Dynamic-Stochastic Vehicle Routing and Inventory Problem

by

Rodrigo Rubio Maqueo

Electrical-Mechanical Engineer, UNAM (1989)

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1995

Author ....
Sloan School of Management
August 16, 1995

Certified by ....
Lawrence M. Wein
Professor of Management Science
Thesis Supervisor

Accepted by ....
Thomas L. Magnanti
Co-director, Operations Research Center

Dynamic-Stochastic Vehicle Routing and Inventory Problem

by

Rodrigo Rubio Maqueo

Submitted to the Sloan School of Management
on August 16, 1995, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The vehicle routing and inventory problem (VRIP) is concerned with the joint decision, relevant to many distribution systems, of how much inventory to keep at each retail site and how to route the deliveries to these locations. This thesis presents the analysis of two types of dynamic control for a VRIP system subject to stochastic demand and travel times. The system consists of a central depot where an infinite supply of a standard product is kept at no cost. A single finite-capacity truck is used to deliver the product to m geographically dispersed retailers. Demand is served from these sites in a make-to-stock fashion. Costs are incurred for holding/backordering inventory and for operating the truck. The objective is to minimize the long-run average cost per unit time.

For the first problem, we choose to concentrate on the inventory aspect of the VRIP and consider the case where the truck travels along a predefined sequence of trips. We analyze two such sequences:(1) a tour of all retailers (TSP); (2) a fixed sequence of full-load direct shipments (DS). The second problem incorporates the routing aspects of the VRIP by allowing for the dynamic choice among the TSP and DS routing schemes. In each case we assume that the system operates under specific heavy traffic conditions, and obtain a decomposition of the problem into a non-linear program for the delivery allocation and a diffusion control problem.

For the fixed route cases, our results fully characterize a dynamic control policy that is asymptotically optimal. The limiting diffusion control problem for the dynamic routing case is considerably more difficult and does not lend itself to a general solution. We solve this latter problem for a specific double threshold policy. For all cases, numerical results show that the heavy traffic approximation is quite accurate over a broad range of system parameters. In addition, considerable insight is gained into the nature of the VRIP. In particular, our results indicate a high sensitivity of the system performance to the dynamic-stochastic behavior of the system.

Thesis Supervisor: Lawrence M. Wein
Title: Professor of Management Science

Acknowledgments

It seems as if writing these few lines will be the next to last thing I will do as a student at MIT (I still need to run to the printer and then rush the thesis to the Sloan School in time for my deadline). Some how it is harder than I expected. But there are good reasons for that. These have been four extraordinary years, mostly because during my time at MIT I have been fortunate to have the friendship and guidance of some extraordinary people. There are many names that come to me, and many more that I am sure I will forget. So I will relax the non-aggravation constraint and ask for your forgiveness if you do not find your name explicitly mentioned here, it probably should be if you have gotten this far.

I am deeply indebted to my advisor, Larry Wein. If I am here today doing heavy traffic analysis it is thanks to his being a constant source of wisdom and encouragement during my four years as his student. It has been a privilege to work with him. It is an honor to become his colleague.

I am grateful to my other committee members Dimitris Bertsimas and Vien Nguyen. Not only for valuable comments which helped improve this work, but also for going beyond the call of duty in accommodating the tight scheduling constraints for the final stretch of this project. I also want to thank Marty Reiman whose help proved invaluable in getting all the heavy traffic details right.

It would be hard to find a better place to have done this work than at the ORC. The support and friendship from each of you have made all the difference. Thank you Paulette, Laura and Cheryl for keeping this place running. Special thanks to my comrades Beril, Dave and Stefanos, I have learned a lot from being in a team with you. Thanks also to Edie, Alan, Sarah, Elaine, Rob, Eriko, Rich, Mitch, Kerry, Guy, Joe, Ram, Jim, Leon, Keely, Cristina,... I have so many more of you to thank that I should probably just include the ORC phone list as Appendix C. But I think I'll

rather take it with me in hope that we will stay in touch!

Dedication:

If I have made it here, it is in great part because of the support and love of my family. Though I have been so far away, you were always with me. I specially want to thank my Mother Patricia to whose courage and joy of life I dedicate this thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation and Objectives

On any given day, millions of customers buy a myriad different mass-produced goods from their local retailers. Whether it be individuals looking for their favorite snack, or industrial enterprises shopping for the raw materials required for their endeavor, they all take it as a given fact of today's modern life that they will find the item they seek in the required quantity and at the right price. Consequently, in an economic environment of fierce global competition, no company can afford the aggravation caused to the customer by frequent stock-outs nor the carrying cost of consistently high levels of inventory. Neither can the required availability be achieved at the expense of an inefficient transportation of the goods to the retail sites. Rather, to adequately address these requirements, corporations must jointly decide how to route their vehicles to reach the sites from which the customers are served, and how much inventory to keep at each of these places. The tradeoff between these competing inventory and transportation costs in the distribution chain is the essence of the vehicle routing and inventory problem (VRIP).

