PROBABILISTIC ANALYSES AND PRACTICAL ALGORITHMS FOR INVENTORY-ROUTING MODELS

LAP MUI ANN CHAN

Philips Laboratories, Briarcliff Manor, New York

AWI FEDERGRUEN

Columbia University, New York

DAVID SIMCHI-LEVI

Northwestern University, Evanston, Illinois

We consider a distribution system consisting of a single warehouse and many geographically dispersed retailers. Each retailer faces demands for a single item which arise at a deterministic, retailer specific rate. The retailers' stock is replenished by a fleet of vehicles of limited capacity, departing and returning to the warehouse and combining deliveries into efficient routes. The cost of any given route consists of a fixed component and a component which is proportional with the total distance driven. Inventory costs are proportional with the stock levels. The objective is to identify a combined inventory policy and a routing strategy minimizing system-wide infinite horizon costs. We characterize the asymptotic effectiveness of the class of so-called Fixed Partition policies and those employing Zero Inventory Ordering. We provide worst case as well as probabilistic bounds under a variety of probabilistic assumptions. This insight is used to construct a very effective algorithm resulting in a Fixed Partition policy which is asymptotically optimal within its class. Computational results show that the algorithm is very effective on a set of randomly generated problems.

In many distribution systems, important cost savings can be achieved by integrating inventory control and routing decisions, i.e., by determining simultaneously the timing and sizes of the retailer deliveries as well as efficient vehicle schedules so as to minimize total transportation and inventory carrying costs. In this type of systems the "warehouse" and the "retailers" may represent (part of) consecutive layers in the distribution network of a single company; alternatively, customers may be external, as in the increasingly popular "vendor managed" or "direct replenishment" arrangements in which vendors assume the responsibility of maintaining their customers' inventories instead of responding to customer generated orders.

The impact of integrated inventory and routing strategies was recently emphasized by Stalk et al. (1992) who review the evolution of the discount retailing industry. They observe Wal-Mart developing into the largest and highest profit retailer in the world. This success story was attributed by Stalk et al. to a relentless focus on satisfying retailer needs by *efficient logistical design and planning.* "*The key to achieving these goals was to make the way the company replenished inventory the centerpiece of its competitive strategy.*" Stalk et al. identify a number of major components in this strategic vision, most importantly, a logistics technique referred to as "cross-docking." This refers to a distribution strategy in which the stores are supplied by central warehouses which act as coordinators of the supply process, and as transshipment points for incoming orders from outside vendors, but which do not keep stock themselves.

The distribution planning problem associated with a cross-docking strategy can be modeled as follows: a single warehouse serves many retailers which are geographically dispersed in a given area. Stock for a single item is delivered to the retailers by a fleet of vehicles of limited capacity. Each retailer faces a deterministic, retailer specific, demand rate. Inventory holding costs are accrued at a constant rate, which is assumed to be identical for all retailers. No inventory is kept at the warehouse. Each time a vehicle is sent out to replenish inventory, it incurs a fixed cost (independent of the specific route driven) plus a cost proportional to the total distance traveled by the vehicle. The objective is to determine an inventory policy and a routing strategy such that each retailer can meet its demands and the long-run average transportation and inventory costs are minimized.

In a distribution system of this type, one may have an additional constraint limiting the frequency with which each retailer is visited. Such a constraint may, for example, be due to limited material handling capacity and/or due to the set-up time required for unloading deliveries at the retailers.

It is highly improbable that an optimal strategy will ever be identified for this model; such attempts have long been abandoned even for far simpler models, e.g., the special case where the cost of dispatching a vehicle to a group of

retailers consists only of the fixed component and is independent of the distance traveled. Models, with joint replenishment costs of this type, are often referred to as Joint Replenishment Problems, see Jackson et al. (1985) and Federgruen and Zheng (1992). Most importantly, the structure of a (fully) optimal strategy is so complex that it would fail to be implementable even if it could be determined in a reasonable amount of time. As a consequence, various authors have restricted themselves up front to specific classes of strategies and developed methods to identify optimal or asymptotically optimal rules within the chosen class.

To date, nothing appears to be known on how much is lost by restricting oneself to any of these classes of strategies. It is noteworthy that all of the proposed classes of policies are subsets of the class of **Zero Inventory Ordering** (ZIO) policies, under which a retailer is replenished if and only if its inventory is down to zero. In the absence of constraints on the vehicle capacity or the frequency with which retailers can be served, it is easily verified that a ZIO policy is optimal. However, in the presence of these constraints, ZIO policies may fail to be optimal, as we shall demonstrate shortly.

Even the structure of an optimal ZIO policy is too complex to permit implementation or identification by a reasonable algorithm; this is why all the literature on this model has restricted itself to specific subclasses of the ZIO policies. One attractive such class are the **Fixed Partition Policies** (FP) introduced by Bramel and Simchi-Levi (1992). A FP strategy partitions the set of retailers into a number of regions such that each region is served separately and independently from all other regions. Moreover, whenever a retailer in a set is visited by a vehicle, all other retailers in the set are visited as well. FP policies are easy to implement: they allow for an easy integration of the distribution, marketing, and customer service functions.

The main objective of this paper is to characterize the asymptotic effectiveness of the class of ZIO and the class of FP policies. Interestingly enough, the cost of solutions produced by an optimal ZIO policy is directly related to the optimal solution of an associated bin-packing problem in which the retailers need to be packed into unit size bins and their "sizes" are proportional to their relative demand rates. Indeed, we demonstrate that as the number of retailers grows, the cost of an optimal FP policy as well as that of an optimal ZIO policy exceeds a lower bound for the minimum cost value under *any* strategy by no more than a factor $\sqrt{\alpha}$: here α denotes the so-called *packing constant* in the associated bin packing problem. The packing constant $\alpha \geq 1$ is defined as the asymptotic reciprocal of the average utilization of a bin in an optimal solution. Since for any sequence of the retailers demand rates, $1 \leq \alpha \leq 2$, $0.41 = \sqrt{2} - 1$ represents a *worst case* bound for the (asymptotic) optimality gap for FP (ZIO) policies. In particular, when demand rates are generated independently from a common distribution, we have for many distributions of the demand rates allowing for so-called "perfect"

packing, that $\alpha = 1$, in which case an FP (ZIO) strategy is asymptotically optimal while for other common distributions α is close to one, see Coffman and Lueker (1991) and Rhee (1988) for a more detailed discussion and characterization of the packing constant.

To put these results in perspective we now review the different existing approaches developed for the model considered here. This model was first introduced by Anily and Federgruen (1990). These authors, restricted their analysis to a class of replenishment strategies Ψ with the following properties: a replenishment strategy in Ψ specifies a collection of regions (subset of retailers); if a retailer belongs to several regions a specific fraction of its sales is assigned to each of these. Each time one of the outlets in a given region gets a delivery, this delivery is made by a vehicle which visits all other outlets in the region as well. Observe that a large amount of flexibility is preserved by allowing for overlapping regions of retailers but this may overestimate inventory costs for split retailers. The generated partitions have an asymptotically insignificant number of split retailers; see Hall (1991) and Anily and Federgruen (1991) for details. With n retailers, Anily and Federgruen show that regions can be formed by a simple regional partitioning scheme and a combined inventory and routing strategy can thus be computed in $O(n \log n)$ time, which is asymptotically optimal within the class Ψ.

