions from one demand-generating state to another take place. The actual demand observed in a period is used to compute a posterior probability distribution over the different states of demand types. The model to be presented below may be viewed as a special case of their more general framework.

All of the research described above can be characterized in a broad manner as allowing for Bayesian updating of prior demand estimates, based on actual demands recorded to date (in Hertz, Shaffir and related cases the initial prior distribution may be considered infinitely diffuse, so that only sample information is used subsequently). In related research, this writer [6] found that it was generally possible to characterize both mechanistic extrapolating forecasting methods and complex, ambiguous "black box" human forecasters in a useful statistical manner. Specifically, a small number of different forecast revision processes and associated historical data were studied, with the following conclusion: "The data are generally, although not entirely, consistent

---

[3] For example, if we are dealing with a manufacturer's replacement part demands, the manufacturer may be the sole source of supply; moreover, wearout considerations may dominate the demand pattern. In this situation the "underlying market-share" approach of [14] is not a natural representation of the situation.

with the hypothesis that ratios of successive forecasts [of the same unknown] are independent variates; their distribution appears to be ... Lognormal."[4] Only one of the forecasting mechanisms out of the several studied was a pure mechanical extrapolation; the others involved human beings who presumably processed a huge amount of potentially-relevant information between forecasts, and produced a revised point estimate on the basis of their previous experience in attempting to deal with the complex problem at hand.[5] Thus, although the model to be presented below, like previous ones described earlier, uses recent demand data to update forecasts in a purely mechanical manner, in the light of [6] it may also be of some use in cases where a more complex "black-box" forecaster is active and/or in cases where continuing Bayesian weight is placed on the initial forecast.

II. A MODEL OF THE FORECAST DATA-GENERATING PROCESS

Following [3], [6], [10], [11], and [14], we presume knowledge of the shape (but not the level) of some underlying pattern of average demand over the entire product life cycle involving N periods. Let:

R_j = cumulative expected fraction of total demand received through period j; $j = 1, \ldots, n$; $1 \geq R_j \geq R_{j-1} \geq 0$.

X_j = revised forecast of total demand over the entire planning horizon, with the forecast made at the beginning of period j; $j = 1, \ldots N$

---

[4] Reference [6], p. 93.

[5] The various individuals were forecasting (and reforecasting) agricultural vegetable crop supply of a particular company, total U.S. vegetable crop supply, and earnings-per-share for firms in three industries (electricity, oil and drugs). The mechanical forecasting scheme was forecasting total wholesale demands of cruise-season women's dresses. See [6] and [4] for further details.

X_{N+1} = actual total demand for the product over periods 1 through N.

D_j = actual demand during period j; j = 1, ..., N; $D_j \geq 0$.

S_j = cumulative demand received through period j; j = 1, ..., N;

$$(1) \quad S_j = \sum_{i=1}^{j} D_i$$

Also, by definition $X_{N+1} = S_N$.

Pure Demand Extrapolation: A Completely Diffuse Prior

Now, following [6],[10] and[11], consider a simple demand-to-date extrapolation as our forecasting scheme; then

$$(2) \quad X_j = S_{j-1}/R_{j-1}, \qquad j = 1, \ldots, N + 1.$$

Also let:

$$(3) \quad Z_j = X_{j+1}/X_j, \qquad j = 1, \ldots, N.$$

Substantial evidence has been presented in [6] to assume, as a first approximation, that the variates $\{Z_j\}$ representing ratios of successive forecasts are independent two-parameter Lognormal variates; i.e.,

$$Z_j \sim f_{LN}(Z_j|\mu_j, \sigma_j) = \frac{1}{\sqrt{2\pi}\,\sigma_j Z_j} e^{-(\log Z_j - \mu_j)^2/2\sigma_j^2},$$

$$0 < Z_j < \infty, \quad j = 1, \ldots, N,$$

with $\mathrm{Cov}(Z_j, Z_k) = 0$ if $j \neq k$.

We shall subsequently see that some (fortunately minor) approximation must necessarily be involved with this assumed distribution.