Operations Research practitioners have long been interested in gaining insight into the proper design and operation of distribution systems. Most of the approaches to date have concentrated on the routing aspects of the problem. The basic problem in this area is the standard vehicle routing problem (VRP) which consists of finding a set of vehicle routes that start and end at a central depot in order to minimize the travel cost while satisfying capacity constraints and meeting customer demands. The customer requirements and travel costs are assumed to be exogenous deterministic quantities. Since it includes the TSP as a subproblem, the VRP is NP-hard. It also possesses a very particular structure and many heuristic optimization approaches have been devised to exploit this. There is a huge literature on the VRP (and several extensions) and, from an applications point of view, it is one of the big success stories in the field. For an excellent sample of the results in this area, see Golden and Assad (1988).

However, in the make-to-stock context of the VRIP, customer demand will often be subject to considerable stochastic variation. Correspondingly, a direct extension of the deterministic models in the VRP literature would provide the wrong answer in terms of the inventory level required at each retailer (since it would fail to prevent backorders due to surges in demand). Though at a lesser extent than the demand, the vehicle travel times are also subject to random variation, thus adding to the stochasticity of the system and reducing the accuracy of an answer found by a deterministic model. The importance of this discrepancy depends, of course, on the relative magnitude of the inventory and transportation costs. In this thesis we take the philosophical stand that, in many instances, the inventory cost component will be very important and will in fact dominate the transportation costs. Therefore a dynamic stochastic model is not only appropriate but necessary for the VRIP.

Accordingly, in this thesis we will formulate a dynamic stochastic VRIP, and then use heavy traffic results to solve it with the following objectives: (1) to gain insights into the solution to the general VRIP, and (2) to develop effective policies for the operation of these systems.

The particular instance of the VRIP that we shall consider could represent the challenge faced by a regional branch of a large oil company as it distributes gasoline to its various gas stations. In general, we shall have a system which operates in a make-to-stock fashion as follows: a single central warehouse, which holds an infinite amount of inventory of a particular item (gasoline) at no cost, serves a set of retailers (stations); randomly arriving customers (cars) consume the product at these retail sites; a fleet of finite capacity vehicles (tanker trucks) is used to transport the product from the warehouse to the various retailers; traveling times along the route being random.

The management decisions involved in the set-up 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 (see e.g. Simchi-Levi (1992)). At the top strategic level, the managers of this system face decisions such as how many warehouses and retailers to have and where to locate them, as well as how to assign retailers to warehouses. At a tactical level, the managers face the decision of how many trucks to operate, and possibly of how to assign retailers to a particular truck or subset of trucks serving a certain district. At the operational level the decisions are: when to send each truck out (as opposed to keeping it idle at the warehouse), how much of the capacity of the truck to use (i.e. should the trucks always leave the warehouse full?), which of the retailers should each truck visit, and how much of its load should a truck deliver to each of the retailers on its route. The objective is to identify the set of decisions that will maximize the profitability of the operation. Our treatment of the VRIP will be concerned with the operational aspects of this general problem in a dynamic-stochastic setting.

We shall hence assume that somewhere higher up in the hierarchy, the decision has been taken to assign a single truck (with fixed capacity and operating cost) to serve all retailers in a particular region. This is an assumption often made in the literature (e.g. Kumar, et al. (1994) and Anily and Federgruen (1990)) and is often the case in practice. For this case we model the demand as arbitrary stochastic processes (that satisfy a functional central limit theorem), and the travel times as arbitrary random variables (independent of the demand). We shall be concerned with the minimization of the total steady-state operating cost per unit time. The main operating costs are inventory holding and backordering costs at the retailers, and travel costs for the trucks (including driver and fuel costs).

In our initial analysis we shall take a further step down the hierarchy an assume that a fixed routing policy has been selected for the system. Such a model will allow us to concentrate on the inventory component of the steady state cost. Also, this is a necessary step towards the analysis of the full VRIP (i.e. tactical and strategic levels), since if we are to select the best routing scheme we must be able to characterize the performance of the system under an arbitrary fixed route.