Subsequent work considers restrictions to other classes of strategies. Gallego and Simchi-Levi (1990) show that Direct Shipping policies, i.e., policies in which each vehicle visits a single retailer, are within 6% of optimality under certain restricted parameter settings. Herer and Roundy (1997) and Viswanathan and Mathur (1997) show good empirical performance for the so-called power-of-two strategies under which each retailer is replenished at constant intervals which are power-of-two multiples of a common base planning period. Power-of-two policies have been shown (See Federgruen et al. 1992) to be within 2% of optimality for general submodular joint replenishment cost structures, but vehicle routing costs may fail to be submodular, as shown in Anily and Federgruen (1990). Finally, as mentioned above, Bramel and Simchi-Levi (1992) analyzed the class of Fixed Partition strategies. They show good empirical performance for medium size problems in the absence of frequency constraints. For a more detailed literature review on inventory-routing problems see Federgruen and Simchi-Levi (1992a and 1992b).

The remainder of this paper is organized as follows. In Section 1 we specify the model assumptions, introduce the notation and provide an example in which ZIO policies may fail to be optimal (even in an asymptotic sense). In Section 2 we develop a lower bound B^* for the cost under any feasible policy. In Section 3 we construct a close-to-optimal Fixed Partition Policy and develop worst-case bounds for the gap between its cost and the lower bound B^* derived in Section 2. Section 4 is devoted to probabilistic analyses of this optimality gap and of the optimal cost

value Z^*. A by-product of our analysis is a practical algorithm for the combined inventory routing problem. For this purpose we describe, in Section 5, an alternative, mathematical programming based heuristic for identifying close-to-optimal FP policies which is of superior performance. A numerical study reported in Section 6, demonstrates that this heuristic comes close to the lower bound B^* even for problems with a moderate number of retailers, n. Finally, in Section 7, we discuss generalizations and variations of our model.

1. NOTATION AND MODEL ASSUMPTIONS

Consider a distribution system with a set $N = \{1, 2, \ldots, n\}$ of geographically dispersed retailers. A central warehouse with an unlimited supply of a given product serves the retailers using vehicles of limited capacity, Q. Retailer i, located at a distance d_i from the warehouse, faces a deterministic demand rate D_i per unit of time and accrues a linear holding cost at a constant rate, h, per unit of product stored there per unit of time. Demand at each retailer must be met over an infinite horizon without shortages or backlogging. The frequency with which a given retailer can be visited is bounded from above by f, i.e., the time that elapses between two consecutive deliveries to a retailer should be at least $1/f$. This upper bound on the delivery frequency to each retailer may be due to the set-up time required for unloading at the retailers or may be due to other material handling constraints.

Following Anily and Federgruen (1990), we assume that the demand rates are rational, i.e., for all $i = 1, \ldots, n$, D_i is an integer multiple of some common quantity D, or $D_i = k_i D$ with k_i a positive integer. We refer to the quantity k_i as the multiplier of retailer i, $i \in N$. Assuming that Q and f are rational as well we choose D sufficiently small that the vehicle capacity Q is an integer multiple of $q = D/f$, the smallest possible delivery quantity for any retailer. As a consequence, $\bar{b} = Qf/D$ is integer.

Each time a vehicle is sent out to replenish inventory to a set of retailers S, it incurs a fixed cost c plus a cost proportional to the total distance traveled by the vehicle, i.e., a cost proportional to $L(S)$, the length of the optimal traveling salesman tour through the warehouse and the retailers in the set S. Without loss of generality, we set the cost per mile equal to one. We seek a combined inventory control and routing strategy that procures retailers in time to meet their demands and minimizes the long-run average total inventory holding and transportation cost per unit of time. As in traditional joint replenishment inventory models, it is not clear that an "optimal" policy always exists. So, let Z^* denote the infimum of the long-run average cost values over all feasible policies. Similarly, let Z_{zi}^* denote the infimum of the long-run average cost over all Zero-Inventory Ordering policies. The following example shows that Z^* may be strictly smaller than Z_{zi}^* even in an asymptotic sense, i.e., in a sequence of problem instances in which $n \to \infty$, we may have $\lim_{n\to\infty} Z^*/n < \lim_{n\to\infty} Z_{zi}^*/n$.

An Example

Consider an inventory routing problem in which there are $3n$ retailers, each one with demand rate 2, located at the same point, at a distance $d = 1$ from the warehouse. Let $f = 1$ and $Q = 3$. The fixed cost of sending out a vehicle, c, equals 1 and similarly the holding cost rate, h, is 1.

Lemma 1. *There exists a feasible policy with long-run average cost $Z = Z_{zi}^* - 0.5n$.*

Proof. Consider policies which satisfy the Zero-Inventory property. Let w be the size of a single delivery to a retailer in a policy of this type. The frequency constraint implies that $w \geq 2/f = 2$, and hence each delivery to a retailer must be made by a separate vehicle. Since $2(2d_i + c)D/h = 12 > 9 = Q^2$, the optimal ZIO policy delivers a full truck load (three units) to each retailer every 1.5 units of time. The long-run average transportation cost of this policy is $(3n)(2d + c)/1.5 = 6n$ while the long-run average holding cost is $3n(1.5) = 4.5n$.

Consider now a different policy which fails to satisfy the Zero-Inventory Ordering property. Under this policy, each retailer receives a delivery every unit of time. The frequency constraint is clearly satisfied. Without loss of generality, assume the system starts with zero-inventory at each retailer. Partition the retailers into groups of three retailers each. For each such group of three retailers, let the delivery sizes be $(2, 2, 3)$ at time 0, $(2, 3, 1)$ at time $2t - 1$ and $(2, 1, 3)$ at time $2t$ for each $t = 1, 2, 3, \ldots$. Hence, for each $t = 1, 2, 3, \ldots$, only two fully loaded vehicles are needed to visit each group of three retailers. It is easy to see that the long run average transportation cost of this policy is $(2d + c)2n = 6n$ while the long run average holding cost is $n[1 + 1.5 + 1.5] = 4n$. $\square$

2. A LOWER BOUND FOR THE COST OF ANY FEASIBLE POLICY

In this section we develop a lower bound for the minimum long run average cost of every policy.

Lemma 2.

$$B^* = \sum_{i=1}^{n} \left[\frac{k_i D(2d_i + c)}{Q} + \frac{hk_i \underline{q}}{2} \right],$$

is a lower bound for the minimum long-run average cost among all feasible policies.

Proof. Let $I_i \geq 0$ be the initial inventory level at retailer i for every i. Consider an arbitrary policy $\mathcal{P}$ over an infinite horizon. Let $\mathcal{C}(\mathcal{P}, t)$ be the average cost per unit of time incurred by this policy over the interval $[0, t)$. It suffices to show that $\mathcal{C}(\mathcal{P}, t) \geq (t/(t + 1/f)) B^* - c'/t$ for some constant c' for all $t > \max_i I_i/k_i D$.

Assume the retailers are ordered such that $d_1 \geq d_2 \geq \cdots \geq d_n$. Let M be the number of vehicles sent out from the warehouse during the interval $[0, t)$; S_j be the set of

retailers visited by vehicle j, $j = 1, 2, \ldots, M$; and w_i^j the number of units of product received by retailer i from vehicle j during $[0, t)$. Let Q_j be the amount of product delivered by the jth vehicle during the interval $[0, t)$, i.e., $Q_j = \sum_{i=1}^n w_i^j$.

