Heavy Traffic Analysis of the Dynamic Stochastic Inventory-Routing Problem

MARTIN I. REIMAN

Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974

RODRIGO RUBIO

McKinsey and Company, Mexico City, Mexico

LAWRENCE M. WEIN

Sloan School of Management, M.I.T., Cambridge, Massachusetts 02139

We analyze three queueing control problems that model a dynamic stochastic distribution system, where a single capacitated vehicle serves a finite number of retailers in a make-to-stock fashion. The objective in each of these vehicle routing and inventory problems is to minimize the long run average inventory (holding and backordering) and transportation cost. In all three problems, the controller dynamically specifies whether a vehicle at the warehouse should idle or embark with a full load. In the first problem, the vehicle must travel along a prespecified (TSP) tour of all retailers, and the controller dynamically decides how many units to deliver to each retailer. In the second problem, the vehicle delivers an entire load to one retailer (direct shipping) and the controller decides which retailer to visit next. The third problem allows the additional dynamic choice between the TSP and direct shipping options. Motivated by existing heavy traffic limit theorems, we make a time scale decomposition assumption that allows us to approximate these queueing control problems by diffusion control problems, which are explicitly solved in the fixed route problems, and numerically solved in the dynamic routing case. Simulation experiments confirm that the heavy traffic approximations are quite accurate over a broad range of problem parameters. Our results lead to some new observations about the behavior of this complex system.

A prototypical example of the inventory-routing problem (IRP) is the challenge faced by a large oil company as it distributes gasoline to its various gas stations: several warehouses hold inventory of a particular item (gasoline) and serve a set of retailers (stations) in a make-to-stock fashion; arriving customers (automobiles) consume the product at these retail sites, and a fleet of finite capacity vehicles (tanker trucks) is used to transport the product from the warehouse to the various retailers.

The management decisions involved in the design and operation of such a system are many-fold and complex. Traditionally, a hierarchical decomposition of the problem is used to allow for a solvable model at each of the levels (e.g., SIMCHI-LEVI 1992). At the strategic level, the managers of this system must determine the location and number of warehouses and retailers, as well as the assignment of retailers to warehouses. At a tactical level, they must decide on the number of vehicles to operate, and possibly on the assignment of vehicles to service districts. At the operational level, the decisions include: whether to send a particular vehicle out or let it idle, how much of the capacity of the vehicle to use, which of the retailers should each vehicle visit, and how much of its load should a vehicle deliver to each of the retailers on its route.

At the tactical and operational levels, the essence

Transportation Science, © 1999 INFORMS
Vol. 33, No. 4, November 1999

0041-1655/99/3304-0361 $05.00
pp. 361–380, 1526-5447 electronic ISSN

of the IRP is the tradeoff between inventory costs and transportation costs: to reduce inventory levels at the retail sites without affecting the service level, more frequent replenishment deliveries are required, thereby increasing the transportation costs. In many applications, customer demand (and to a lesser extent, vehicle travel times) is subject to considerable stochastic variation. In such cases, a stochastic model is required to accurately capture the inventory costs, and, in this paper, we focus on the operational aspects of the IRP in a dynamic stochastic setting.

The field of operations research is sometimes criticized because real-world applications have lagged behind theoretical progress (e.g., ACKOFF 1987). The IRP is an important counterexample to this perception: although heuristics for this notoriously difficult problem have led to some spectacularly successful industrial applications at both the operational (e.g., the Edelman Prize-winning work of BELL et al. 1983, GOLDEN et al. 1984) and tactical (LARSON 1988) levels, a concomitant mathematical theory for the IRP in a dynamic and stochastic environment has not been forthcoming. FEDERGRUEN and ZIPKIN (1984) analyze a single-period IRP with stochastic demand, and DROR and BALL (1987) develop a heuristic technique to reduce the long-run average problem to a single period problem. Recent studies that consider the operational aspects of the stochastic IRP include TRUDEAU and DROR (1992), who develop heuristics for the case of an external supplier, where retailer inventories are only observable at delivery times; MINKOFF (1993), who constructs a decomposition heuristic for a Markov decision model that dispatches vehicles on a prespecified set of itineraries, where each itinerary is characterized by an inventory allocation to a subset of customers; and KUMAR, SCHWARZ, and WARD (1995), who develop myopic static and dynamic strategies for allocating the contents of a vehicle to the various retailers on a predetermined tour. CHAN, FEDERGRUEN, and SIMCHI-LEVI's (1998) probabilistic analysis of random instances of the deterministic IRP is useful for addressing tactical and strategic issues, but has no bearing on the operational aspects of the IRP with stochastic demand. Readers are also referred to BERTSIMAS and SIMCHI-LEVI (1996) for a recent survey of some related inventory-routing problems and dynamic vehicle routing problems.

Our system model has one warehouse and one capacitated vehicle; hence, we effectively assume that the higher level decisions have been made to assign a single warehouse and a single vehicle to serve all retailers in a particular region. An ample amount of inventory is available at the warehouse, and the cost of holding this inventory is not included in the model. Retailer demand and vehicle travel times are random, unsatisfied demand is backordered, and the objective is to minimize costs due to holding and backordering inventory (cost rates may differ by retailer) and operating the vehicle. One of the crucial decisions in our problem is the *vehicle idling policy*: when the vehicle is at the warehouse, the controller can either send the vehicle out with a full load or let the vehicle sit idle; because we use a long-run average cost criterion, if the vehicle does not idle and the traffic intensity of the system is less than one, then an infinite amount of retailer inventory will build up over the long run.

Two types of IRPs are analyzed: the first assumes fixed routing and the second allows dynamic routing. We consider two variants of the fixed routing IRP: in the IRP with TSP routing, when a vehicle leaves the warehouse, it uses a preoptimized tour of the m retailers, which we refer to as the traveling salesman problem (TSP) tour. In addition to the vehicle idling policy, the controller must decide how many units to deliver to each retailer, and this decision is based on the current inventory levels at all retailers and on the remaining number of units in the vehicle. The second variant is the IRP with direct shipping; in this case, each time the vehicle leaves the warehouse, it delivers all of its contents to a single retailer, and the controller dynamically specifies which retailer to visit next. In the IRP with dynamic routing, the controller decides, based upon the current inventory levels, whether to use a TSP tour or direct shipping.

By only allowing TSP tours or direct shipping, we avoid an assault on the combinatorial aspects of the embedded routing problem, and model these problems as queueing control problems. Because these control problems appear to be analytically intractable, heavy traffic analysis is used to make further progress. Guided by the heavy traffic limit theorems in COFFMAN, PUHALSKII, and REIMAN (1995, 1998), we assume that a time scale decomposition holds in the heavy traffic limit. The resulting diffusion control problems are solved analytically for the fixed route IRPs and numerically for the dynamic IRP. A computational study is also carried out that confirms the accuracy of the heavy traffic analysis and allows us to obtain insights into the relative importance of the various operational decisions (e.g., the vehicle idling policy, static versus dynamic allocation, TSP versus direct shipping, fixed versus dynamic routing).

In Sections 1 and 2, we analyze the IRP with TSP routing and direct shipping, respectively. The performance of these two routing schemes is compared in Section 3 and the IRP with dynamic routing is analyzed in Section 4. The computational study is described in Section 5 and our key findings are summarized in Section 6.

1. THE IRP WITH TSP ROUTING

1.1 Problem Formulation

Consider a system where a single vehicle with capacity V is used to distribute a standard product to m geographically dispersed retailers. An infinite supply of the product is kept at the central warehouse at no cost. Customers are served from the retailer inventories in a make-to-stock fashion, and demand that cannot be served immediately is backordered. When the vehicle is operating, the following policy is used: the vehicle leaves the warehouse (indexed as station 0) with a full load and then visits all the retailers in a predefined sequence before returning empty. Alternatively, the vehicle may idle at the depot. Though the order in which retailers are visited could be arbitrary, we assume that it is the solution to the implied TSP, and refer henceforth to this service scheme as the TSP policy. Without loss of generality, we assume that retailers are indexed from 1 through m according to their position in the TSP tour.

Two sources of variability are considered: customer demand and travel times. For $i = 1, \ldots, m$, customer demand at retailer i occurs according to an independent renewal process $\{D_i(t), t \geqslant 0\}$ with rate λ_i and squared coefficient of variation c_{di}^2 (variance of the interdemand time divided by the square of the mean). The cumulative total demand in $[0, t]$ is denoted by $D(t) = \Sigma_i D_i(t)$, and $\lambda = \Sigma_i \lambda_i$ is the total demand rate. (In all summations of this paper, the index runs over the set of retailers $\{1, 2, \ldots, m\}$, unless explicitly indicated otherwise.) Our results easily generalize to cases with correlated compound renewal processes; see Section 6 of REIMAN (1984) for details. The sequence of travel times between facilities i and j is given by iid samples of the random variable T_{ij}, which has mean θ_{ij} and squared coefficient of variation c_{ij}^2 (i, j run from 0 to m). These travel times are independent of the demand streams and of each other. Keeping with the convention in the literature, we assume that pickup and delivery of units occur instantaneously; in practice, load/unload times tend to be dwarfed by the travel times. (Although nonzero load/unload times can be incorporated in a straightforward manner, the analysis becomes more tedious and its inclusion would cloud the basic issues). Hence, the mean and variance of the total time required to complete the TSP tour are given by $\theta_{\mathrm{T}} = \Sigma_{j=0}^{m-1} \theta_{j,j+1} + \theta_{m0}$ and $s_{\mathrm{T}}^2 = \Sigma_{j=0}^{m-1} \theta_{j,j+1}^2 c_{j,j+1}^2 + \theta_{m0}^2 c_{m0}^2$, respectively, where the subscript T is mnemonic for TSP. For later use, we define the squared coefficient of variation of the tour completion time as $c_{\mathrm{T}}^2 = s_{\mathrm{T}}^2/\theta_{\mathrm{T}}^2$, and let $\{S_{\mathrm{T}}(t), t \geqslant 0\}$ denote the counting process for TSP tour completions up to time t assuming the vehicle is continuously active in $[0, t]$.

Because the route is fixed, only two operating control decisions remain: (i) whether the vehicle should be busy or idle; (ii) while the vehicle is busy, how much of the load to leave at each retailer. The busy/idle control is expressed in terms of the cumulative process $B_{\mathrm{T}}(t)$, which represents the amount of time the vehicle is busy in $[0, t]$. We do not allow tours to be interrupted, and so the sequence τ_k, $k = 1, 2, \ldots$ of tour completion epochs is given by $\tau_k = \inf\{t | S_{\mathrm{T}}(B_{\mathrm{T}}(t)) \geqslant k\}$. The delivery allocations are modeled by the m-dimensional control process $L_i(t)$, which represents the cumulative amount delivered to retailer i up to time t. In anticipation of future developments, let us express this control in terms of a nominal delivery size for retailer i, denoted by V_i, and a dynamic allocation process $\epsilon_i^{\mathrm{T}}(t)$. We let $V_i = \lambda_i V/\lambda$ for all i, so that the nominal delivery size corresponds to allocating the vehicle capacity V among the retailers according to their relative demands. The load allocation process is defined by

$$\epsilon_i^{\mathrm{T}}(t) = L_i(t) - V_i S_{\mathrm{T}}(B_{\mathrm{T}}(t)) \quad \text{for} \quad t \geqslant 0, \qquad (1)$$

which represents the cumulative deviations from the nominal delivery size over past tours, plus the amount delivered during the current cycle for retailer i. Because the tour completion history can be observed, we need only specify the value of $\epsilon_i^{\mathrm{T}}(t)$ to determine the total deliveries to retailer i up to time t. Notice that deviations from the nominal allocation cancel out across the retailers and the process $\epsilon_{\mathrm{T}}(t) = \Sigma_i \epsilon_i^{\mathrm{T}}(t)$ represents the total amount delivered during the current cycle. Because we assume that the vehicle leaves the warehouse with a full load and returns empty, the dynamic load allocation process must satisfy

$$\epsilon_i^{\mathrm{T}}(0) = 0 \quad \text{for all } i, \qquad (2)$$

$$\epsilon_i^{\mathrm{T}}(t^+) > \epsilon_i^{\mathrm{T}}(t^-) \quad \text{only if retailer } i \text{ is visited at time } t, \qquad (3)$$

$$\epsilon_i^{\mathrm{T}}(t) \geqslant \epsilon_i^{\mathrm{T}}(\tau_{k-1}) \quad \text{for } t \in (\tau_{k-1}, \tau_k) \quad \text{and all } i, \qquad (4)$$

$$\epsilon_{\mathrm{T}}(\tau_k^-) = V \qquad (5)$$

and

$$\epsilon_{\mathrm{T}}(\tau_k) = 0, \qquad (6)$$

where the superscripts − and + denote the times just before and after an epoch.