1.2 Literature Review

While extensions of the deterministic VRP literature have been developed to consider the inventory aspects of the problem (see Federgruen and Simchi-Levi (1992) for a detailed literature review), little attention has been given to the VRIP with random demand. In the first such effort of which we are aware, Federgruen and Zipkin (1984) consider the single-period problem with stochastic demand. Since the expected inventory holding and backordering cost are calculable for this problem instance, they construct a (non-linear) mathematical programming formulation for the problem, and observe that, once a fixed assignment of delivery points (retailers) to routes (trucks) is made, the problem decomposes into an inventory allocation problem and one TSP problem per truck. Based on this decomposition approach, they apply some interexchange optimization heuristics from the traditional VRP (such as 2-opt) to find near optimal solutions to their case. Our approach differs from theirs (and those in the traditional VRP literature) in several ways: (1) We are concerned with a steady state analysis and so model the stochastic variation of the demand over an infinite horizon and not just over one period; (2) we allow for the travel times between sites in the system to vary randomly, and correspondingly model the transportation cost as an expense incurred per unit time the truck is busy; (3) we formally constraint the system capacity by having only one truck available. This contrasts with the model in Federgruen and Zipkin (1984), which assumes that trucks will cover their assigned routes with certainty. While the authors mention that their decomposition approach still works if one adds a total travel time constraint to the truck routing subproblem, this is still fundamentally different from the approach in this thesis due to the different planning horizons considered. As readers familiar with queueing theory will doubtlessly expect (and as our results will show), the system steady state performance is highly sensitive to the relative congestion in the system (i.e. the long run fraction of time that the truck must remain busy to satisfy the average requirements at its assigned retailers). This is not necessarily the case in the single period model where even a heavily congested system will only seem to experience high levels of backorders once (and not as an ordinary phenomenon that calls for a high base stock level).

More recently Chan, et al. (1994) have done a probabilistic analysis of the VRIP: that is, they consider particular policies for the deterministic problem, and then analyze their average behavior when instances of the deterministic problem are generated by sampling from probability distributions for the retailer location and demand. This work does not explicitly consider the stochastic nature of the demand since, once each instance is generated the model makes a delivery plan based on the future orders. Also, their model assumes that there is no limit on the total distance that any truck can travel, and therefore ignores the congestion effects in the system. As a consequence, the possibility of unfilled demand (that is either lost or backordered at a penalty) is not considered.

Perhaps the most closely related paper in the literature to the work presented here is Kumar et al. (1994). Their work is concerned with the comparison of fixed versus dynamic inventory allocations along a fixed delivery route: They assume that per-period demand at each retailer is given by iid normal random variables, that identical linear holding and backordering costs are incurred at all the retailers, and that travel times along the route are deterministic. A central depot acts as a transshipment point where no inventory is kept but where system-wide replenishment orders are placed periodically and received after a fixed lead time (for immediate shipment to the retailers). The authors observe that the optimal system-wide replenishment policy will depend on the rule used to allocate the truck cargo to the retailers. They solve both the static (allocation fixed at the depot) and dynamic (allocation for the remaining sites in the route reviewed at every retailer) cases under specific allocation assumptions (i.e. they solve the problem when the requirement that the allocation be positive and less than the current load is relaxed), and find that optimal replenishment policies are base stock policies, that the optimal base stock level can be found from a normalized composite retailer, and that the dynamic allocation policy will always perform better than its static counterpart. Like these authors, we shall initially consider instances of the VRIP where the truck is restricted to a predetermined routing scheme, thereby concentrating on the inventory aspects of the problem. However, unlike them we shall explicitly consider that a finite capacity single truck provides the required transportation service (they implicitly take both the number of trucks available and their capacities to be infinite, since neither appears as a constraint in their problem formulations). The vehicle characteristics (capacity and speed) are a key component of the system utilization level which, as mentioned before, plays an important role in determining the inventory cost. Furthermore, explicit consideration of the truck size is required to obtain a steady state transportation cost, which is necessary for us in order to later incorporate the routing aspects of the problem. Also, our assumptions in terms of the stochastic nature of the system shall be much less onerous; both their demand distribution and allocation assumptions do not seem realistic for systems with more than just moderate demand coefficients of variation.

1.3 Organization of the Thesis

The remainder of this thesis is organized as follows. Chapters 2 and 3 present the analyses of two fixed route instances of the VRIP: one where, on all trips, the truck visits every retailer assigned to it in a fixed sequence; and another where direct full-load shipments are made to the retailers. In both cases, we present a general dynamic-stochastic formulation for the particular instance of the VRIP. As these problems appear to be intractable, we use the machinery of heavy traffic analysis to solve a limiting control problem. The problem is made tractable by the decomposition and state-space collapse allowed by the heavy traffic averaging principle, as established in Coffman, Puhalskii and Reiman (1993). Under the heavy traffic regime, we find explicit state-dependent solutions for the optimal inventory allocation among the retailers and the system-wide base stock level. A proof of the optimality of base stock control for system-wide replenishment is included as an appendix.

Chapter 4 presents the main results for the fixed-route problems. The first part of the chapter is concerned with interpreting the heavy traffic solutions in terms of the original system, an interesting and often quite challenging task. These results are then used to compare the steady state performance of the system under the two routing schemes considered. One outcome of these analyses is that direct shipping is expected to dominate system-wide tours in many instances. This is an interesting revelation, since most of the literature so far has been concerned with the difficult task of finding optimal (or near optimal) tours. Finally, a simulation study is performed to analyze the performance of the proposed control policy over a range of values for the system utilization. The results are very good over a broad range of system parameter values, even those that are far from the heavy traffic regime assumed in the analytical work underlying the proposed policy.