We first construct a lower bound for the total transportation cost incurred by policy $\mathcal{P}$. Consider the jth vehicle and a retailer $i \in S_j$. Clearly, $L(S_j) + c \geq 2d_i + c$ and hence

$$Q_j[L(S_j) + c] = \sum_{i \in S_j} w_i^j[L(S_j) + c] \geq \sum_{i \in S_j} w_i^j(2d_i + c).$$

Since $Q_j \leq Q$,

$$L(S_j) + c \geq \sum_{i \in S_j} \frac{w_i^j}{Q}(2d_i + c).$$

Hence the total transportation cost is no smaller than

$$\sum_{j=1}^M [L(S_j) + c] \geq \sum_{j=1}^M \sum_{i \in S_j} \frac{w_i^j}{Q}(2d_i + c)$$

$$= \sum_{i=1}^n \sum_{j: i \in S_j} \frac{w_i^j}{Q}(2d_i + c)$$

$$\geq \sum_{i=1}^n \frac{k_i Dt - I_i}{Q}(2d_i + c).$$

Consider now the holding cost for each retailer i. Let r_i be the number of deliveries received by retailer i over the interval $[0, t)$. Due to the upper bound for the frequency with which each retailer receives deliveries, $r_i \leq (t + 1/f) f$. Hence, the holding cost incurred by retailer i is no smaller than when the total delivery quantity to retailer i in $[0, t)$ is the minimum required, i.e., $k_i Dt - I_i$, and the quantity is delivered at r_i equidistant epochs when inventories are down to zero (see Carr and Howe 1962 for a rigorous proof). In this case the average inventory level equals $(k_i Dt - I_i)/2r_i$. The total holding costs incurred by retailer i in $[0, t)$ are thus bounded from below by

$$th\frac{k_i Dt - I_i}{2r_i} \geq \frac{hk_i t}{2}\frac{tD}{(t + 1/f)f} - \frac{k_i Dt - I_i}{2r_i}$$

$$= \frac{hk_i t^2 q}{2(t + 1/f)} - \frac{k_i Dt - I_i}{2r_i}.$$

Let

$$c' = \sum_{j=1}^M \frac{I_i}{Q}(2d_i + c).$$

Combining the lower bounds on the transportation and the holding costs, we have

$$\mathcal{C}(\mathcal{P}, t) \geq \frac{t}{t + 1/f}\sum_{i=1}^n \left[\frac{k_i D(2d_i + c)}{Q} + \frac{hk_i q}{2}\right]$$

$$- \frac{c'}{t} - \frac{h\sum_i I_i}{2f}\frac{1}{(t + 1/f)}$$

$$= \left(\frac{t}{t + 1/f}\right)B^* - \frac{c'}{t} - \frac{h\sum_i I_i}{2f}\frac{1}{(t + 1/f)}. \qquad \square$$

3. A CLOSE-TO-OPTIMAL FIXED PARTITION POLICY

In this section we construct a FP policy which comes close to being optimal. In particular, we show that the cost of this FP policy asymptotically (as $n \supset \infty$) exceeds the lower bound B^* by no more than a factor $\sqrt{\alpha}$, where $1 \geq \alpha \geq 2$ denotes the so-called **packing constant** associated with packing customers of "size" $\{k_i; i = 1, 2, \ldots, n\}$ into bins of size $\bar{b} = Q/q = Qf/D$ (see below as well as the introduction for precise definitions.)

We construct the FP policy using the following two-step procedure. In the first step, we partition the given area A where the retailers are distributed into subregions. The retailers in each such subregion are then partitioned into sets of retailers by solving the bin-packing problem defined by the multipliers of the retailers and bins of size $\bar{b}$. Each such set is then served in an efficient way.

The Region Partitioning Scheme

Let $G(u)$ be an infinite grid of squares with edges parallel to the coordinate axes and side length $u/\sqrt{2}$. Intersecting each of these squares with A, let $\{A_1, A_2, \ldots, A_m\}$ denote the resulting collection of nonempty intersections. Accordingly, each subregion A_j, $j = 1, 2, \ldots, m$, is either a square of side $u/\sqrt{2}$ or the intersection of such a square with A.

Let N_j be the set of retailers in subregion A_j with $n_j = |N_j|$, $j = 1, 2, \ldots, m$. Given subregion A_j, let $\underline{d}_j$ be the distance from the warehouse to its closest point in A_j, $j = 1, 2, \ldots, m$.

To construct the fixed partition policy, we group all the retailers in subregion A_j, $j = 1, 2 \ldots, m$, into sets by solving the bin-packing problem defined by the multipliers (the numbers k_i) of the retailers in N_j and bins of capacity $\bar{b}$. Each such set S of retailers is served together and is visited using a reorder interval that depends on $k(S) \equiv \sum_{i \in S} k_i$ and the subregion where the retailers in S are located. If S is in the subregion A_j for some $j = 1, 2, \ldots, m$, then the reorder interval is

$$t_S$$

$$= \begin{cases} \dfrac{1}{f}, & \text{if } \sqrt{2k(S)D(2\underline{d}_j + c)/h} \leq k(S)\underline{q} \\[2ex] \sqrt{\dfrac{2(2\underline{d}_j + c)}{k(S)Dh}}, & \text{if } k(S)\underline{q} < \sqrt{2k(S)D(2\underline{d}_j + c/h)} \leq Q, \\[2ex] \dfrac{Q}{k(S)D}, & \text{otherwise.} \end{cases}$$

That is, the reorder interval is chosen so that $q_S = k(S)Dt_S$ is the value of q achieving

$$\min_{k(S)\underline{q} \leq q \leq Q}\left\{\frac{k(S)D(2\underline{d}_j + c)}{q} + \frac{hq}{2}\right\}. \qquad (1)$$

Consequently, these reorder intervals satisfy the capacity as well as the frequency constraints.

For any set of retailers S, $S \subseteq N_j$, we use the following routing strategy. The vehicle travels from the warehouse to its closest point in A_j, visits the retailers in S in any order,

and then returns to the warehouse. It is clear that the total distance traveled is no more than $2\underline{d}_j + (|S| + 1)u$.

Analysis of the Upper Bound

For each subregion A_j, let $b(N_j)$ be the optimal solution to the bin-packing problem defined by the multipliers of the retailers in N_j, $j = 1, 2, \ldots, m$. Let $S_j(l)$, $l = 1, 2, \ldots,$ $b(N_j)$ be the lth set of retailers assigned to one bin in this optimal solution.

We first need the following technical lemma.

Lemma 3. (a) *The function*

$$F(b, d) = \min_{\underline{b}\underline{q} \leqslant q \leqslant Q} \left[\frac{bD(2d + c)}{q} + \frac{hq}{2} \right],$$

is concave in b for all $b \in [1, \bar{b}]$.

(b) $F(b, d) \leqslant F(\bar{b}, d)\sqrt{b/\bar{b}} = \sqrt{b/\bar{b}}[f(2d + c) + hQ/2]$ for all $b \in [1, \bar{b}]$.

Proof. (a) We consider two cases:

Case (i): $2D(2d + c)/h \geqslant Q\underline{q}$. This implies that for any $1 \leqslant b \leqslant \bar{b}$, $2bD(2d + c)/h \geqslant (b\underline{q})^2$ and hence we have

$$F(b, d) = \begin{cases} \dfrac{bD(2d + c)}{Q} + \dfrac{hQ}{2}, & \text{for } \dfrac{2bD(2d + c)}{h} > Q^2, \\ & \text{i.e. } b > b^0 = \dfrac{Q^2 h}{2D(2d + c)}, \\ \sqrt{2bD(2d + c)h}, & \text{otherwise.} \end{cases}$$

Thus, $F(b, d)$ is piecewise concave and since its left derivative in the breakpoint $b = b^0$ equals its right derivative, we conclude that $F(b, d)$ is concave in b.

Case (ii): $2D(2d + c)/h < Q\underline{q}$. Since $2bD(2d + c)/h < Q^2$ for any $1 \leqslant b \leqslant \bar{b}$, we have

$F(b, d)$

$$= \begin{cases} \dfrac{D(2d + c)}{\underline{q}} + \dfrac{hb\underline{q}}{2}, & \text{for } \dfrac{2bD(2d + c)}{h} \\ & < (b\underline{q})^2, \\ & \text{i.e. } b > b^0 = \dfrac{2D(2d + c)}{h\underline{q}^2}, \\ \sqrt{2bD(2d + c)h}, & \text{otherwise.} \end{cases}$$

The concavity proof of $F(b, d)$ is analogous to that in the first case.

