SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK 14850

Technical Report No. 634

August, 1984
Revised, June, 1986
Revised, September, 1987

Risk Pooling in a Two-Period, Two-Echelon Inventory Stocking and Allocation Problem

by

Peter L. Jackson & John A. Muckstadt

This research was supported in part by NSF Grants MEA-81-212429, ECS-81-06455 and DMC-85-09131.

ABSTRACT

The object of this paper is to investigate the risk pooling effect of depot stock in a two-echelon distribution system in which the depot serves n retailers in parallel, and to develop computationally tractable optimization procedures for such systems. The depot manager has complete information about stock levels and there are two opportunities to allocate stock to the retailers within each order cycle. We identify first and second order aspects to the risk pooling effect. In particular, the second order effect is the property that the minimum stock available to any retailer after the second allocation converges in probability to a constant as the number of retailers in the system increases, assuming independence of the demands. This property is exploited in the development of efficient procedures to determine near-optimal values of the policy parameters.

ACKNOWLEDGEMENTS

The authors extend their sincere thanks to Professor Linus Schrage for his help in formulating the model, to Professor Howard Taylor for suggesting the approach taken in Appendix 1 and to the reviewers for their valuable suggestions.

1. Problem Description

This paper considers an inventory stocking problem in a two-echelon distribution system consisting of a depot that serves n retailers in parallel. Eppen and Schrage [1981] demonstrated that backorder costs are reduced in such systems if the depot acts as a centralized ordering facility and delays the assignment of stock to the retailers from the time at which an order is placed with the outside supplier to the time at which the depot actually receives the stock. This phenomenon is referred to as risk pooling over the supplier lead time. Our focus is on the next stage in the allocation process: should stock be immediately distributed to retailers upon receipt at the depot or should some depot stock be held in reserve, to be distributed to the retailers later in the cycle? Jackson [1983] demonstrated circumstances under which the use of depot reserve stock significantly reduces system backorder costs. This phenomenon could be termed the risk pooling effect of depot stock. Badinelli and Schwarz [1984] have demonstrated other circumstances in which this risk pooling effect appears to be very small. Consequently, it becomes important to understand the factors that contribute to its existence. For example, Zipkin [1981] has demonstrated analytically that the risk pooling benefit of depot stock decreases as demand at the retailers becomes more positively correlated. In this paper, we develop a model that focusses on the risk pooling motive for holding depot reserve stock and conduct an intensive qualitative analysis of the model.

In particular, by restricting attention to two allocations per order cycle, we identify two aspects of the risk pooling phenomenon. The first aspect, termed the first order effect of risk pooling, is that by holding stock in reserve the depot can increase, in expectation, the minimum stock available to any retailer after the second allocation. That is, the distribution of stock in the system will be more balanced as a result of holding stock in reserve. The second aspect, termed the second order effect of risk pooling, is that the minimum stock available to any retailer after the second allocation converges in probability to a constant as the number of retailers in the system increases, assuming independence of demands. Hence, as the term suggests, risk pooling removes some of the uncertainty involved in planning stock levels.

Whether or not these technical benefits of depot stock are sufficient to outweigh the disadvantages of holding depot stock depends on the cost parameters involved. The model we develop involves holding costs and backorder costs. By exploiting the second order effect of risk pooling we develop approximate cost functions and computational techniques to determine near optimal amounts of depot reserve stock. The model ignores the additional fixed costs of using more than one allocation per cycle including, for example, the cost of the information system that would be required to implement the proposed allocation policy. However, the computational techniques are quite practical and could be included as part of a broader empirical study of the tradeoffs involved in implementing such a policy.

Our analysis of the effect of risk pooling is based on a number of important and simplifying assumptions, which are as follows:

Assumptions:

(1) All lead times are zero. We have restricted lead times to be zero only to simplify notation and to focus on risk pooling due to allocation rules. The extension of the model to allow positive lead times is straightforward. Eppen and Schrage [1981] have already drawn attention to the risk pooling phenomenon that takes place over the supplier lead time.

(2) Transshipment of stock among retailers is not permitted.

(3) As in the Eppen and Schrage model, the depot order cycle is of fixed length and at the beginning of each cycle, the depot places a single procurement order sufficient to raise total system stock to a fixed level. (The order cycle length and the system stock level are set as policy parameters.)

(4) There are two allocation periods within each cycle: that is, there are two opportunities to ship stock from the depot to the retailers in each cycle. One opportunity occurs at the beginning of the cycle, after receipt of the initial procurement order, and the second occurs at a later, predetermined point in the cycle. The lengths of the two allocation periods need not be identical. This assumption facilitates a simple, yet practical analysis. The concern with stock imbalance in the system occurs towards the end of the cycle as stock levels become depleted. The second shipment can be used to correct imbalances. While we could consider more periods, the method of analysis becomes more cumbersome. Furthermore, the two period model

demonstrates the existence and importance of risk pooling. We also note that the flexibility gained by more frequent shipments may not justify the additional shipping costs.

(5) Allocation Assumption: each retailer begins the cycle with less stock than would be otherwise optimal if costless transshipments were allowed in the first allocation. This implies that the initial allocation from the depot to the retailers will be positive for each retailer. To simplify the notation, we also make the stronger assumption that there is no stock in the system at the beginning of the cycle.

(6) A holding cost of h_t is charged for each unit of inventory held at the end of period t, t = 1,2, regardless of which location it is held. This assumption is made so that there is no holding cost advantage to holding stock at the depot rather than at the retailers. If the demand process is deterministic, Roundy [1985] and Maxwell and Muckstadt [1985] showed that it is desirable to allocate all inventory in the first period. Hence we make this assumption to focus on the risk pooling reasons for holding inventory at the depot during period 1.

(7) Excess demand in each period is backordered. A backorder cost, π_t, is incurred for all outstanding customer demands at any retailer that could not be satisfied by the end of period t.

(8) The demand at location i in period t is denoted by d_{it}, i = 1,...,n,

$t = 1,2$. A known joint probability distribution function $G_t(d_{1t},\ldots,d_{nt})$ is assumed for demand in period t. Demand during the second period is independent of demand during the first period. To focus attention on the risk pooling motive for centralized stocking, we assume that $G_t(d_{1t},\ldots,d_{nt}) < 1$ for all finite $(d_{1t},\ldots,d_{nt})$ in period t, $t = 1,2$. That is, there is a positive probability of stockout in each period, regardless of how much stock is allocated to the retailers. Because of backorder costs, this assumption ensures that stock will not be held at the depot unless there is an economic advantage to do so.

The net effect of the assumptions we have made concerning the holding costs and demand distribution is that the only motive for holding stock at the depot in this model is the risk pooling motive.

A number of papers have been written on the control of inventory in multi-echelon systems. Reviews of this literature can be found in Veinott [1966], Iglehart [1967], Clark [1972], Aggarwal [1974], Nahmias [1981], and Silver [1981]. In an early seminal paper, Clark and Scarf [1960] analyzed serial multi-echelon problems by decomposing them into a series of single echelon problems. They also discussed the difficulties associated with the allocation of inventory in a single depot, n-retailer system. Subsequently, others extended their work including Hochstaedter [1970], Federgruen and Zipkin [1982] and [1984], and Zipkin [1981].

Eppen and Schrage [1981] examined a depot, n-retailer system in which the depot receives inventory each cycle and immediately allocates all of it to the n retailers. Their analysis reveals the effect of risk pooling over the depot's supplier lead time. Erkip, Hausman and Nahmias [1984] extend their model to allow for correlated demand over time and among retailers.

Jackson [1983] also extended the Eppen and Schrage model by allowing shipments to be made to each retailer from the depot in each of m periods per depot reorder cycle. He considers a policy in which the depot shipments raise the inventory position at each retailer to a fixed level, called the ship-up-to levels, in each period until a period is reached in which the depot stock is depleted. This policy permits the risk of imbalance to be eliminated during the pre-runout period. He derives and uses an approximation to the m-period cost function for calculating the ship-up-to levels and the depot reserve stock. Later, Erkip [1984a] compared this policy with an "α-policy" using simulation. In this policy, the depot allocates a fraction α of its total stock to the retailers at the beginning of a cycle and allocates the remainder at a single time later in the cycle. Erkip [1984b] then used an approximating dynamic programming model to pre-determine the timing of this second allocation.

Jonsson and Silver [1985] consider a similar model to the one devloped here and in Jackson and Muckstadt [1984]. Their approach to approximating the cost function is considerably different from ours.

The remainder of the paper is organized as follows. In Section 2, we develop the cost and demand models based on the stated assumptions. Section 3 presents a characterization of the second period's allocation problem in terms of a constrained minimum fractile allocation. In Section 4 we assume the n retailers have independent and identically distributed demands and costs. We begin the section by further studying the second period allocation problem, showing that risk pooling occurs even when there are only two retailers, and then by analyzing the asymptotic behavior of the second period allocation. We turn to computational considerations in Section 5 where we develop algorithms for computing approximately optimal values for the policy parameters in two cases: (1) the demands are independent among the retailers, but the demand distributions need not be identical, and (2) the retailers have identically distributed but correlated demands. Section 6 summarizes the conclusions of the paper.

2. Modeling the Two-Echelon, Two-Period Allocation Problem

2.1 The Cost Model

The basic quantities in the cost function for this model are acquisition, holding and shortage costs, and salvage values. Let c denote the unit acquisition cost, which is paid at the beginning of the first period. Let s be the unit salvage cost, which represents the value of excess stock remaining in the system after the second period, and assume $s < \pi_2$.

Let Y be the quantity of stock ordered by the depot at the beginning of period 1. After the depot makes its first period allocation, let S_{i1} represent the quantity of stock at retailer i, and let Q represent the amount of stock held in reserve at the depot. Hence, we have the equation $Y = Q + \sum_{i=1}^{n} S_{i1}$.

Demands in a period are assumed to occur following the allocation decision. After observing a demand of d_{i1} in period 1 at retailer i, the net inventory at that location is $I_i = S_{i1} - d_{i1}$. Let S_{i2} represent the retailer net inventory level at retailer i following the second period allocation by the depot. The quantity allocated to retailer i at the beginning of period 2 from the depot reserve stock may be expressed as $S_{i2} - I_i$. Hence, S_{i2} is a random variable that depends on the solution to the allocation problem in the second period.