Chapter 5 extends the model to allow for dynamic choice among the two routing schemes analyzed before. Again, we present a dynamic-stochastic general formulation for this problem which appears to be intractable. Suitable extensions of the heavy traffic analyses performed for the fixed route cases allow us to obtain a limiting control problem formulation. While easily solved numerically, this problem does not allow for a closed form solution for the routing control decision. Numerical experimentation and a simulation study are used to gage the performance of the proposed policy. As in the static cases, the policy obtained from the heavy traffic analysis performs quite well, even when the system parameters have values that do not correspond to the heavy traffic regime. Also, as one would expect, overall system performance is improved when dynamic routing is allowed (compared to the fixed route systems), but in most cases the improvement is of moderate size. In fact, there seems to be a rather narrow band (in system parameter space) over which the dynamic policy is different from either of the two static policies. Over most of the space of parameter values, the selected dynamic policy is to repeatedly use the best fixed routing scheme. These results support our position that finding the appropriate dynamic inventory control is probably more important than finding the best dynamic route selection policy.

Finally, Chapter 6 presents a summary of the main results obtained in this thesis and a discussion of some direct extensions. We place special emphasis on the insights gained into the VRIP and propose an agenda for further research on this area.

Chapter 2

Heavy Traffic Analysis of a Fixed Route VRIP: The Traveling Salesman Tour Case

2.1 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 depot at no cost. Customers are served from the retailers in a make-to-stock fashion, and demand that cannot be served immediately due to lack of inventory at the corresponding retailer is backordered. When the truck is operating the following policy is used: the truck leaves the warehouse (which we shall index as station 0) with a full load and then visits all the retailers in a predefined sequence before returning empty. Alternatively, the truck may idle at the depot. Though the order in which retailers are visited could be arbitrary, we will assume that it is the solution to the implied Traveling Salesman Problem, and refer henceforth to this service scheme as the fixed-route TSP policy. In this same order, it will prove advantageous in the forthcoming analyses to index the retailers from 1 through m according to their position in the TSP tour. The system just described is

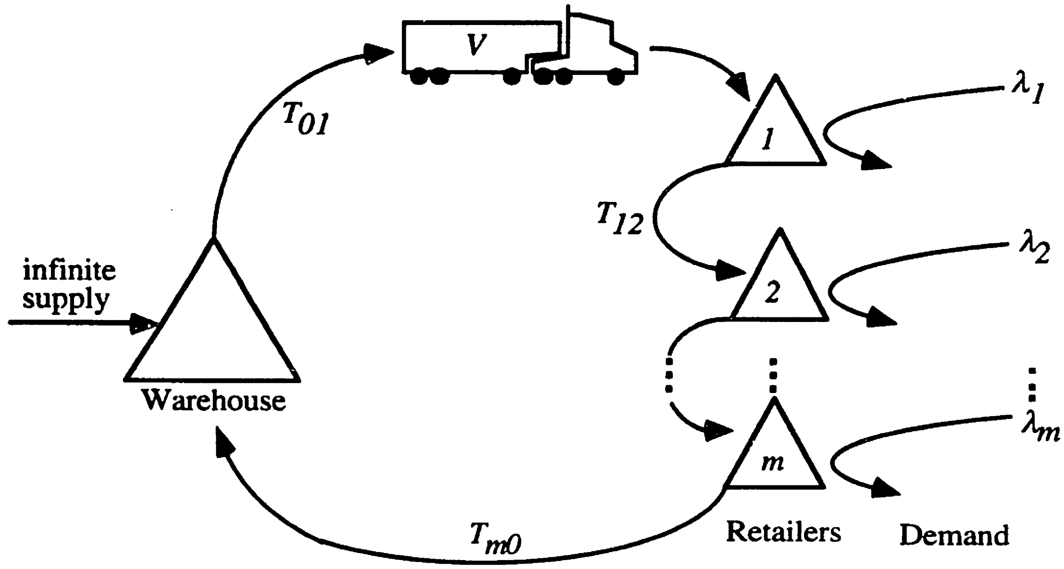

Figure 2-1: The Fixed-Route TSP VRIP.

schematically represented in Figure 2-1.