It is presumed that some method to produce an initial forecast X_1 is available; however, we assume in this section that this method has so little to recommend it that as soon as one period has been completed, equation (2) is the sole source of the revised forecast X_2, and the initial forecast is completely disregarded subsequently. The quantities $\{R_j\}$ are specified constants; moreover, at any period j, previous cumulative actual demands $\{S_{j-1}\}$ are known. Substituting (2) into (3) and using the identity

$S_j = S_{j-1} + D_j$,

$$(4)\quad Z_j = \frac{(S_{j-1}+D_j)/R_j}{S_{j-1}/R_{j-1}} = \frac{R_{j-1}}{R_j} + \frac{R_{j-1}}{R_j S_{j-1}} D_j,$$

where $S_0/R_0 = X_1$, thereby defining the ratio R_0/S_0. From equation (4) there is a linear relationship ($Z_j = a_j + b_j D_j$) between the random variables D_j and Z_j, with both parameters a_j and b_j in the range (0,1) inclusive. Since we assume no returns, D_j is non-negative; therefore, in each period j, the constant term in (4) is a lower bound on Z_j, and thus Z_j cannot be a variate distributed precisely according to the two-parameter Lognormal distribution indicated above. However, since the constant term ($a_j = R_{j-1}/R_j$) is known, the <u>three-parameter</u> Lognormal distribution described in reference [1] may be considered. In this distribution the quantity ($Z_j - a_j$) is distributed as two-parameter Lognormal, with the third parameter merely shifting the distribution upward so that the lower bound of the distribution of Z_j is a_j rather than zero. Intuitively, the reason for the excellent fit of the two-parameter Lognormal distribution to the apparel forecasting data cited in [6] is due to the fact that the dispersion of the distribution of Z_j was not excessive relative to the lower bound. For example, in the apparel data the ratio (R_{N-1}/R_N) was approximately .90, but the dispersion in Z_{N-1} was consistent with .90 being an approximate lower bound.[6] In other words, because of the limited dispersion of Z_j,it turned out that the lower bound was very close to the zeroth percentile of the distribution in its two-parameter form. This state of affairs may not always exist; for example, if the planning horizon were divided into a very large number of periods so that the ratio R_{j-1}/R_j were very close to unity, then the lower bound may be more binding, and Z_j may not fit the two-parameter Lognormal distribution.

---

[6] See [6], Figure 10 (p. 103).

It should be emphasized that as long as the value of the ratio R_{j-1}/R_j is known, one may always work with the three-parameter Lognormal by making the transformation

(5) $$Z_j' = Z_j - (R_{j-1}/R_j)$$

and then estimating new[7] parameters μ_j' and σ_j' from the transformed data. We present the above discussion solely to illustrate that previous results demonstrating approximate two-parameter Lognormality and independence of ratios of successive forecasts remain approximately valid.

Proceeding with equation (4), if Z_j is three-parameter Lognormally distributed with parameters μ_j', σ_j' and shift parameter (R_{j-1}/R_j), then it can be easily derived[8] that D_j is distributed as two-parameter Lognormal with parameters $(\mu_j' + \log(R_{j-1}/(R_j S_{j-1})), \sigma_j')$:

$$D_j \sim f_{LN}(D_j | \mu_j' + \log(R_{j-1}/(R_j S_{j-1})), \sigma_j')$$

Also, the forecast-revision equation (2) may be rewritten to be a funetion of D_j:

(6) $$X_{j+1} = S_j/R_j = (S_{j-1} + D_j)/R_j$$

We also presume that sufficient historical demand data of relevant comparability is available so that the necessary estimates of R_j, μ_j' and σ_j' can be made .

---

[7] These parameters are not equal to the previously-cited parameters μ_j and σ_j.

[8] See reference [1] for such a derivation.

At this point we have fully described the statistical data-generating process associated with a pure demand-to-date extrapolation forecast. Now this process is integrated with the decision and cost framework posed at the outset.