The model that we will now state considers the two-period nature of the problem. The consequences of the first period purchasing and allocation decisions directly affect the cost and allocation possibilities in the second period. We begin by describing the second period allocation problem. Suppose

$K_2(Q, I_1, \ldots, I_n)$ denotes the minimal cost associated with the second period allocation decision. As stated, the second period cost depends on the retailer net inventory levels at the end of period 1 and the depot reserve stock. Then,

$$K_2(Q, I_1, \ldots, I_n) = \min_{S_{12}, \ldots, S_{n2}} \Bigg\{ (h_2 - s)\left[Q - \sum_{i=1}^{n} (S_{i2} - I_i)\right] + \sum_{i=1}^{n} \Bigg[(h_2 - s)\, \mathcal{E}[(S_{i2} - d_{i2})^+] + \pi_2\, \mathcal{E}[(d_{i2} - S_{i2})^+] \Bigg] \Bigg\} , \tag{1}$$

subject to:

$$\sum_{i=1}^{n} S_{i2} \leq Q + \sum_{i=1}^{n} I_i \ , \tag{1a}$$

$$S_{i2} \geq I_i \ , \quad i=1,\ldots,n. \tag{1b}$$

The objective function (1) reflects the stock levels at the beginning of period 2, the depot reserve stock, the cost of holding inventory at each facility, the salvage value of left-over stock and the penalty incurred for backorders in period 2. Constraint (1a) implies that no additional system stock may be acquired during the second period; the problem is purely allocational. Constraints (1b) represent the restriction that transshipments between retailers and returns to the depot are not permitted.

The overall two-period system cost function is defined in terms of the period 2 cost function. Let $K(Q, S_{11}, \ldots, S_{n1})$ denote the minimal expected two-period cost associated with a given vector $(Q, S_{11}, \ldots, S_{n1})$ of initial period 1 inventory levels:

$$K(Q, S_{11}, \ldots, S_{n1}) = c\left[Q + \sum_{i=1}^{n} S_{i1}\right] + h_1 Q$$
$$+ \sum_{i=1}^{n} \left[h_1\, \mathcal{E}[(S_{i1} - d_{i1})^+] + \pi_1\, \mathcal{E}[(d_{i1} - S_{i1})^+]\right] \qquad (2)$$
$$+ \mathcal{E}[K_2(Q, S_{11} - d_{11}, \ldots, S_{n1} - d_{n1})] \; .$$

This cost function reflects period 1 purchase, holding and backorder charges as well as the period 2 costs captured by the function K_2. We have chosen to represent holding and backorder costs so that they are charged only at the end of each period. For a more accurate accounting of these costs, we could easily subdivide period 1 and period 2 into m_1 and m_2 accounting sub-periods, respectively. We could then charge holding and backorder costs in each of these subperiods. The two-period nature of the decision problem would remain; the method of analysis described in the following sections would be essentially unaffected by this change.

The proposed model was developed to emphasize the economic impact of risk pooling. It ignores fixed costs among other factors. We could easily compare the total cost of operating a system in which there is only one allocation made to retailers with the total cost of operating the system when two allocations are made by including the fixed costs. We could also more accurately account for the holding and backorder costs in this comparison by subdividing the cycle length into shorter periods and assessing these costs in each subperiod. However, we have chosen to exclude these terms from our cost model so that we can concentrate on the effects of risk pooling on the depot allocation policy.

The two-period, two-echelon stocking and allocation problem that results from our assumptions can then be stated as follows:

$$\begin{aligned} &\text{minimize } K(Q,S_{11},\ldots,S_{n1}) \\ &\text{subject to} \\ &\qquad Q \geq 0,\ S_{i1} \geq 0,\ i = 1,\ldots,n. \end{aligned} \tag{3}$$

By Theorem 5.7 of Rockafellar [1970], $K_2(\cdot,\cdot,\ldots,\cdot)$ is jointly convex in $(Q,I_1,\ldots,I_n)$. It follows easily that $K(\cdot,\cdot,\ldots,\cdot)$ is jointly convex in $(Q,S_{11},\ldots,S_{n1})$.

The allocation problem is made difficult by the restrictions against returns and transshipments (constraints (1b)). In the absence of these restrictions, an ideal post-allocation stock level exists at the retailers that depends only on the total stock in the system at the beginning of period 2, the allocation period.

The distribution of stock is said to be unbalanced if one or more of the retailers begins the second period with more stock than is called for by its ideal stock level. Such retailers are overprotected against period 2 demand relative to the other retailers.

The function of depot reserve stock, Q, is to ensure a more equitable, or balanced distribution of stock in period 2. In the event that high demand is observed at one location and low demand at another, the reserve stock may be applied to rebalance the stock distribution in the system. The risk of imbalance is pooled over period 1, while stock is retained at the depot. This allocation policy will thereby reduce the expected level of backorders in the system without requiring additional system stock.

We may now restate the inventory problem given by (1) and (2). Let $V(Q,I_1,\ldots,I_n)$ denote a minimal weighted function of backorders in period 2, expressed as:

$$V(Q,I_1,\ldots,I_n) = \min_{S_{12},\ldots,S_{n2}} \sum_{i=1}^{n} \bar{\pi}_2 \, \mathcal{E}[(d_{i2}-S_{i2})^+]$$

subject to:

$$\sum_{i=1}^{n} S_{i2} \le Q + \sum_{i=1}^{n} I_i \,,$$

$$S_{i2} \ge I_i, \quad \text{for} \quad i = 1,\ldots,n, \tag{4}$$

where $\bar{\pi}_2 = (h_2 - s + \pi_2)$, $i = 1,\ldots,n$. Hence, $K_2(Q,I_1,\ldots,I_n)$ and $K(Q,S_{11},\ldots,S_{n1})$ from systems (1) and (2) can be written as:

$$K_2(Q,I_1,\ldots,I_n) = (h_2 - s)\left[Q + \sum_{i=1}^{n} I_i - \sum_{i=1}^{n} \mathcal{E}[d_{i2}]\right] + V(Q,I_1,\ldots,I_n) \,, \tag{5}$$

and

$$K(Q,S_{11},\ldots,S_{n1}) = \bar{c}\left[Q + \sum_{i=1}^{n} S_{i1}\right] + \sum_{i=1}^{n} \bar{\pi}_1 \mathcal{E}[(d_{i1} - S_{i1})^+]$$

$$+ \mathcal{E}[V(Q,S_{11}-d_{11},\ldots,S_{n1}-d_{n1})] - R \ , \qquad (6)$$

where:

$$\bar{c} = c + h_1 + h_2 - s \ , \qquad (7a)$$

$$\bar{\pi}_1 = h_1 + \pi_1 \ , \quad \text{for} \quad i = 1,\ldots,n, \text{ and} \qquad (7b)$$

$$R = (h_1 + h_2 - s) \sum_{i=1}^{n} \mathcal{E}[d_{i1}] + (h_2 - s) \sum_{i=1}^{n} \mathcal{E}[d_{i2}] \ . \qquad (7c)$$

2.2. The Demand Model

In each period, t, let $\{z_{it}, \ i = 1,\ldots,n\}$ be a collection of independent, identically distributed random variables with zero mean and unit variance and let δ_t be another zero mean, unit variance random variable such that δ_t, z_{1t}, ..., z_{nt}, t = 1,2, are mutually independent. Assume the demand in period t at retailer i, denoted d_{it}, is given by

$$d_{it} = \mu_{it} + \alpha_{it} z_{it} + \beta_{it} \delta_t , \qquad (8)$$

where μ_{it}, α_{it}, and β_{it} are parameters satisfying

$$\alpha_{it} = \sigma_{it}\sqrt{(1 - \rho_t)} \ , \text{ and}$$
$$\beta_{it} = \sigma_{it}\sqrt{\rho_t} \ , \qquad (9)$$

for some given nonnegative parameters σ_{it} and ρ_t $(0 \le \rho_t \le 1)$. For example, the common demand factor, δ_t, could represent random variation in general market share of total demand and z_{it} could represent local random variation. The model lends itself to econometric estimation of the parameters.

Assume $\sigma_{it} \neq 0$, for all i and t. Observe that the normalized demand variables, $(d_{it}-\mu_{it})/\sigma_{it}$, are identically distributed:

$$\frac{d_{it} - \mu_{it}}{\sigma_{it}} = \sqrt{1-\rho_t}\, z_{it} + \sqrt{\rho_t}\, \delta_t, \tag{10}$$

but are independent only if $\rho_t = 0$. Let $F_t(\cdot)$ denote the common marginal cumulative distribution function of the normalized demand variables and $\bar{F}_t(\cdot)$ denote the complementary cumulative distribution function ($\bar{F}_t(x) = 1-F_t(x)$, $\forall x$). Assume $F_t(\cdot)$ is absolutely continuous on the interval $(-\infty,+\infty)$ and assume the inverse complementary cumulative, denoted $\bar{F}_t^{-1}(\cdot)$, exists on $(0,1)$. Also observe that the unit variance assumption implies

$$\mathrm{Cov}\left[\frac{d_{it} - \mu_{it}}{\sigma_{it}}, \frac{d_{jt} - \mu_{jt}}{\sigma_{jt}}\right] = \begin{cases} 1, & \text{if } i = j, \\ \rho_t, & \text{if } i \neq j. \end{cases} \tag{11}$$

Thus, ρ_t is the correlation coefficient for any pair of normalized demand variables within a period. Negative correlation could be modelled by allowing σ_{it} to take on negative values for some of the retailers. In this case, demand for retailer i should be normalized by the absolute value of σ_{it}.

To clarify some of the expectation operations in what follows, let $\gamma_1(z_1)$ denote the common probability density function of the random variables $\{z_{i1},\ i = 1,\dots,n\}$ and let $\theta_1(\delta_1)$ denote the probability density function of the random variable δ_1. Let $\Gamma_1(z_1)$ and $\Theta_1(\delta_1)$ be the corresponding cumulative distribution functions. Note that $F_t(\cdot) = \Gamma_t(\cdot)$ if $\rho_t = 0$.

Let $\mathcal{L}(s)$ be the loss function associated with the random variables z_{i1}:

$$\begin{aligned}\mathcal{L}(s) &= \mathcal{E}[(z_{11}-s)^+] \\ &= \int_{-\infty}^{+\infty}(z_1 - s)^+\gamma_1(z_1)dz_1 \\ &= \int_s^{\infty}[1-\Gamma_1(y)]\,dy \qquad (12)\end{aligned}$$

If ρ_1 then the loss function simplifies to

$$\mathcal{L}(s) = \int_s^{\infty}[1-F_1(y)]\,dy.$$

Let $\mathcal{L}^{-1}(x)$ denote the inverse loss function for $x > 0$.

3. The Period 2 Minimum Fractile Allocation

The characterization of solutions to problems of the form (4) is well-known (see Zipkin [1980] for a general treatment of such problems):

Proposition 3.1: Let $(S^*_{12},\ldots,S^*_{n2})$ solve the period 2 allocation problem (4). Then there exists a common allocation factor k^* such that

$$S^*_{i2} = \max\,(\mu_{i2} + \sigma_{i2}k^*,\ I_i)\ , \qquad i = 1,\ldots,n. \tag{13}$$

The factor k^* is referred to as the minimum fractile.

Corollary 3.2: Under the same assumptions, the period 2 allocation factor, k^*, solves the following equation:

$$\sum_{i=1}^{n} \left(\mu_{i2} + \sigma_{i2}k^* - I_i\right)^+ = Q\ . \tag{14}$$

Proof: The constraint in (4) involving Q is binding. Substitution of (13) into this constraint yields (14).