We shall consider two sources of uncertainty in this problem: demand arrivals and travel times. The arrival of customer demand at retailer i, $i = 1, \ldots, m$ is assumed to be an independent renewal process with rate λ_i and squared coefficient of variation c_{di}^2 (variance divided by the square of the mean). The sequence of travel times between facilities i and j is is assumed to be given by iid samples of the random variable T_{ij}, which has mean θ_{ij} and squared coefficient of variation c_{ij}^2 (the indexes i, j run from 0 to m). These travel times are independent of the demand arrival streams and of each other. Keeping with the convention in the literature, we will assume that all load (at the warehouse) and unload (at the retailers) times are zero; in practice, these times tend to be dwarfed by the travel times. Hence, we can obtain θ_T and c_T^2, the mean and squared coefficient of variation respectively of the total time required to complete the TSP tour by

$$\theta_T = \sum_{j=0}^{m-1} \theta_{j,j+1} + \theta_{m0}$$

$$c_T^2 = \frac{\sigma_{TS}^2}{\theta_T^2} = \frac{\sum_{j=0}^{m-1} \theta_{j,j+1}^2 c_{j,j+1}^2 + \theta_{m0}^2 c_{m0}^2}{\theta_T^2}.$$

As illustrated by these calculations, the assumption of uncorrelated demand and travel times clearly simplifies the process of obtaining moments for sums of the basic random variables, and thus is useful in keeping the exposition simple. The assumption however is not necessary and our results generalize to cases with correlated compound renewal processes; see §6 of Reiman (1984) for details.

At the operational level we shall be concerned with two types of costs for this system: transportation costs, and inventory holding and backordering costs. The travel cost rate per unit time, which includes vehicle depreciation, fuel and driver cost, is f. Note that these costs can be combined because we are ignoring the load/unload times (only the driver, but not the truck is busy while loading and unloading). Inventory costs are assumed to be piece-wise linear, with the holding cost rate (per unit in inventory per unit time) at retailer i being denoted by h_i and the backorder cost rate by b_i.

Following the approach in Harrison (1988), we characterize the state of the system as the vector process $Q_i(t)$, the number of units in inventory (or backordered if this quantity is negative) at retailer i, so that the total inventory at the retailers is $Q(t) = \sum_i Q_i(t)$ (note: in this and all summations of this thesis the index runs over the set of retailers $\{1, 2, \ldots, m\}$, except when explicitly indicated otherwise). Also, we take as primitives for the system the following stochastic processes: $D_i(t)$, which denotes the cumulative demand at retailer i, so that $D(t) = \sum_i D_i(t)$ represents the total cumulative demand in $[0, t]$; and $S(t)$, the counting process for TSP tour completions up to time t assuming the truck is *continuously active* in $[0, t]$.

Once the route is fixed, two control decisions remain for the operation of the system: (1) when the truck is back at the warehouse, we must decide whether to send it out immediately with a new load or to let it idle; (2) as the truck visits each retailer, we have to determine how much of the load to leave there. In this order, express the busy/idle control in terms of the cumulative process $T(t)$, which represents the amount of time the delivery truck is active in $[0, t]$. The sequence $\tau_k,\ \ k = 1, 2, \ldots$ of tour completion epochs (i.e. the times at which the truck returns to the warehouse after visiting the retailers) will therefore depend on this control, and will be given by

$$\tau_k = \inf\{t \mid S(T(t)) \geq k\}.$$

In terms of the delivery size control, let the m-dimensional vector of processes $L_i(t)$ represent 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 $\varepsilon_i(t)$. We set the nominal deliveries to the amount necessary to maintain material flow balance over the long run. That is, we let

$$V_i = \frac{\lambda_i}{\lambda} V \text{ for all } i, \tag{2.1}$$

where $\lambda = \sum_i \lambda_i$, the total demand arrival rate. So that the nominal allocation is to split the vehicle capacity V among the retailers according to their demand share. The load allocation process is then defined as

$$\varepsilon_i(t) = L_i(t) - V_i S(T(t)). \tag{2.2}$$

Notice that, since we can observe the tour completion history, we need only specify the process $\varepsilon_i(t)$ to determine the total deliveries to the retailer up to time t. As illustrated in Figure 2-2, the size of the upward jump of process $\varepsilon_i(t)$ at the epoch when retailer i is visited determines the size of the delivery to make. The value of process $\varepsilon_i(t)$ is then always reduced by the nominal load allocation V_i at every tour completion epoch τ_k. The m delivery controls $\varepsilon_i(t)$ may equivalently be taken to represent the cumulative deviations from the nominal delivery size over past tours, plus the amount delivered during the current cycle for retailer i. In the proposed problem set-up, this control cannot be exercised without limit. Recall that, by assumption, the truck leaves the warehouse with a full load and returns empty to the depot, having delivered all of its cargo at the retailers. To guarantee this, the delivery control must satisfy