The number of units in inventory (or backordered if this quantity is negative) at retailer i at time t is denoted by $Q_i(t)$, and the total inventory at the retailers is $Q(t) = \Sigma_i Q_i(t)$. If we assume that $Q_i(0) = 0$ (a long-run average cost criterion is being used, and later we restrict our analysis to a class of policies that lead to an ergodic system) then the current inventory $Q_i(t)$ equals the cumulative deliveries minus the cumulative demand, which, by Eq. 1, is given by

$$Q_i(t) = V_i S_{\mathrm{T}}(B_T(t)) - D_i(t) + \epsilon_i^{\mathrm{T}}(t) \quad \text{for} \quad i = 1, \ldots, m, \quad t \geqslant 0. \tag{7}$$

Define the cumulative vehicle idle time process $I(t)$ by

$$I(t) = t - B_{\mathrm{T}}(t) \quad \text{for } t \geqslant 0, \tag{8}$$

so that the control policy $B_{\mathrm{T}}(t)$, $\epsilon_i^{\mathrm{T}}(t)$ must satisfy

$$B_{\mathrm{T}},\ \epsilon_i^{\mathrm{T}} \quad \text{are nonanticipating with respect to } Q, \tag{9}$$

$$B_{\mathrm{T}} \quad \text{is nondecreasing and continuous with} \quad B_{\mathrm{T}}(0) = 0, \tag{10}$$

$$I \quad \text{is nondecreasing with } I(0) = 0. \tag{11}$$

Our objective function includes transportation costs and inventory holding and backordering costs. The travel cost rate per unit of time traveled, which includes vehicle depreciation, fuel, and driver cost, is r. Note that these costs can be combined because we are ignoring the load/unload times (only the driver, but not the vehicle, is busy while loading and unloading). Inventory costs are assumed to be piecewise-linear, with the holding cost rate (per unit in inventory per unit time) at retailer i denoted by h_i and the backorder cost rate by b_i. Because travel costs are incurred whenever the vehicle is busy, the travel cost rate r can be equivalently treated as a reward for exerting idleness. Hence the problem reduces to finding a control policy $(B_{\mathrm{T}}(t),\ \epsilon_i^{\mathrm{T}}(t))$ to minimize

$$\limsup_{T\to\infty} \frac{1}{T} \mathrm{E}\left[\int_0^T \sum_i \left(h_i\{Q_i(t)\}^+ + b_i\{Q_i(t)\}^-\right) \mathrm{d}t - rI(T) \right] \tag{12}$$

subject to Eqs. 2–11, where the + and − denote the positive and negative parts.

The dynamic stochastic IRP, as formulated in Eqs. 2–12, does not seem to be tractable. Even under Markovian assumptions for the underlying random processes, the action space is enormous and the state space has $m + 2$ dimensions: the inventory/backorder level at each retailer and the location and total contents of the vehicle. To gain further understanding of the problem, we analyze it when the system operates in the heavy traffic regime.

1.2 Heavy Traffic Normalizations

We begin our heavy traffic development by centering the service completion and demand processes; define the centered processes $\mathcal{S}_{\mathrm{T}}(t) = S_{\mathrm{T}}(t) - \theta_T^{-1} t$ and $\mathcal{D}(t) = D(t) - \lambda t$. It is convenient to define the process

$$\chi(t) = \left(\frac{V}{\theta_{\mathrm{T}}} - \lambda\right)t + V\mathcal{S}_{\mathrm{T}}(B_{\mathrm{T}}(t)) - \mathcal{D}(t); \tag{13}$$

we refer to this quantity as the netput process, although it does not correspond precisely to the netput processes constructed in the heavy traffic analysis of conventional queueing networks (e.g., PETERSON 1991). Summing the inventory evolution Eqs. 7 over all retailers and substituting the relevant definitions yields

$$Q(t) = \chi(t) - \frac{V}{\theta_T} I(t) + \epsilon_{\mathrm{T}}(t). \tag{14}$$

In heavy traffic analysis, one typically constructs a sequence of systems indexed by the heavy traffic parameter n. Even though no weak convergence proofs will be undertaken here, because some of the scalings that we introduce are nontraditional, we index quantities with n (in an appropriate place) to make the scalings clear. This indexing will be confined to this subsection; for the rest of the paper, we leave off the index, with the understanding that we are considering a single system that has an associated value of n. The parameter n can be thought of as a large integer (e.g., 100) but (as is typically the case) the policy recommendations that emerge from our heavy traffic analysis are independent of n. The parameter n is used to normalize the various processes according to standard heavy traffic conventions (notice that only the process B_{T} undergoes a fluid scaling):

$$W_i^{(n)}(t) = \frac{Q_i^{(n)}(nt)}{\sqrt{n}} \quad \text{for all } i,$$
$$W^{(n)}(t) = \sum_i W_i^{(n)}(t) = \frac{Q^{(n)}(nt)}{\sqrt{n}}, \tag{15}$$

$$Y^{(n)}(t) = \frac{I^{(n)}(nt)}{\sqrt{n}}, \qquad \hat{\chi}^{(n)}(t) = \frac{\chi^{(n)}(nt)}{\sqrt{n}},$$

$$\hat{\epsilon}_{\mathrm{T}}^{(n)}(t) = \frac{\epsilon_{\mathrm{T}}^{(n)}(nt)}{\sqrt{n}}, \tag{16}$$

$$\hat{\mathcal{D}}^{(n)}(t) = \frac{\mathcal{D}^{(n)}(nt)}{\sqrt{n}}, \qquad \hat{\mathcal{S}}_{\mathrm{T}}^{(n)}(t) = \frac{\mathcal{S}_{\mathrm{T}}^{(n)}(nt)}{\sqrt{n}}$$

and (17)

$$\hat{B}_{\mathrm{T}}^{(n)}(t) = \frac{B_{\mathrm{T}}^{(n)}(nt)}{n}.$$

The processes $W^{(n)}$ and $Y^{(n)}$ represent the normalized inventory and idleness, respectively. To reduce the amount of notation, the normalized versions of the remaining processes contain a hat (ˆ). Using these scalings in Eqs. 13 and 14 yields expressions for the normalized netput process

$$\hat{\chi}^{(n)}(t) = \sqrt{n}\left(\frac{V^{(n)}}{\theta_{\mathrm{T}}^{(n)}} - \lambda^{(n)}\right)t + V^{(n)}\hat{\mathcal{S}}_{\mathrm{T}}^{(n)}(\hat{B}_{\mathrm{T}}^{(n)}(t)) - \hat{\mathcal{D}}^{(n)}(t), \tag{18}$$

and the normalized inventory process

$$W^{(n)}(t) = \hat{\chi}^{(n)}(t) - \frac{V^{(n)}}{\theta_{\mathrm{T}}^{(n)}} Y^{(n)}(t) + \hat{\epsilon}_{\mathrm{T}}^{(n)}(t). \tag{19}$$

To obtain a nontrivial control problem in heavy traffic, we normalize the system parameters in a particular fashion. The demand and inventory cost parameters are not scaled, and the other parameters are normalized as follows:

$$\hat{V}^{(n)} = \frac{V^{(n)}}{\sqrt{n}}, \tag{20}$$

$$\hat{\theta}_{\mathrm{T}}^{(n)} = \frac{\theta_{\mathrm{T}}^{(n)}}{\sqrt{n}}, \tag{21}$$

$$\mu_{\mathrm{T}}^{(n)} = \sqrt{n}\left(\frac{V^{(n)}}{\theta_{\mathrm{T}}^{(n)}} - \lambda^{(n)}\right) > 0, \tag{22}$$

$$\hat{c}_{\mathrm{T}}^{2}(n) = \sqrt{n}\; c_{\mathrm{T}}^{2}(n), \tag{23}$$

$$\hat{r}^{(n)} = \frac{r^{(n)}}{n}. \tag{24}$$

We assume that all the quantities on the left side of definitions 20–24 converge to finite and positive limits as $n \to \infty$. Eqs. 20–24 are the heavy traffic conditions and they specify, in a unified manner via the heavy traffic parameter n, the relative magnitudes of the various system parameters. These conditions are more extensive than those enforced in traditional queueing systems, and therefore warrant some discussion. Because the natural definition of the traffic intensity is $\rho_{\mathrm{T}} = \lambda\theta_{\mathrm{T}}/V$, condition 22 is the traditional heavy traffic condition, which requires that ρ_{T} be close to, but less than, unity.

Now we turn to conditions 20–21. Because the state space is compressed by a factor of $\sqrt{n}$ in the heavy traffic normalization, the vehicle capacity, in terms of scaled inventory units, is $V^{(n)}/\sqrt{n}$. Hence, if $V^{(n)}$ were $O(1)$, it would vanish in the limit, and our system would reduce to a variant of the multiclass make-to-stock queue analyzed in WEIN (1992). Although such a model would be tractable, a limit that uses infinitesimal vehicle sizes fails to capture the essence of the behavior of the original system. Therefore, we enforce condition 20, so that $V^{(n)}$ is $O(\sqrt{n})$, and the bulkiness of the retailer deliveries is retained in the limit. However, because the demand rate $\lambda^{(n)}$ is unscaled, we need to also scale the tour lengths according to Eq. 21 to ensure that the ratio $V^{(n)}/\theta_{\mathrm{T}}^{(n)}$ converges to a finite and positive limit.

Turning to Eq. 23, note that, because the vehicle capacity is $O(\sqrt{n})$, if $c_{\mathrm{T}}^2(n)$ is not scaled, then a standard calculation shows that the variance term for the normalized netput process $\hat{\chi}^{(n)}$ is $O(n)$ and, hence, approaches infinity in the heavy traffic limit. Because $s_{\mathrm{T}}^2(n) = c_{\mathrm{T}}^2(n)\theta_{\mathrm{T}}^2(n)$, by Eqs. 21 and 23, we obtain $s_{\mathrm{T}}^2(n) = \sqrt{n}\hat{s}_{\mathrm{T}}^2(n)$, where $\hat{s}_{\mathrm{T}}^2(n) = \hat{\theta}_{\mathrm{T}}^2\hat{c}_{\mathrm{T}}^2(n)$. Thus, by enforcing condition 23, we assume that the variance of the tour completion time is $O(\sqrt{n})$; in contrast, this quantity would be $O(n)$ if travel times were simply multiplied by $\sqrt{n}$. One way to achieve Eq. 23 is to assume that the travel time of the tour is the sum of $\sqrt{n}$ iid finite variance travel times. This construction could arise by superimposing the warehouse and retailer locations on a two-dimensional map with a fine grid, in such a way that the tour passes through approximately $\sqrt{n}$ grid points. However, this modeling artifice is problematic (because adjacent travel times would not likely be independent and the necessary data would be tedious to collect) and is not pursued here; see RUBIO (1995) for further details.

Finally, as is standard for heavy traffic optimization problems, we need to normalize the cost parameters to account for distortions in the relative magnitudes of the transportation and inventory costs that result from the heavy traffic scaling. The appropriate scaling is to allow the travel cost rate $r^{(n)}$ to be approximately n times larger than the inventory cost rates, as in condition 24; see Rubio for a detailed explanation.

In summary, although there may be alternative heavy traffic conditions that can be constructed, we use the set of conditions that naturally arise as a consequence of retaining the bulkiness of retailer deliveries in the heavy traffic limit. The heavy traffic conditions assume that the vehicle must be busy the great majority of the time to meet average demand, the vehicle capacity must be large, the tour completion time must be large and nearly deterministic, and the travel cost rate must be very large relative to the inventory cost rates. The computational study in Section 5 reveals that our results are rather insensitive to these conditions. Finally, although we assume that the system is heavily loaded and the vehicle capacity is large, our analysis explicitly accounts for the capacity restriction and the fact that the vehicle occasionally idles; in fact, our proposed policy is expressed in terms of the traffic intensity ρ_T and the vehicle capacity V.