The corollary permits us to characterize the solution to the period 2 allocation problem in terms of $(Q,S_{11},\ldots,S_{n1})$ and the period 1 demand variables. That is, k^* is a random variable satisfying

$$\sum_{i=1}^{n} \left(\mu_{i2} + \sigma_{i2}k^* - S_{i1} + d_{i1}\right)^+ = Q\ , \tag{15a}$$

and

$$S^*_{i2} = \max\,(\mu_{i2} + \sigma_{i2}k^*,\ S_{i1} - d_{i1})\ , \qquad i = 1,\ldots,n. \tag{15b}$$

4. Identical Retailers, Independent Demands

For the case in which retailer demands are independent and identically distributed, it is possible to derive analytically a number of results that cast light on the nature of the risk pooling effect. Assume $\rho_t = 0$, $\mu_{it} = \mu_t$ and $\sigma_{it} = \sigma_t$ for $i = 1,\ldots,n$, and $t = 1,2$. Also, assume that $F_t(\cdot)$ has support on $[-\mu_t/\sigma_t, \infty)$ to ensure that demands are nonnegative. We will restrict attention to policies satisfying $S_{i1} = S_1$, for all $i = 1,\ldots,n$, since it is easily shown that an optimal policy exists that provides the same period 1 allocation to all retailers.

4.1 Characterizations of the Period 2 Solution

For this sub-section only, random variables are indicated by a tilde (~) to emphasize their nature. The expectations in this section involve random functions of random variables.

The characterization of the period 2 solution, (15a) and (15b), can be specialized in the identical retailer case to

$$\sum_{i=1}^{n} [\tilde{S}_2^* - S_1 + \tilde{d}_{i1}]^+ = Q \,, \tag{16a}$$

and

$$\tilde{S}_{i2}^* = \max\,(\tilde{S}_2^*, S_1 - \tilde{d}_{i1}) \,, \qquad i = 1,\ldots,n, \tag{16b}$$

where $\tilde{S}_2^* = \mu_2 + \sigma_2\tilde{k}^*$, the common minimum period 2 stock level, referred to here as the *period 2 stock floor*.

Let $\tilde{d}_{[1]} \le \tilde{d}_{[2]} \le \cdots \le \tilde{d}_{[n]}$ denote the ordered period 1 demands. Then (16a) and (16b) can be rewritten as

$$\sum_{i=1}^{n} \left(\tilde{S}_2^* - S_1 + \tilde{d}_{[i]}\right)^+ = Q \ , \tag{17a}$$

and

$$\tilde{S}_{[i]2}^* = \max\ (\tilde{S}_2^*,\ S_1 - \tilde{d}_{[i]}) \ , \qquad i = 1,\ldots,n, \tag{17b}$$

where $[i]$ is the index of the retailer with the i'th smallest period 1 demand.

Let $\tilde{N}(x)$ denote the number of retailers that experience demands in period 1 not exceeding x, for $x \ge 0$. Define $\tilde{N}(x) \equiv 0$, for all $x < 0$. The mapping $\tilde{N}:\mathcal{R} \longrightarrow \{0,1,2,\ldots,n\}$ is nondecreasing, increases by jumps only, and is right-continuous. Assuming that the demand distribution function $F_1(\cdot)$ is absolutely continuous, we note that, with probability one, $\tilde{N}(0) = 0$ and each jump of $\tilde{N}(\cdot)$ is of unit magnitude.

Yet another characterization of the period 2 stock floor is provided by the following lemma. This characterization will prove useful when investigating the limiting behavior of $\tilde{S}_2^*$.

Lemma 4.1: In the identical retailer, independent demand case, the value of $\tilde{S}_2^*$ is determined by

$$\int_{S_1 - \tilde{S}_2^*}^{\infty} \left[n - \tilde{N}(x)\right] dx = Q \ . \tag{18}$$

Proof: Since $\tilde{N}(\cdot)$ increases only by unit jumps, with probability one, we have

$$\int_y^\infty \left[n - \tilde{N}(x)\right] dx = \sum_{i=0}^{n} (\tilde{d}_{i1} - y)^+ .$$

Letting $y = S_1 - \tilde{S}_2^*$ and noting (16a) results in (18).■

The relationship between $\tilde{d}_{[1]}, \ldots, \tilde{d}_{[n]}$, $\tilde{N}(\cdot)$, Q, and $S_1 - \tilde{S}_2^*$, for a particular realization of period 1 demands, is illustrated in Figure 1 and Figure 2.

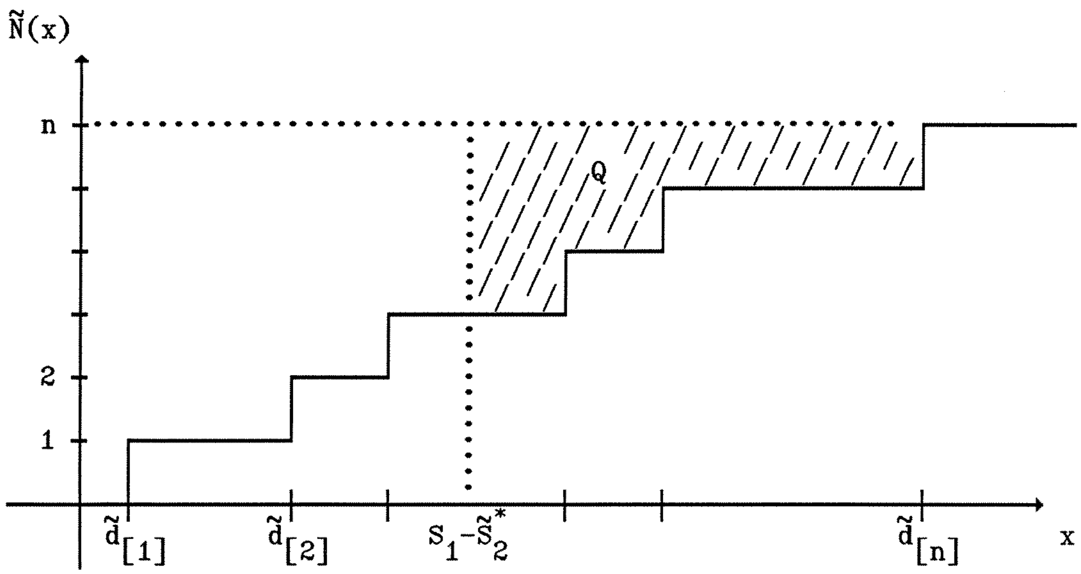

Figure 1. Characterization of $\tilde{S}_2^*$ in terms of $\tilde{N}(x)$

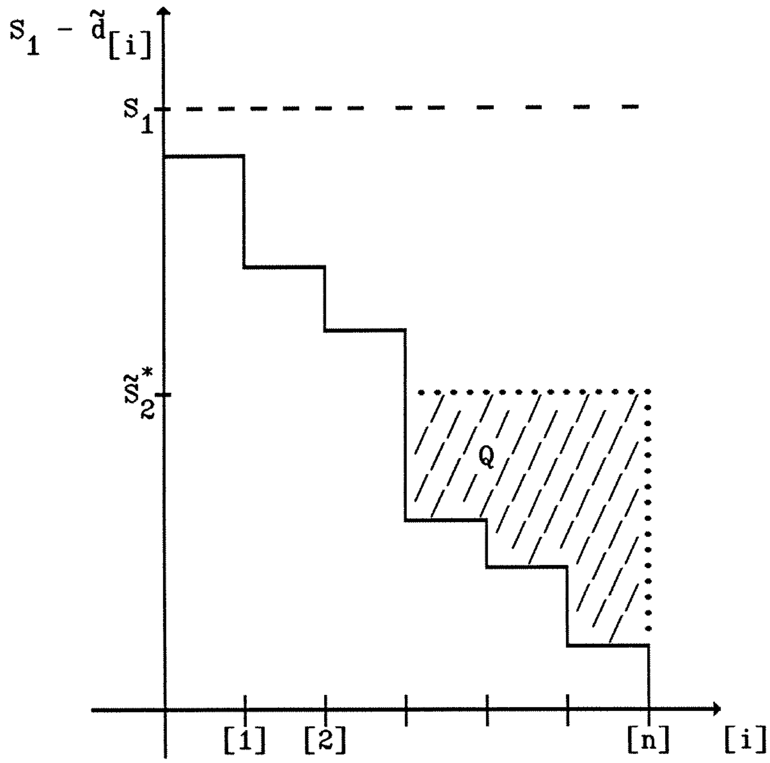

Figure 2. Characterization of $\tilde{S}_2^*$ in terms of $S_1-\tilde{d}_{[i]}$

Building on the above characterization of the period 2 solution, Jackson and Muckstadt [1984] develop the first order optimality conditions for the period 1 decision variables. Those conditions involve the convolutions of truncated random variables. Solving the conditions exactly is computationally feasible only for the exponential distribution. In this paper, we explore the behavior of this characterization and use it to suggest an approximation to the cost function and optimization techniques.

4.2 Risk Pooling in Two Identical Retailer Systems

For a single retailer, it is easily shown that $S_{12}^* = Y - d_{11}$, where Y is the initial total system stock. That is, the amount of depot reserve stock has no effect on the period 2 stock level after allocation. There is no risk pooling if there is only one retailer.

On the other hand, if there are two identical retailers the period 2 stock floor is given by

$$S_2^* = \min \left\{ S_1 - d_{11} + Q,\ S_1 - d_{21} + Q,\ \frac{Y - d_{11} - d_{21}}{2} \right\} .$$

Noting that $S_1 = (Y - Q)/2$ simplifies this expression to

$$S_2^* = \frac{Y}{2} + \min \left\{ \frac{Q}{2} - d_{11},\ \frac{Q}{2} - d_{21},\ \frac{-(d_{11} + d_{21})}{2} \right\} .$$

Differentiation reveals that

$$\frac{d\mathcal{E}[S_2^*]}{dQ} = \mathcal{P}\ \{d_{11} - d_{21} \geq Q\} . \tag{19}$$

Since this derivative is positive for sufficiently small Q, it is clear that it is technically possible to see a benefit in period 2 from holding depot stock in reserve even with as few as two retailers. Whether the improvement in period 2 system performance outweighs the degradation in period 1 performance (since $dS_1/dQ < 0$) or not depends on the magnitude of the effect and the various cost parameters involved. Observe that at $Q = 0$, this derivative is exactly 1/2 and that this is an upper bound on the derivative. That is, an increase in depot reserve stock would increase the expected period 2 stock floor at each of the retailers by at most one half of the increase in Q.

