MANAGEMENT SCIENCE
Vol. 31, No. 4, April 1985
Printed in U.S.A.

FILL-RATE OPTIMIZATION IN A ONE-WAREHOUSE N-IDENTICAL RETAILER DISTRIBUTION SYSTEM*

LEROY B. SCHWARZ, BRYAN L. DEUERMEYER AND RALPH D. BADINELLI
Krannert Graduate School, Purdue University, West Lafayette, Indiana 47907
Department of Industrial Engineering, Texas A & M University, College Station, Texas 77843
School of Business, University of Kentucky, Louisville, Kentucky 40506

This paper examines the system fill-rate of a one-warehouse N-identical retailer distribution system as a function of warehouse and retailer safety stock. Using the approximation model of Deuermeyer and Schwarz (1981) we examine the problem of maximizing system fill-rate subject to a constraint on system safety stock. Optimal safety stock policy is characterized to be the intersection of a *fill-rate policy line* and the *safety stock budget line*. Properties of fill-rate policy lines are given. These properties may be used to provide managerial insight into system optimization and as the basis for heuristics. One heuristic, the *vertical heuristic*, based on the corresponding deterministic version of the problem, yields trivially simple and near-optimal policies.
(INVENTORY/PRODUCTION; MULTI-ECHELON INVENTORY SYSTEM; FILL-RATE OPTIMIZATION)

1. Introduction

This paper is concerned with the optimization of system service level for a one-warehouse N-retailer distribution system. In particular, we examine system fill-rate as a function of retailer and warehouse safety stocks. The system examined is a simple one: N identical retailers each facing independent, stationary unit Poisson demand with known rate λ_R. Each retailer follows a continuous review (Q_R, r_R) stock replenishment policy; i.e., whenever inventory position (inventory on hand plus on order less backorders) falls to r_R, an order for Q_R is placed. Retailer orders are filled by a single warehouse following a similar (Q_w, r_w) policy, where Q_w and r_w are integer multiples of Q_R. The warehouse receives its supply after a fixed lead time, L_w, from an unlimited source outside the system. The retailers receive their orders from the warehouse after a fixed nominal lead time, L_R, provided the warehouse has adequate on-hand inventory. If the warehouse inventory is not adequate, we assume that the entire retailer order is backlogged and filled on a first-come-first-served basis when stock becomes available. Thus, *effective retailer lead time* has a fixed component, L_R, plus a stochastic component, $d \geqslant 0$, where $d > 0$ is the delay time between receipt of the given retailer order at the warehouse and subsequent shipment of warehouse stock adequate to fill the order. Unsatisfied customer demand at the retailers is backlogged. Consequently, the service level performance of the system depends upon: (1) the safety stock at the warehouse, S_w, which buffers effective retailer lead time against delays at the warehouse due to out-of-stock conditions, $d > 0$; and (2) total retail safety stock, NS_R, which buffers the system against unknown demand during the effective retailer lead time.

The goal of our analysis is to describe general properties of the fill-rate maximizing safety stock policy. Such general properties aid the development of insights into system optimization and provide a basis for heuristics. In what follows we shall characterize the fill-rate maximizing safety stock policy as a *fill-rate policy line* in (S_w, NS_R) space.

*Accepted by Warren H. Hausman, former Departmental Editor; received September 1981. This paper has been with the authors 8 months for 3 revisions.

0025-1909/85/3104/0488$01.25

Several general properties of the fill-rate maximizing safety stock policy will be determined based on the location and shape of the fill-rate policy line.

The model used in the analysis is that of Deuermeyer and Schwarz (1981). The Deuermeyer and Schwarz (*DS*) model predicts retailer fill rate, F_R, expected backorders, B_R, and expected warehouse delay, T_w, given a particular safety stock policy, (S_w, NS_R), for fixed values of the problem parameters: N, λ_R, L_R, L_w, Q_R, and Q_w. Tests using computer simulation have demonstrated that the *DS* model provides adequate approximations to F_R, B_R, and T_w for a wide range of test problems for which (Q, r) policies are appropriate; i.e., "high" demand items.

The paper is organized as follows: Following a short literature review, we state the fill-rate optimization problem, summarize relevant portions of the *DS* model, define *policy lines* and *budget lines*, and interpret the fill-rate optimization problem for a fixed budget as the intersection of the policy line and fixed budget line in (S_w, NS_R) space. More generally, the *form* of the optimal policy as a function of the problem parameters is shown to correspond to the shape and location of the policy line. Several properties of the policy line are described and interpreted in §3. *An approximate policy line* applicable to all problem parameterizations is also described and interpreted. §4 describes two heuristic policies, the *midband heuristic* and the *vertical heuristic*, and summarizes the results of empirical tests on them. The *vertical heuristic* provides outstanding results with a remarkably simple policy. Important formulas are provided in the Appendix. Detailed proofs are provided in Schwarz et al. (1982), available from the authors.