1.3 System Behavior in Heavy Traffic

This subsection considers the limiting behavior of Eqs. 18–19. Following HARRISON (1988), we replace $\hat{B}_T(t)$, which is the fluid-scaled busy time process, by $\rho_T t$; the justification for this substitution is that any policy that does not utilize the vehicle for a fraction ρ_T of the time over a sufficiently long time interval will generate extremely large inventory costs. In addition, we consider the normalized netput process embedded at tour completion epochs. Without some embedding or averaging, the limit of the normalized netput process would not exist because it varies (after normalization) by $O(1)$ on a time of length $O(1/\sqrt{n})$. The process we consider is thus defined as

$$\tilde{\chi}(t) = \sqrt{n}\left(\frac{V}{\theta_T} - \lambda\right)\tau_{k-1} + V\hat{\mathcal{S}}_T(\rho_T \tau_{k-1}) - \hat{\mathcal{D}}(\tau_{k-1}) \quad \text{for} \quad t \in [\tau_{k-1}, \tau_k).$$

With this definition, the standard tools of weak convergence (the functional central limit theorem for renewal processes, the random time change theorem and the continuous mapping theorem; see BILLINGSLEY 1968) can be used to show that the normalized netput process embedded at tour completion epochs $\tilde{\chi}$ is well approximated by a Brownian motion X with drift μ_T and variance $\sigma_T^2 = \lambda(c_d^2 + Vc_T^2)$.

Now we turn our attention to the process $\hat{\epsilon}_T$. This process equals zero at tour completion epochs and has jumps of size $O(1)$ whenever a delivery occurs and at the end of the cycle. In addition, because the tour length is $O(\sqrt{n})$ by Eq. 21, a tour takes only $O(1/\sqrt{n})$ time units under the heavy traffic normalization (where time is compressed by the factor n); hence, tours occur instantaneously in the heavy traffic limit. Consequently, neither $\hat{\epsilon}_T(t)$ nor the m-dimensional normalized inventory process converge to a limit in the usual sense. However, if we start with the heavy traffic normalization and expand time by a factor of $\sqrt{n}$, then a fluid scaling is obtained, where both time and space are compressed by the factor $\sqrt{n}$. At this faster time scale, the Brownian motion X remains constant, and the individual inventory levels move in a deterministic fashion, decreasing at a finite rate between the jumps at delivery epochs. The process $\hat{\epsilon}_T$ traverses through many tours before X changes value, and equals zero at each tour completion epoch.

This is similar to the state of affairs in the heavy traffic results of Coffman, Puhalskii and Reiman (1995). In their exhaustive polling system, the total queue length process behaves as a one-dimensional diffusion under the slow time scale associated with the heavy traffic scaling, and the individual queues move as a fluid under the faster time scale associated with the fluid limit. This time scale decomposition gives rise to a heavy traffic averaging principle (HTAP) that implies the following: for purposes of calculating performance measures for the individual queues, one can analyze the deterministic fluid cycle for each fixed value of the diffusion process.

There are three key differences between the HTAP in the polling system and in the IRP. First, the time scale decomposition in the polling system emerged as a consequence of the standard heavy traffic normalization, whereas, in the IRP, it follows from the scaling assumptions 20–23. Second, the fluid trajectories are different. In the polling problem, the fluid paths look like those for the economic production quantity model: they go up and down at a finite rate. In the IRP, the paths look like those from the economic order quantity model: they go down at finite rate but jump up at delivery epochs. The third key difference relates to the issue of control. In the polling system, the exhaustive discipline guarantees that, whenever the server switches from a queue, that queue is empty. This exerts a type of control that keeps the multidimensional process well behaved. There is no such natural mechanism in the IRP, and we must introduce a dynamic allocation scheme to keep the multidimensional process well behaved.

The proof of the HTAP in the polling context is difficult, involving a threshold queue. A proof for the IRP would not need the threshold queue, but would have to deal with the dynamic allocation scheme. Although we conjecture that the HTAP holds for the IRP, we do not attempt a proof. Rather, we construct, in Section 1.5, a dynamic class of allocation policies and assume that the HTAP holds for this class of policies.

1.4 The Limiting Control Problem

Under this HTAP assumption, the analysis of our limiting control problem decomposes onto two time scales as described above. On the slow time scale associated with the diffusion limit, we can average out the effects of the controlled allocation process $\hat{\epsilon}_T$ and choose the vehicle idling policy. This policy is generated by the normalized cumulative idleness Y, which we assume is nondecreasing and right continuous. Let $Z(t) = X(t) - (V/\theta_T)Y(t)$; this is the process that would be obtained if one were to observe the total inventory only at tour completion epochs. We refer to this process as the total embedded inventory process, to differentiate it from W.

At the faster time scale, where the total embedded inventory is fixed at $Z(t) = x$ and the individual inventories behave as a fluid, we must find the optimal allocation policy that minimizes inventory costs per unit time. The limit cycle associated with an allocation policy can be viewed as a closed m-dimensional path (see Section 1.5), and the optimal allocation policy reduces to the problem of optimally placing a deterministic cycle in $\mathbf{R}^m$. Let $g(x)$ represent the inventory cost per unit time that is achieved by optimally locating a cycle when $Z(t) = x$.

We can now state the limiting stochastic control problem for the IRP with TSP routing: (i) find the optimal cycle placement for a given total embedded inventory level $Z(t) = x$, and its corresponding inventory cost rate $g(x)$; and (ii) choose the nondecreasing right continuous process Y to minimize

$$\limsup_{T\to\infty} \frac{1}{T} \mathrm{E}\left[\int_0^T g(Z(t))\, \mathrm{d}t - \hat{r}Y(T)\right] \qquad (25)$$

$$\text{subject to} \quad Z(t) = X(t) - \frac{V}{\theta_T} Y(t). \qquad (26)$$

The cycle placement problem is a nonlinear program and problem 25–26 is a singular control problem for Brownian motion; these two problems are solved in the next two subsections.

1.5 Optimal Cycle Placement and Dynamic Allocation

To optimally place the limit cycle, we follow the approach used in MARKOWITZ, REIMAN, and WEIN (1999) for the stochastic economic lot scheduling problem (ELSP). Let us fix $Z(t) = x$, and denote the individual fluid inventory levels by $\bar{W}_i(t) = Q_i(\sqrt{n}t)/\sqrt{n}$ (an overbar will be used to denote quantities introduced for the fluid limit). The cycle placement can be defined in many ways, and we choose to specify it by

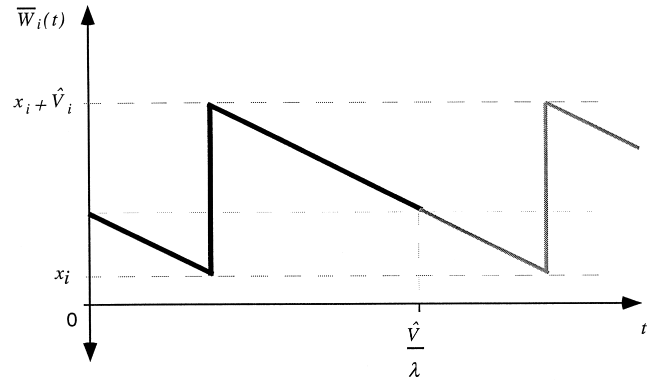

Fig. 1. The fluid inventory evolution at retailer i during a nominal allocation cycle.

the vector $(x_1, x_2, \ldots, x_m)$, where x_i represents the lowest point during the cycle of $\bar{W}_i(t)$.

The choice of optimal $(x_1, \ldots, x_m)$ is a constrained optimization problem: we want to choose $(x_1, \ldots, x_m)$ to minimize the inventory cost rate subject to consistency with the total embedded inventory level. The inventory cost rate will come from the averaging principle to be described below. We first deal with the consistency issue.

To establish the relationship between the cycle placement variables x_i and the total embedded inventory level x, we need to introduce some new notation. Denote the mean travel time along the TSP path between any two sites $i, j \in 0, 1, \ldots, m$ by $\theta_{ij}^{\mathrm{TSP}}$; in terms of the intersite mean travel times θ_{ij}, these quantities are defined by $\theta_{ij}^{\mathrm{TSP}} = \Sigma_{k=\mathrm{i}}^{j-1} \theta_{k,k+}$ for $j > i$ and $\theta_{ii}^{\mathrm{TSP}} = 0$. Because time is compressed by a factor of $\sqrt{\mathrm{n}}$ in the fluid limit, define the corresponding travel times for the fluid model by $\bar{\theta}_{ij}^{\mathrm{TSP}} = \theta_{\mathrm{TSP}}^{\mathrm{ij}}/\sqrt{n}$. If we measure time over a cycle so that the vehicle leaves the warehouse at $t = 0$, then $\bar{W}_i(0)$ is related to its corresponding cycle placement value x_i by (see Fig. 1) $\bar{W}_i(0) = x_i + \lambda_i \bar{\theta}_{0i}^{\mathrm{TSP}}$ for $i = 1, \ldots, m$. Summing these inventory levels over all retailers, we obtain

$$\sum_i x_i = x - \sum_i \lambda_i \bar{\theta}_{0i}^{\mathrm{TSP}}. \qquad (27)$$

Given a vector $(x_1, \ldots, x_m)$ satisfying Eq. 27, we use the time scale decomposition assumption to determine the associated inventory cost rate. Long term stability requires that, in the long run, the average amount delivered to retailer i per cycle be $V_i = \lambda_i V/\lambda$; under the diffusion and fluid scalings, this delivery size is given by $\hat{V}_i = V_i/\sqrt{n}$. Viewing the fluid inventory of retailer i in isolation with a delivery of $\hat{V}_i$ on each tour cycle, we see a fluid starting out at $x_i + \hat{V}_i$ immediately after delivery, decreasing at a constant rate until x_i is reached just prior to the next delivery (see Figure 1). The inventory cost component associated with retailer i can then be calculated by considering the normalized inventory process W_i to be uniformly distributed on the interval $[x_i, x_i + \hat{V}_i]$. It is important to note that a dynamic (state-dependent) delivery size is needed to keep the long-run average cost finite. Simply delivering $\hat{V}_i$ to retailer i on every visit will result in an infinite long-run average cost because this allocation leads to a null recurrent process. To see this, note that, under this simple allocation scheme, the drift of $W_i(t)$ does not depend on $W_i(t)$ and equals zero, because the inventory arrival rate $\hat{V}_i/\hat{\theta}_T$ equals the demand rate λ_i when $\rho_T = 1$; with $\rho_T < 1$, a similar result is generated with the effective arrival rate being $\rho_T\hat{V}_i/\hat{\theta}_T$. With a zero drift, central limit theorem arguments indicate that the inventory or backlog will grow as $\sqrt{t}$. In fact, similar arguments can be used to explain some numerical results in FEDERGRUEN and KATALAN (1996) and Wein (1992), where a state-independent policy performs poorly in a stochastic setting.

A simple dynamic allocation policy avoids this difficulty. We determine delivery sizes at the warehouse as follows. Given a fluid inventory level $(w_1, \ldots, w_m)$ when the vehicle is at the warehouse, the fluid limit of the inventory immediately before delivery is $w_i - \lambda_i\bar{\theta}_{0i}^{\text{TSP}}$. If possible, we would like to deliver $d_i = x_i + \hat{V}_i - w_i + \lambda_i\bar{\theta}_{0i}^{\text{TSP}}$ to retailer i to bring the fluid inventory level immediately after delivery to $x_i + \hat{V}_i$. If $d_i \geq 0$ for $i = 1, \ldots, m$, then this delivery allocation is feasible. If $d_i < 0$ for some i, then, because $\Sigma_i\, d_i = \hat{V}$, we must have $\Sigma_{\{i:d_i>0\}}\, d_i > \hat{V}$. This is a transient state for the fluid limit; within a finite number of cycles we will have $d_i \geq 0$ for all i. This transient interval has no effect on the long-run average inventory cost, which is our heuristic justification for assuming that an averaging principle holds under this dynamic allocation scheme. In summary, the essence of the averaging principle here is that, under this dynamic allocation scheme, when $Z(t) = x$, $W_i(t)$ can be treated as if it is uniformly distributed between x_i and $x_i + \hat{V}_i$.

The average inventory cost per unit time is equal to the cost incurred over a cycle divided by the corresponding cycle length. The cost at retailer i may be obtained by simple geometric arguments for any cycle placement x_i. When the cycle placement is sufficiently high (low) so that the inventory remains positive (negative) for the duration of the cycle, the cost is simply the holding (backordering) rate multiplied by the absolute value of $x_i + \hat{V}_i/2$, which is the average inventory level over a cycle. When the inventory changes sign during the cycle, the total holding (backordering) cost over a cycle equals the area of one of the triangles above (below) the time axis multiplied by h_i (b_i). To obtain the time average inventory cost when there is a sign change, we sum the areas of these two triangles and divide by the cycle length. In the heavy traffic limit, the amount of fluid delivered per cycle, $\hat{V}$, equals the amount demanded per cycle, which is $\lambda\hat{\theta}_T$; hence, we set the cycle length in the fluid model equal to $\hat{V}/\lambda$, rather than $\hat{\theta}_T$. In summary, we have the following expression for retailer i:

$$g_i(x_i) = \begin{cases} h_i\left(x_i + \dfrac{\hat{V}_i}{2}\right) & \text{if } x_i \geq 0 \\[2mm] \dfrac{h_i + b_i}{2\hat{V}_i}\, x_i^2 + h_i x_i + \dfrac{h_i\hat{V}_i}{2} & \text{if } -\hat{V}_i < x_i < 0 \\[2mm] -b_i\left(x_i + \dfrac{\hat{V}_i}{2}\right) & \text{if } x_i \leq -\hat{V}_i. \end{cases} \tag{28}$$

Notice that $g_i(x_i)$ is a convex function of x_i. With Eq. 28 in hand, the cycle placement problem is to minimize $\Sigma_i\, g_i(x_i)$ subject to Eq. 27.