4.3 Limiting Behavior of the Period 2 Minimum Stock Allocation

For a known value of $S_2 \in (-\infty,+\infty)$, let $X_i(S_2)$ be given by

$$X_i(S_2) = [S_2 - S_1 + d_{i1}]^+ ,$$

for $i = 1,\ldots,n$. Suppose the period 1 stock level, S_1, is fixed, for all large values of n. Then, by the strong law of large numbers [Chung, 1974, Theorem 5.4.1, p. 124], as $n \to \infty$,

$$\frac{1}{n}\sum_{i=1}^{n} (X_i(S_2)-\mathcal{E}[X_i(S_2)]) \longrightarrow 0 \qquad \text{a.e.} \tag{20}$$

This simple result suggests that we work with expected values in (16) for large values of n. Accordingly, let $\hat{S}_2$, a function of (Q,S_1) for $Q > 0$, denote the unique solution for S_2 to the following equation:

$$\sum_{i=1}^{n} \mathcal{E}\left[(S_2 - S_1 + d_{i1})^+\right] = Q . \tag{21}$$

Let $\underline{S}_2$ denote the infimum of values S_2 such that the left hand side of (21) is positive. That a unique solution exists to (21) follows from the fact that the left hand side is a strictly increasing function of S_2 on $(\underline{S}_2,\infty)$ with range $(0,+\infty)$, and $Q > 0$.

Rewriting (21) in terms of the loss function (12) yields:

$$\sum_{i=1}^{n} \mathcal{L}\left[\frac{S_1 - \hat{S}_2 - \mu_1}{\sigma_1}\right] = Q/\sigma_1 .$$

Hence,

$$\hat{S}_2 = -\sigma_1 \Phi^{-1}(Q/(n\sigma_1)) + S_1 - \mu_1. \tag{22}$$

The convergence in (20) is based on the assumption of a known value of S_2. However, by Lemma 4.1, S_2^* is a random variable satisfying (18). That is, S_2^* is a random variable depending on the decisions Q and S_1 and on the period 1 demand variables $d_{11},\ldots,d_{n1}$ (which determine the function $N(\cdot)$). In this sub-section, we develop conditions under which it can be shown that $S_2^* \longrightarrow \hat{S}_2$, in probability, as $n \longrightarrow \infty$.

We assume that total expected system demand grows linearly with the number of retailers and that demand at each of the retailers is independent. Under this assumption we argue that S_2^* converges in probability to a constant.

Let $F_1^n(y)$ denote the proportion of n retailers observing normalized demand less than or equal to y. $F_1^n(y)$ is called the *empirical distribution function* of the normalized demand, $\sigma_1^{-1}(d_{11}-\mu_1)$, ..., $\sigma_1^{-1}(d_{n1}-\mu_1)$. Dividing through (18) by n and employing a change of variable shows that S_2^* must satisfy

$$\int_{\sigma_1^{-1}(S_1-S_2^*-\mu_1)}^{\infty} \left[1 - F_1^n(y)\right] dy = \sigma_1^{-1}Q/n \; . \tag{23}$$

For large values of n, the empirical distribution function is well approximated by $F_1(y)$, the true normalized distribution function. We observe that if Q and S_1 are well-behaved as functions of n, then so is

S_2^*. In what follows, we will assume that $Q = qn$, for fixed $q > 0$, and that S_1 is fixed. An intuitive justification can be made if we start with the weaker assumption that total system stock grows linearly with the number of retailers. Since there are limits to the usefulness of increasing S_1 (for large values of S_1 the period 1 stockout probability is negligible), any growth in system safety stock for large numbers of retailers must take place at the depot level.

For technical reasons, we further assume that the distribution of normalized demand is concentrated on some compact interval $[-a,a]$ with $F_1(-a) = 0$ and $F_1(a) = 1$. This contradicts the earlier assumption of unbounded demands used in the proof of Proposition 3.1; but, a version of that proposition can be made to hold provided a is sufficiently large. The spirit of these results should still apply for unbounded demands; but, there are difficult technical issues to resolve when presenting the proofs in this case. These issues detract from the main purpose of this paper and so are avoided. As before, the distribution function, F_1, is assumed to be absolutely continuous.

Define the *empirical loss function*, $\mathcal{L}_n(x)$ by

$$\mathcal{L}_n(x) = \int_x^a \left[1 - F_1^n(y)\right] dy \ . \tag{24}$$

The real-valued random function $\mathcal{L}_n$ is nonincreasing and continuous on $[0,\infty)$. Let $x_0 = \sup \{x:\mathcal{L}_n(x)>0\}$ and note that $\mathcal{L}_n$ is strictly decreasing, continuous, and unbounded on $(-\infty,x_0)$. Hence, $\mathcal{L}_n$ has a unique inverse function, denoted by $\mathcal{L}_n^{-1}$, defined on $(0,\infty)$. For $q > 0$, solving (23) yields:

$$S_2^* = -\sigma_1 [\mathcal{L}_n^{-1}(q/\sigma_1)] + S_1 - \mu_1 \ . \tag{25}$$

We are thus led to consider the limiting behavior of $\mathcal{L}_n^{-1}$ as $n \to \infty$.

Let m denote the maximum total demand that can occur at any retailer over two periods. By assumption, $m < \infty$.

Theorem 4.2: Under the assumptions of this section, for $q < m$,

$$\mathcal{L}_n^{-1}(q/\sigma_1) \to \mathcal{L}^{-1}(q/\sigma_1) \ ,$$

in probability as $n \to \infty$.

Proof: Appendix 1.

Corollary 4.3: Under the assumptions of this section, for $q < m$, $S_2^* \to \hat{S}_2$ in probability as $n \to \infty$.

Corollary 4.3 is used as the basis for approximate computational techniques developed in the next section.

Proposition 4.4 For fixed Y,

$$\frac{d\hat{S}_2}{dQ} = n^{-1} \left[\bar{F}_1 \left[\frac{S_1 - \hat{S}_2 - \mu_1}{\sigma_1} \right]^{-1} - 1 \right] . \tag{26}$$

Proof: Let $\hat{k} = (\hat{S}_2-\mu_2)/\sigma_2$. By lemma 5.1, below, the partial derivative $\partial\hat{S}_2/\partial Q$ $(=\sigma_2\partial\hat{k}/\partial Q)$ is given by

$$\frac{\partial\hat{S}_2}{\partial Q} = n^{-1}\,\overline{\Gamma}_1\left[\frac{S_1-\hat{S}_2-\mu_1}{\sigma_1}\right]^{-1},$$

where $\overline{\Gamma}_1(\cdot) = \overline{F}_1(\cdot)$ in the case $\rho_1 = 0$.

Together with (22) and the fact that $dS_1/dQ = -n^{-1}$, since $S_1 = (Y-Q)/n$, this implies (26).■

Relation (26) is a many-retailer analog of (19). Observe from the definition of $\hat{S}_2$ in (21) that

$$\lim_{Q \to 0} \overline{F}_1\left[\frac{S_1-\hat{S}_2-\mu_1}{\sigma_1}\right] = 0 .$$

Consequently, $d\hat{S}_2/dQ \to +\infty$ as $Q \to 0$, in contrast to the two retailer case in which $d\mathcal{E}[S_2^*]/dQ$ is bounded by 1/2.

We can now identify two aspects to risk pooling in the current model. The first is given by relations (19) and (26) indicating that at sufficiently small values of depot reserve stock, Q, there is a positive technical benefit to increasing Q in terms of the period 2 stock floor. We could refer to this as the *first order effect of risk pooling* since it relates to the expected value of the period 2 stock floor. That the derivative in the many-retailer case dominates the derivative in the two retailer case, at least for small values of Q, suggests that the degree of risk pooling in this sense increases with the number of retailers in the system.

The second aspect of risk pooling is given by Corollary 4.3, namely that as the number of retailers increases, the period 2 stock floor stabilizes, at any positive value of Q. In a system with many retailers, each experiencing independent demands, the depot can practically guarantee a fixed period 2 stock floor to each retailer. This could be referred to as the *second order effect of risk pooling* since it relates to the variance of the period 2 stock floor.

Whether either of these technical effects of risk pooling justifies the use of depot reserve stock depends on the cost parameters. In the next section, we develop computational procedures for determining near-optimal values of Q and the vector of period 1 allocations.

5. Computational Procedures

5.1. An Approximate Cost Function

A computational procedure for finding the optimal period 1 decisions for the two retailer case is given in Brown [1984]. For more than two retailers, the problem of finding the exact optimal solution is computationally intractable for all but the simplest demand models (Jackson and Muckstadt [1984]). It is the goal of the remainder of this paper to develop computationally efficient procedures to find good solutions. The approach is to approximate the cost function in the many retailer case using the results of the preceding section. For sufficiently many retailers the approximation should be quite good. It is not known how accurate it is for small numbers of retailers.

Let $\underset{\sim}{S}_1 = (S_{11},\ldots,S_{n1})$, the vector of period 1 allocations, and return to the general demand model of Section 2.

Substituting for d_{i1} from (8) in (15a) yields

$$\sum_{i=1}^{n} [\mu_{i1} + \mu_{i2} + \alpha_{i1}z_{i1} + \beta_{i1}\delta_1 + \sigma_{i2}k^* - S_{i1}]^+ = Q \ .$$

Analogous to (21), for a given vector of period 1 decisions, $(Q,S_{11},\ldots,S_{n1})$, with $Q > 0$, and for a given value of the market variable δ_1, let $\hat{k}$ denote the unique solution for k to the following equation:

$$\sum_{i=1}^{n} \mathcal{E}\Big[[\mu_{i1} + \mu_{i2} + \alpha_{i1}z_{i1} + \beta_{i1}\delta_1 + \sigma_{i2}k - S_{i1}]^+ \Big| \ \delta_1\Big] = Q \ . \qquad (27)$$

Let

$$\hat{\ell}_i = \mu_{i1} + \mu_{i2} + \sigma_{i2}\hat{k} + \beta_{i1}\delta_1 \ , \tag{28}$$

for $i = 1,\ldots,n$, and rewrite (27) using the loss function, $\mathcal{L}(\cdot)$:

$$\sum_{i=1}^{n} \alpha_{i1}\mathcal{L}\left(\frac{S_{i1}-\hat{\ell}_i}{\alpha_{i1}}\right) = Q \ . \tag{29}$$

The event that retailer i receives a shipment in the optimal period 2 allocation can be approximated by E_i:

$$E_i = \left\{ z_{i1} > \frac{S_{i1}-\hat{\ell}_i}{\alpha_{i1}} \right\} .$$

Similarly, the event that retailer i receives no shipment in period 2 can be approximated by the complementary event $\bar{E}_i$. The conditional probability of event E_i, conditioned on the common demand factor, δ_1, is given by

$$\mathcal{P}\{ E_i \mid \delta_1\} = \bar{\Gamma}_1\left(\frac{S_{i1}-\hat{\ell}_i}{\alpha_{i1}}\right) ,$$

where $\bar{\Gamma}_1$ is the complementary cumulative distribution of $z_{i1}, i = 1,\ldots,n$.

Lemma 5.1: For a given common demand factor, δ_1, the partial derivatives of $\hat{k}(Q,\underset{\sim}{S}_1,\delta_1)$ are given by

$$\frac{\partial \hat{k}}{\partial Q}(Q,S_{11},\ldots,S_{n1}) = \frac{1}{\sum_{i=1}^{n} \sigma_{i2}\overline{\Gamma}_1\left[\frac{S_{i1}-\hat{\ell}_i}{\sigma_{i1}}\right]}\ , \tag{30}$$

and

$$\frac{\partial \hat{k}}{\partial S_{i1}}(Q,S_{11},\ldots,S_{n1}) = \overline{\Gamma}_1\left[\frac{S_{i1}-\hat{\ell}_i}{\sigma_{i1}}\right]\frac{\partial \hat{k}}{\partial Q}(Q,\underset{\sim}{S}_1,\delta_1)\ , \tag{31}$$

for $i = 1,\ldots,n$.