Before we begin, perhaps some comments regarding the relevance of (Q, r) policies for warehouse/retailer systems are in order. This paper assumes that (Q, r) policies are used at all inventory locations. From a practical standpoint, (Q, r) policies are the most commonly used policies. This may be because of their simple structure, their modest information requirements, their ease of implementation, etc. However, the implementation of these policies is usually based on single-level inventory theory; i.e., that the retailers order from a warehouse that is always in stock and/or that the warehouse supplies "outside" customer demand. See Muckstadt and Thomas (1980), Rosenbaum (1981), and Schwarz (1977). Our analysis is intended to provide some multi-level (Q, r) theory. It is important to note, however, that policies other than (Q, r) are in use and may perform better in some circumstances. For example, Schwarz (1977) developed a modified (Q, r) policy which limited warehouse shipments when the warehouse was "low" on stock. In this manner, warehouse safety stock is more effectively managed. Schwarz determined performance and control parameters via multiple simulations of the system. A somewhat broader approach would be to base warehouse ordering policies on information concerning the inventory position at each of the retailers. Such information-based policies should perform better than a (Q, r) policy at the warehouse since more information is used. At the current time little appears to be known about such information-based policies. In particular, neither the form of such policies nor the relevant tradeoffs in policy determination are yet known.

Finally, this research addresses the problem of positioning safety stock of a *single* item in a warehouse/retailer system. Most real systems handle *several products* simultaneously. An important problem for the manager of such a system is to determine, for a given total investment in safety stock, how much should be allocated to each item and how this safety stock should be distributed throughout the system. Currently, no computationally feasible (Q, r)-based mathematical model of this problem exists. In fact, no model exists which adequately describes the response of a *single* item system to: (1) a particular investment in safety stock, and (2) the positioning policy used to distribute its safety stock in the system. Such a model and the insight provided by its analysis seem to be prerequisite to solving the multi-item problem. This research is a step in that direction.

Literature Review

Multi-level distribution systems are commonly found in practice and frequently modeled in the management science literature, but, in fact, very little is known in general about the service-level performance of such systems. In particular, little is known about the manner in which system service level depends on safety stock location, lot sizes, lead times, etc. Hannsman (1959) has examined a two-stage serial production system (one semifinished and one finished product) and points out the trade-off involved in locating fixed system inventories.

The classic model of a continuous review multi-level inventory system is METRIC (Sherbrooke 1968). This model assumes a two-level depot/base reparable inventory system following an $(S-1, S)$ policy. Base-level customer demand for "good" items is unit Poisson; return of a "failed" item accompanies each customer demand. Routine repairs are done at the bases; other repairs are performed at the depot. The METRIC model can be used to obtain the inventory policy which minimizes expected customer backorders subject to a constraint on total system inventory investment. Although METRIC is only an approximate model (see Shanker 1977 and Gross 1982), the insights it provides are valuable in concept and useful in practice (Demmy and Presutti 1981). MOD–METRIC (Muckstadt 1973) extends the METRIC model to include inventory item components and subassemblies.

The desirability of extending the insights of METRIC-type models to more general systems and to policies other than $(S-1, S)$ is well recognized. However, the analysis of such extensions is quite difficult. For example, consider a single retail facility facing stationary unit-Poisson demand and using a (Q_R, r_R) replenishment policy. Assuming backordering of excess demand and fixed lead times, an exact model of the retailer may be constructed and optimized. See Hadley/Whitin (1963). Such a model is a natural building block for analysis of a multi-level system. However, several difficulties arise in its use. In a system of N such identical retailers supplied by a single warehouse the demand process at the warehouse is the superposition of N identical Erlang processes, each with mean Q_R/λ_R. This process is computationally quite complex.[1] See Deuermeyer and Schwarz (1981). Despite this complexity, Muckstadt (1977) derives the steady-state distribution of the number of backorders for a one-warehouse two-identical retailer system. Deuermeyer and Schwarz (1981) approximate the demand process at the warehouse using a single Erlang process with mean $(Q_R/(N\lambda_R))$ and, with this approximation (and one other), construct a model for predicting system service level.

Two empirical studies are also relevant to the work presented here. Muckstadt and Thomas (1980) compare the performance of a single-level heuristic model and multi-level heuristic model for determining stock levels for low-demand items in a large spare parts supply system. The multi-level heuristic provided superior results. Rosenbaum (1981) describes the field test of a multi-level heuristic for determining safety stock location for high demand items in a two-level distribution system. See Clark (1974) for a survey of multi-level inventory control. See Schwarz (1981) for a sample of more recent research and practice.