Let us make the innocuous assumption that $b_i \geq h_i$ for all i, and define the labeling conventions $h_\ell = h = \min_i h_i$ and $b_p = b = \min_i b_i$, where $\ell = p$ is allowed. A closed-form solution to the cycle placement problem is found by using constraint 27 to turn the problem into one of unconstrained optimization over $m - 1$ variables; readers are referred to an analogous optimization in Markowitz, Reiman and Wein (1995) for further details. The solution yields the vector of optimal placements x_i^* and $g(x)$, the inventory cost as a function of the total embedded inventory x. Not surprisingly, $g(x)$ is quadratic with linear edges in the inventory level x.

PROPOSITION 1. *The solution to the cycle placement problem is*

Region 1:

$$x < \hat{\alpha}_T = \sum_i \lambda_i\bar{\theta}_{0i}^{\text{TSP}} - \sum_i \frac{b + h_i}{b_i + h_i}\hat{V}_i,$$

$$x_i^* = -\frac{b + h_i}{b_i + h_i}\hat{V}_i \quad \textit{for } i \neq p,$$

$$x_p^* = x - \sum_i \lambda_i\bar{\theta}_{0i}^{\text{TSP}} + \sum_{i\neq p} \frac{b + h_i}{b_i + h_i}\hat{V}_i,$$

$$g(x) = -bx + \hat{a}_1,$$

$$\hat{a}_1 = b\sum_i \lambda_i\bar{\theta}_{0i}^{\text{TSP}} + \frac{1}{2}\sum_i h_i\hat{V}_i - \frac{1}{2}\sum_i \frac{(b + h_i)^2}{b_i + h_i}\hat{V}_i;$$

Region 2:

$$\hat{\alpha}_T \leqslant x \leqslant \hat{\beta}_T = \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} - \sum_i \frac{h_i - h}{b_i + h_i} \hat{V}_i,$$

$$x_i^* = \frac{2\hat{a}_2 \hat{V}_i}{h_i + b_i} \left(x - \sum_k \lambda_k \bar{\theta}_{0k}^{TSP} - \sum_k \frac{(h_i - h_k)\hat{V}_k}{b_k + h_k} \right),$$

$$g(x) = \hat{a}_2 x^2 + \hat{a}_3 x + \hat{a}_4,$$

$$\hat{a}_2 = \frac{1}{2} \left(\sum_i \frac{\hat{V}_i}{b_i + h_i} \right)^{-1},$$

$$\hat{a}_3 = 2\hat{a}_2 \left(\sum_i \frac{h_i \hat{V}_i}{b_i + h_i} - \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} \right),$$

$$\hat{a}_4 = \hat{a}_2 \left(\sum_i \frac{h_i \hat{V}_i}{b_i + h_i} - \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} \right)^2 + \frac{1}{2} \sum_i \frac{b_i h_i \hat{V}_i}{b_i + h_i};$$

Region 3:

$$x > \hat{\beta}_T,$$

$$x_i^* = -\frac{h_i - h}{b_i + h_i} \hat{V}_i \quad \text{for } i \neq \ell,$$

$$x_\ell^* = x - \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} + \sum_i \frac{h_i - h}{b_i + h_i} \hat{V}_i,$$

$$g(x) = hx + \hat{a}_5,$$

$$\hat{a}_5 = -h \sum_i \lambda_i \bar{\theta}_{0i}^{TSP} + \frac{1}{2} \sum_i h_i \hat{V}_i - \frac{1}{2} \sum_i \frac{(h_1 - h)^2}{b_i + h_i \, \hat{V}_i}.$$

In the region where the total inventory is much greater (smaller) than zero, the optimal cycle holds (backorders) most of the inventory at retailer ℓ (p), where it is cheapest to do so, while the cycle for the rest of the retailers remains close to zero. The exact level for each site depends upon two factors: the difference between its holding (backordering) cost h (b) and its nominal delivery size (or equivalently, the proportion of demand that the particular retailer represents). In the region where the total inventory is close to zero, the cycle placement at each retailer varies linearly with the embedded inventory in the system.

1.6 Optimal Base Stock Level

Now that $g(x)$ is known, we can proceed with the solution to the one-dimensional Brownian control problem. The following proposition is proved in Appendix A of RUBIO (1995).

PROPOSITION 2. *The optimal solution to Eqs.* 25–26 *is* $Y^*(t) = \sup_{0 \leqslant s \leqslant t}\{X(s) - z_T^*\}^+$ *for some base stock level* z_T^*.

Hence, the optimal solution is the local time of the Brownian motion at the barrier z_T^*, and the optimally controlled process Z is a reflected Brownian motion (RBM) on $(-\infty, z_T^*]$ (see Section 2.2 of HARRISON, 1985 for a definition). The remainder of this subsection is devoted to the derivation of z_T^*, which can be found by using two well known facts regarding a reflected Brownian motion on $(-\infty, z]$. First, for Y defined in Proposition 2, we have $\lim_{t\to\infty} t^{-1} \mathrm{E}_x[Y(t)] = \theta_T \mu_T / V$ for $\mu_T > 0$, which is independent of the base stock level z. Hence, the transportation cost does not affect the selection of z, and the problem simplifies to minimizing $\limsup_{T\to\infty} (1/T)\mathrm{E}[\int_0^T g(Z(t))\, \mathrm{d}t]$ subject to Eq. 26.

The steady-state density for Z is given by $p_Z(x) = \hat{\nu}_T e^{\hat{\nu}_T(x-z)}$ if $x \leqslant z$ and $p_Z(x) = 0$ if $x > z$, where $\hat{\nu}_T = 2\mu_T/\sigma_T^2 > 0$. Therefore, the optimal base stock level can be found by minimizing

$$\hat{F}_T(z) = \int_{-\infty}^{\hat{\alpha}_T} (-bx + \hat{a}_1)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}\, \mathrm{d}x + \int_{\hat{\alpha}_T}^{\hat{\beta}_T} (\hat{a}_2 x^2 + \hat{a}_3 x + \hat{a}_4)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}\, \mathrm{d}x + \int_{\hat{\beta}_T}^{z} (hx + \hat{a}_5)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}\, \mathrm{d}x \quad \text{for} \quad z \geqslant \hat{\beta}_T, \tag{29}$$

and

$$\hat{F}_T(z) = \int_{-\infty}^{\hat{\alpha}_T} (-bx + \hat{a}_1)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}\, \mathrm{d}x + \int_{\hat{\alpha}_T}^{z} (\hat{a}_2 x^2 + \hat{a}_3 x + \hat{a}_4)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}\, \mathrm{d}x \tag{30}$$

for $\hat{\alpha}_T < z < \hat{\beta}_T$. The constants in Eqs. 29 and 30 have the same definitions as in Section 1.5. Note that, while the optimal base stock level z_T^* always satisfies $z_T^* > \hat{\alpha}_T$ (this is easily seen by the fact that $g(x)$ is linear and has a negative slope for $x < \hat{\alpha}_T$), it need not be larger than $\hat{\beta}_T$.

The following proposition is derived by using integration by parts on Eqs. 29 and 30, and then taking the first two derivatives of $\hat{F}_T(z)$ with respect to z.

PROPOSITION 3. *The value that minimizes $\hat{F}_T(z)$ is*

$$z_T^* = -\frac{1}{\hat{\nu}_T}\ln\left[\left(\frac{h}{b+h}\right)\left(\frac{\hat{\nu}_T(\hat{\beta}_T - \hat{\alpha}_T)}{e^{\hat{\nu}_T(\hat{\beta}_T-\hat{\alpha}_T)} - 1}\right)\right] + \hat{\alpha}_T \quad \text{if} \quad z_T^* \geq \hat{\beta}_T; \quad (31)$$

otherwise, z_T^* is the solution to

$$\frac{2\hat{a}_2}{\hat{\nu}_T} e^{-\hat{\nu}_T(z_T-\hat{\alpha}_T)} + 2\hat{a}_2 z_T + \hat{a}_3 - \frac{2\hat{a}_2}{\hat{\nu}_T} = 0. \quad (32)$$

Furthermore, the predicted optimal cost is $\hat{F}_T(z_T^*) = hz_T^* + \hat{a}_5$ if $z_T^* \geq \hat{\beta}_T$, and $\hat{F}_T(z_T^*) = \hat{a}_2(z_T^*)^2 + \hat{a}_3 z_T^* + \hat{a}_4$ otherwise.

One can show (from the fact that $\hat{F}_T(z)$ is convex and continuously differentiable) that there is a unique optimum base stock level; that is, either there exists a solution z_T^* to Eq. 31 that satisfies $z_T^* \geq \hat{\beta}_T$ or a solution z_T^* to Eq. 32 that satisfies $z_T^* < \hat{\beta}_T$, but not both.

1.7 The Proposed Policy

In this subsection, we map the solution of the approximating heavy traffic control problem into a policy for the original IRP with TSP routing. The control concerns two decisions: whether the vehicle should be busy or idle, and how to assign the load among the retailers during a tour. We address the load allocations first.

Because the system evolves dynamically in time, the decision of how much of the load to leave at each retailer is best delayed until the vehicle arrives at the site. Let t_0 correspond to the epoch at which the vehicle leaves the warehouse with a full load, and consider the epoch $t_i^- > t_0$, which is the point in time just before the vehicle arrives at retailer i. At time t_i^-, the state of the system is given by the inventory levels at the retailers, $(Q_1(t_i^-), \ldots, Q_m(t_i^-))$, and the size of the remaining load, $L(t_i^-)$. The mapping from heavy traffic solution to proposed policy is straightforward: the proposed policy attempts to track the heavy traffic solution (in particular, the optimal cycle placement) as closely as possible. The key issue to be addressed is that the heavy traffic solution is expressed in terms of normalized space and time and in terms of the total embedded inventory process, whereas the proposed policy must be expressed in terms of $(Q_1(t_i^-), \ldots, Q_m(t_i^-), L(t_i^-))$.

Recall that Eq. 27 relates the scaled cycle placement vector x_i and the normalized total embedded inventory $Z(t) = x$. Because the load allocation decision is taken when each retailer is reached, we first establish a relationship between the current total system inventory $Q(t_i^-) = \sum_j Q_j(t_i^-)$ and the corresponding embedded inventory level. Because we need to reverse the scalings in the solution to the heavy traffic control problem, let us define the unscaled embedded inventory $q = \sqrt{n}x$ and the unscaled cycle placement vector $q_i = \sqrt{n}x_i$. In keeping with the behavior predicted by the heavy traffic averaging principle, we develop this relation under a deterministic evolution for the retailer inventories over the course of a cycle. If the vehicle leaves the warehouse at time t_0, then it arrives at retailer i at time t_i^-, where $t_i = t_0 + \theta_{0i}^{TSP}$. Therefore, the retailer inventories relate to the cycle placement parameters by $Q_j(t_i^-) = q_j + \lambda_j\theta_{ij}^{TSP}$ for $j \geq i$ and $Q_j(t_i^-) = q_j + V_j - \lambda_j\theta_{ji}^{TSP}$ for $j < i$. Summing over all retailers, we get

$$Q(t_i^-) = \eta_i + \sum_j q_j, \quad (33)$$

where $\eta_i = \sum_{j<i}(V_j - \lambda_j\theta_{ji}^{TSP}) + \sum_{j\geq i}\lambda_j\theta_{ij}^{TSP}$ is an epoch locator constant for retailer i.