(b) We consider three cases depending on the quantity $\sqrt{2bD(2d + c)/h}$, the unconstrained minimizer of the function $bD(2d + c)/q + hq/2$.

Case 1: $b\underline{q} \leqslant \sqrt{2bD(2d + c)/h} \leqslant Q$.

$$F(b, d) = \sqrt{2bD(2d + c)h} = \sqrt{\frac{b}{\bar{b}}}\sqrt{2\bar{b}D(2d + c)h}$$

$$\leqslant \sqrt{\frac{b}{\bar{b}}}F(\bar{b}, d),$$

since $b \leqslant \bar{b}$ and $\sqrt{2\bar{b}D(2d + c)h}$ represents the unconstrained minimum of the function $\bar{b}D(2d + c)/q + hq/2$.

Case 2: $b\underline{q} > \sqrt{2bD(2d + c)/h}$.

$F(b, d) = x + ab$, where $x = D(2d + c)/\underline{q}$ and $a = h\underline{q}/2$. Note that

$$\bar{b}\underline{q} > \frac{\bar{b}}{b}\sqrt{\frac{2bD(2d + c)}{h}} \geqslant \sqrt{\frac{\bar{b}}{b}}\sqrt{\frac{2bD(2d + c)}{h}}$$

$$= \sqrt{\frac{2\bar{b}D(2d + c)}{h}},$$

and $F(\bar{b}, d) = x + a\bar{b}$. Thus

$$\frac{F(b, d)}{F(\bar{b}, d)} = f(x) = \frac{x + ab}{x + a\bar{b}}.$$

Observe that $f(x)$ achieves its maximum subject to the constraint $x \leqslant ab$ (implied by case 2) at $x = ab$. Thus,

$$\frac{F(b, d)}{F(\bar{b}, d)} \leqslant \frac{2ab}{a(b + \bar{b})} = \frac{2}{1 + \bar{b}/b} = \frac{2\sqrt{b/\bar{b}}}{\sqrt{b/\bar{b}} + \sqrt{\bar{b}/b}} \leqslant \sqrt{b/\bar{b}},$$

since $(1 - \sqrt{b/\bar{b}})^2 \geqslant 0$.

Case 3: $\sqrt{2bD(2d + c)/h} > Q$.
$F(b, d) = x + ab$, where $x = hQ/2$ and $a = D(2d + c)/Q$. Note that

$$Q \leqslant \sqrt{\frac{\bar{b}}{b}}\, Q < \sqrt{\frac{2\bar{b}D(2d + c)}{h}},$$

and $F(\bar{b}, d) = x + a\bar{b}$ with $x \leqslant ab$ (implied by case 3). Thus

$$\frac{F(b, d)}{F(\bar{b}, d)} = f(x) = \frac{x + ab}{x + a\bar{b}} \leqslant \sqrt{\frac{b}{\bar{b}}},$$

by the argument used in Case 2. □

We are now able to derive an upper bound for the cost of the above defined FP policy and hence for Z^{FPP}, the *infimum* of the cost values among all FP policies. This bound depends on the number of routes $b(N_j)$ into which each of the customer sets $\{N_j; j = 1, 2, \ldots, m\}$ in the collection of subregions $\{A_j; j = 1, 2, \ldots, m\}$ is partitioned. For each subregion $j = 1, 2, \ldots, m$, we express the number of routes generated in the subregion *relative* to the minimum possible number of routes, i.e., the number of routes required if the demand multipliers $\{k_i: i \in N_j\}$ allow for *perfect* packing; in other words, we express the number of routes employed by the FP policy in terms of

$$\beta_j = \frac{b(N_j)}{\sum_{l=1}^{b(N_j)} k(S_j(l))/\bar{b}} \geqslant 1.$$

Theorem 1.

$$Z^{\text{FPP}} \leqslant \sum_{j=1}^{m} \sqrt{\beta_j} \sum_{i \in N_j} \left[\frac{k_i D(2d_i + c)}{Q} + \frac{hk_i\underline{q}}{2} \right] + 2nuf.$$

Proof. We bound Z^{FPP} by the cost value of the above described FP policy. Under this policy, the reorder interval for every subset of retailers $S_j(l)$, $l = 1, 2, \ldots, b(N_j)$, is $t_{S_j(l)} \geqslant 1/f$. Hence, Z^{FPP} is bounded by

$$Z^{\text{FPP}} \leqslant \sum_{j=1}^{m} \sum_{l=1}^{b(N_j)} \{[2\underline{d}_j + c + u(|S_j(l)| + 1)]/t_{S_j(l)}$$
$$+ hk(S_j(l)) Dt_{S_j(l)}/2\}$$

$$\leqslant \sum_{j=1}^{m} \sum_{l=1}^{b(N_j)} \left\{ \frac{2\underline{d}_j + c}{t_{S_j(l)}} + \frac{hk(S_j(l)) Dt_{S_j(l)}}{2} \right\} + 2nuf$$

$$= \sum_{j=1}^{m} \sum_{l=1}^{b(N_j)} \min_{k(S_j(l))\underline{q} \leqslant q \leqslant Q} \left\{ \frac{k(S_j(l)) D(2\underline{d}_j + c)}{q} + \frac{hq}{2} \right\}$$
$$+ 2nuf$$

$$= \sum_{j=1}^{m} \sum_{l=1}^{b(N_j)} F(k(S_j(l)), \underline{d}_j) + 2nuf.$$

By Lemma 3 (a), $F(b, \underline{d}_j)$ is a concave function of b for every $j = 1, 2, \ldots, m$, and therefore we have for every j,

$$\sum_{l=1}^{b(N_j)} F(k(S_j(l)), \underline{d}_j) \leqslant b(N_j) F\left(\sum_{l=1}^{b(N_j)} \frac{k(S_j(l))}{b(N_j)}, \underline{d}_j \right)$$
$$= b(N_j) F\left(\frac{\bar{b}}{\beta_j}, \underline{d}_j \right).$$

Hence, we have by Lemma 3 (b),

$$Z^{\text{FPP}} \leqslant \sum_{j=1}^{m} b(N_j)(\bar{b}/\beta_j) \frac{F\left(\frac{\bar{b}}{\beta_j}, \underline{d}_j \right)}{\bar{b}/\beta_j} + 2nuf$$

$$= \sum_{j=1}^{m} \sum_{l=1}^{b(N_j)} k(S_j(l)) \frac{F\left(\frac{\bar{b}}{\beta_j}, \underline{d}_j \right)}{\bar{b}/\beta_j} + 2nuf$$

$$\leqslant \sum_{j=1}^{m} \sum_{i \in N_j} k_i \sqrt{\beta_j} \frac{F(\bar{b}, \underline{d}_j)}{\bar{b}} + 2nuf$$

$$\leqslant \sum_{j=1}^{m} \sum_{i \in N_j} k_i \sqrt{\beta_j} \frac{F(\bar{b}, \underline{d}_i)}{\bar{b}} + 2nuf$$

$$= \sum_{j=1}^{m} \sqrt{\beta_j} \sum_{i \in N_j} \left[\frac{k_i D(2\underline{d}_i + c)}{Q} + \frac{hk_i\underline{q}}{2} \right] + 2nuf. \quad \square$$

We now relate the upper bound for Z^{FPP} to the lower bound for Z^* obtained in Section 2. Note that, in case all $\beta_j = 1$, $j = 1, 2, \ldots, m$, the two bounds coincide except for the term $2nuf$ which can be made arbitrarily small by employing a small enough grid size u. The upper bound in Theorem 2 also exhibits the fundamental tradeoff to be considered in implementing the above regional partitioning scheme: the second term in the upper bound ($2nuf$) can be reduced to zero by adopting an increasingly small grid size u in the first step of the procedure and hence by dividing the total retailer population into an increasingly large set of subregions $\{A_j : j = 1, 2, \ldots, m\}$. On the other hand, the larger the subregions $\{A_j: j = 1, 2, \ldots, m\}$ are, the more efficient solutions can be obtained for the bin packing problem to be solved in the *second* step of the procedure, thus reducing the values of $\{\beta_j\}$.