2. The Optimization Problem

The optimization problem considered is the maximization of retailer fill rate, F_R, subject to a system-wide budget, S, on the total amount of safety stock at the

[1] Note that if $Q_R = 1$, then the warehouse demand process is a superposition of N identical Poisson processes, which is Poisson. In this case (Q, r) policies and $(S-1, S)$ policies are identical.

warehouse, S_w, or at the N retailers, NS_R; that is:

$$\text{Max} \quad F_R \qquad \text{w.r.t.}\,(S_w, S_R) \tag{1}$$

$$\text{s.t.} \quad S_w + NS_R \leqslant S. \tag{2}$$

It is assumed that all problem parameters, $N, \lambda_R, L_w, L_R, Q_w$ and Q_R are fixed, positive quantities. In order to highlight the general properties of optimal policies with respect to safety stock, both S_w and S_R are allowed to vary continuously over the range $(-\infty, \infty)$. Indeed one general result of our research has been that the solution to (1)–(2), in both deterministic and stochastic model formulations, involves warehouse safety stock, S_w, in the vicinity of $-Q_W$. Although many managers have trouble accepting the notion of negative safety stocks (and reorder points), they are perfectly valid mathematically and nothing prevents their use in practice. See Snyder (1980). A reduction (increase) in safety stock implies a reduction (increase) in on-hand inventory even when safety stock is negative, provided that the reorder point is greater than $-Q$. Within these limits, negative warehouse safety stock may be used via constraint (2) to increase retail safety stock above the nominal system budget. Hence a budget for safety stock *approximates* the tradeoffs associated with a budget for on-hand inventory.

Reorder points and safety stocks are related to each other in the usual manner; that is:

Safety Stock = Reorder Point − Expected Lead Time Demand.

However, a complication arises at both the retailer and the warehouse with respect to whether nominal or effective lead time demand should be used in the safety stock/reorder point definition. For simplicity's sake, we shall define safety stocks S_R and S_w in nominal terms; i.e.,

$$S_R = r_R - \lambda_R \cdot L_R, \tag{3}$$

$$S_w = r_w - N \cdot \lambda_R \cdot L_w. \tag{4}$$

The corresponding effective safety stocks will be defined and commented on below.

The Deuermeyer and Schwarz Model

Retail fill rate, F_R, given S_w and S_R is as specified by the Deuermeyer and Schwarz (DS) model (1981). In the DS model the *effective* retailer lead time demand is Poisson with mean $\lambda_R(L_R + T_w)$, where T_w is the steady-state expected delay due to out-of-stock conditions at the warehouse. Given this approximation, the corresponding *effective retailer safety stock*, S_R^e, is:

$$S_R^e = r_R - \lambda_R(L_R + T_w) = S_R - \lambda_R T_w. \tag{5}$$

That is, effective retailer safety stock is less than nominal retailer safety stock by an amount equal to the retailer's average demand during the expected warehouse delay.

The expected delay at the warehouse, T_w, in (5) depends upon warehouse safety stock and the warehouse lead time demand process. As noted above, the warehouse demand process is the superposition on N identical Erlang processes. In the DS model, the effective demand during the fixed warehouse lead time L_w is approximated by a normal distribution with mean μ_w and variance σ_w^2 given by:

$$\mu_w = N\lambda_R L_w, \tag{6}$$

$$\sigma_w^2 = N\lambda_R L_w. \tag{7}$$

Hence, the *effective warehouse safety stock* S_w^e, is given by:

$$S_w^e = r_w - \mu_w = S_w. \tag{8}$$

Given the above, the DS model provides closed-form expressions for F_R and T_w. See Appendix.

Determination of the Optimal Policy

Since F_R is increasing in both S_w and S_R, which are assumed to vary continuously, (2) is binding at optimality. Thus, for a *fixed* budget S, the optimal policy, (S_w, NS_R) may be found by a one-dimensional search of, say, S_w, using $NS_R = S - S_w$. However, our interest here is in the general characteristics, or *form* of the optimal policy for (1)–(2) as a func*tion* of the budget S, and as a function of the other problem parameters. The following definitions aid our analysis:

Policy Line. The (fill-rate) policy line is the locus of points (S_w, NS_R), each one of which is a solution to (1)–(2) for some fixed budget S.

Budget Line. The (safety stock) budget line is the locus of points (S_w, NS_R) satisfying $S_w + NS_R = S$ for a fixed budget S; i.e., a straight line in (S_w, NS_R) space with slope of -1.

Given these definitions, the optimal policy for a fixed budget S may be viewed as the intersection of the policy line and the corresponding budget line in (S_w, N_{SR}) space. Figure 1 presents an example policy line. The optimal safety stock policy for a fixed budget S, denoted $(S_w^*(S), NS_R^*(S))$, is shown by the intersection of the policy line with the corresponding budget line. Given this, the policy line may be defined in terms of S as the locus of points $(S_w^*(S), NS_R^*(S))$ for $-\infty \leq S \leq \infty$. Alternatively, the policy line may be defined in terms of S_w as the locus of optimal retailer safety stocks, denoted $NS_R^*(S_w)$, corresponding to a given S_w, $-\infty \leq S_w \leq \infty$. In this case, of course, the corresponding system safety stock is determined by $S = S_w + NS_R^*(S_w)$. In what follows we use these two equivalent definitions interchangeably.