Making the substitutions $q_i/\sqrt{n} = x_i$, $q/\sqrt{n} = x$ and $\theta_{ij}^{TSP}/\sqrt{n} = \bar{\theta}_{ij}^{TSP}$ into Eq. 27 yields the unscaled version of constraint 27,

$$\sum_i q_i = q - \sum_i \lambda_i\theta_{0i}^{TSP}. \quad (34)$$

Using Eqs. 33 and 34, we can express the total inventory at time t_0 (i.e., when the vehicle was at the warehouse) as a translation of the inventory vector at time t_i^-:

$$Q(t_0) = q = \sum_j Q_j(t_i^-) - \eta_i + \sum_j \lambda_j\theta_{0j}^{TSP} \quad \text{for} \quad i = 0, \ldots, m. \quad (35)$$

This equation maps the current inventory levels $Q_j(t_i^-)$ into the one-dimensional quantity q that is required to interpret the heavy traffic results. In particular, for a given value of q, we can find $q_i^* = \sqrt{n}\, x_i^*$, the corresponding optimal (unscaled) cycle placement parameters, by reversing the normalizations in Proposition 1. In particular, if we make the substitutions $q_i/\sqrt{n} = x_i$, $q/\sqrt{n} = x$, $V_i/\sqrt{n} = \hat{V}_i$, and $\theta_{ij}^{TSP}/\sqrt{n} = \bar{\theta}_{ij}^{TSP}$ in Proposition 1, then the scaling parameter n vanishes and the solution q_i^* in terms of the state q is given by Proposition 1 with q_i^*, q, V_i, and θ_{ij}^{TSP} in place of x_i^*, x, $\hat{V}_i$, and $\bar{\theta}_{ij}^{TSP}$, respectively.

We can now use these results to determine the delivery size at retailer i. Under the deterministic inventory evolution for the optimal cycle placement,

$Q_i(t_i^+)$ (i.e. the inventory level at retailer i just after the delivery is made) satisfies

$$Q_i(t_i^+) = q_i^* + V_i. \tag{36}$$

If we deliver d_i units to retailer i, then the actual inventory after the delivery is made is $Q_i(t_i^-) + d_i$; equating this quantity to Eq. 36 yields the desired delivery size,

$$d_i = q_i^* + V_i - Q_i(t_i^-). \tag{37}$$

Because we cannot allocate more than the available load and do not want to make negative deliveries, the proposed delivery size is given by

$$d_i^* = \max[d_i,\ 0] + \min[0,\ L(t_i^-) - d_i] \quad \text{for} \quad i = 1, 2, \ldots, m-1. \tag{38}$$

Finally, to guarantee that the vehicle returns to the warehouse empty, we set

$$d_m^* = L(t_m^-). \tag{39}$$

We could, in principle, allow negative deliveries as long as there is inventory available at retailer i and the total amount of load the vehicle carries as it leaves this retailer is kept under its total capacity. However, because the vehicle returns empty, the items accrued from negative deliveries would most likely be shifted to the last few retailers of the tour, which will not necessarily bring the state of the system closer to the optimal cycle; hence, we disallow negative deliveries.

To recapitulate, the proposed dynamic delivery allocations are derived by the following procedure: (i) observe the current inventory levels $(Q_1(t_i^-), \ldots, Q_m(t_i^-))$ and compute the unscaled embedded inventory q using Eq. 35, (ii) use the unscaled version of Proposition 1 to derive the optimal unscaled cycle placement parameters q_i^*, and (iii) observe the current remaining load $L(t_i^-)$ and compute the proposed delivery size using Eqs. 37–39.

We now turn our attention to the busy/idle policy, which has decision epochs when the vehicle is at the warehouse. At these points in time, the vehicle starts a new tour if the total inventory level $\Sigma_j Q_j(t)$ is below the unscaled aggregate base stock level $w_T^* = \sqrt{n}\, z_T^*$; otherwise, it idles. If we define $\nu_T = 2(1-\rho_T)V/[\lambda\theta_T(c_d^2 + Vc_T^2)]$ then substituting w_T^*, ν_T, and $F_T(w_T^*)$ into Proposition 3 (and using the unscaled version of Proposition 1) yields the optimal unscaled base stock level w_T^* and the predicted optimal cost $F_T(w_T^*)$ solely in terms of the original problem parameters.

We have completely characterized a dynamic control policy that depends exclusively on the original system parameters. The policy specifies the two controllable aspects of the system: the aggregate base stock level defined by the unscaled version of Proposition 3 determines the vehicle idling policy, and the delivery sizes d_i^* defined in Eqs. 38–39 characterize the allocation of units to retailers.

2. THE IRP WITH DIRECT SHIPPING

THE ONLY DIFFERENCE between the direct shipping (DS) case and the TSP case is the routing scheme used when the vehicle is operating. We retain all notation from Section 1, occasionally using the subscript D (for direct shipping) in place of the subscript T (for TSP). Moreover, because the procedure is very similar in both cases (as before, it is assumed that the HTAP holds for a certain class of policies), we omit nearly all of the details for the DS case, describing only the distinctive aspects of the analysis. The most significant difference between DS and TSP is that the DS case does not have a cyclic structure. This has several consequences, one of which is that the results are not as theoretically solid as in the TSP case.

2.1 Heavy Traffic Analysis

In the DS case, the vehicle always leaves the warehouse with a full load, visits a single retailer, and returns empty, so that, every time a retail site is visited, its inventory level increases by V units. As before, it is convenient to express the dynamic allocation as deviations from a nominal policy. The nominal policy we consider is not achievable: we assume that, under the nominal policy, an amount V_i is delivered to retailer i in every delivery. We let $S_D(t)$ denote the number of DS deliveries made by a continuously active vehicle during $[0, t]$. We then let $\epsilon_i^D(t)$ be defined by the analog of Eq. 1 obtained by replacing T by D. Then, $\epsilon_D(t) = \Sigma_i\, \epsilon_i^D(t)$ is the total amount delivered during the current trip. These processes satisfy Eqs. 2–6. Equation 7 still holds as well, once Ts are replaced by Ds. The problem formulation is thus nearly identical to Eqs. 2–12.

The DS case lacks the natural cyclic structure of the TSP. Tour times in the TSP are i.i.d., and each tour results in the delivery of V units. The DS case would have a cyclic structure if the sequence of retailers visited followed a cyclic pattern (such as a polling table), or had a regenerative structure (such as a Markov chain). Neither of these can be used in the DS case for exactly the same reason that fixed delivery sizes could not be used in the TSP case: inventory costs would be infinite over the long run. A dynamic policy, described in Section 2.2, is used. For this policy, the fraction of total shipments that go to retailer i does not vary over times of order n. To satisfy average demand at all retailers, this fraction must be λ_i/λ for retailer i. The traffic intensity for the DS system is thus $\rho_D = 2 \sum_i \lambda_i \theta_{0i}/V$.

We use the same heavy traffic normalizations as in Section 1.2. The heavy traffic conditions are given by Eqs. 20–24, with Eqs. 21 and 23 understood to hold, respectively, for θ_{0i} and c_{0i}^2, $1 \leq i \leq m$, and Eq. 22 replaced by

$$\mu_{\mathrm{D}} = \sqrt{n}\left(\frac{\lambda V}{2 \sum_i \lambda_i \theta_{0i}} - \lambda\right) > 0. \tag{40}$$

As mentioned above, the DS case does not have the cyclic structure of the TSP case. Thus, we cannot use functional central limit theorems based on renewal processes to prove convergence to a Brownian motion. In the DS case, the parameter corresponding to s_{T}^2 in the TSP case, s_{D}^2, is given by $s_{\mathrm{D}}^2 = \lambda^{-1} \sum_i \lambda_i \theta_{0i}^2 c_{0i}^2$. This expression would clearly arise if the next retailer were chosen in an i.i.d. manner. It can be shown to hold for any policy where the fraction of times that a retailer is visited does not vary over times of order n using the random time change theorem.

The lack of a cyclic structure makes it impossible to consider an embedded normalized netput process. Indeed, if we embed at epochs during which the vehicle is at the warehouse, we will not obtain a meaningful process. We must thus average the normalized netput process to obtain a meaningful limit. By arguments similar to those in Section 1.3, this averaged normalized netput process is well approximated by a Brownian motion X with drift μ_{D} and variance

$$\sigma_{\mathrm{D}}^2 = \lambda\left(c_d^2 + \frac{\lambda V \sum_i \lambda_i \theta_{0i}^2 c_{0i}^2}{2(\sum_i \lambda_i \theta_{0i})^2}\right).$$

The averaged total inventory process is defined by

$$Z(t) = X(t) - \frac{\lambda V}{2 \sum_i \lambda_i \theta_{0i}} Y(t). \tag{41}$$

Now, we slow down time by a factor of $\sqrt{n}$ and turn to the fluid model. Again we face the problem of optimally placing the limit cycles for the deterministic evolution of the retailer inventories in $\mathbf{R}^m$. In contrast to the TSP case, an approximation is introduced to facilitate this optimization. We motivate this approximation by use of an example. Suppose that $V = 60$, $\lambda_1 = 3$, $\lambda_2 = 2$ and each retailer was exactly six time units away from the warehouse. Then a continually busy vehicle using the polling table (12121) could visit, on average, retailer 1 every 20 time units and retailer 2 every 30 time units, thereby satisfying average demand. Although optimally placing a limit cycle for a small polling table such as this one is manageable, the optimization problem gets unwieldy very quickly as the size of the table grows. Our approximation assumes the existence of an idealized policy that would make a delivery to retailer i every V/λ_i time units in the fluid model; in the context of our example, we assume that retailer 1 (retailer 2) receives a delivery exactly every 20 (30) time units, even though the (12121) polling table cannot achieve such perfect regularity in the fluid model.

Hence, our approach is to optimally place an idealized fluid cycle at the fast time scale, and then track this cycle as closely as possible with our proposed policy. Because deliveries are perfectly regular in the idealized cycle, the use of this approximation causes us to underestimate the inventory cost incurred over a cycle; however, simulation results in Section 5 show that the heavy traffic analysis incorporating this approximation appears to be very accurate, at least for the five-retailer cases considered there. Moreover, the use of an idealized cycle allows us to avoid the task of determining the actual behavior of the individual inventory levels on the fluid time scale. This task is more than just tedious: due to the lack of a cyclic structure it appears to be extremely difficult.

Now we turn to the optimal placement of the idealized cycle. We still define the cycle placement by the vector $(x_1, x_2, \ldots, x_m)$, where x_i represents the lowest point during the cycle of the fluid inventory level at retailer i. Under the idealized policy, the fluid inventories are similar to the TSP paths pictured in Figure 1; the only differences are the delivery size (in this case we deliver a full load of $\hat{V}$ units on each visit to a retailer) and the visit frequency, which equals $\hat{V}/\lambda_i$ to maintain a balanced flow.

The next step is to establish the relationship between the cycle placement variables x_i and the averaged total inventory level $Z(t) = x$. The constraint related to consistency between individual and total inventory levels takes the form that the averaged total inventory equals the sum of the average individual inventory levels. The average fluid inventory at retailer i over an idealized cycle is $x_i + \hat{V}/2$. Hence, when the averaged total inventory $Z(t) = x$, the cycle placement parameter must satisfy

$$\sum_i x_i = x - \frac{m\hat{V}}{2}. \tag{42}$$

Because the fluid delivery size equals $\hat{V}$ under the DS case, the inventory cost function $g_i(x_i)$ is given as in Eq. 28, except that $\hat{V}_i$ is replaced by $\hat{V}$. Comparing constraints 27 and 42, it is clear that the optimal cycle placement is precisely the solution given in the TSP case, except that we replace $\hat{V}_i$ by $\hat{V}$ and $\Sigma_i \lambda_i \bar{\theta}_{0i}^{\mathrm{TSP}}$ by $m\hat{V}/2$.

The computation of the optimal vehicle idling policy is identical to that in Section 1, except that the parameter of the exponential stationary density of the RBM is $\hat{\nu}_{\mathrm{D}} = 2\mu_{\mathrm{D}}/\sigma_{\mathrm{D}}^2$. Hence, Proposition 3 characterizes the optimal base stock level for the DS case, with $\hat{\nu}_{\mathrm{D}}$ replacing $\hat{\nu}_{\mathrm{T}}$, and with the substitutions described above in the corresponding definitions of the constants $\hat{a}_i$ and the thresholds $\hat{\alpha}$ and $\hat{\beta}$.