By comparing the upper bound in Theorem 1 and the lower bound in Section 2, we immediately obtain the following *asymptotic worst-case* bound for the optimality gap of Z^{FPP} and hence for Z^*_{zi}.

Theorem 2. *Consider an arbitrary sequence of retailer locations* $\{x_1, x_2, \ldots\}$ *and associated retailer multipliers* $\{k_1, k_2, \ldots\}$. *Let* $Z^*(n)$, $Z^*_{zi}(n)$ *and* $Z^{\text{FPP}}(n)$ *denote the infimum of the costs incurred to serve the first n retailers among all possible strategies, all zero-inventory strategies and all FP policies, respectively. Then*

$$\overline{\lim_{n \to \infty}} \frac{Z^*_{zi}(n)}{Z^*(n)} \leqslant \overline{\lim_{n \to \infty}} \frac{Z^{\text{FPP}}(n)}{Z^*(n)} \leqslant \sqrt{2} = 1.41.$$

Proof. Consider the FP policy obtained by the above two-step procedure with a given grid size u. Let $J^+ = \{j: \lim_{n \to \infty} |N_j| = \infty\}$ and $J^0 = \{1, 2, \ldots, m\}\setminus J^+ = \{j: \lim_{n \to \infty} |N_j| < \infty\}$ so that

$$Z^0(n) \equiv \sum_{j \in J^0} \sqrt{\beta_j} \sum_{i \in N_j} \left[\frac{k_i D(2\underline{d}_i + c)}{Q} + \frac{hk_i\underline{q}}{2} \right],$$

is bounded in n. Observe from Lemma 2 that $Z^* \leqslant n [Dc/Q + h\underline{q}/2]$. Thus,

$$\overline{\lim_{n \to \infty}} \frac{Z^{\text{FPP}}(n)}{Z^*(n)} \leqslant \overline{\lim_{n \to \infty}} \frac{Z^0(n)}{n[Dc/Q + h\underline{q}/2]}$$

$$+ \overline{\lim_{n \to \infty}} \frac{\sum_{j \in J^+} \sqrt{\beta_j} \sum_{i \in N_j} [k_i D(2\underline{d}_i + c)/Q + hk_i\underline{q}/2]}{\sum_{j=1}^{m} \sum_{i \in N_j} [k_i D(2\underline{d}_i + c)/Q + hk_i\underline{q}/2]}$$

$$+ \frac{2uf}{Dc/Q + h\underline{q}/2}$$

$$\leqslant \overline{\lim_{n \to \infty}} \{\max_{j \in J^+} \sqrt{\beta_j}\} + \frac{2uf}{Dc/Q + h\underline{q}/2} \qquad (2)$$

$$\leqslant \sqrt{2} + \frac{2uf}{Dc/Q + h\underline{q}/2},$$

since $\beta_j \leqslant 2$ (the latter follows from the well-known fact that the optimal solution to the bin-packing problem is no more than twice the sum of the fraction of the bin capacity taken by each item). Since Z^{FPP} denotes the infimum over *all* fixed partition policies, the theorem follows by considering a sequence of FP policies generated by the two-step procedure corresponding to a sequence of grid sizes $\{u_l\}$ with $\lim_{l \to \infty} u_l = 0$. $\square$

4. PROBABILISTIC ANALYSIS OF OPTIMALITY GAPS AND OPTIMAL COST VALUES

Significantly sharper bounds for the asymptotic optimality gap may be obtained if the sequence of retailer locations and sizes (multipliers) can be assumed to arise from a specific probabilistic pattern. A basic probabilistic model assumes that the sequences of retailer locations $\{x_1, x_2, \ldots\}$ and retailer multipliers $\{k_1, k_2, \ldots\}$ are both independent and identically distributed, and independent of each other.

In this probabilistic model we have for all subregions j with $\lim_{n\to\infty} |N_j| = \infty$, i.e., for all $j \in J^+$, that the bin-packing problem to be solved in the second of the two-step construction procedure for the FP policy, deals with a sequence of retailer sizes $\{k_i\}$ which is i.i.d. and with the same *common* multiplier distribution in *all* subregions. It is well known from the theory of subadditive processes by Kingman (1976) (see also Rhee and Talagrand (1987)) that for all $j \in J^+$,

$$\lim_{n\to\infty} b(N_j)/|N_j| = \gamma \quad \text{(a.s.)},$$

while by the strong law of large numbers,

$$\lim_{n\to\infty} \frac{1}{|N_j|} \sum_{i \in N_j} \frac{k_i}{\bar{b}} = \frac{E(k)}{\bar{b}} \quad \text{(a.s.)}.$$

We conclude that for all $j = J^+$,

$$\lim_{n\to\infty} \beta_j = \alpha \equiv \gamma\bar{b}/E(k), \tag{3}$$

and obtain the following corollary from Theorem 2.

Corollary 1. *Consider a sequence of retailer locations* $\{x_1, x_2, \ldots\}$ *and retailer multipliers of* $\{k_1, k_2, \ldots\}$ *which are both i.i.d and independent of each other. Let* $Z^*(n)$, $Z^*_{zi}(n)$ *and* $Z^{FPP}(n)$ *be defined as in Theorem 2. Then*

(a) $\quad \overline{\lim_{n\to\infty}} \dfrac{Z^*_{zi}(n)}{Z^*(n)} \leqslant \lim_{n\to\infty} \dfrac{Z^{FPP}(n)}{Z^*(n)} \leqslant \sqrt{\alpha} \quad$ *(a.s.)*.

(b) $\quad \dfrac{DE(k)[2E(d) + c]}{Q} + \dfrac{hE(k)\underline{q}}{2} \leqslant \lim_{n\to\infty} \dfrac{Z^*(n)}{n}$

$\quad \leqslant \sqrt{\alpha}\left[\dfrac{DE(k)[2E(d) + c]}{Q} + \dfrac{hE(k)\underline{q}}{2} \right] \quad$ *(a.s.)*.

(c) *If* $\alpha = 1$, *i.e., if the distribution of retailer multipliers allows for perfect packing:*

$$\lim_{n\to\infty} \frac{Z^*_{zi}(n)}{Z^*(n)} = \lim_{n\to\infty} \frac{Z^{FPP}(n)}{Z^*(n)} = 1 \quad \text{(a.s.)},$$

i.e., Fixed Partition policies are (a.s.) asymptotically optimal, and

$$\lim_{n\to\infty} \frac{Z^*(n)}{n} = \frac{DE(k)[2E(d) + c]}{Q} + \frac{hE(k)\underline{q}}{2}.$$

Proof. Parts (a) and (c) are immediate from Lemma 2, (2), and (3); the existence of $\lim_{n\to\infty} Z^*(n)/n$ follows from $Z^*(n_1 + n_2) \leqslant Z^*(n_1) + Z^*(n_2)$ and the fact that the minimum cost to cover a group of retailers $\{x_{n+1}, \ldots, x_{n+n}\}$ is identically distributed for all $n \geqslant 1$, see Kingman (1973). The upper and lower bound for $\lim_{n\to\infty} Z^*(n)/n$ follow again from Lemma 2, (2), and (3). $\square$

Remarks. Part (c) of the corollary deals with the case where the distribution of retailer multipliers allows for perfect packing, i.e., the wasted space in the bins is asymptotically insignificant. Karmarkar (1982) first proved that any nonincreasing probability density function (with some mild regularity conditions) allows for perfect packing.