More important, the *form* of the optimal policy; i.e., $(S_w^*(\cdot), NS_R^*(\cdot))$ as a function of S, is determined by the *shape* and *location* of the policy line in (S_w, NS_R) space. For example, for a given increase in the budget ΔS in Figure 1, the relative size of the allocations, ΔS_w and ΔNS_R, depends on the slope of the policy-line between the two budget lines S and $S + \Delta S$; i.e., the larger the slope the larger the retailer's share. In the same manner the sensitivity of the optimal policy to changes in other problem

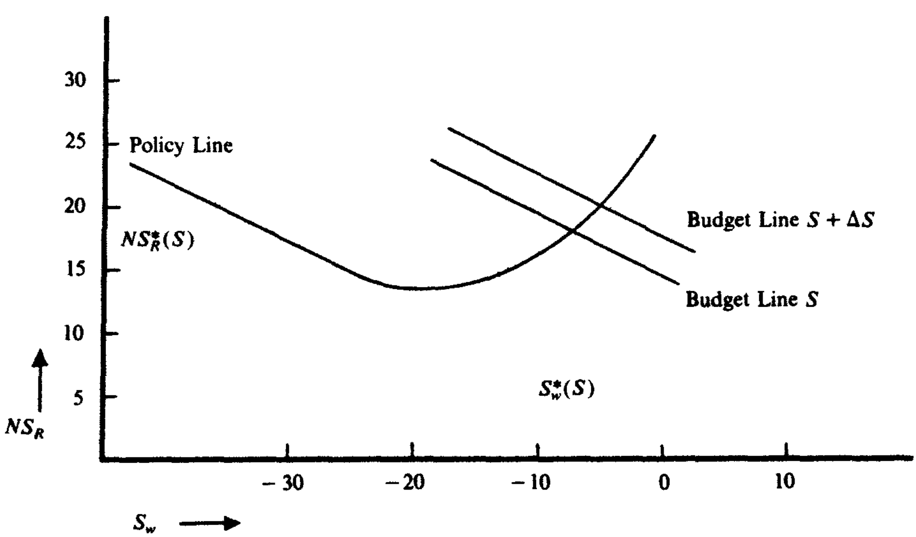

FIGURE 1. Example Policy Line and Budget Line.

parameters (e.g., Q_w, L_R, N) corresponds to the effect of these parameters on the policy line shape and location. If the location and shape of the policy line are insensitive to changes in a given parameter, the optimal safety stock policy is insensitive to that parameter. See Schwarz et al. (1982) for further discussion and examples.

3. General Properties of Policy Lines

In what follows we will describe and interpret several general properties of policy lines. Proofs and more detailed interpretations are provided in Schwarz et al. (1982). In addition, where appropriate, we will summarize our empirical findings. Although our knowledge about policy lines is far from comprehensive, the properties below do provide some basic knowledge about fill rate optimization and sensitivity. They also provide bases for two very simple, near-optimal heuristics.

For notational convenience, we will denote the policy line as $NS_R^*(S_w)$ as defined above.

Property 1. The policy line satisfies:

(a) $NS_R^*(S_w) \geqslant -NQ_R$;

(b) $NS_R^*(S_w) \to \infty$ as $S_w \to \pm\infty$.

Discussion. This property provides some crude limits on the location and shape of the policy line. Property 1(a) states that the policy line lies in the half space $NS_R \geqslant -NQ_R$; i.e., at optimum each retailer's safety stock will be at least $-Q_R$ *regardless* of the system budget, the warehouse safety stock, or any other problem parameters. Property 1(b) says that in the limit retail safety stocks should be "large" if the corresponding warehouse safety stock is "large" or "small" (and negative). The intuition is straightforward.

The crude limits on the policy line provided by Property 1 may be considerably improved, at some computational expense, by Property 2.

Property 2. For a given value of S_w, the policy line, $NS_R^*(S_w)$, lies between a lower bound, denoted NS_R^L, and an upper bound, denoted NS_R^U. The closed-form expressions for these bounding values are:

$$NS_R^U = N(\lambda_R L_R + \lambda_R T_w)\left[\frac{2Q_w}{\sigma_w h(S_w)} - 1\right] - N\lambda_R L_R \quad \text{and} \tag{9}$$

$$NS_R^L = NS_R^U - NQ_R. \tag{10}$$

Expressions for T_w and $h(S_w)$ are provided in the Appendix.