$$\sum_i L_i(\tau_k) = kVS(T(t))$$

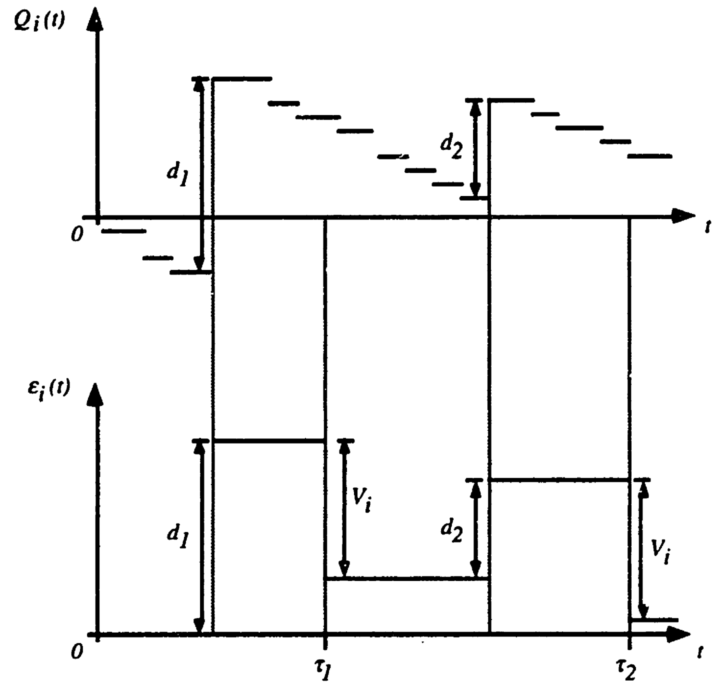

Figure 2-2: Joint Evolution of Inventory and Load Allocation Control at Retailer i.

$$L_i(t) \qquad \text{is non-decreasing with } L_i(0) = 0$$

which, in terms of the vector of dynamic load allocations, implies that

$$\varepsilon_i(0) = 0, \text{ for all } i \tag{2.3}$$

$$\varepsilon_i(t) \geq \varepsilon_i(\tau_{k-1}), \text{ for } t \in (\tau_{k-1}, \tau_k) \text{ and all } i \tag{2.4}$$

$$\sum_i \varepsilon_i(\tau_k^-) = V, \tag{2.5}$$

$$\sum_i \varepsilon_i(\tau_k) = 0, \tag{2.6}$$

where τ_k is the time of the k-th tour completion, and τ_k^- is the epoch an infinitesimal amount of time before this completion. Therefore, deviations from the nominal allocation will cancel out across the retailers and the process $\varepsilon(t) = \sum_i \varepsilon_i(t)$ represents the total amount delivered during the current cycle.

We are finally in a position to state the equations that govern the dynamic behavior of the system in terms of the primitive processes and the controls. By the above definitions, assuming that $Q_i(0) = 0$, we have that (see Figure 2-2)

$$Q_i(t) = V_i S(T(t)) - D_i(t) + \varepsilon_i(t) \quad \text{for} \quad i = 1, \ldots, m, \quad t \geq 0. \tag{2.7}$$

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

$$I(t) = t - T(t) \quad \text{for} \quad t \geq 0 \tag{2.8}$$

so that the control policy $T(t), \varepsilon_i(t)$ must satisfy

$$T, \varepsilon_i \quad \text{are nonanticipating with respect to } Q \tag{2.9}$$

$$T \quad \text{is nondecreasing and continuous with} \quad T(0) = 0 \tag{2.10}$$

$$I \quad \text{is nondecreasing with} \quad I(0) = 0. \tag{2.11}$$

The constraint in (2.9) formalizes the stochastic nature of the system by precluding the development of controls that require knowledge of the future state of the system. For their part, constraints (2.10) and (2.11) formalize the truck capacity constraint.

Recall that travel costs are incurred whenever the truck is busy. Equivalently, we can consider the travel cost rate as a *reward* for exerting idleness (i.e., as a reward for exerting the control $I(t)$, as opposed to a cost for the use of control $T(t)$). Hence the problem reduces to finding a control policy $(T(t), \varepsilon_i(t))$ to

$$\min \limsup_{T \to \infty} \frac{1}{T} E\left[\int_0^T \sum_i (h_i\{Q_i(t)\}^+ + b_i\{Q_i(t)\}^-)\, dt - fI(T)\right] \tag{2.12}$$

subject to (2.3) - (2.11).

The dynamic-stochastic VRIP, as formulated in (2.3) - (2.12), does not seem to be tractable. Even under memoryless (i.e. exponential) assumptions for the underlying random processes the state space becomes unmanageably large. One reason for this is that we need to consider $m + 1$ dimensions: the inventory/backorder level at each retailer plus the position of the truck along the tour. Furthermore, the inventory levels at each retailer are not independent, they are inter-related by the load allocation policy since the amount delivered over the nominal allocation at one retailer must represent a reduction with respect to the nominal delivery for at least one other site. In cases like this, one naturally resorts to some approximate method in order to gain further understanding of the problem. In this thesis we shall solve a control for a limiting system obtained under the regime of heavy traffic. The solution to this approximation will provide insight for the control of the actual system.