Rhee completely characterizes the class of distributions which allow for perfect packing, see also Coffman and Lueker.

More generally, in many distribution problems we encounter significant correlations between the retailer locations and their sales volumes, e.g., retailers in urban areas or in specific states or regions may have larger demands. We may thus wish to generalize the above probabilistic model to one in which the sequence of *pairs* $\{(x_i, k_i)\}$ is i.i.d with a common joint distribution, characterized by the marginal distribution of the locations $\mu(\cdot)$ and the *conditional* multiplier distributions $(k|x = x_0)$. Let $g(x)$ denote the asymptotic average number of bins (routes) required for retailers with multipliers independently distributed as $(k|x)$ and let $\alpha(x) = g(x)\bar{b}/E(k|x)$. Following, once again, the proof of Theorem 2 we obtain:

Corollary 2. *Consider a sequence of pairs of retailer locations and multipliers* $\{(x_1, k_1,); (x_2, k_2,); \ldots\}$ *which are independent and identically distributed with a common joint distribution, characterized by* $\mu(\cdot)$, *the marginal cdf of the retailer locations, and the conditional distributions* $(k|x = x_0)$. *Let* $Z^*(n)$, $Z^*_{zi}(n)$, $Z^{FPP}(n)$ *be defined as in Theorem 2. Then, almost surely*

(a) $\quad \overline{\lim_{n\to\infty}} \dfrac{Z^*_{zi}(n)}{Z^*(n)} \leqslant \lim_{n\to\infty} \dfrac{Z^{FPP}(n)}{Z^*(n)} \leqslant \displaystyle\int \sqrt{\alpha(x)}\, d\mu(x)$.

(b) $\quad \dfrac{D[2E(kd) + cE(k)]}{Q} + \dfrac{hE(k)\underline{q}}{2} \leqslant \lim_{n\to\infty} \dfrac{Z^*(n)}{n}$

$\quad \leqslant \displaystyle\int \sqrt{\alpha(x)}\, d\mu(x) \left[\dfrac{D[2E(kd) + cE(k)]}{Q} + \dfrac{hE(k)\underline{q}}{2} \right]$.

(c) *If all conditional multiplier distributions* $(k|x)$ *allow for perfect packing:*

$$\lim_{n\to\infty} \frac{Z^*_{zi}(n)}{Z^*(n)} = \lim_{n\to\infty} \frac{Z^{FPP}(n)}{Z^*(n)} = 1 \quad \text{(a.s.), and}$$

$$\lim_{n\to\infty} \frac{Z^*(n)}{n} = \frac{D[2E(kd) + cE(k)]}{Q} + \frac{hE(k)\underline{q}}{2} \quad \text{(a.s.)}.$$

We observe that the above probabilistic analysis is based on the sequence of retailer multipliers satisfying two elementary limit results:

(i) in each subregion j, the sequence of multipliers satisfies the strong law of large numbers;

(ii) in each subregion, the bin-packing problem associated with the sequence of multipliers, has an asymptotic almost sure average value, i.e., for all $j \in J^+$, $\lim_{n\to\infty} b(N_j)/|N_j| = \gamma_j$(a.s.) for some γ_j.

Both limit results can be established under conditions far more general than those of Corollaries 1 and 2. Both limit results apply, e.g., when the sequence of retailer attributes $\{x_i, k_i)\}$ is *stationary*, i.e., the *joint* distribution of any m-tuple $\{(x_n, k_n); (x_{n+1}, k_{n+1}), \ldots, (x_{n+m}, k_{n+m})\}$ is *independent* of n; see Kingman (1976).

The above probabilistic models are suitable to characterize the relative performance of FP and ZIO policies for sequences of progressively expanding retailer chains, with each additional retailer acquiring an *incremental* clientele. A different model is needed when a *fixed* customer market is covered by a progressively larger and denser retailer chain, e.g., one with a given customer demand rate density $w(\cdot)$ in the plane. If n retailers with i.i.d. locations offer identical service and merchandise, it is reasonable to assume that any particular retailer attracts the customers in the region of those locations for which this retailer is the closest among all of the n available retailers. Alternatively, in an expanding market with n i.i.d. customer locations $\{y_1, \ldots, y_n\}$ and associated i.i.d. demand rates each retailer i attracts the subset of customers to which it is closest. The problem in analyzing this model is that the retailer size (multiplier) k_i of any *given* retailer i $(1 \leq i \leq n)$ has a distribution which varies as n is increased, i.e., the retailer sizes need to be described by a *tableau* of random variables $\{k_{in}; 1 \leq i \leq n\}$ rather than a single *sequence*. The asymptotic behavior of the packing constants $\{\beta_j\}$ appears to be unknown under this type of probabilistic model.

5. AN EFFICIENT ALGORITHM FOR INVENTORY ROUTING MODELS

The effectiveness of the Fixed Partition Policies suggests a new algorithm for general inventory-routing problems similar to the one developed by Bramel and Simchi-Levi for the Capacitated Vehicle Routing Problem with Unsplit Demands. The algorithm is based on formulating the inventory-routing model as a Capacitated Concentrator Location Problem (CCLP), for the purpose of generating a partition of regions. Each of these regions is assigned a vehicle which visits all retailers in the region at equidistant epoches. The CCLP is subsequently solved, and its solution provides a policy whose cost is, asymptotically, no larger than Z^{FPP}.

5.1. The Capacitated Concentrator Location Problem

The Capacitated Concentrator Location Problem (CCLP) can be described as follows: given m possible sites for concentrators of fixed capacity C, we would like to locate concentrators at a subset of these m sites and connect n terminals, where terminal i uses w_i units of a concentrator's capacity, in such a way that each terminal is connected to *exactly one* concentrator, the concentrator capacity is not exceeded and the total cost is minimized. A site-dependent cost is incurred for locating each concentrator; that is, if a concentrator is located at site j, the *set-up* cost is v_j, for $j = 1, 2, \ldots, m$. The cost of connecting terminal i to concentrator j is c_{ij} (the *connection* cost), for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$. No assumptions need to be made on the costs $\{c_{ij}\}$ and $\{v_j\}$. We assume that there is enough capacity so that a feasible solution exists.

The CCLP can be formulated as the following integer linear program. Let

$$y_j = \begin{cases} 1, & \text{if a concentrator is located at site } j, \\ 0, & \text{otherwise,} \end{cases}$$

and let

$$x_{ij} = \begin{cases} 1, & \text{if terminal } i \text{ is connected to concentrator } j, \\ 0, & \text{otherwise.} \end{cases}$$

Problem P: Min $\sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij}x_{ij} + \sum_{j=1}^{m} v_jy_j$

s.t. $\sum_{j=1}^{m} x_{ij} = 1 \quad \forall\, i,$ (4)

$\sum_{i=1}^{n} w_ix_{ij} \leq C \quad \forall\, j,$ (5)

$x_{ij} \leq y_j \quad \forall\, i, j,$ (6)

$x_{ij} \in \{0, 1\} \quad \forall\, i, j,$ (7)

$y_j \in \{0, 1\} \quad \forall\, j.$ (8)

Constraints (4) ensure that each terminal is connected to exactly one concentrator, and constraints (5) ensure that the concentrator's capacity constraint is not violated. Constraints (6) guarantee that if a terminal is connected to site j, then a concentrator is located at that site. Constraints (7) and (8) ensure the integrality of the variables.