Discussion. For a given value of S_w, NS_R^L and NS_R^U provide a finite interval (of size NQ_R) in which $NS_R^*(S_w)$ is known to lie. The existence of this interval, or *policy band*, and its closed form facilitates the search for policy lines. Our empirical tests have indicated two additional general characteristics of policy lines and bands: the policy line is approximately centered within the policy band; and both are convex. The "centering" of the policy lines within the policy band, a property which can be proven analytically for $S_w \leqslant -N\lambda_R L_w - Q_w$ (see Schwarz et al.), provides the basis for our first heuristic policy, the *mid-band heuristic*, to be discussed below.

The policy band provides a closed-form approximation to the policy line throughout (S_w, NS_R) space. However, for "small" values of S_w, this approximation is unnecessary because the policy line itself has a simple analytical form.

Property 3. For $S_w \leqslant -N\lambda_R L_w - Q_w$ and $S_R \geqslant -\lambda_R L_R$ the fill-rate policy line is given by:

$$NS_R^*(S_w) = N\lambda_R T_w - NQ_R/2 = -S_w - (Q_w + NQ_R)/2. \tag{11}$$

Discussion. Property 3 states that for "small" values of S_w the fill-rate policy line is linear (in parameters N, Q_R, and Q_w) with a slope of -1. This implies that for *every* problem parameterization there exists a safety stock budget $S = S_w + NS_R^*(S_w)$ for which the budget line and the policy line overlap. This budget is denoted the critical budget, S_c, where:

$$S_c = -Q_w/2 - NQ_R/2. \tag{12}$$

The parallel/overlap nature of policy and budget lines in this region has two important implications: (1) for safety stock budgets equal to the critical budget *all* safety stock policies are equally good; and (2) for safety stock budgets $S \neq S_c$ there is no intersection with the policy line in this region; i.e., there is no finite optimal safety stock policy in this region.

Property 4 provides a relationship between optimal system fill-rates and the critical budget S_c. Let $F_R(S)$ denote the fill-rate corresponding to an arbitrary allocation of the system budget S; let $F_R^*(S)$ denote the fill-rate corresponding to the optimal safety stock policy for budget S.

Property 4. (a) $F_R^*(S_c) = 0.5$;

(b) $F_R(S) < 0.5$ for $S < S_c$ given any finite safety stock policy;

(c) $F_R^*(S) > 0.5$ for $S > S_c$, and the optimal policy is finite.

Discussion. Property 4(a) states that optimal system fill-rate is 0.5 at a safety stock budget of S_c. Recall from the discussion above that the "intersection" of the critical budget with the policy line is an "overlap" with the linear portion of the policy line. For budgets smaller than S_c no finite optimal policy exists; i.e, the policy and budget lines do not intersect for finite (S_w, NS_R). However, it can be shown (see Schwarz et al.) that for budgets below S_c, $F_R(S)$ increases with decreasing S_w, reaching a limit of 0.5 as S_w approaches $-\infty$. Hence Property 4(b). Although such an *asymptotic policy* is of little practical value (since it implies that the warehouse is driven into an arbitrarily large backordered position) it does imply that, for $S < S_c$, the more heavily backordered the warehouse the better. For finite budgets $S > S_c$, we know that the policy line is below the budget line for $S_w \leqslant -N\lambda_R L_W - Q_W$ (Property 3) and rises above the budget line as $S_w \to \infty$ (Property 1). Thus a finite intersection exists for which $F_R^*(S) > 0.5$.

An Approximate Policy Line/An Approximately Optimal Policy

The properties described above suggest the construction of the following piecewise linear approximation to the policy line. Although this approximation is somewhat simplistic, it is quite general; i.e., independent of any problem parameters, and provides the basis for heuristics. When viewed as a function of S_w, the policy line, $NS_R^*(S_w)$, is decreasing with slope of -1 for small values of S_w and increasing for large values of S_w (Properties 3 and 1(b)). Further, the policy line is contained within a policy band of height NQ_R. Our empirical tests indicate further that the policy lines are convex and "centered" vertically in the (observed) convex band. In addition, the "elbow" or minimum of the policy line has been consistently *observed* to be in the neighborhood of $S_w = -Q_w$. We will interpret this observation below.

The approximate policy line is presented in Figure 2. It is piecewise linear with minimum at $S_w = -Q_w$. To the left of the minimum it is linear with slope -1, overlapping the critical budget line S_c and the linear portion of the exact policy line (Property 3). To the right of the minimum the approximate policy line is linear with positive slope (Property 1(b)).