2.2 Heavy Traffic Normalizations and Averaging Principle

2.2.1 Normalizations and Diffusion Limit for the System Netput

A tractable and coherent formulation for the VRIP may be found as the limit of a sequence of systems, which we index by n. We shall obtain a limiting control problem by letting $n \to \infty$ and making the appropriate normalizations for the parameters and processes of each system in the sequence. Readers familiar with the heavy traffic literature should be warned that, in order to adequately model the VRIP, we shall require a scaling which is different from most of the traditional work in the area. Having issued this warning, we shall initially proceed with the normalizations without a detailed justification of the particular scaling chosen. This is done in the interest of brevity, since the motivation of our choice will be much easier once we have obtained the desired limit.

To this order, take $V^{(n)}$ (notation note: the superscript (n) denotes the index for the sequence of systems, and is not an exponent), the vehicle capacity, and $(\theta_T^{(n)}, \sigma_{TS}^{2\,(n)})$, the mean and variance of the tour completion time in the n-th system, and make the following definitions:

$$C^{(n)} = \frac{V^{(n)}}{\sqrt{n}}, \tag{2.13}$$

$$\vartheta_T^{(n)} = \frac{\theta_T^{(n)}}{\sqrt{n}}, \tag{2.14}$$

$$\varsigma_{TS}^{2\,(n)} = \frac{\sigma_{TS}^{2\,(n)}}{\sqrt{n}}. \tag{2.15}$$

Also, consider the sequence of iid random variables $s_i^{(n)}$, which have mean $\vartheta_T^{(n)}$ and variance $\varsigma_{TS}^{2\,(n)}$ with its corresponding counting process

$$\mathcal{S}^{(n)}(t) = \min\{k \mid \sum_{i=0}^{k+1} s_i^{(n)} > t\}.$$

Notice that, in terms of these definitions, we can generate the service completion process for the n-th VRIP system, $S^{(n)}(t)$, by considering a tour completion to occur for every $\lceil\sqrt{n}\,\rceil$ arrivals to $\mathcal{S}^{(n)}(t)$. The tour completion times thus generated would have mean and variance $\theta_T^{(n)}\lceil\sqrt{n}\,\rceil/\sqrt{n}$ and $\sigma_{TS}^{2\,(n)}\lceil\sqrt{n}\,\rceil/\sqrt{n}$ respectively, which are asymptotically (as $n \to \infty$) indistinguishable from $\theta_T^{(n)}$ and $\sigma_{TS}^{2\,(n)}$. To highlight this relationship, we shall henceforth refer to $\mathcal{S}^{(n)}(t)$ as the partial tour completion process.

The next step towards the limiting VRIP system is to define centered versions of the partial service completion and demand arrival processes. This is done by subtracting the interarrival rate times the elapsed time from the corresponding counting process, which leads to the following definitions:

$$\bar{\mathcal{S}}^{(n)}(t) = \mathcal{S}^{(n)}(t) - \frac{t}{\vartheta_T^{(n)}},$$

and

$$\bar{D}^{(n)}(t) = D^{(n)}(t) - \lambda^{(n)}t.$$

With these constructs in hand, we now define the n-th system's netput process as follows

$$\chi^{(n)}(t) = \left(\frac{C^{(n)}}{\vartheta_T^{(n)}} - \lambda^{(n)}\right)t + C^{(n)}\bar{\mathcal{S}}^{(n)}(T^{(n)}(t)) - \bar{D}^{(n)}(t). \tag{2.16}$$

This process represents the total virtual input to the system minus the demand. We can now obtain an expression for the dynamics of the total inventory at the retailers in terms of the system netput by summing the inventory evolution equations in (2.7) over all retailers and substituting the relevant definitions. The expression is:

$$Q^{(n)}(t) = \chi^{(n)}(t) - \frac{V^{(n)}}{\theta_T^{(n)}} I^{(n)}(t) + \xi^{(n)}(t), \tag{2.17}$$

where

$$\xi^{(n)}(t) = \varepsilon^{(n)}(t) + V^{(n)} S^{(n)}(T^{(n)}(t)) - C^{(n)} \mathcal{S}^{(n)}(T^{(n)}(t)). \tag{2.18}$$

The process $\xi^{(n)}(t)$ captures the inventory evolution over the current cycle and makes the necessary adjustment from virtual deliveries to true ones. To see this, let us consider the $\varepsilon(t)$ term first. This is quite straightforward since, as a consequence of our assumption that every tour delivers a full load to the retailers, the load allocation control $\varepsilon(t)$ does not affect the evolution of the total inventory beyond the current tour. Now, by construction,

$$V^{(n)} S^{(n)}(T^{(n)}(\tau_k^{(n)})) = C^{(n)} \mathcal{S}^{(n)}(T^{(n)}(\tau_k^{(n)})) + o\left(\frac{1}{\sqrt{n}}\right)$$

at all service completion epochs $\tau_k^{(n)}$. The $o\,(1/\sqrt{n})$ error is due to the non-integrality of $\sqrt{n}$ and may be ignored in the limit. The difference

$$V^{(n)} S^{(n)}(T^{(n)}(t)) - C^{(n)} \mathcal{S}^{(n)}(T^{(n)}(t))$$

in (2.18) makes the necessary adjustment from the virtual deliveries during the current tour implied by the $C^{(n)}\bar{\mathcal{S}}(T^{(n)}(t))$ term in the netput process.