2.2 The Proposed Policy

The mapping from heavy traffic solution to proposed policy uses the same philosophy as in the TSP case. We begin by establishing a relationship between the total inventory in the system at the current decision epoch and the cycle placement parameters. Although there are several possible ways to do this, we keep track of the vector process $(r_1(t), \ldots, r_m(t))$, which specifies the time of the most recent visit to each retailer. Hence, if we denote the current time by t, then $t - r_i(t)$ represents the elapsed time since the vehicle last visited retailer i. The inventory at retailer i at time t relates to the unscaled cycle placement parameter $q_i = \sqrt{n}\, x_i$ via

$$Q_i(t) = q_i + V - \lambda_i(t - r_i(t)). \tag{43}$$

Because stochastic effects can lead to unusually long intervisit periods, it is possible to have $V - \lambda_i(t - r_i(t)) < 0$, which would make the cycle placement of the retailer higher than the current inventory level, thereby contradicting the definition of q_i. Because one would expect that $Q_i(t) \geq q_i + \lambda_i\theta_{0i}$, we modify Eq. 43 to $Q_i(t) = q_i + \max[V - \lambda_i(t - r_i(t)), \lambda_i\theta_{0i}]$; this modification leads to considerable improvements in system performance in the simulation study. Summing over all retailers, we obtain $\Sigma_i q_i = Q(t) - u(t)$, where $u(t) = \Sigma_i \max[V - \lambda_i(t - r_i(t)), \lambda_i\theta_{0i}]$. Combining this equation with the unscaled version of Eq. 42 yields

$$q = \sum_i Q_i(t) - u(t) + \frac{mV}{2}, \tag{44}$$

which relates the unscaled averaged inventory q (this quantity represents the unscaled average total inventory in the DS case) to the current inventory level. Once again, the heavy traffic parameter n vanishes when the heavy traffic normalizations are reversed, and the formulas for the unscaled optimal cycle placement vector q_i^* given q can be obtained by substituting q, q_i^*, V and $mV/2$ for x, x_i^*, $\hat{V}_i$ and $\Sigma_i \lambda_i \bar{\theta}_{0i}^{\mathrm{TSP}}$, respectively, into Proposition 1.

We use this cycle placement as an ideal m-dimensional inventory state for a given q and $u(t)$, and choose the next retailer to bring the current inventory vector as close as possible to this ideal state. Let t_0 be the time epoch at which the vehicle is ready to depart from the warehouse, and consider the inventory evolution over the next delivery trip. Under a deterministic inventory evolution, the vehicle will reach retailer i (if it chooses to go there next) at time $t_i = t_0 + \theta_{0i}$. In the deterministic tour corresponding to the optimal cycle placement, the retailer inventory levels right after a delivery is made to retailer i is given by $Q_i^*(t_i^+) = q_i^* + V$ and

$$Q_j^*(t_i^+) = \max[q_j^* + V - \lambda_j(t_0 + \theta_{0i} - r_j(t_0)),\; q_j^* + \lambda_j(\theta_{0i} + \theta_{0j})] \quad \text{for} \quad j \neq i,$$

where the maximization makes the adjustment for long intervisit times as discussed above. In contrast, under the deterministic evolution, the actual inventory vector after a delivery to retailer i is given by $Q_i(t_i^+) = Q_i(t_0) + V - \lambda_i\theta_{0i}$ and $Q_j(t_i^+) = Q_j(t_0) - \lambda_j\theta_{0i}$ for $j \neq i$. Therefore, the resulting Euclidean distance between the ideal and actual inventory vectors after a delivery to retailer i is $\Delta(i) = \sqrt{\Sigma_j (Q_j(t_i^+) - Q_j^*(t_i^+))^2}$. The proposed control sends the vehicle to retailer k, where $k = \arg\min_i \Delta(i)$.

Finally, as in the TSP case, the busy/idle control is found by unscaling the heavy traffic results. The only added complexity is that, for the DS case, the vehicle visits the warehouse after every delivery and so has many possible idling decision epochs. By Eq. 44, our proposed policy idles a vehicle at the warehouse whenever $Q(t) - u(t) + mV/2 > w_{\mathrm{D}}^*$, where $w_{\mathrm{D}}^* = \sqrt{n}\, z_{\mathrm{D}}^*$ is the unscaled idling threshold. As in the TSP case, if we define $\nu_{\mathrm{D}} = (1 - \rho_{\mathrm{D}})\lambda V/[\sigma_{\mathrm{D}}^2 \Sigma_i \lambda_i\theta_{0i}]$ then substituting w_{D}^*, ν_{D}, and $F_{\mathrm{D}}(w_D^*)$ into Proposition 3 (and using the unscaled DS version of Proposition 1) yields the optimal unscaled base stock level w_{D}^* and the predicted optimal cost $F_{\mathrm{D}}(w_{\mathrm{D}}^*)$ in terms of the original problem parameters.

3. COMPARISON OF TSP AND DS ROUTING

IN THIS SECTION, we compare the relative performance of the two fixed routing schemes (TSP and DS). The predicted cost functions F_{T}, F_{D} derived in Sections 1 and 2 represent only the inventory component of the system cost. Denote a generic fixed routing scheme by $\Re \in \{\mathrm{TSP}, \mathrm{DS}\}$, and by $C(\Re)$ the total cost for the system under this scheme. The total system cost is obtained by adding the transportation cost (or equivalently, subtract the idleness reward) to the inventory cost; that is, we set $C(\Re) = F_{\Re}(w^*_{\Re}) - r(1 - \rho_{\Re})$.

A crucial observation is that $\rho_D < \rho_T$; this fact (see Rubio (1995) for a proof) is a simple consequence of the triangle inequality. Although trivial to prove, this inequality has several important implications. First, the DS policy achieves lower transportation costs than does the TSP policy. This is quite interesting because, at first glance, one might expect the converse to hold. However, minimization of the steady-state transportation cost in the IRP context is equivalent to maximizing the amount of items delivered per unit time traveled; hence, full-load direct shipping provides the highest transportation efficiency of any fixed routing scheme. More importantly, for any given problem instance, the demand rate can be increased until $\rho_T = 1$ and $\rho_D < 1$; that is, there exist some demand levels where the DS policy would be stable while the TSP policy would not. We should note that $\rho_{\Re} < 1$ is a necessary condition for stability of any fixed routing scheme $\Re$ but it is not sufficient. In particular, having $\rho_{\Re} < 1$ will keep the total inventory stable but, in the absence of adequate dynamic load allocation, it is possible to accumulate inventory at one retail site while backorders grow without bound at another. Hence, DS will dominate TSP routing as $\rho_T \to 1$ as long as some form of stable dynamic allocation is used in the DS case.

The remainder of this section investigates the relative performance of the DS and TSP schemes as a function of the cost parameters r and b. However, readers should keep in mind that, although the qualitative statements below are true, these results are not exact (even if the HTAP assumption is valid), because our calculation of $F_D(w^*_D)$ is approximate and represents a slight underestimate of the true heavy traffic cost under the DS policy. The inequality $\rho_D < \rho_T$ implies that DS is preferred to the TSP policy if the transportation cost is high enough. In particular, the DS policy achieves a lower overall cost for any $r > (F_T(w^*_T) - F_D(w^*_D))/(\rho_T - \rho_D)$. Although the value of this threshold cost may be found numerically for any particular problem instance, a more precise characterization requires a better understanding of the relationship between the inventory costs in both systems. Unfortunately, it is hard to make simple inventory cost performance comparisons for the different routing schemes, primarily because the base stock levels, and hence the predicted inventory costs, are not in closed form (see Proposition 3).

To study the relative inventory cost performance of the TSP and DS schemes, let us consider the case where the inventory costs at the retailers are symmetric (i.e., $h_i = h$ and $b_i = b$ for all i) and b becomes large. Because the values of w^*_T and w^*_D in the unscaled version of Proposition 3 are increasing in b/h, one expects that there exist some critical values b_T, b_D such that, if b is increased above them (while leaving h fixed), the optimal base stock is given in closed form. These critical values indeed exist and, for the case of symmetric costs, have the following closed form expressions:

$$b_T = h\left[\frac{\nu_T V e^{\nu_T V}}{e^{\nu_T V} - 1} - 1\right]$$

and

$$b_D = h\left[\frac{\nu_D m V e^{\nu_D m V}}{e^{\nu_D m V} - 1} - 1\right].$$

For $b > \max\{b_T, b_D\}$, the inventory cost difference $F_T(w^*_T) - F_D(w^*_D)$ can be expressed as

$$h\left(\frac{\nu_D - \nu_T}{\nu_D \nu_T} \ln\left[1 + \frac{b}{h}\right] + \frac{1}{\nu_T} \ln\left[\frac{e^{\nu_T V} - 1}{\nu_T V}\right] - \frac{1}{\nu_D} \ln\left[\frac{e^{\nu_D m V} - 1}{\nu_D m V}\right] + \frac{(m-1)V}{2}\right). \quad (45)$$

As $b \to \infty$, the value of Eq. 45 is dominated by the term $(\nu_T^{-1} - \nu_D^{-1}) \ln(1 + b/h)$, whose sign will be the same as the sign of $\nu_D - \nu_T$. Define the critical value

$$b_c = h\left(\frac{e^{\nu_T V} - 1}{\nu_T V}\right)^{\nu_D/(\nu_T - \nu_D)} \cdot \left(\frac{e^{\nu_D m V} - 1}{\nu_D m V}\right)^{\nu_T/(\nu_D - \nu_T)} \exp\left[\frac{\nu_D \nu_T (m-1)V}{2(\nu_T - \nu_D)}\right] - h,$$

where $\exp[x] = e^x$. Then for $b > \max\{b_T, b_D, b_c\}$, the DS policy achieves the lower inventory cost if and only if $\nu_D - \nu_T > 0$, where $\nu_{\Re}$ is the exponential parameter for the steady-state distribution of the RBM associated with routing scheme $\Re$. Because $\rho_D < \rho_T$, the condition $\sigma^2_D > \sigma^2_T$ is required for the TSP policy to be preferred. For the case of deterministic travel times, it follows that $\sigma^2_D = \sigma^2_T$, and so DS dominates in the high backorder case. Moreover, if both μ_D and μ_T are finite, then the difference in mean distance traveled must be $O(n^{-1/2})$; see Eq. 51. Hence, in the heavy traffic limit, we actually have $\sigma^2_D = \sigma^2_T$.

This result is somewhat counterintuitive: because the TSP policy makes smaller and more frequent deliveries to each retailer, it might be expected to outperform the DS scheme in terms of inventory cost. However, for large backorder penalties, the TSP policy sacrifices efficiency over the long run through its smaller drift, and causes the total inventory to spend too much time in the expensive back-order regions.

4. THE IRP WITH DYNAMIC ROUTING

CONSIDER NOW a situation where, once the vehicle is loaded at the warehouse, it can embark on either a full load TSP tour or a direct shipment to some retailer. All other aspects of the problem (e.g., sources of uncertainty, cost structure) remain the same as in the fixed routing cases. Because the formulation of the dynamic problem is a natural extension of the two fixed routing problems described earlier, we only briefly summarize our approach, using the notation introduced earlier. Readers are referred to REIMAN, RUBIO, and WEIN (1997) for a sketch of the arguments and to Rubio (1995) for a detailed treatment.

We start by defining the limiting dynamic control problem. The heavy traffic conditions are a union of the conditions for the two fixed routing problems. Let $\Re(t) \in \{\text{TSP, DS}\}$ denote the routing mode that is used at time t in the limiting control problem, and let $\hat{\delta}(t) = B_{\text{D}}(nt)/(B_{\text{T}}(nt) + B_{\text{D}}(nt))$ denote the cumulative fraction of busy time that the DS service has been used. Define the diffusion process $X(t, \Re(t))$ with control-dependent drift and variance given by

$$\mu(\Re(t)) = \begin{cases} \mu_{\text{D}} & \text{if} \quad \Re(t) = \text{DS} \\ \mu_{\text{T}} & \text{if} \quad \Re(t) = \text{TSP} \end{cases}$$

and

$$\sigma^2(\Re(t)) = \begin{cases} \sigma_{\text{D}}^2 & \text{if} \quad \Re(t) = \text{DS} \\ \sigma_{\text{T}}^2 & \text{if} \quad \Re(t) = \text{TSP}, \end{cases}$$

respectively. Then the averaged total inventory level for the system is given by

$$Z(t, \Re(t)) = X(t, \Re(t)) - \left(\frac{V}{\theta_{\text{T}}}(1 - \hat{\delta}(t)) + \frac{\lambda V}{2\sum_i \lambda_i \theta_{0i}}\hat{\delta}(t)\right) Y(t). \quad (46)$$

If we assume that the time scale decomposition holds for the dynamic IRP, then the problem decomposes into: (i) given $Z(t) = x$ and routing mode $\Re(t) \in \{\text{TSP, DS}\}$, use the results in Sections 1 and 2 to determine the optimal cycle placement and the corresponding inventory cost function $g_{\Re(t)}(x)$; (ii) choose the nonanticipating control $(Y(t), \Re(t))$ (where Y is nondecreasing and right continuous) to minimize

$$\limsup_{T\to\infty} \frac{1}{T} \mathrm{E}\left[\int_0^T g_{\Re(t)}(Z(t, \Re(t)))\, \mathrm{d}t - \hat{r}Y(T)\right], \quad (47)$$

subject to Eq. 46.