Corresponding to this approximate policy line is the following heuristic policy:

(a) For $S < S_c$, the more backordered the warehouse the better;

(b) For $S = S_c$ set $S_w = -Q_w$ and $NS_R = S + Q_w$;

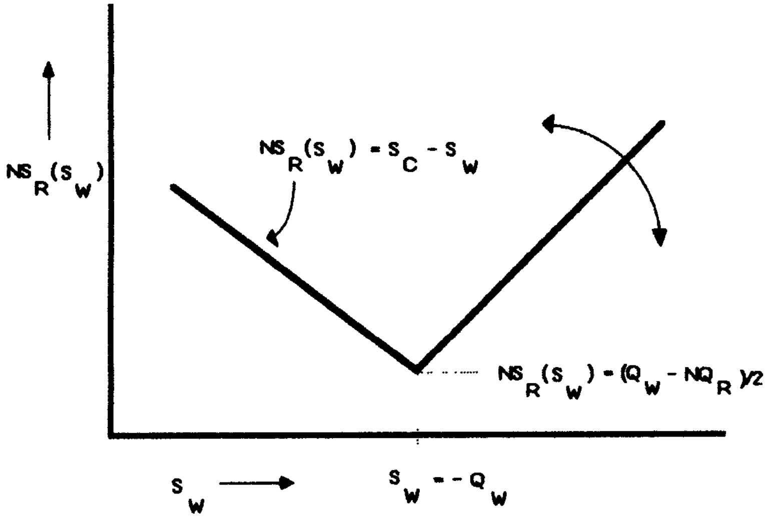

FIGURE 2. A Piecewise Linear Approximation to the Policy Line.

(c) For $S > S_c$, allocate the first S_c units as in (b). Allocate the remaining $(S - S_c)$ units according to the approximate policy line's slope; the larger the slope, the larger the retailers' share. In particular, if the slope is zero (i.e., a horizontal policy line) the warehouse receives all remaining $(S - S_c)$ units. If the approximate policy line is vertical then all remaining $(S - S_c)$ units go to the retailers. The latter is the basis of the vertical heuristic.

For $S \geqslant S_c$; i.e., the "interesting" budgets, the approximate policy line minimum at $S_w = -Q_w$ assures the warehouse that its share of the safety stock budget is at least $S_w = -Q_w$. Schwarz (1981b) has shown that in the *deterministic* version of the DS model that the optimal warehouse safety stock is $S_w = -Q_w$ for *all* interesting budgets. Thus we may intuitively justify the approximate policy line minimum at $S_w = -Q_w$ if we accept, intuitively, the notion that the optimal allocation of safety stock in the stochastic case has "roots" in the corresponding deterministic case.

4. Heuristics: Empirical Tests

Property 2 of §3, that the policy line is contained in a closed-form policy band, and our general observation that the policy line is approximately centered within these bands motivates the midband heuristic:

The Midband Heuristic. Given any system safety stock budget, S, let (S_w, NS_R) be the point of intersection of the midband (the average of NS_R^U and NS_R^L; see (9)–(10)) and the budget line.

The approximate policy line described above and our empirical tests suggest the following heuristic:

The Vertical Heuristic. Given any safety stock budget, S, set $S_w = -Q_w$; set $NS_R = S - S_w = S + Q_w$.

The name *vertical heuristic* is derived from the assumption that the policy line is vertical (slope $= \infty$) at $S_w = -Q_w$. This corresponds to the optimal policy for the *deterministic* version of the DS model. See Schwarz (1981b).

Test Problems

The performance of these heuristics was evaluated using the ten problem parameterizations specified in Table 1. Note that these parameterizations represent a wide range of parameters: L_w/L_R ranges from 0.625 to 8, Q_w/Q_R ranges from 4 to 10. In

TABLE 1

Parameters for Test Problems

Problem No.	Q_R	L_R	Q_w	L_w	L_w/L_R	Q_w/Q_R
1	5	0.1	20	0.1	1	4
2	15	0.3	90	0.45	1.5	6
3	40	0.8	400	2	2.5	10
4	15	0.3	150	0.75	2.5	10
5	15	0.3	60	0.3	1	4
6	40	0.1	160	0.8	8	4
7	24	0.5	96	0.5	1	4
8	40	0.1	160	0.4	4	4
9	40	0.8	400	1	1.25	10
10	40	0.8	400	0.5	0.625	10

the results described below $N = 2$ and $\lambda_R = 100$. Exploratory tests for $N = 1$ and $N = 5$, for $\lambda_R = 10$ and $\lambda_R = 300$ showed comparable results. For each parametrization optimal fill-rates were computed over a range of budgets such that $0.5 \leqslant F_R^*(S) \leqslant 0.99$ with $\Delta S = 1$. Then for each budget the corresponding heuristic policy was determined and evaluated. The evaluation measure used in these tests was the percent deviation of the heuristic fill-rate from the optimal fill-rate.

The midband heuristic performed remarkably well. On the average, the midband fill-rate was within one percent of optimum and was always within two percent of the optimum. (See Badinelli 1982 for details.) The vertical heuristic performed almost as well. Its average error was within one percent of optimum while the worst case error was three percent. Of course, care must be taken in interpreting these test results, particularly for high optimal fill rates; e.g., if $F_R^*(S) = 0.99$, a policy with $F_R(S)$ as low as 0.9702 would be within 2% of optimum. Nonetheless, the heuristic results are, in the words of one referee, "good and surprising".