Proof: The proposition is a straightforward application of the implicit function theorem [eg. Benavie, 1972, Theorem 1.9, p. 26] applied to (27). Totally differentiating (27) yields

$$\sum_{i=1}^{n} \sigma_{i2}\mathcal{P}\left\{z_{i1} > \frac{S_{i1}-\mu_{i1}-\mu_{i2}-\sigma_{i2}\hat{k}-\beta_{i1}\delta_1}{\alpha_{i1}} \;\middle|\; \delta_1\right\}d\hat{k}$$

$$+ \sum_{i=1}^{n} \mathcal{P}\left\{z_{i1} > \frac{S_{i1}-\mu_{i1}-\mu_{i2}-\sigma_{i2}\hat{k}-\beta_{i1}\delta_1}{\alpha_{i1}} \;\middle|\; \delta_1\right\}dS_{i1} = dQ\ .$$

Simplifying, using (28), yields:

$$\sum_{i=1}^{n} \sigma_{i2} \bar{\Gamma}_1 \left[\frac{S_{i1} - \hat{\ell}_i}{\sigma_{i1}} \right] d\hat{k} + \sum_{i=1}^{n} \bar{\Gamma}_1 \left[\frac{S_{i1} - \hat{\ell}_i}{\sigma_{i1}} \right] dS_{i1} = dQ \quad .$$

Formally, this implies (30) and, together with (30), implies (31). ∎

Let the optimal period 2 decision S_{i2}^* be approximated by $\hat{S}_{i2}$:

$$\hat{S}_{i2} = \begin{cases} \mu_{i2} + \sigma_{i2}\hat{k} \ , & \text{if } \ E_i, \\ S_{i1} - \mu_{i1} - \alpha_{i1} z_{i1} - \beta_{i1}\delta_1, & \text{if } \ \bar{E}_i, \end{cases} \tag{32}$$

for $i = 1,\ldots,n$. Let $\hat{B}(Q,\underset{\sim}{S}_1)$ denote the following approximation to the minimal period 2 weighted backorder function:

$$\begin{aligned} \hat{B}(Q,\underset{\sim}{S}_1) &= \sum_{i=1}^{n} \bar{\pi}_2 \mathcal{E}[(d_{i2} - \hat{S}_{i2})^+] \\ &= \sum_{i=1}^{n} \bar{\pi}_2 \Big\{ \mathcal{E}[1_{\{E_i\}}(d_{i2} - \mu_{i2} - \sigma_{i2}\hat{k})^+] \\ &\qquad + \mathcal{E}[1_{\{\bar{E}_i\}}(d_{i2} - S_{i1} + \mu_{i1} + \alpha_{i1} z_{i1} + \beta_{i1}\delta_1)^+] \Big\}, \end{aligned} \tag{33}$$

where $1_{\{E\}}$ is the indicator function of event E (that is, $1_{\{E\}} = 1$ if E is true, 0 otherwise). Similarly, let $\hat{K}(Q,\underset{\sim}{S}_1)$ denote the approximate total two period cost function:

$$\hat{K}(Q,\underset{\sim}{S}_1) = \bar{c}\left[Q + \sum_{i=1}^{n} S_{i1}\right] + \sum_{i=1}^{n} \bar{\pi}_1 \mathcal{E}\left[(d_{i1} - S_{i1})^+\right] - R$$

$$+ \sum_{i=1}^{n} \bar{\pi}_2 \Big\{ \mathcal{E}\left[1_{\{E_i\}}(d_{i2} - \mu_{i2} - \sigma_{i2}\hat{k})^+\right] \tag{34}$$

$$+ \mathcal{E}\left[1_{\{\bar{E}_i\}}(d_{i2} - S_{i1} + \mu_{i1} + \alpha_{i1} z_{i1} + \beta_{i1}\delta_1)^+\right]\Big\} .$$

In general, even this approximate cost function is difficult to optimize. However, there are two special cases in which an efficient optimization procedure can be developed. These are considered in the next two sub-sections.

5.2. Non-identical Retailers, Independent Demands

In this sub-section, an optimization procedure for minimizing $\hat{K}(Q,\underset{\sim}{S}_1)$ is developed for the case of independent demands ($\rho_t = 0$, $t = 1,2$). The location and scale parameters (μ_{it},σ_{it}; $i = 1,\ldots n$) are not required to be identical across retailers. In the next sub-section, the correlation coefficient is allowed to be nonzero but the demands are assumed to be identically distributed.

Proposition 5.2: Assume $\rho_t = 0$. Let $(\hat{Q},\hat{\underset{\sim}{S}}_1)$ minimize the approximate cost function $\hat{K}$ in (34), ignoring nonnegativity restrictions. The first order necessary conditions for a minimum reduce to

$$\hat{k}(\hat{Q},\hat{\underset{\sim}{S}}_1) = \bar{F}_2^{-1}\left[\frac{\bar{c}}{\bar{\pi}_2}\right] , \tag{35}$$

and

$$\bar{\pi}_1\bar{\Gamma}_1\left[\frac{\hat{S}_{i1}-\mu_{i1}}{\sigma_{i1}}\right] + \bar{\pi}_2\int_{-\infty}^{\frac{\hat{S}_{i1}-\hat{\ell}_i}{\sigma_{i1}}} \bar{F}_2\left[\frac{\hat{S}_{i1}-\mu_{i1}-\sigma_{i1}z_{i1}-\mu_{i2}}{\sigma_{i2}}\right]\Gamma_1(dz_{i1})$$

$$= \bar{c}\ \Gamma_1\left[\frac{\hat{S}_{i1}-\hat{\ell}_i}{\sigma_{i1}}\right], \tag{36}$$

for each $i = 1,\ldots,n$.

Proof: Setting $\partial\hat{K}/\partial Q = 0$ yields

$$0 = \bar{c} - \bar{\pi}_2\sum_{i=1}^{n}\sigma_{i2}\cdot\frac{\partial\hat{k}}{\partial Q}\cdot\mathcal{E}\left[1_{\{E_i\}}1_{\{d_{i2}>\mu_{i2}+\sigma_{i2}\hat{k}\}}\right]$$

$$= \bar{c} - \bar{\pi}_2\sum_{i=1}^{n}\sigma_{i2}\cdot\frac{\partial\hat{k}}{\partial Q}\cdot\mathcal{P}\left\{z_{i1}>\frac{\hat{S}_{i1}-\hat{\ell}_i}{\sigma_{i1}}\right\}\cdot\mathcal{P}\left\{d_{i2}>\mu_{i2}+\sigma_{i2}\hat{k}\right\},$$

since z_{i1} and d_{i2} are independent and $\hat{k}$ and $(\hat{\ell}_i;\ i = 1,\ldots,n)$ are probabilistically constant when $\rho_1 = 0$. Hence,

$$0 = \bar{c} - \bar{\pi}_2\cdot\frac{\partial\hat{k}}{\partial Q}\cdot\sum_{i=1}^{n}\sigma_{i2}\cdot\bar{\Gamma}_1\left[\frac{\hat{S}_{i1}-\hat{\ell}_i}{\sigma_{i1}}\right]\bar{F}_2(\hat{k}) .$$

Substituting for $\partial\hat{k}/\partial Q$ from (30) shows that $\bar{\pi}\bar{F}_2(\hat{k}) = \bar{c}$, which implies (35).

Similarly, setting $\partial\hat{K}/\partial S_{i1} = 0$ for each i yields

$$0 = \bar{c} - \bar{\pi}_1 \mathcal{P}\{d_{i1} > \hat{S}_{i1}\} - \bar{\pi}_2 \sum_{j=1}^{n} \sigma_{j2} \cdot \frac{\partial \hat{k}}{\partial S_{i1}} \cdot \mathcal{E}\left[1_{\{E_j\}} 1_{\{d_{j2} > \mu_{j2} + \sigma_{j2}\hat{k}\}}\right]$$

$$- \bar{\pi}_2 \mathcal{E}\left[1_{\{\bar{E}_i\}} 1_{\{d_{i2} > S_{i1} - \mu_{i1} - \sigma_{i1} z_{i1}\}}\right]$$

$$= \bar{c} - \bar{\pi}_1 \mathcal{P}\{d_{i1} > \hat{S}_{i1}\} - \bar{\pi}_2 \cdot \frac{\partial \hat{k}}{\partial S_{i1}} \cdot \sum_{j=1}^{n} \sigma_{j2} \cdot \bar{\Gamma}_1\left(\frac{\hat{S}_{j1} - \hat{\ell}_j}{\sigma_{j1}} \right) \bar{F}_2(\hat{k})$$

$$- \bar{\pi}_2 \mathcal{E}\left[1_{\{\bar{E}_i\}} 1_{\{d_{i2} > S_{i1} - \mu_{i1} - \sigma_{i1} z_{i1}\}}\right]$$

$$= \bar{c} - \bar{\pi}_1 \mathcal{P}\{d_{i1} > \hat{S}_{i1}\} - \bar{\pi}_2 \cdot \bar{\Gamma}_1\left(\frac{\hat{S}_{i1} - \hat{\ell}_i}{\sigma_{i1}} \right) \bar{F}_2(\hat{k})$$

$$- \bar{\pi}_2 \mathcal{E}\left[1_{\{\bar{E}_i\}} 1_{\{d_{i2} > S_{i1} - \mu_{i1} - \sigma_{i1} z_{i1}\}}\right] ,$$

after substituting for $\partial\hat{k}/\partial S_{i1}$ from (31) and for $\partial\hat{k}/\partial Q$ from (30) in the resulting expression and cancelling terms. Now, $\bar{\pi}_2 \bar{F}_2(\hat{k}) = \bar{c}$, by (35), so,

$$0 = \bar{c} - \bar{\pi}_1 \mathcal{P}\{d_{i1} > \hat{S}_{i1}\} - \bar{c} \cdot \bar{\Gamma}_1\left(\frac{\hat{S}_{i1} - \hat{\ell}_i}{\sigma_{i1}} \right)$$

$$- \bar{\pi}_2 \mathcal{E}\left[1_{\{\bar{E}_i\}} 1_{\{d_{i2} > S_{i1} - \mu_{i1} - \sigma_{i1} z_{i1}\}}\right] ,$$

which is equivalent to (36). ∎

Remarks: Equation (35) implies that at the (approximate) optimum in period 2, the probability of a stockout at retailer i should not exceed $\bar{c}/\pi_2$, with equality holding if retailer i receives stock as a result of the period 2 allocation. Equation (36) has the following economic interpretation: the first term on the left hand side is the marginal benefit in period 1 of an extra unit of stock assigned to retailer i in period 1. The second term is the marginal benefit of that extra unit of stock to retailer i in period 2: the cost-weighted probability that retailer i does not receive stock in the period 2 allocation, event $\bar{E}_i$, but does experience a stockout at the end of period 2. The right hand side is the marginal cost of an additional unit of stock at retailer i, $\bar{c}$, weighted by the probability of event E_i — that is, retailer i effectively receives a credit for any additional stock purchased if happens that it shares in the equal fractile allocation of the second period.