The diffusion control problem, Eqs. 46–47, appears to be difficult to tackle for two reasons: the coefficient in front of the control $Y(t)$ in Eq. 46 depends upon the routing control $\Re(t)$, and the control-dependent cost function $g_{\Re}(x)$ is very complex. We analyze this problem in two ways. First, we use the Markov chain approximation procedure pioneered by KUSHNER (1977) (see KUSHNER and DUPUIS 1992 for a recent account) to numerically compute the optimal solution. We also derive the optimal policy within a class of triple threshold policies that is described in Section 6. Computational results for the dynamic IRP also appear in Reiman, Rubio and Wein (1997), and the main observations from these results are summarized in Section 6.

5. COMPUTATIONAL RESULTS

THIS SECTION CONTAINS a description of a series of computational experiments aimed at assessing the accuracy of the heavy traffic analysis and determining what aspects of the control policy are most important for good system performance. A discussion of these results appears in Section 6.

The Monte Carlo simulation experiments performed in this subsection consider systems that have five retailers and Poisson demand processes. We also set the transportation cost rate r equal to zero and concentrate on the inventory cost. The total arrival rate λ is varied to obtain different utilization rates; however, the fraction of demand represented by retailer i is fixed so that $\lambda_1 = \lambda/5$, $\lambda_2 = \lambda/10$, $\lambda_3 = \lambda/10$, $\lambda_4 = \lambda/5$, and $\lambda_5 = 2\lambda/5$. The travel time random variables T_{ij} are i.i.d. second-order Erlang. The mean travel times are adjusted so that $10\theta_{\text{T}} = V$ always holds; this allows us to consider several vehicle sizes while maintaining the traffic intensity at $\rho_{\text{T}} = 0.1\lambda$.

We perform four simulation experiments aimed at various aspects of system performance.

EXPERIMENT 1. The first set of simulation runs quantifies the cost improvement obtained under the TSP policy by recalculating the cycle placement at each retailer, as opposed to determining these values only once per cycle (i.e., when the vehicle is at the warehouse). We let $b_i = b = 5$, $h_i = h = 1$ for $i = 1, \ldots, 5$ and consider the mean travel times

TABLE I
Cost Increase when Placement Calculation is Only at Warehouse

	$\rho_T = 0.5$	$\rho_T = 0.7$	$\rho_T = 0.9$
		%	
$V = 100$	4.0	5.0	1.1
$V = 10$	7.4	18.6	0.7
$V = 5$	7.3	18.1	1.1

$\theta_{01} = \theta_{50} = \theta_T/10$, $\theta_{12} = \theta_{23} = \theta_{34} = \theta_{45} = \theta_T/5$. These travel times are consistent with a pentagon structure, where the five retailers are placed at the vertices of a pentagon, and the warehouse is located midway between stations 1 and 5.

We consider nine different scenarios, which are generated by the different combinations of three vehicle sizes (100, 10, 5) and three traffic intensities (0.5, 0.7, 0.9); notice that some of these scenarios grossly violate the heavy traffic conditions. For all cases with $\rho_T \leq 0.7$, we simulated three replications of 36,000 time units (starting with an empty system and discarding the first 2000 time units) with cycle placement recalculation at the retailers and three more with calculation only at the warehouse; for the $\rho_T = 0.9$ instances, the length of each replication was increased to 240,000 time units (discarding the first 20,000 time units). This simulation design was used throughout our study and allowed us to keep the standard deviation of the cost estimate under 1% of its mean.

Table I summarizes the results of the experiment. The entries in the table represent the increase in the average inventory cost when delivery sizes are calculated only at the warehouse, and not adjusted over the course of the tour. All subsequent TSP simulations use the cycle placement recalculation at the retailers.

EXPERIMENT 2. The second simulation experiment assesses the accuracy of the heavy traffic analysis by comparing the cost incurred under the derived base stock levels with the cost incurred under the best possible base stock level. We maintain the same set-up as in the first experiment, except that asymmetric cost cases are also considered, where the holding rates are (1, 1, 2, 2, 2) and the backorder rates are (5, 10, 5, 10, 5) for the five retailers, respectively. For each of these 18 cases (three vehicle sizes, three traffic intensities, and two cost structures), we performed an exhaustive search in a series of simulations (each consisting of three replications with the length described in Experiment 1) to find the base stock level that provides the lowest system cost. Table II summarizes the results; each entry represents the suboptimality (within the class of base stock policies) of the cost incurred by using the derived base stock level instead of the optimal base stock level found by exhaustive search.

TABLE II
Cost Suboptimality of Derived Base Stocks for Pentagon TSP

	$\rho_T = 0.5$	$\rho_T = 0.7$	$\rho_T = 0.9$
		%	
$V = 100$			
Symm.	0.0	0.9	2.6
Asym.	0.6	2.2	0.0
$V = 10$			
Symm.	19.1	4.3	1.7
Asym.	6.7	2.2	1.8
$V = 5$			
Symm.	14.6	1.1	1.5
Asym.	11.6	1.1	0.6

EXPERIMENT 3. Now we study the performance of the direct shipping policy, and compare it to the performance of the TSP policy on the same system. As before, this is done by comparing the average cost obtained under the derived base stock levels with that under the optimal base stock level found by exhaustive search. Because the DS policy has a huge drift advantage over the TSP policy in the pentagon topology used for Experiments 1 and 2, this experiment uses the travel times $\theta_{0i} = 0.45\theta_T$ for $i = 1, \ldots, 5$ and $\theta_{12} = \theta_{23} = \theta_{34} = \theta_{45} = \theta_T/40$, so that $\rho_T = 0.1\lambda$ and $\rho_D = 0.09\lambda$. This case will be referred to as the wedge topology, because these travel times are consistent with such a shape. In practice, TSP tours are often generated by placing the warehouse at the center of the pie, dividing the pie into wedges and solving a TSP on each wedge; see Figure 1 of Bell et al. (1983) and Figure 3 of FEDERGRUEN and SIMCHI-LEVI (1992). The other problem parameters remain as in the symmetric cost scenarios in Experiment 2, except for the fact that we also simulate the DS policy for the case when $\lambda = 10$. The TSP policy is not simulated for this case because it corresponds to $\rho_T = 1$, and the system is not stable under this scheme. Hence, we consider 12 cases (four traffic intensities and three vehicle sizes) for the DS policy and nine cases for the TSP.

Tables III and IV summarize the results of this experiment. The entries in Table III compare the performance of the proposed base stock level to the cost obtained under the best base stock level for the same policy. Table IV presents a comparison of the average inventory cost for the DS and TSP policies. The percentage difference between the TSP cost and the DS cost is given by

$$\frac{\text{TSP cost} - \text{DS cost}}{\text{DS cost}} \times 100\%.$$

TABLE III

Cost Suboptimality of Derived Base Stocks for Wedge Topology

	$\rho_T = 0.50$ $\rho_D = 0.45$	$\rho_T = 0.70$ $\rho_D = 0.63$	$\rho_T = 0.90$ $\rho_D = 0.81$	$\rho_T = 1.00$ $\rho_D = 0.90$
		%		
$V = 100$				
TSP	2.41	5.31	6.10	N.A.
DS	4.13	2.42	1.74	0.52
$V = 10$				
TSP	11.04	0.00	3.67	N.A.
DS	4.30	3.01	2.43	0.08
$V = 5$				
TSP	17.48	0.00	1.19	N.A.
DS	3.70	1.75	1.40	0.00

The entries labeled "Sim." in Table IV represent the percentage difference in inventory cost when base stock levels are found by exhaustive search. The entries labeled "Pred." represent the difference in inventory costs as predicted by the heavy traffic analysis. Recall that the percentage differences in Table IV only assess the inventory costs, and the DS policy will always incur lower transportation costs than the TSP policy. Hence, the desired policy is a function of the transportation cost rate r.

EXPERIMENT 4. The last experiment in this subsection measures the increase in cost incurred by using a base stock level different from the one proposed in the heavy traffic analysis. We already have the required data for this analysis from the exhaustive search performed in the simulation experiments above. Figure 2 plots three examples of the cost increase with respect to the proposed policy, as a function of the base stock level (expressed in units of vehicle size). These three cases correspond to the DS system on the wedge topology for $V = 100$ and $\lambda \in \{5, 7, 9\}$. The behavior illustrated here is typical of all other instances analyzed in our simulation experiments.

TABLE IV

Inventory Cost Comparison (TSP − DS)/DS: Wedge Topology

	$\rho_T = 0.50$ $\rho_D = 0.45$	$\rho_T = 0.70$ $\rho_D = 0.63$	$\rho_T = 0.90$ $\rho_D = 0.81$	$\rho_T = 1.00$ $\rho_D = 0.90$
		%		
$V = 100$				
Sim.	−80.2	−72.5	−29.1	N.A.
Pred.	−78.4	−71.8	−22.1	∞
$V = 10$				
Sim.	−66.5	−58.6	−3.0	N.A.
Pred.	−76.7	−65.0	2.5	∞
$V = 5$				
Sim.	−56.7	−64.5	12.6	N.A.
Pred.	−74.3	−57.8	25.2	∞

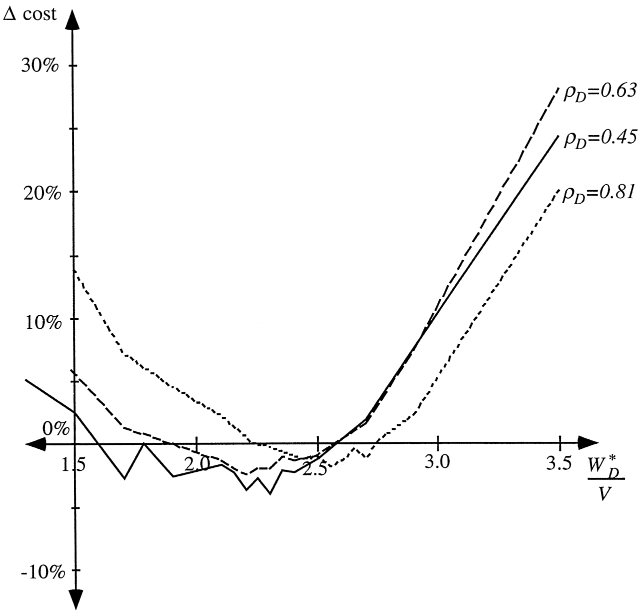

Fig. 2. Sensitivity of inventory cost to base stock level.

6. SUMMARY AND CONCLUSIONS

THE IRP IS ONE of the more challenging problems in operations research, especially when considered from a dynamic and stochastic viewpoint. We focus on the operational aspects of the problem and consider a system with a single capacitated vehicle that operates out of a single warehouse and services a finite set of retailers. By restricting an outgoing vehicle to deliver full loads to either a single retailer (direct shipping or DS) or along a prespecified (TSP) tour, we avoid the combinatorial complexities inherent in the problem and maintain a sharp focus on the crucial tradeoff between inventory costs and transportation costs that lies at the heart of the IRP. Our modeling of the dynamic stochastic IRP as a queueing control problem offers a new perspective on the problem: rather than view the IRP as a variant of the vehicle routing problem, we see it as a variant of a production/inventory control problem (where the capacitated vehicle plays the role of the production system); as such, this paper is a natural descendant of WEIN (1992) and MARKOWITZ, REIMAN and WEIN (1999), which consider more conventional production/inventory control problems.

By assuming that the system is operating in the (suitably defined) heavy traffic regime and that a heavy traffic time scale decomposition holds, we approximate the queueing control problem by a diffusion control problem. When only TSP tours are allowed, this modeling approach allows us to fully characterize the solution to the diffusion control problem, thereby generating an operating policy for the original system. By assuming the existence of a fixed sequence of retailer visits that can achieve constant interdelivery times to each retailer in the fluid limit, we perform a similar analysis for the DS case. The control policy in both cases is characterized by a vehicle idling policy, which dictates whether a vehicle at the warehouse should sit idle or set out with a full load, and a dynamic allocation policy, which specifies how many units to leave off at each retailer under a TSP scheme, and which retailer to visit next in the DS scheme.

We also consider the case where dynamic route selection (either TSP or DS) is allowed. The diffusion control problem is solved numerically and a class of triple-threshold policies is analyzed. Finally, a series of simulation studies is performed to complement the heavy traffic analysis.