The performance of the vertical heuristic is truly remarkable not only from the standpoint of its simplicity, but from its managerial implications: $S_w = -Q_w$ implies that the average on-hand inventory at the warehouse will be small and most retail orders will be backordered. In fact, a *deterministic* model at the warehouse $S_w = -Q_w$ yields a warehouse fill rate of 0 and expected backorders of $Q_w/2$. See Badinelli (1982) and Schwarz (1981b) for details. In short, under the vertical heuristic the warehouse virtually serves as a clearing house for retailer orders rather than a stocking location. These results and interpretations, of course, apply only to (Q, r) based policies with $Q_R > 1$, and should not be intrepreted to apply to "low" demand situations, in which the warehouse may have an inventory role to play.

Appendix

$$F_R = \begin{cases} 1 - \dfrac{\sigma_R}{Q_R}\left[\hat{\alpha}\left(\dfrac{S_R^e}{\sigma_R}\right) - \hat{\alpha}\left(\dfrac{S_R^e + Q_R}{\sigma_R}\right)\right], & \text{if } \; S_R^e > -\lambda_R(L_R + T_w); \\ 1 - \dfrac{\sigma_R}{Q_R}\left[\hat{\alpha}\left(-\dfrac{\mu_R}{\sigma_R}\right) - \hat{\alpha}\left(\dfrac{S_R^e + Q_R}{\sigma_R}\right) - \dfrac{(S_R^e + \lambda_R(L_R + T_w))}{\sigma_R}\cdot\Phi\left(-\dfrac{\mu_R}{\sigma_R}\right)\right], & \\ & \text{if } \; -\lambda_R(L_R + T_w) - Q_R < S_R^e < -\lambda_R(L_R + T_w); \\ 1 - \Phi(-\mu_R/\sigma_R), & \text{if } \; S_R^e < -\lambda_R(L_R + T_w) - Q_R, \quad \text{where} \end{cases} \tag{A1}$$

$$\hat{\alpha}(x) = \phi(x) - x\Phi(x) \tag{A2}$$

and $\phi(\cdot)$ and $\Phi(\cdot)$ are the unit normal density and complementary cumulative distribution, respectively.

Equation (A1) is based directly on (1*) in Deuermeyer and Schwarz (1981). However, in order to highlight its role in the optimization, (A1) is expressed in terms of effective retailer safety stock, S_R^e, whereas in (1*) of Deuermeyer and Schwarz (1981) F_R is expressed in terms of the corresponding reorder point, r_R. Also $\hat{\alpha}(\cdot)$ above is the unit-normal version of $\alpha^*(\cdot)$ in Deuermeyer and Schwarz (1981). Finally, (A1) extends the domain of F_R to $r_R < -Q_R$, or, equivalently $S_R^e < -\lambda_R(L_R + T_w) - Q_R$. See Badinelli (1982).

$$T_w = \begin{cases} \dfrac{\sigma_w^2}{N\lambda_R Q_w}\left[\hat{\beta}\left(\dfrac{S_w^e}{\sigma_w}\right) - \hat{\beta}\left(\dfrac{S_w^e + Q_w}{\sigma_w}\right)\right] & \text{if } S_w^e > -\mu_w; \\ \dfrac{\sigma_w^2}{N\lambda_R Q_w}\left[\hat{\beta}\left(-\dfrac{\mu_w}{\sigma_w}\right) - \hat{\beta}\left(\dfrac{S_w^e + Q_w}{\sigma_w}\right) + \dfrac{(S_w^e + \mu_w)^2}{2\sigma_w^2}\Phi\left(-\dfrac{\mu_w}{\sigma w}\right)\right. & \\ \qquad \left. - \dfrac{(S_w^e + \mu_w)}{\sigma_w}\hat{\alpha}\left(-\dfrac{\mu_w}{\sigma w}\right)\right] & \text{if } -\mu_w - Q_w < S_w^e < -\mu_w; \\ \dfrac{\sigma_w}{N\lambda_R}\hat{\alpha}\left(\dfrac{-\mu_w}{\sigma_w}\right) - \dfrac{[2(S_w^e + \mu_w) + Q_w]}{2N\lambda_R}\Phi\left(\dfrac{-\mu_w}{\sigma_w}\right) & \text{if } S_w < -\mu_w - Q_w \quad \text{where} \end{cases} \tag{A3}$$

$$\hat{\beta}(x) = \left(\frac{1 + x^2}{2}\right)\Phi(x) - \frac{x}{2}\phi(x). \tag{A4}$$

Equation (A3) is based directly on (7) and (11) in Deuermeyer and Schwarz (1981). However, in order to highlight its role in the optimization, (11) is expressed in terms of warehouse safety stock, S_w, whereas in Deuermeyer and Schwarz (1981), T_w is expressed in terms of the corresponding warehouse reorder point. Note also that $\hat{\beta}(\cdot)$ above is the unit-normal version of $\hat{\beta}(\cdot)$ in Deuermeyer and Schwarz (1981). Finally, the domain of T_w has been extended in (A3) above to $r_w < -Q_w$, or, equivalently, $S_w < -\mu_w - Q_w$. See Badinelli (1982).