Observe that the approximately optimal decision vector $(\hat{Q}, \hat{S}_{11}, \ldots, \hat{S}_{n1})$ can be determined as the result of a sequence of n one-dimensional searches:

Algorithm 5.3 (Non-Identical Retailers, Independent Demands)

1. Set $\hat{k} \leftarrow \bar{F}_2^{-1}(\bar{c}/\pi_2)$.
2. For $i = 1,\ldots,n$:
   - 2a. Determine $\hat{\ell}_i$ by (28) using $\hat{k}$ from step 1.
   - 2b. Solve (36) for $\hat{S}_{i1}$ using a one dimensional search technique.
3. Substitute $\hat{k}, \hat{S}_{11}, \ldots,$ and $\hat{S}_{n1}$ into (27) to determine $\hat{Q}$.

5.3. Identical Retailers, Correlated Demands

In this sub-section we relax the assumption of independence to allow correlated demands ($\rho_t > 0$, $t = 1,2$) but we restrict attention to the special case in which retailers face identically distributed demands ($\mu_{it} = \mu_t$, $\sigma_{it} = \sigma_t$, $\alpha_{it} = \alpha_t$, and $\beta_{it} = \beta_t$, for $i = 1,\ldots,n$, and $t = 1,2$). We derive an algorithm analogous to Algorithm 5.3 to determine an approximately optimal solution.

As in Section 4, we limit attention to policies that satisfy $S_{i1} = S_1$ for some S_1 and for all $i = 1,\ldots,n$. Observe from (28) that for identical retailers, $\hat{\ell}_i = \hat{\ell}$, for all i. For $Q > 0$, the solution to (29) yields

$$\hat{\ell} = -\alpha_1 \mathcal{I}^{-1}(Q/(\alpha_1 n)) + S_1 \quad . \tag{36}$$

Note that $\hat{\ell}$ does not depend on δ_1. Also, letting $\hat{S}_2 = \mu_2 + \sigma_2\hat{k}$, note that by (28) and (36)

$$\hat{S}_2 = -\alpha_1 \mathcal{I}^{-1}(Q/(\alpha_1 n)) + S_1 - \mu_1 - \beta_1\delta_1,$$

which agrees with (22) when $\rho_1 = 0$.

Proposition 5.4: Let $(\hat{Q},\hat{S}_1)$ minimize the approximate cost function $\hat{K}$ in (34), ignoring nonnegativity restrictions. The first order necessary conditions for a minimum reduce to

$$\int_{-\infty}^{+\infty} \bar{F}_2\left[\ \hat{k}(\hat{Q},\hat{S}_1,\delta_1)\right]\theta_1(\delta_1)d\delta_1 = \frac{\bar{c}}{\bar{\pi}_2} \quad , \tag{37}$$

and

$$\bar{\pi}_1\bar{F}_1\left[\frac{\hat{S}_1-\mu_1}{\sigma_1}\right] + \bar{\pi}_2 \int_{-\infty}^{+\infty}\int_{-\infty}^{\frac{\hat{S}_1-\hat{\ell}}{\sigma_1}} \bar{F}_2\left[\frac{\hat{S}_1-\mu_1-\alpha_1 z_1-\beta_1\delta_1-\mu_2}{\sigma_2}\right]\gamma_1(z_1)\theta(\delta_1)dz_1 d\delta_1$$

$$= \bar{c}\Gamma_1\left[\ \frac{\hat{S}_1-\hat{\ell}}{\alpha_1}\ \right] \ , \tag{38}$$

Proof: Setting $\partial\hat{K}/\partial Q = 0$ yields

$$0 = n\bar{c} - n\sigma_2\bar{\pi}_2\cdot\int_{-\infty}^{+\infty}\frac{\partial\hat{k}}{\partial Q}\cdot\mathcal{E}\left[1_{\{E_1\}}1_{\{d_2>\mu_2+\sigma_2\hat{k}\}}|\delta_1\right]\theta_1(\delta_1)d\delta_1$$

$$= n\bar{c} - n\sigma_2\bar{\pi}_2\cdot\int_{-\infty}^{+\infty}\frac{\partial\hat{k}}{\partial Q}\cdot\bar{\Gamma}_1\left[\ \frac{\hat{S}_1-\hat{\ell}}{\alpha_1}\ \right]\bar{F}_2\left[\ \hat{k}\ \right]\theta_1(\delta_1)d\delta_1 \quad ,$$

since $\hat{\ell}$ does not depend on δ_1. Substituting for $\partial\hat{k}/\partial Q$ from (30) and cancelling terms results in (37).

Similarly, setting $\partial\hat{k}/\partial S_1 = 0$ yields

$$0 = n\bar{c} - \bar{\pi}_1 \mathcal{P}\left\{d_1 > \hat{S}_1\right\} - n\sigma_2\bar{\pi}_2 \int_{-\infty}^{+\infty} \frac{\partial\hat{k}}{\partial S_1} \mathcal{E}[1_{\{E_1\}} 1_{\{d_2 > \mu_2 + \sigma_2\hat{k}\}} | \delta_1] \theta_1(\delta_1) d\delta_1$$

$$- n\bar{\pi}_2 \int_{-\infty}^{+\infty} \mathcal{E}[1_{\{\bar{E}_1\}} 1_{\{d_2 > \hat{S}_1 - \mu_1 - \alpha_1 z_1 - \beta_1\delta_1\}} | \delta_1] \theta_1(\delta_1) d\delta_1$$

$$= n\bar{c} - \bar{\pi}_1 \mathcal{P}\left\{d_1 > \hat{S}_1\right\} - n\bar{\pi}_2 \bar{\Gamma}_1\left[\frac{\hat{S}_1 - \hat{\ell}}{\alpha_1}\right] \int_{-\infty}^{+\infty} \bar{F}_2[\,\hat{k}\,] \theta_1(\delta_1) d\delta_1$$

$$- n\bar{\pi}_2 \int_{-\infty}^{+\infty} \mathcal{E}[1_{\{\bar{E}_1\}} 1_{\{d_2 > \hat{S}_1 - \mu_1 - \alpha_1 z_1 - \beta_1\delta_1\}} | \delta_1] \theta_1(\delta_1) d\delta_1 \quad ,$$

after substituting for $\partial\hat{k}/\partial S_1$ from (31). By (37), this becomes

$$0 = n\bar{c} - \bar{\pi}_1 \mathcal{P}\left\{d_1 > \hat{S}_1\right\} - n\bar{c}\cdot\bar{\Gamma}_1\left[\frac{\hat{S}_1 - \hat{\ell}}{\alpha_1}\right]$$

$$- n\bar{\pi}_2 \int_{-\infty}^{+\infty} \mathcal{E}[1_{\{\bar{E}_1\}} 1_{\{d_2 > \hat{S}_1 - \mu_1 - \alpha_1 z_1 - \beta_1\delta_1\}} | \delta_1] \theta_1(\delta_1) d\delta_1 \quad ,$$

which is equivalent to (38). ∎

Substituting (28) into (37) yields:

$$\int_{-\infty}^{+\infty} \bar{F}_2\left[\frac{\hat{\ell} - \mu_1 - \mu_2 - \beta_1\delta_1}{\sigma_2}\right] \theta_1(\delta_1) d\delta_1 = \frac{\bar{c}}{\bar{\pi}_2} \quad , \tag{39}$$

Consequently, only two dimensional searches are required to find an approximate optimal solution in this case:

Algorithm 5.5 (Identical Retailers, Correlated Demand)

1. Solve (39) for $\hat{\ell}$ using a one-dimensional search technique.
2. Solve (38) for $\hat{S}_1$ using a one-dimensional search technique.
3. Determine $\hat{Q}$ according to

$$\hat{Q} = n\alpha_1 \mathcal{I}\left[\frac{\hat{S}_1 - \hat{\ell}}{\alpha_1}\right] .$$

6. Conclusions

The main contributions of this paper are the development of a model to explore the risk pooling effect of depot stock in a two-echelon distribution system, the identification of both a first order and a second order aspect of the risk pooling effect, and the development of computational procedures to find near-optimal values for the policy parameters in two special cases.

Preliminary computational experience indicates that the algorithms proposed here yield solution values for the policy parameters that are within a few units of the optimal values for the case of two identical retailers. Since the approximation technique is based on the assumption of a large number of retailers, this two retailer comparison suggests that the procedure may be quite robust. The experiments conducted to date are too limited to report here. Further computational experimentation in this area is clearly needed.

References

Aggarwal, S.C., "A Review of Current Inventory Theory and its Applications", *International Journal of Production Research*, Vol. 12, No. 4 (1974), pp. 443-482.

Agnihothri, S., U.S. Karmarkar and P. Kabat, "Stochastic Allocation Rules", *Operations Research*, Vol. 30, No. 3 (1982), pp. 545-555.

Allen, S.G., "Redistribution of Total Stock Over Several User Locations", *Naval Research Logistics Quarterly*, Vol. 5 (1958), pp. 337-345.

———————, "A Redistribution Model with Set-up Charge", *Management Science*, Vol. 8 (1961), pp. 99-108.

Allen, S.G., "Computation for the Redistribution Model with Set-up Charge", *Management Science*, Vol. 8, No. 4 (1962), pp. 482-489.

Badinelli, R.D. and L.B. Schwarz, "Backorders Optimization in a One-Warehouse, N-Identical Retailer Distribution System," Working Paper in Management, No. MGT-2, College of Business and Economics, University of Kentucky, Lexington, Kentucky, April 1984.

Benavie, A., *Mathematical Techniques for Economic Analysis*, Prentice Hall, Englewood Cliffs, N.J. (1972).

Bessler, S.A., and A.F. Veinott, Jr., "Optimal Policy for a Dynamic Multi-Echelon Inventory Model", *Naval Research Logistics Quarterly*, Vol. 13 (1966), pp. 355-389.

Billingsley, P., *Convergence of Probability Measures*, Wiley, New York (1968).

Brown, K.A., "A Two-Period, Two-Echelon, Two-Retailer Inventory Stocking Problem", unpublished M.S. thesis, School of Operations Research and Industrial Engineering, Cornell University, 1984.

Chung, K.L., *A Course in Probability Theory*, 2nd edition, Academic Press, New York (1974), pp. 124-125.

Clark, A.J., "An Informal Survey of Multi-Echelon Inventory Theory", *Naval Research Logistics Quarterly*, Vol. 19 (1972), pp. 621-650.

———————, and H. Scarf, "Optimal Policies for a Multi-Echelon Inventory Problem", *Management Science*, Vol. 6, No. 4 (1960), pp. 475-490.

Connors, M.M., and W.I. Zangwill, "Cost Minimization in Networks with Discrete Stochastic Requirements", *Operations Research*, Vol. 19 (1971), pp. 794-821.