III. DYNAMIC PROGRAMMING FORMULATION OF COMPLETE PROBLEM

For ease of reference we include previously-defined symbols in the following list. Let:

N = number of periods into which total product life-cycle is divided

A_j = setup cost in period j; $A_j \le A_{j+1}$, $j = 1, \ldots, N$

c_j = unit production cost in period j; $c_j \le c_{j+1}$, $j = 1, \ldots, N$

X_j = revised forecast of total demand over the entire planning horizon, with the forecast made at the beginning of period j; $j = 1, \ldots, N$

Y_j = inventory on hand minus backorders at beginning of period j; $j = 1, \ldots, N$

$$b_j(Y_j) = \begin{cases} 0 \text{ if } Y_j \ge 0 \\ \text{cost of } (-Y_j) \text{ units backordered at beginning of period } j, \text{ if } Y_j < 0 \end{cases}$$

D_j = actual number of units demanded during period j

S_j = cumulative demand through period j inclusive;

$$S_j = \sum_{i=1}^{j} D_i$$

R_j = cumulative expected fraction of total demand received through period j inclusive; $0 \le R_j \le R_{j+1} \le 1$

α = present value discount factor for one time period (risk-free rate)

e = unit salvage value of item after period N

P_j = number of units produced in period j (the decision variable):

$P_j \geq 0.$

$$\delta(P_j) = \begin{cases} 0 \text{ if } P_j = 0 \\ 1 \text{ if } P_j > 0 \end{cases}$$

Note: backorders must be filled one period after they are incurred.

The Demand Data-Generating Process

Based on earlier discussion we shall assume that demand in each period is a conditionally independent two-parameter Lognormal variate with parameters as shown (R_j, μ_j' and σ_j' are estimated from other relevant historical data):

$$D_j \sim f_{LN}(D_j | \mu_j' + \log(R_{j-1}/(R_j S_{j-1})), \sigma_j').$$

The Forecast Revision Data-Generating Process

Again from earlier discussion we assume that successive forecasts $\{X_{j+1}\}$ of total demand are generated by equation (6), reproduced here for convenience:

$$(6) \qquad X_{j+1} = S_j/R_j = (S_{j-1} + D_j)/R_j \text{ , } j = 1, \ldots, N.$$

Dynamic Programming Formulation

At the beginning of each period j the state variable will be the two-dimensional vector (X_j, Y_j) whose first element is the latest revised forecast of total demand, and whose second element represents current inventory

less backorders. (We assume all production is completed within a period so there is never any inventory on order.) Following the usual dynamic programming approach, we define the return function as follows:

$g_j(X_j,Y_j)$ = minimum discounted expected cost from period j on through period N+1, given state (X_j,Y_j) obtains at the beginning of period j and optimal decisions are made subsequently.

At period N+1 any remaining inventory contributes a negative cost (salvage value) of e per unit. Recalling that all backorders must be filled within one period, if demand in period N is larger than inventory plus any production in that period, then a final production quantity of size $(-Y_{N+1})$ must be produced in period N+1 to satisfy backorders incurred. Thus at period N+1,

$$g_{N+1}(X_{N+1},Y_{N+1}) = \begin{cases} -e(Y_{N+1}) \text{ if } Y_{N+1} \geq 0 \\ A_{N+1} + c_{N+1}(-Y_{N+1}) \text{ if } Y_{N+1} < 0 \end{cases} \tag{7}$$

Now in general, for period j, the following recurrence relation applies:

$$g_j(X_j,Y_j) = \min_{\substack{P_j \geq 0, \\ P_j \geq -Y_j \text{ if } Y_j < 0}} \{b_j(Y_j) + \delta(P_j)(A_j + c_jP_j) + \alpha \int_{D_j=0}^{\infty} g_{j+1}(X_{j+1},Y_{j+1})\, f_{LN}(D_j|\,.\,,\,.)dD_j\}, \tag{8}$$

where from equation (6)

$$X_{j+1} = (S_{j-1} + D_j)/R_j \tag{9}$$

and by definition, the following inventory balance equation holds:

$$(10) \qquad Y_{j+1} = Y_j + P_j - D_j.$$

The first term on the right-hand side of (8) represents a charge for any backorders if present; the next term charges setup and production costs for any production in period j; and the final term represents the discounted expected[9] cost associated with all future periods, the expectation being performed with respect to the random variable D_j. By (9) and (10), (8) may be rewritten as follows:

$$(11) \qquad g_j(X_j, Y_j) = \underset{\substack{P_j \geq 0, \\ P_j \geq -Y_j \text{ if } Y_j < 0}}{\text{Min}} \{ b_j(Y_j) + \delta(P_j)(A_j + c_j P_j)$$

$$+ \alpha \int_{D_j=0}^{\infty} g_{j+1}((S_{j-1}+D_j)/R_j,\ Y_j+P_j-D_j)\ f_{LN}(D_j|\cdot\ ,\ \cdot)\ dD_j\},$$

for $j = 1, \ldots, N$.

In the usual manner equations (11) and (7) may be solved recursively, starting with equation (7) and working backward until j = 1. The optimal decision rule at the j^{th} stage, $P_j^*(X_j, Y_j)$, is tabulated as a function of the state vector and stored for reference. The <u>set</u> of optimal decision rules $\{P_j^*(\cdot\ ,\ \cdot)\}$ constitutes an <u>optimal</u> <u>policy</u>. Finally, given an initial forecast X_1 and initial inventory $Y_1=0$, the return function $g_1(X_1, Y_1=0)$ represents minimum discounted expected costs associated with following the optimal policy from periods 1 through N.

[9] Note that the parameters of the distribution of D_j include R_{j-1} and R_j, which are known constants, and S_{j-1}, which equals $X_j R_{j-1}$.

Computational Aspects

The state space of above formulation has dimensionality two; this represents a tractable[10] dynamic programming formulation. That is, one may form a grid over the two-dimensional state space, perform the indicated optimization operation of equations (11) and (7) by complete enumeration, and program the entire operation on a computer (see [5] for further discussion). Further, if specific assumptions concerning the various cost elements are made, one may take advantage of previous work (see reference [15]) to prove rigorously that the form of the optimal policy is of the (s,S) type. Specifically, under appropriate assumptions the optimal policy would be as follows: if current inventory less backorders Y_j is less than some quantity $s_j(X_j)$, then one should "order up to" the quantity $S_j(X_j)$; otherwise, do not produce any units in period j. Such a proof will not be pursued here, however.

Estimation of the $\{R_j\}$ quantities could be performed directly by a manager familiar with related product demand growth curves, as suggested in [14]. Alternatively, depending on the nature of the product, a Gompertz curve or a Weibull curve for total demand may be applicable.[11]

Finally, it should be emphasized that the forecasting model studied in detail here assumed a completely diffuse prior distribution about the initial point forecast X_1 of total demand for the product over its entire life cycle. There may be situations in which it is appropriate to maintain some reliance

---

[10] In contrast to a computationally-infeasible many-dimensional state space formulation for a related multiproduct problem; see [9].

[11] Se references [13] and [7], [8] respectively (and their bibliographies) for situations in which each of these distributions have been successfully used to describe demand curves over a product's life cycle.

on the initial estimate X_1 even when current demands depart substantially from its implications. In this case the Bayesian procedures of Murray and Silver [14], Chang and Fyffe [3], or of Brown, Lu, and Wolfson [2] could potentially be adapted to our equations (7) and (11). It is important to recognize that in this situation, the mechanics for updating the forecast based on current demand D_j must be completely specified.[12] That is, in equation (11) one must be able to express the next forecast $\{X_{j+1}\}$ as some function of D_j. Only when this is possible can one deal solely with the distribution of D_j as is done in equation (11). Alternatively, if there exists a "black box" forecaster in the spirit of [6] whose ratios of successive forecasts $Z_j = X_{j+1}/X_j$ follow some Lognormal relationship, it would be necessary to estimate the joint probability density function over the variables (Z_j, D_j). This could be a difficult estimation problem to perform in practice, since the two variates would clearly not be independent, and the amount of dependency is a function of how much Bayesian weight is continued to be placed on the initial forecast X_1 made at the outset.