$$h(S_w) = \begin{cases} \hat{\alpha}\left(\dfrac{S_w}{\sigma_w}\right) - \hat{\alpha}\left(\dfrac{S_w + Q_w}{\sigma_w}\right) & \text{if } S_w > -N\lambda_R L_w; \\ \hat{\alpha}\left(-\dfrac{\mu_w}{\sigma_w}\right) - \hat{\alpha}\left(\dfrac{S_w + Q_w}{\sigma_w}\right) - \dfrac{(S_w + N\lambda_R L_w)}{\sigma_w}\Phi\left(-\dfrac{\mu_w}{\sigma_w}\right), & \\ & \text{if } -N\lambda_R L_w - Q_w < S_w < -N\lambda_R L_w; \\ \dfrac{Q_w}{\sigma_w}\Phi\left(-\dfrac{\mu_w}{\sigma_w}\right) & \text{if } S_w < -N\lambda_R L_w - Q_w. \end{cases} \tag{A5}$$

References

BADINELLI, R. D., "Optimal Inventory Policies for the One-Warehouse N-Retailer System," Ph.D. Dissertation, Purdue University, 1982.

CLARK, A. J., "An Informal Survey of Multi-Echelon Inventory Theory," *Naval Res. Logist. Quart.*, 19, 4 (1974), 621–650.

DEMMY, W. S., AND V. J. PRESUTTI, "Multi-Echelon Inventory Theory in the Air Force Logistics Command," Chapter 13 in *Multi-Level Production/Inventory Control Systems: Theory and Practice*, North-Holland, Amsterdam, 1981.

DEUERMEYER, B. L. AND L. B. SCHWARZ, "A Model for the Analysis of System Service Level in Warehouse/Retailer Distribution Systems: The Identical Retailer Case," Chapter 8 in *Multi-Level Production/Inventory Control Systems: Theory and Practice*, North-Holland, Amsterdam, 1981.

GRAVES, S. C. AND L. B. SCHWARZ, "Single Cycle Continuous Review Policies for Arborescent Production /Inventory Systems," *Management Sci.*, 23, 5 (1/77), 529–540.

GROSS, D., "On the Amplé Service Assumption of Palm's Theorem in Inventory Modeling," *Management Sci.*, 28, 9 (September 1982), 1065–1079.

HADLEY, G. AND T. M. WHITIN, *Analysis of Inventory Systems*, Prentice-Hall, Englewood Cliffs, N. J., 1963.

HANNSMAN, F., "Optimal Inventory Location and Control in Production and Distribution Networks," *Oper. Res.*, 7, 4 (1959), 483–498.

MUCKSTADT, J. A., "On the Probability Distribution for Inventory Position in Two-Echelon Continuous Review Systems," Technical Report No. 336, School of Operations Research and Industrial Engineering, Cornell University, 1977.

———, "A Model for a Multi-Item, Multi-Echelon, Multi-Indenture Inventory System," *Management Sci.*, 20, 4 (1973), 472–481.

——— AND L. J. THOMAS, "Are Multi-Echelon Inventory Models Worth Implementing for Systems with Low Demand Rates," *Management Sci.*, 26, 5 (1980), 483–494.

ROSENBAUM, B. A., "Inventory Placement in a Two-Echelon Inventory System: An Application," Chapter 9 in *Multi-Level Production/Inventory Control Systems: Theory and Practice*, North-Holland, Amsterdam, 1981.

SCHWARZ, L. B. (ED.), *Multi-Level Production/Inventory Control Systems: Theory and Practice*, North-Holland, Amsterdam, 1981a.

———, "Physical Distribution: The Analysis of Inventory and Location," *AIIE Trans.*, 13, 2 (1981b), 138–150.

———, "An Examination of United States Air Force (Q, R) Policies for Managing Depot-Base Inventories," Working Paper No. 7640, Graduate School of Management, University of Rochester, 1977.

———, B. L. DEUERMEYER AND R. B. BADINELLI, "Detailed Analysis of Fill-Rate Optimization in a One-Warehouse N-Identical Retailer Distribution System", Technical Report, Krannert Graduate School of Management, Purdue University, West Lafayette, Indiana 47907, 1982.

SHANKER, K., "An Analysis of a Two-Echelon Inventory System for Recoverable Items," Technical Report No. 341, School of Operations Research and Industrial Engineering, Cornell University, 1977.

SHERBROOKE, C. C., "METRIC: A Multi-Echelon Technique for Recoverable Item Control," *Oper. Res.*, 16, 1 (1968), 122–141.

SNYDER, R. D., "The Safety Stock Syndrome," *J. Oper. Res. Soc.*, 31, 9 (1980), 833–837.