5.2. Formulation

To formulate the region partitioning part of the inventory-routing problem as an instance of the Capacitated Concentrator Location Problem, we refer to each retailer as a terminal whose weight is k_i, i.e., we set $w_i = k_i$. Each retailer is also a possible site for a concentrator with capacity $\bar{b}$, i.e., $C = \bar{b}$. Thus, in our formulation of the inventory-routing problem as a CCLP, $m = n$. The set-up cost for installing a concentrator at site j, (where site j corresponds to retailer j) is $2d_j$. Similar to Bramel and Simchi-Levi, we have used two possible connection costs, c_{ij}:

direct cost: $c_{ij} = 2d_{ij}$,

nearest insertion cost: $c_{ij} = d_i + d_{ij} - d_j$.

The solution to the CCLP provides the grouping of the retailers into subsets. Each such subset is served together.

We have implemented two versions of the algorithm corresponding to this pair of connection cost specifications which we refer respectively as the *Star-Connection Heuristic* (ST) and the *Nearest Insertion Heuristic* (NI).

The next Theorem shows that asymptotically, the cost of the solution produced by the ST heuristic approaches the value of Z^{FPP}.

Theorem 3. *Under the assumptions of Corollary* 1 *and for any distribution* Φ *of the retailer multipliers which allows for perfect packing, we have*

$$\lim_{n \to \infty} Z^{ST}/Z^{FPP} = 1 \text{ (a.s.)}.$$

Table I
List of Parameter Values for Each Instance

Parameter Set	Category	$\underline{q}$	Q	c	h	D
[1]	I	5	74	2	6	10
[2]	I	5	74	10	6	10
[3]	I	5	74	100	6	10
[4]	II	5	74	2	1	10
[5]	II	5	74	2	10	10
[6]	II	5	74	2	100	10
[7]	III	5	74	2	6	5
[8]	III	5	74	2	6	20
[9]	III	5	74	2	6	50
[10]	III	5	74	2	6	500

Proof. We omit the details of the proof since it is similar to the proof of the upper bound from Section 3. See Chan (1995) for details. □

6. A NUMERICAL STUDY

In this section we report our computational experience with the Location Based Heuristic using randomly generated problems. Clearly, computing the optimal cost for even small size problems is intractable. We therefore report the heuristic's performance relative to the two lower bounds developed in the previous sections: B_{zi}^* and B^*.

We have generated 80 problem instances, partitioned into eight different sets. All the instances in a set share the same combination of retailer and warehouse locations and the same retailer multipliers and value of $\bar{b}$. In each set, the retailer and warehouse locations are independently and uniformly located in a square of size [100, 100]. Across each set, the number of retailers varies from 30 to 200. In all cases $\bar{b} = 14$, and the retailer multipliers are uniformly distributed on the integers $\{1, 2, \ldots, 14\}$. For every set, we have generated 10 different problem instances differing in the values of the parameters $\underline{q}$, Q, c, h, and D only. The characteristics of each problem are reported in Table I.

The instances in each set can be subdivided into three categories. In the first category (instances 1–3), we investigate the impact of an increase in c, the fixed set-up cost. In the second category (instances 4–6), we investigate the impact of increasing the holding cost h. In the third category (instances 7–10), we investigate the impact of increasing the demand rate D.

Table II reports the ratios between the heuristic cost and the lower bound on the cost of every policy (i.e., Z^H/B^*).

We observe that the algorithm produces solutions relatively close to the lower bound; the Optimality gap with respect to the lower bound B^* is always less than 16% and in most cases no more than 10%. The results also show that increasing the fixed set-up cost, c, tends to improve the performance of the algorithm; in category I the relative error decreases as c increases. A similar behavior is

Table II
Heuristic Cost Over Lower Bound on the Cost of Every Policy

Parameter Set	30	50	80	100	120	150	180	200
[1]	1.140	1.111	1.110	1.100	1.096	1.094	1.086	1.090
[2]	1.137	1.098	1.109	1.099	1.096	1.094	1.085	1.089
[3]	1.103	1.089	1.086	1.079	1.079	1.076	1.067	1.070
[4]	1.153	1.111	1.115	1.120	1.105	1.100	1.090	1.093
[5]	1.112	1.089	1.088	1.080	1.078	1.075	1.070	1.072
[6]	1.017	1.013	1.013	1.012	1.011	1.011	1.010	1.010
[7]	1.100	1.080	1.078	1.071	1.069	1.067	1.062	1.064
[8]	1.150	1.112	1.118	1.114	1.106	1.103	1.092	1.096
[9]	1.153	1.112	1.116	1.119	1.105	1.100	1.091	1.093
[10]	1.154	1.104	1.111	1.122	1.101	1.096	1.088	1.089

Table III
Heuristic Cost Over Lower Bound on the Cost of Every Zero-Inventory Ordering Poicy

Parameter Set	30	50	80	100	120	150	180	200
[1]	1.113	1.085	1.085	1.075	1.071	1.070	1.061	1.065
[2]	1.110	1.084	1.083	1.074	1.070	1.069	1.060	1.064
[3]	1.081	1.068	1.064	1.057	1.057	1.054	1.045	1.049
[4]	1.143	1.101	1.105	1.110	1.095	1.090	1.081	1.083
[5]	1.090	1.068	1.067	1.059	1.057	1.055	1.049	1.051
[6]	1.014	1.010	1.010	1.009	1.008	1.008	1.007	1.008
[7]	1.080	1.060	1.060	1.052	1.050	1.049	1.043	1.045
[8]	1.129	1.096	1.097	1.093	1.085	1.082	1.072	1.075
[9]	1.140	1.101	1.104	1.108	1.094	1.089	1.080	1.082
[10]	1.153	1.103	1.109	1.120	1.100	1.094	1.086	1.088

Table IV
Computational Result for Nonperfect Packing Instances

	30	50	80	100	120	150	180	200
Z^H/B_{zi}^*	1.151	1.126	1.115	1.107	1.100	1.091	1.094	1.091
Z^H/B^*	1.188	1.161	1.149	1.142	1.134	1.125	1.128	1.125

observed in category II; increasing the holding cost h tends to decrease the relative error.

We have also compared the lower bound B^* which applies to *all* strategies and the bound B_{zi}^* which applies to the ZIO policies only. B_{zi}^* is extremely close to B^*; the ratio between these two values is no more than $Q/\bar{b}q$. In our problem sets B_{zi}^*/B^* is no more than 1.025. Table III reports the ratios $2^H/B_{zi}^*$ for all problem sets.

The next set of experiments is designed to estimate the effect of multipliers distributions that do not allow for perfect packing on the performance of the location based heuristic. For this purpose, Table IV reports the results of the eight instances when $\bar{b} = 10$ and the retailer multipliers are uniformly distributed on the integers $\{3, 4, 5, 6, 7, 8\}$. The values of the parameters are $\underline{q} = 5$, $Q = 54$, $c = 2$, $h = 6$ and $D = 10$.

Finally, we investigate the impact of the frequency constraints on our algorithm. Since the frequency constraint plays a role only through $\underline{q}$, we have changed this parameter, and hence the value of $\bar{b}$ in the last set of parameters. The next two tables reports our computational experience with this set of problems. Observe, that the results are similar to the previous ones; the error decreases as the number of retailer increases and for problems with at least 80 retailers the optimality gap between the solution produced by our algorithm and the lower bound B^* is no more than 19% and the gap with respect to B_{zi}^* is no more than 15%.

7. FURTHER RESULTS AND ALTERNATIVE MODELS

The analysis performed in this paper can be carried over to more general versions of our model. For instance, a somewhat restrictive assumption in the model analyzed in Sections 2–4 is the assumption that the common quantity D used to measure the retailer demand rates $\{d_i = k_i D: i = 1, \ldots, n\}$ is chosen small enough that Q, the vehicle capacity, is an integer multiple of D/f. Such a choice considerably simplifies the analysis. However, tighter bounds and tighter characterizations of optimality gaps may be obtained under a larger (and often more natural) common quantity D; see Chan.