Das, C., "Supply and Redistribution Rules for Two-Location Inventory Systems: One-Period Analysis", *Management Science*, Vol. 21, No. 7 (1975), pp. 765-776.

Eppen, G., "Effects of Centralization on Expected Costs in a Multi-Location Newsboy Problem", *Management Science*, Vol. 25, No. 5 (1979), pp. 498-501.

———, and L. Schrage, "Centralized Ordering Policies in a Multi-Warehouse System with Lead Times and Random Demand", in Schwarz, L.B. (Ed.), *TIMS Studies in the Management Sciences*, Vol. 16, North-Holland Publishing Company, Amsterdam (1981), pp. 51-67.

Erkip, N., "Approximate Policies in Multi-Echelon Inventory Systems", unpublished Ph. D. dissertation, Department of Industrial Engineering and Engineering Management, Stanford University, June 1984.

———, "A Restricted Class of Allocation Policies in a Two-Echelon Inventory System", Technical Report No. 628, School of Operations Research and Industrial Engineering, Cornell University, 1984.

———, W. Hausman, and S. Nahmias, "Optimal and Near Optimal Policies in Multi-Echelon Inventory Systems with Correlated Demands", presented at the TIMS XXVI International Meeting, Copenhagen Denmark, June 17-21, 1984.

Federgruen, A., and P. Zipkin, "Approximations of Dynamic Multilocation Production and Inventory Problems", *Management Science*, Vol. 30, No. 1 (1984), pp. 69-84.

———, and P. Zipkin, "Computational Issues in an Infinite-Horizon, Multi-Echelon Inventory Model", (1982) to appear in *Operations Research*.

Fukuda, Y., "Optimal Disposal Policies", *Naval Research Logistics Quarterly*, Vol. 8 (1961), pp. 221-227.

Gross, D., "Centralized Inventory Control in Multi-Location Supply Systems", Chapter 3 in Scarf, H., D. Gilford and M. Shelly (Eds), *Multistage Inventory Models and Techniques*, Stanford University Press, Stanford, California, 1963.

Hadley, G., and T.M. Whitin, "A Model for Procurement, Allocation, and Redistribution for Low Demand Items", *Naval Research Logistics Quarterly*, Vol. 8 (1961), pp. 395-414.

Hadley, G., and T.M. Whitin, "An Inventory Transportation Model with N Locations", Chapter 5 in Scarf, H., D. Gilford and M. Shelly (Eds), *Multistage Inventory Models and Techniques*, Stanford University Press, Stanford, California, 1963.

Hochstaedter, D., "An Approximation of the Cost Function for a Multi-Echelon Inventory Model", *Management Science*, Vol. 16 (1970), pp. 716-727.

Iglehart, D.L., "Recent Results in Inventory Theory", *Journal of Industrial Engineering*, Vol. 18 (1967), pp. 48-51.

____________, and A.P. Lalchanandi, "An Allocation Model", *SIAM Journal of Applied Mathematics*, Vol. 15, No. 2 (1967), pp. 303-323.

Ignall, E., and A.F. Veinott, Jr., "Optimality of Myopic Inventory Policies for Several Substitute Products", *Management Science*, Vol. 15 (1969), pp. 284-304.

Jackson, P.L., "Stock Allocation in a Two Echelon Distribution System or What to Do Until Your Ship Comes In", to appear in *Management Science*.

____________, and J.A. Muckstadt, "A Two-Period, Two-Echelon Inventory Stocking and Distribution Problem", Technical Report No. 616, School of Operations Research and Industrial Engineering, Cornell University, 1984.

Jonsson, H., and E.A. Silver, "Stock Allocation Among a Central Warehouse and Identical Regional Warehouses in a Particular Push Inventory Control System," Working Paper, Faculty of Management, The University of Calgary, Calgary Alberta, 1985.

Karmarkar, U.S., "Convex/Stochastic Programming and Multi-Location Inventory Problems", *Naval Research Logistics Quarterly*, Vol. 26, No. 1 (1979), pp. 1-19.

____________, "Multi-Period Multi-Location Inventory Problems", *Operations Research*, Vol. 29, No. 2 (1981), pp. 215-228.

____________, and N. Patel, "The One-Period, N-Location Distribution Problem", *Naval Research Logistics Quarterly*, Vol. 24, No. 4 (1977), pp. 559-575.

Krishnan, K.S., and V.R.K. Rao, "Inventory Control in N Warehouses", *Journal of Industrial Engineering*, Vol. 16 (1965), pp. 212-215.

Rockafellar, R.T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey (1970), p. 38.

Maxwell, W.L. and J.A. Muckstadt, "Establishing Consistent and Realistic Reorder Intervals in Production-Distribution Systems," *Operations Research*, Vol. 33, No. 6 (1985), pp. 1316-1341.

Mendelson, H., J.S. Pliskin and U. Yechiali, "A Stochastic Allocation Problem", *Operations Research*, Vol. 28, No. 3 (1980), pp. 687-693.

Miller, B.L., "Dispatching from Depot Repair in a Recoverable Item Inventory System: On the Optimality of a Heuristic Rule", *Management Science*, Vol. 21, No. 3 (1974), pp. 316-325.

Nahmias, S., "Managing Reparable Item Inventory Systems: A Review", in Schwarz, L.B. (Ed.), *TIMS Studies in the Management Sciences*, Vol. 16, North-Holland Publishing Company, Amsterdam (1981), pp. 253-277.

Prastacos, G.P., "Allocation of a Perishable Product Inventory", *Operations Research*, Vol. 29, No. 1 (1981), pp. 95-107.

Roundy, R.O., "98%-Effective Integer-Ratio Lot-Sizing for One-Warehouse Multi-Retailer Systems," *Management Science*, Vol. 31, No. 11 (1985), pp. 1416-1430.

Rosenman, B., and D. Hockstra, "A Management System for High-Value Army Aviation Components", U.S. Army, Advanced Logistics Research Office, Frankfort Arsenal, Report No. TR64-1, Philadelphia, Pennsylvania (1964).

Sherbrooke, C.C., "METRIC: A Multi-Echelon Technique of Recoverable Item Control", *Operations Research*, Vol. 16 (1968), pp. 122-141.

Silver, E.A., "Operations Research in Inventory Management", *Operations Research*, Vol. 29 (1981), pp. 628-645.

Simpson, K.F., Jr., "A Theory of Allocations of Stock to Warehouses", *Operations Research*, Vol. 7 (1959), pp. 797-805.

Spencer, F.W., "An Application of Weak Convergence Theory to the Study of the Stochastic Failures in Parallel Mechanical Systems", unpublished Ph.D. dissertation, School of Operations Research and Industrial Engineering, Cornell University, 1978.

Veinott, A.F., Jr., "Optimal Policy for a Multi-Product, Dynamic, Nonstationary Inventory Problem", *Management Science*, Vol. 12, No. 3 (1965), pp. 206-222.

——————————, "The Status of Mathematical Inventory Theory", *Management Science*, Vol. 12, No. 11 (1966), pp. 745-777.

Zipkin, P., "Simple Ranking Methods for Allocation of One Resource," *Management Science*, Vol. 26 (1980), p. 34.

Zipkin, P., "On the Imbalance of Inventories in Multi-Echelon Systems", *Mathematics of Operations Research*, Vol. 9 (1984), p. 402.

Appendix 1

In this appendix, we prove Theorem 4.2 under the assumptions of Section 4. By assumption, F_1 and F_1^n have their support on a compact interval $[-a,a]$. By a simple change of variable, these functions could be redefined to have their support on $[0,1]$ without materially affecting any of the results of this paper. Without loss in generality, assume that the functions have been so defined.

Let $J_n(x)$ be given by

$$J_n(x) = \int_{1-x}^{1} [1 - F_1^n(y)]\,dy \ . \qquad \text{(A1)}$$

$$= \mathcal{L}_n(1-x)$$

The real-valued function J_n is nondecreasing and continuous on $[0,\infty)$. Let $x_0 = \inf \{x : J_n(x) > 0\}$ and note that J_n is strictly increasing, continuous, and unbounded on (x_0,∞). Hence, J_n has a unique inverse function, denoted by J_n^{-1}, defined on $(0,\infty)$. Furthermore, $J_n^{-1}(y) = 1 - \mathcal{L}_n^{-1}(y)$.

Let m denote the maximum total demand that can occur at any retailer over two periods (assume $m < \infty$). If $q/\sigma_1 > m/\sigma_1$, then $Q > nm$, which is n times the maximum two period demand. Purchasing such a large amount of depot stock is clearly sub-optimal, so we can limit attention to J_n^{-1} defined on $(0,m/\sigma_1]$. An upper bound on $J_n^{-1}(m/\sigma_1)$ is given by $x_1 = m/\sigma_1$, since $J_n(x) \le x$ for all $x \ge 0$.

Consequently, we can limit attention to J_n defined on $[0,m/\sigma_1]$.

Let $D = D([0,1])$ be the space of real-valued functions defined on $[0,1]$ that are right-continuous and have left-hand limits. Then, $F_1(\cdot)$ and each random occurrence of $F_1^n(\cdot)$ are elements of D. We assume D is endowed with the Skorohod topology. The Skorohod topology and the topic of convergence in distribution of random elements of $D([0,1])$ are treated in detail by Billingsley [1968]. It is a simple matter to extend the theory to the space $D([0,b])$ for $0 < b < \infty$, so when quoting his results we will occassionally do so in terms of the latter space, for $b = m/\sigma_1$.

Let X_n be a random element of D given by

$$X_n(y) = \sqrt{n}\left(F_1^n(y) - F_1(y)\right) , \tag{A2}$$

at each $y \in [0,1]$.

Theorem A1.1 (Billingsley, 1968, Theorem 16.4): Under the assumptions of this section,

$$X_n \xrightarrow{\mathcal{D}} X , \tag{A3}$$

where X is a Gaussian process specified by

$$\begin{cases} \mathcal{E}[X(t)] = 0 \\ \mathcal{E}[X(s)X(t)] = F_1(s)\left(1 - F_1(t)\right), \quad 0 \le s \le t \le 1 . \end{cases} \tag{A4}$$

Remark: Each random occurrence of X is an element of D, since the sample paths of a Gaussian process are continuous with probability one.

Restrict J_n to the domain $[0,1]$ and denote the restricted function by $\hat{J}_n$. Let $J(x) = \int_{1-x}^{1}(1-F_1(y))dy$ for $x \in [0,\infty)$, and let $\hat{J}$ denote its restriction to $[0,1]$. Then, $\hat{J}$ and each random occurrence of $\hat{J}_n$ are elements of D. Let $\hat{Y}_n \in D$ be given by

$$\hat{Y}_n(x) = \sqrt{n}\left[\hat{J}_n(x) - \hat{J}(x)\right] , \tag{A5}$$

at each $x \in [0,1]$.