Our key findings can be summarized as follows:

- The inventory component of the total long-run average cost depends on the stochastic characteristics of the system, whereas the transportation component for a fixed routing scheme is determined solely from first-moment information. Moreover, for the two fixed routing IRPs, the proposed solution is independent of the transportation cost rate because all base stock policies incur the same long-run average transportation cost.
- The vehicle idling policy is characterized by an aggregate base stock level: the vehicle idles at the warehouse whenever the total retailer inventory exceeds a certain threshold level. Although the existing IRP literature does not typically address the vehicle idling issue, our simulation results show that system performance is quite sensitive to the value of the base stock level, deteriorating rapidly when the base stock level differs from the optimal value by more than the vehicle capacity. Moreover, simulation results also confirm that the system cost under our derived base stock levels are typically within several percent of the cost achieved by the best (found by exhaustive search using simulation) base stock level, unless the heavy traffic conditions are grossly violated (e.g., traffic intensity = 0.5 and vehicle capacity $\leqslant 10$ units).
- The allocation of load among the retailers is dictated by the desire to concentrate most of the total inventory (backorders) at the site with the smallest holding (backorder) cost rate.
- Dynamic (i.e., state-dependent or closed-loop) delivery allocations greatly outperform their static (state-independent or open-loop) counterparts in a stochastic environment. In fact, central limit theorem arguments indicate that static delivery allocations lead to unbounded costs over the long run.
- The relative advantage of recalculating the load allocation at each retailer within a TSP tour, as opposed to setting it once at the beginning of each tour, decreases as utilization increases, and vanishes in the heavy traffic limit. This observation complements those in Kumar, Schwarz, and Ward (1995), who focus on this issue (calculating delivery allocations once per cycle or at each retailer) using a much different model.
- The policy that achieves the lowest transportation cost is the one that delivers the largest amount per unit time traveled (subject to meeting average demand). Therefore, direct shipping is the most transportation-efficient routing scheme. This fact helps highlight the basic cost tradeoff in the IRP: DS leads to smaller transportation costs, but TSP routing, by making smaller and more frequent deliveries, may lead to smaller inventory costs. This result also implies that DS has a larger stability region than TSP; that is, for any given problem instance, one can increase the demand rates to a level where the TSP routing scheme has a traffic intensity greater than or equal to one and the DS scheme has an intensity less than one.
- Heavy traffic analysis shows that, for cost-symmetric systems with sufficiently high backorder costs, DS will be preferred to TSP routing.
- Simulation results show that there is often a large difference in performance between the DS and TSP policies. Although the traffic intensity, backorder costs and transportation costs all play a significant role, the topology probably plays the largest role in the relative attractiveness of each policy. For systems with relatively high loads, it appears that TSP could only be a desirable alternative when the tour is wedge-shaped, as is often the case in practice.
- The heavy traffic analysis accurately predicts the relative cost of using the fixed DS or fixed TSP schemes. Hence, our procedure can be used as an aid in higher-level decisions, as discussed at the end of this paper.
- If one can dynamically choose between the DS and TSP options, we conjecture that the most general form of the solution is a triple-threshold policy characterized by $w_1 \leqslant w_2 \leqslant w_3$: if the total retailer inventory is less than w_1, then DS is preferred; if it is in the interval $[w_1, w_2)$, then TSP is preferred; and if it is in the interval $[w_2, w_3)$, where w_3 is the idling threshold, then

DS is preferred. Our rationale is as follows: if the absolute value of the total retailer inventory is large, then the routing scheme may have little effect on the rate at which inventory costs are incurred. In these cases, DS may be preferable because it incurs smaller transportation costs. In addition, the efficiency of DS has a tendency to increase the total inventory level relative to TSP (in the diffusion control problem, the DS option has a larger drift than the TSP option), and so DS will be even more attractive when the total inventory is much less than zero, because it will help to decrease future backorders. However, when the total inventory is in the interval $[w_1, w_2)$, which should contain zero in the nondegenerate case, the frequent deliveries of TSP lead to less backorders and smaller inventory costs, making it the more attractive alternative. Finally, because the effective penalty for using the inefficient TSP policy decreases when the total inventory is large, we believe that, in most cases, the optimal solution will be no more complex than a double threshold policy, where $w_2 = w_3$. This state of affairs is somewhat analogous to the stochastic economic lot scheduling problem with setup times analyzed in Markowitz, Reiman and Wein (1995), where large (small) lot sizes correspond to DS (TSP).

- We computed the numerical solution to the diffusion control problem corresponding to the dynamic IRP for a number of instances, and the results were consistent with our conjectures: the most general optimal policy was of the triple-threshold form, and, in most cases, a degenerate form of the policy was optimal: either the fixed DS case ($w_1 = w_2 = -\infty$ or $w_1 = w_2 = w_3$), the fixed TSP case ($w_1 = -\infty$, $w_2 = w_3$) or the double-threshold policy ($w_2 = w_3$). Moreover, in our limited simulation experiments, we did not find a numerically computed triple threshold policy that outperformed the analytically derived double-threshold policy (although we did not search beyond the computed triple-threshold values). By performing many exhaustive searches using simulation, we also found that the best double-threshold policy differed from the better of the two fixed routing policies in only a narrow range of system parameter space; in this range, the best TSP and DS policies achieve fairly similar performance. Hence, coupling this observation with a previous one suggests that finding the best fixed route policy is very important, whereas allowing for dynamic routing provides a much less substantial benefit; this is particularly true in light of the increased complexity of implementing a dynamic routing scheme.

In summary, the important operational levers for the IRP include the aggregate base stock level, the dynamic allocation policy and the choice of fixed routing scheme, but not the dynamic routing policy. Moreover, these key decisions are interrelated and a unified stochastic control model, such as the one considered here, is required for achieving reliable system performance.

Two topics for future research naturally come to mind. The first is to extend the dynamic routing scheme to allow K different types of routes (where $K > 2$) and/or to consider cyclic routes that use a combination of DS and TSP (e.g., a cycle could consist of a TSP tour through retailers 1, 2, and 3, followed by a direct shipment to retailer 2). Although, in theory, these extensions could be incorporated and system improvements could be achieved, the analysis would be tedious.

Another area for future research would be to develop the necessary steps for a hierarchical approach to the general (multi-vehicle, multi-warehouse) IRP; such an approach would be similar in spirit to the vehicle routing analysis performed by Simchi-Levi (1992), but would also incorporate the inventory cost component. Our results for fixed route policies provide estimates for the operating cost for any system given a particular assignment of retailers to vehicles and vehicles to warehouses. Motivated by our observation that the best fixed route policy performs nearly as well as the best dynamic policy over a broad range of parameters, the first level up in the hierarchy could implement an optimization algorithm (e.g., a k-opt algorithm as used in the deterministic vehicle routing literature) to find the best such route. Higher levels in the hierarchy could then be used to select the best possible assignment of vehicles and retailers, and the total number of vehicles to have in the system. At an even higher level, these results could be used to decide on the number and location of warehouses.

ACKNOWLEDGMENT

THIS RESEARCH WAS supported by a grant from the Leaders for Manufacturing Program at MIT, National Science Foundation grant DDM-9057297, and Engineering and Physical Sciences Research Council grant GR/J71786. The last author thanks the Statistical Laboratory at the University of Cambridge for its hospitality while some of this work was carried out. We are grateful to the referees for providing comments that improved the paper.

REFERENCES

R. L. ACKOFF, "OR, A Post Mortem," *Opns. Res.* **35,** 471–474 (1987).

W. BELL, ET AL., "Improving the Distribution of Industrial Gases with an On-Line Computerized Routing and Scheduling Optimizer," *Interfaces.* **13,** 4–23 (1983).

D. J. BERTSIMAS AND D. SIMCHI-LEVI, "A New Generation of Vehicle Routing Research: Robust Algorithms, Addressing Uncertainty," *Opns. Res.* **44,** 286–304 (1996).

P. BILLINGSLEY, *Convergence of Probability Measures.* John Wiley and Sons, New York, 1968.

L. M. A. CHAN, A. FEDERGRUEN, AND D. SIMCHI-LEVI, Probabilistic Analyses and Practical Algorithms for Inventory-Routing Models, *Opns. Res.* **46,** 96–106 (1998).

E. G. COFFMAN, JR., A. A. PUHALSKII, AND M. I. REIMAN, "Polling Systems with Zero Switchover Times: A Heavy-Traffic Averaging Principle," *Ann. Appl. Probab.* **5,** 681–719 (1995).

E. G. COFFMAN, A. A. PUHALSKII, AND M. I. REIMAN, "Polling Systems in Heavy Traffic: A Bessel Process Limit," *Math. Opns. Res.* **23,** 257–304 (1998).

M. DROR AND M. BALL, "Inventory Routing: Reduction from an Annual to Short Period Problem," *Naval Res. Log. Quart.* **34,** 891–905 (1987).

A. FEDERGRUEN AND Z. KATALAN, "The Stochastic Economic Lot Scheduling Problem: Cyclic Base-Stock Policies with Idle Times," *Management Sci.* **42,** 783–796 (1996).

A. FEDERGRUEN AND D. SIMCHI-LEVI, "Analytical Analysis of Vehicle Routing and Inventory-Routing Problems," in *Handbooks in Operations Research and Management Science*, Networks and Distribution, M. Ball, T. Magnanti, C. Monma and G. Nemhauser, (eds.), 1992.

A. FEDERGRUEN AND P. ZIPKIN, "A Combined Vehicle Routing and Inventory Allocation Problem," *Opns. Res.* **32,** 1019–1037 (1984).

B. GOLDEN, A. ASSAD, AND R. DAHL, "Analysis of a Large Scale Vehicle Routing Problem with an Inventory Component," *Large Scale Systems.* **7,** 181–190 (1984).

B. L. GOLDEN AND A. A. ASSAD (Editors), *Vehicle Routing: Methods and Studies*, North-Holland Publishers, 1988.

J. M. HARRISON, "Brownian Models of Queueing Networks with Heterogeneous Customer Populations," in *Stochastic Differential Systems, Stochastic Control Theory and Applications*, IMA Vol. 10, W. Fleming and P. L. Lions (eds.), Springer-Verlag, New York, 147–186, 1988.

J. M. HARRISON, *Brownian Motion and Stochastic Flow Systems*. John Wiley and Sons, New York, 1985.

D. L. IGLEHART AND W. WHITT, "Multiple Channel Queues in Heavy Traffic I," *Adv. Appl. Probab.* **2,** 150–177 (1970).

A. KUMAR, L. B. SCHWARZ, AND J. E. WARD, "Risk-Pooling Along a Fixed Delivery Route Using a Dynamic Inventory-Allocation Policy," *Management Sci.* **41,** 344–362 (1995).

H. J. KUSHNER, *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*, Academic Press, New York, 1977.

H. J. KUSHNER AND P. G. DUPUIS, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag. New York, 1992.

R. C. LARSON, "Transporting Sludge to the 106-Mile Site: An Inventory/Routing Model for Fleet Sizing and Logistics System Design," *Transp. Sci.* **22,** 186–198 (1988).

D. M. MARKOWITZ, M. I. REIMAN, AND L. M. WEIN, The Stochastic Economic Lot Scheduling Problem: Heavy Traffic Analysis of Dynamic Cyclic Policies, To appear in *Opns. Res.* 1999.

A. S. MINKOFF, "A Markov Decision Model and Decomposition Heuristic for Dynamic Vehicle Dispatching," *Opns. Res.* **41,** 77–90 (1993).

W. P. PETERSON, "A Heavy Traffic Limit Theorem for Networks of Queues with Multiple Customer Types," *Math. Opns. Res.* **16,** 90–118 (1991).

M. I. REIMAN, "Open Queueing Networks in Heavy Traffic," *Math. Opns. Res.* **9,** 441–458 (1984).

M. I. REIMAN, R. RUBIO, AND L. M. WEIN, Heavy Traffic Analysis of the Dynamic Stochastic Inventory-Routing Problem, Working Paper, Sloan School of Management, MIT, Cambridge, MA, 1997.

R. RUBIO, Dynamic-Stochastic Vehicle Routing and Inventory Problem, Ph.D. thesis, Operations Research Center, MIT, Cambridge, MA, 1995.

D. SIMCHI-LEVI, "Hierarchical Planning for Probabilistic Distribution Systems in Euclidean Spaces," *Management Sci.* **38,** 198–211 (1992).

P. TRUDEAU AND M. DROR, "Stochastic Inventory Routing: Route Design with Stockouts and Route Failures," *Transp. Sci.* **26,** 171–184 (1992).

L. M. WEIN, "Dynamic Scheduling of a Multiclass Make-to-Stock Queue," *Opns. Res.* **40,** 724–735 (1992).

(Received: February 1996; revisions received: April 1997, June 1998; accepted: July 1998)