We now extend our model to include *discounted costs*. For this purpose observe that the model introduced in the introduction, considers the *average* cost criterion, capturing the capital costs associated with system-wide inventories as part of the holding costs. An alternative model, perhaps more directly reflecting the company's cash flows, considers the total, continuously *discounted*, value of all out-of-pocket expenses (i.e., the routing costs and inventory carrying charges beyond the cost of capital, if any).

The model with the discounted cost criterion is significantly more complex to analyze. This applies even for the simple EOQ model which corresponds with the special case of a *single* retailer ($n = 1$), and even when the constraints on vehicle capacities and delivery frequencies are ignored. For this basic, discounted EOQ-model, it is possible to show, see, e.g., Jesse et al. (1983), Porteus (1985), or Lee and Nahmias (1993), that analogous to the classical EOQ-model with the average cost criterion, a stationary ZIO policy is optimal with a constant delivery quantity q. The optimal value of q is the (unique) cost of a nonlinear equation, which cannot be obtained in closed form. However, by ignoring third and higher degree terms in the Taylor series expansion of the nonlinear components of the cost expression, we obtain an approximation for the optimal value of q which is *identical* to the well-known (Harris-Wilson) EOQ-formula in the *average* cost case,

Table V
Z^H/B_{zi}^* for Different Values of $\underline{q}$

$\bar{b}$	30	50	80	100	120	150	180	200
8	1.104	1.073	1.078	1.070	1.063	1.062	1.066	1.067
10	1.151	1.126	1.115	1.107	1.100	1.091	1.094	1.091
12	1.140	1.132	1.119	1.116	1.117	1.108	1.099	1.096

Table VI
Z^H/B^* for Different Values of $\underline{q}$

$\bar{b}$	30	50	80	100	120	150	180	200
8	1.142	1.108	1.113	1.105	1.098	1.096	1.101	1.102
10	1.188	1.161	1.149	1.142	1.134	1.125	1.128	1.125
12	1.174	1.165	1.151	1.149	1.150	1.141	1.131	1.129

with the capital cost component in the unit carrying cost rate h, replaced by the product of the (continuous) discount factor and the item's unit dollar value. Moreover, it has been substantiated that this approximation comes very close to being optimal.

To our knowledge, all continuous time models for *multi-item* or *multi-location* systems have confined themselves to the average cost case, and it has not been possible to adapt to the discounted cost model, any of the numerous characterizations of optimality and accuracy gaps for heuristics and bounds, respectively. This applies even for the *special* case of our model in which the cost of a vehicle route does not depend on the number of miles driven, i.e., where all vehicle routes have an *identical* cost value, c. On the other hand, the analysis in this paper suggests a natural strategy for the discounted cost model: compute a (close-to) optimal FPP for the average cost model (e.g., via the methods described in Section 5): then, for each of the routes generated by the FPP, determine the (constant) delivery quantity q as the value achieving the minimum in (1) with the average cost expression replaced by its discounted cost analogue.

ACKNOWLEDGMENT

REFERENCES

ANILY, S., AND A. FEDERGRUEN. 1990. One Warehouse Multiple Retailer Systems with Vehicle Routing Costs. *Mgmt. Sci.* **36**, 92–114.

ANILY, S., AND A. FEDERGRUEN. 1991. Rejoinder to "Comments on One Warehouse Multiple Retailer Systems with Vehicle Routing Costs." *Mgmt. Sci.* **37**, 1497–1499.

BRAMEL, J., AND D. SIMCHI-LEVI. 1995. A Location Based Heuristic for General Routing Problems. *Opns. Res.* **43**, 649–660.

CARR, C. R., AND C. W. HOWE. 1962. Optimal Service Policies and Finite Time Horizons. *Mgmt. Sci.* **9**, 126–140.

CHAN, L. M. A. 1995. Probabilistic Analysis of Inventory-Routing Models. Ph.D. Thesis, Columbia University.

COFFMAN, E. G. JR., AND G. S. LUEKER. 1991. *Probabilistic Analysis of Packing and Partitioning Algorithms.* John Wiley & Sons Ltd., New York.

FEDERGRUEN, A., M. QUEYRANNE, AND Y. S. ZHENG. 1992. Simple Power of Two Policies Are Close to Optimal in a General Class of Production/Distribution Networks with General Setup Costs. *Math. O. R.* **17**, 951–963.

FEDERGRUEN, A., AND Y. S. ZHENG. 1992. The Joint Replenishment Problem with General Joint Cost Structures: General Solution Methods and Performance Bounds. *Opns. Res.* **40**, 384–404.

FEDERGRUEN, A., AND D. SIMCHI-LEVI. 1992a. Routing and Inventory Control Models for General Distribution Networks. Working Paper, Columbia University.

FEDERGRUEN, A., AND D. SIMCHI-LEVI. (1995), Analytical Analysis of Vehicle Routing and Inventory Routing Problems. *Handbooks in Operations Research and Management Science, Network Routing.* M. Ball, T. Magnanti, C. Monma and G. Nemhauser, Eds., North-Holland, Amsterdam, 297–373.

GALLEGO, G., AND D. SIMCHI-LEVI. 1990. On the Effectiveness of Direct Shipping Strategy for the One Warehouse Multi-Retailer R-Systems. *Mgmt. Sci.* **36**, 240–243.

HALL, R. W. 1991. Comments on "One-Warehouse Multiple Retailer Systems with Vehicle Routing Costs." *Mgmt. Sci.* **37**, 1496–1497.

HERER, Y., AND R. ROUNDY. 1997. Heuristics for a One-Warehouse Multiretailer Distribution Problem with Performance Bounds. *Opns. Res.* **45**, 102–115.

JACKSON, P. L., W. L. MAXWELL, AND J. A. MUCKSTADT. 1985. The Joint Replenishment Problem with Powers of Two Restriction. *AIIE Trans.* **17**, 25–32.

JESSE, R. R., A. MITRA, AND J. F. COX. 1983. EOQ formula: Is it Valid Under Inflationary Conditions. *Decision Sci.* **14**, 370–374.

KARMARKAR, N. 1982. Probabilistic Analysis of Some Bin-Packing Algorithms. Proc. 23rd Annual Symposium. *Foundations Comp. Sci.* 107–111.

KINGMAN, J. F. C. 1976. Subadditive Processes. *Lecture Notes in Math. 539*, Springer-Verlag, Berlin, 168–222.

LEE, H. L., AND S. NAHMIAS. 1993. Single Product, Single Location Models. *Handbooks in Operations Research and Management Science*, the volume on Logistics of Production and Inventory, S. C. Graves, A. H. G. Rinnooy Kan and P. H. Zipkin (Eds.), 3–55.

PORTEUS, E. L. 1985. Undiscounted Approximations of Discounted Regenerative Models. *O. R. Lett.* **3**, 293–300.

RHEE, W. T., AND M. TALAGRAND. 1987. Martingale Inequalities and NP-Complete Problems. *Math. O. R.* **12**, 177–181.

RHEE, W. T. 1988. Optimal Bin Packing with Items of Random Sizes. *Math. O. R.* **13**, 140–151.

STALK, G., P. EVANS, AND L. E. SHULMAN. 1992. Competing on Capabilities: The New Rule of Corporate Strategy. *Harvard Business Review*, 57–69.

VISWANATHAN, S., AND K. MATHUR. 1997. Integrating Routing and Inventory Decisions in One Warehouse Multiretailer Multiproduct Distribution Systems. *Mgmt. Sci.* **43**, 294–312.