Theorem A1.2:

$$\hat{Y}_n \xrightarrow{\mathcal{D}} \hat{Y} , \tag{A6}$$

where $\hat{Y}$ is a Gaussian process specified by

$$\begin{cases} \mathcal{E}[\hat{Y}(t)] = 0 \\ \mathcal{E}[\hat{Y}(s)\hat{Y}(t)] = \int_{1-s}^{1}\int_{1-t}^{1} F_1(x \wedge y)\bar{F}_1(x \vee y)dydx , \quad s, t \in [0,1], \end{cases} \tag{A7}$$

where $x \wedge y = \min\{x,y\}$ and $x \vee y = \max\{x,y\}$.

Proof: Consider the function $\psi: D([0,1]) \to D([0,1])$ given by

$$\psi G(x) = \int_{1-x}^{1} G(y)dy ,$$

for an arbitrary function $G \in D$ and point $x \in [0,1]$. Then,

$$\sqrt{n}(\hat{J}_n - \hat{J}) = \psi\left[\sqrt{n}\ (F - F_n)\ \right]$$

$$= -\psi X_n \quad .$$

That is, $\hat{Y}_n = -\psi X_n$. By (A3), $X_n \xrightarrow{\mathcal{D}} X$ and so $-X_n \xrightarrow{\mathcal{D}} -X$. Now, $-X$ is a continuous process in D, with probability one, so that if $\{G_n\}$ is any sequence of functions in D converging to $-X$ (in the Skorohod metric) then the convergence is uniform; i.e. $G_n \longrightarrow -X$ in the uniform metric [Billingsley, 1968, p.112]. Consequently, for arbitrary $\varepsilon > 0$,

$$\sup_{x\in[0,1]}\left|\psi G_n(x)+\psi X(x)\right| = \sup_{x\in[0,1]}\left|\int_{1-x}^{1}\left[G_n(y)+X(y)\right]dy\right|$$

$$\leq \left[\sup_{y\in[0,1]}\left|G_n(y)+X(y)\right|\right]$$

$$\leq \varepsilon ,$$

for sufficiently large n. Thus, ψ is continous at $-X$ in the uniform metric whenever $-X$ is continuous on [0,1], which event occurs with probability one. Continuity in the uniform metric implies continuity in the Skorohod metric [Billingsley, 1968, p. 150] so we have shown that ψ is continuous in the Skorohod metric at $-X$, with probability one. By the continuous mapping theorem [Billingsley, 1968, Corollary 1, p. 31] we see that $-X_n \xrightarrow{\mathcal{D}} -X$ implies $-\psi X_n \xrightarrow{\mathcal{D}} -\psi X$; i.e. $\hat{Y}_n \xrightarrow{\mathcal{D}} -\psi X$.

Let $\hat{Y} = -\psi X$. Then, $\hat{Y}(t)$ is the integral of a Gaussian process and, therefore, $\hat{Y}$ is itself a Gaussian process. The mean value function is given by

$$\begin{aligned}
\mathcal{E}[\hat{Y}(t)] &= \mathcal{E}[-\psi X(t)] \\
&= \mathcal{E}\left[\int_{1-t}^{1} X(y)\,dy\right] \\
&= \int_{1-t}^{1} \mathcal{E}[X(y)]\,dy \\
&= 0 \quad ,
\end{aligned}$$

for all $t \in [0,1]$. Similarly, the covariance function is given by

$$\begin{aligned}
\mathrm{Cov}[\hat{Y}(s),\hat{Y}(t)] &= \mathcal{E}[\psi X(s)\psi X(t)] \\
&= \mathcal{E}\left[\int_{1-s}^{1}\int_{1-t}^{1} X(x)X(y)\,dy\,dx\right] \\
&= \int_{1-s}^{1}\int_{1-t}^{1} \mathcal{E}[X(x)X(y)]\,dy\,dx \\
&= \int_{1-s}^{1}\int_{1-t}^{1} F_1(x\wedge y)\overline{F}_1(x\vee y)\,dy\,dx \quad .\blacksquare
\end{aligned}$$

Let Y_n denote the unrestricted version of $\hat{Y}_n$. That is,

$$Y_n(x) = \sqrt{n}\big(J_n(x) - J(x)\big) \quad , \tag{A8}$$

for all $x \in [0,m/\sigma_1]$. Note that Y_n, J_n, and $J \in D([0,m/\sigma_1])$. Let Y be a random element of $D([0,m/\sigma_1])$ given by

$$Y(x) = \hat{Y}(x \wedge 1) \ . \tag{28}$$

Corollary A1.3: $Y_n \xrightarrow{\mathcal{D}} Y$ in $D([0,m/\sigma_1])$. (29)

Proof: Since $1-F_1^n(y) = 1$ for $y \le 0$, $J_n(x) = \hat{J}_n(x \wedge 1) + (x-1) \wedge 0$. Similarly, $J(x) = \hat{J}(x \wedge 1) + (x-1) \wedge 0$. Hence, $Y_n(x) = \hat{Y}_n(x \wedge 1)$. Convergence in distribution of $\hat{Y}_n$ to $\hat{Y}$ then establishes the result. ∎

We are now ready to examine the limiting behavior of J_n^{-1}. For an arbitrary function $\lambda \in D([0,m/\sigma_1])$ define its inverse by

$$\lambda^{-1}(y) = \begin{cases} \inf \{x: \lambda(x) > y\} & \text{if it exists,} \\ \lambda(1) & \text{otherwise.} \end{cases} \tag{A11}$$

Since J_n is strictly increasing on (x_0,∞), this definition will cause no problems for $y \in (0,J_n(m/\sigma_1))$. Uniqueness of $J_n^{-1}(y)$ may fail at $y = 0$. However, we are only interested in $J_n^{-1}(q/\sigma_1)$ for positive q, so this definition is adequate.

The following theorem identifies the limiting behavior of J_n^{-1}. It is based on a theorem by Spencer [1978]. Since that result is unpublished we include it in Appendix 2.

Theorem A1.4: Under the assumptions of Section 4,

$$\sqrt{n}\left(J_n^{-1} - J^{-1}\right) \xrightarrow{\mathcal{D}} \frac{Y \circ J^{-1}}{1-F(a-J^{-1})} \tag{A12}$$

in the space $D((0,J(m/\sigma_1)))$. ($\circ$ denotes composition.)

Proof: Checking the conditions of Spencer's theorem 3.17 we note that by the previous corollary, $\sqrt{n}(J_n - J) \xrightarrow{\mathcal{D}} Y$ in $D([0,m/\sigma_1])$. Furthermore, $J_1, J_2, \ldots$ are non-negative, non-decreasing, random elements of $D([0,m/\sigma_1])$. Y has continuous sample paths with probability one. $\mathcal{E}[Y(0)] = 0$ and $\mathrm{Var}[Y(0)] = 0$, so $\mathcal{P}\{Y(0) = 0\} = 1$. Let $\zeta_n = n^{-1/2}$; then $\zeta_n \rightarrow 0$ as $n \rightarrow \infty$. J is a non-random, continuous, nondecreasing element of $D([0,m/\sigma_1])$ with $J(0) = 0$. Its derivative, $\hat{J}'(x) = 1-F_1(a-x)$, is continuous on $(0,m/\sigma_1)$. However, since $J'(0) = 1-F_1(1) = 0$, J fails the condition that there exist a $\delta > 0$ such that $J'(x) > \delta$ for all $x \in [0,m/\sigma_1]$. Assuming $F_1(1-\varepsilon) < 1$ for any $\varepsilon > 0$, we note that $J'(x)$ is bounded away from zero for all $x \in [\varepsilon,m/\sigma_1]$. Hence, all of the conditions of Spencer's theorem can be seen to hold when the domain is restricted to $[\varepsilon,m/\sigma_1]$. The choice of the interval's left hand endpoint in Spencer's result is arbitrary: it applies equally well to convergence in $D([\varepsilon,J(m/\sigma_1))$ as it does to convergence in $D([0,J(m/\sigma_1)))$. Consequently, (A12) holds in the space $D([\varepsilon,J(m/\sigma_1)))$, for all small positive ε. This is the sense of convergence in the space $D((0,J(m/\sigma_1)))$; it is analogous to defining convergence in $D([0,\infty))$ in terms of convergence in $D([0,N])$ for all large N. ■

Corollary A1.5: For large values of n, and for $q < m$, $J_n^{-1}(q/\sigma_1)$ is approximately normally distributed with mean η, $\eta = J^{-1}(q/\sigma_1)$, and variance ζ^2 given by

$$\zeta^2 = \frac{1}{n\left[\bar{F}_1(1-\eta)\right]^2} \int_{1-\eta\wedge 1}^{1} \int_{1-\eta\wedge 1}^{1} F_1(x\wedge y)\bar{F}_1(x\vee y)\,dy\,dx \quad . \tag{A13}$$

Corollary A1.6: For $q < m$,

$$J_n^{-1}(q/\sigma_1) \longrightarrow J^{-1}(q/\sigma_1) \tag{A14}$$

in probability as $n \longrightarrow \infty$.

Proof: By Corollary A1.5, $J_n^{-1}(q/\sigma_1)$ converges in distribution to a constant as $n \longrightarrow \infty$. Convergence in distribution to a constant implies convergence in probability to that constant. ■

Corollary A1.7 (Theorem 4.2): For $q < m$,

$$\mathcal{I}_n^{-1}(q/\sigma_1) \longrightarrow \mathcal{I}^{-1}(q/\sigma_1) \tag{A15}$$

in probability, as $n \longrightarrow \infty$.

Proof: $\mathcal{I}^{-1}(y) = 1 - J^{-1}(y)$. ■

Appendix 2

Let $D = D[0,1]$, $0 < 1 < \infty$, and D_0 be the subset of D of non-decreasing, real valued functions, and $C = C[0,1]$, the set of all continuous, real valued functions defined on the interval $[0,1]$. C will be assumed to have the uniform metric; D has the Skorohod topology. For $x \in D_0$ define the inverse by

$$x^{-1}(t) = \begin{cases} \inf\{u: x(u) > t\} & \text{if it exists,} \\ x(1) & \text{otherwise.} \end{cases}$$

Theorem A2 (Spencer): Let $X_1, X_2, \ldots$ be non-negative random variables in D_0, Y be a random element in C such that $\mathcal{P}\{Y(0) = 0\} = 1$, $\zeta_1, \zeta_2, \ldots$ be positive random variables such that $\zeta_n \xrightarrow{\mathcal{D}} 0$, and f be a fixed element of $C \cap D_0$, such that $f(0) = 0$, f' exists and is continuous on $(0,1)$ and there exists a $\delta > 0$ such that $f'(s) > \delta$ for $s \in [0,1]$, f' possibly taking on the value of $+\infty$ at 0 or 1. If

$$\frac{X_n - f}{\zeta_n} \xrightarrow{\mathcal{D}} Y \text{ in } D[0,1] \text{ ,}$$

then

$$\frac{X_n^{-1} - f^{-1}}{\zeta_n} \xrightarrow{\mathcal{D}} \frac{Y \circ f^{-1}}{f' \circ f^{-1}}$$

in the space $D[0,f(1))$.