IV. CONCLUSION

A model has been formulated for a problem in which a product has a known time-varying <u>pattern</u> of variation in average demand, but whose level of demand is unknown and whose actual demand in any period is stochastic. Two versions of the associated forecasting problem have been considered: One involving a completely diffuse prior distribution about the initial point estimate and mechanical updating by extrapolation, versus one allowing for

---

[12] As they are in references [2], [3] and [14].

Bayesian updating of an initial estimate. The former framework has been integrated with the relevant sequential decision and cost framework to create a dynamic programming formulation which would produce optimal production decisions.

REFERENCES

1. Aitchison, J., and J.A.C. Brown, The Lognormal Distribution. Cambridge, England, Cambridge University Press, 1957.
2. Brown, G.W., Lu, J.Y., and R.J. Wolfson, "Dynamic Modeling of Inventories Subject to Obsolence", Management Science, Vol. 11, No. 1 (Sept. 1964), pp. 51-63.
3. Chang, S.H., and Fyffe, D.E., "Estimation of Forecast Errors for Seasonal-Style-Goods Sales," Management Science (forthcoming).
4. Crowell, R.A., "Earnings Expectations, Security Valuation, and the Cost of Equity Capital," unpublished Ph.D. Dissertation, Sloan School of Management, M.I.T., September 1967.
5. Hadley, G., Nonlinear and Dynamic Programming, Addison-Wesley, Reading, Mass., 1964.
6. Hausman, W.H., "Sequential Decision Problems: A Model to Exploit Existing Forecasters," Management Science, Vol. 16, No. 2 (October 1969), pp. 93-111.
7. Hausman, W.H., and Kamins, M., "Early Failures in Automobile Parts: A Background Study in Reliability," The RAND Corporation, Rand Memorandum RM-4002-PR, March 1964.
8. Hausman, W.H., and Kamins, M., "The Reliability of New Automobile Parts," pp. 863-871 in Annals of Reliability and Maintainability: Volume 4 - Practical Techniques and Application, Spartan Books Inc., Washington, D.C., 1965.
9. Hausman, W.H., and Peterson, R., "Multiproduct Production Scheduling for Style Goods with Limited Capacity, Forecast Revisions and Terminal Delivery," to be published.

10. Hertz, D.B., and Shaffir, K.H., "A Forecasting Method for Management of Seasonal Style-Goods Inventories," Operations Research, Vol. 8, (1960), pp. 45-52.

11. Ichihara, K., L. Neidell, J. Oliver, and B. William-Powlett, "The John P. Jones Company -- A Forecasting Model," umpublished manuscript, January 1965.

12. Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems," ASME Transactions, American Society of Mechanical Engineers, Vol. 82D, March, 1960.

13. Kotler, P., "Competitive Strategies for New Product Marketing Over the Life Cycle," Management Science, Vol. 12, No. 4 (December 1965), pp. 104-119.

14. Murray, G.R., Jr., and Silver, E.A., "A Bayesian Analysis of the Style Goods Inventory Problem," Management Science, Vol. 12, No. 11 (July 1966), pp. 785-797.

15. Scarf, H.E., "The Optimality of (S,s) Policies in the Dynamic Inventory Problem," in Mathematical Methods in the Social Sciences, K.J. Arrow, S. Karlin and P. Suppes (eds.), Stanford University Press, Stanford, California, 1960.

16. Shaw, L.G., "Optimum Stochastic Control," in J. Peschon (ed.), Disciplines and Techniques of Systems Control, Blaisdell, 1965.

17. Theil, H., Economic Forecasts and Policy, North-Holland Publishing Company, Amsterdam, 1965.

18. Wagner, H.M., and Whitin, T.M., "Dynamic Version of the Economic Lot Size Model," Management Science, October, 1958.

Date Due

DEC 05 '75		
APR 18 '76		
May 15, '76		
~~June 14, '76~~		
FEB 3 '84		