We next normalize the processes in each system in the sequence according to the following expressions:

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

$$W^{(n)}(t) = \sum_i W_i^{(n)}(t) = \frac{Q^{(n)}(nt)}{\sqrt{n}},$$

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

$$X^{(n)}(t) = \frac{\chi^{(n)}(nt)}{\sqrt{n}},$$

$$\hat{D}^{(n)}(t) = \frac{\bar{D}^{(n)}(nt)}{\sqrt{n}},$$

$$\hat{\xi}^{(n)}(t) = \frac{\xi^{(n)}(nt)}{\sqrt{n}},$$

$$\hat{S}^{(n)}(t) = \frac{\bar{S}^{(n)}(nt)}{\sqrt{n}},$$

and

$$\tau^{(n)}(t) = \frac{T^{(n)}(nt)}{n}.$$

The processes (W, Y, X) represent the normalized inventory, idleness and netput. We use different notation for them than for their unscaled counterparts (Q, I, χ) so that, in the interest of brevity, we may sometimes refer to them simply as inventory, idleness and netput without confusion.

The dynamics for the normalized n-th system are then found by applying these scalings into equations (2.16) and (2.17). With this procedure we obtain an expression for the normalized netput process $X^{(n)}(t)$ as

$$X^{(n)}(t) = \sqrt{n}\left(\frac{C^{(n)}}{\vartheta_T^{(n)}} - \lambda^{(n)}\right)t + C^{(n)}\hat{S}^{(n)}(\tau^{(n)}(t)) - \hat{D}^{(n)}(t), \qquad (2.19)$$

while the normalized inventory can be expressed as

$$W^{(n)}(t) = X^{(n)}(t) - \frac{V^{(n)}}{\theta_T^{(n)}}Y^{(n)}(t) + \hat{\xi}^{(n)}(t). \qquad (2.20)$$

The limit may now be found by letting the index n go to infinity, under the assumption that the parameters of the sequence of systems satisfy the following heavy traffic conditions:

$$\lim_{n\to\infty} C^{(n)} = C = O(1) \qquad (2.21)$$

$$\lim_{n\to\infty} \vartheta_T^{(n)} = \vartheta_T = O(1) \qquad (2.22)$$

$$\lim_{n\to\infty} \varsigma_{TS}^{2\,(n)} = \varsigma_{TS}^2 = O(1) \tag{2.23}$$

$$\lim_{n\to\infty} \lambda^{(n)} = \lambda \tag{2.24}$$

$$\lambda \vartheta_T = C \tag{2.25}$$

$$\sqrt{n}\left(\frac{V^{(n)}}{\theta_T^{(n)}} - \lambda^{(n)}\right) = \mu_T = O(1), \quad \text{for all } n. \tag{2.26}$$

If we denote the traffic intensity of the n-th system by $\rho_T^{(n)} = \lambda^{(n)}\theta_T^{(n)}/V^{(n)}$, the previous conditions require that this parameter approach unity at the right rate. As a consequence, we expect our model to be a reasonable approximation for systems where the truck must be busy the great majority of the time in order to meet the expected demand.

One aspect of our scaling which differs from that used in most heavy traffic applications is the way in which the original system parameters $(V, \theta_T, \sigma_{TS}^2)$ are being normalized according to the heavy traffic conditions (2.21) to (2.23). In the traditional heavy traffic literature, these quantities are left unscaled in the system sequence, so that as $n \to \infty$ the vehicle size and the tour travel times vanish in scaled inventory and time units. This is the case since, under the scalings used to define (W, Y, X), the state space is reduced by $\sqrt{n}$ and time is sped up by n. For this normalization, the size of the n-th system truck in terms of scaled inventory units is $V^{(n)}/\sqrt{n}$. Hence if V were left unscaled it would vanish in the limit. Under such an assumption our system would reduce to the multi-class single server queue case of Wein (1992). While we would thus be able to obtain a tractable limiting control problem for the VRIP, this would come at a considerable loss in the ability of the model to capture the behavior of the original system. In particular, a limit that obtains instantaneous cycles and infinitesimal truck sizes will not be able to model the inventory evolution during the course of a tour. Since in practice truck sizes and travel times are usually quite significant, this shortcoming would very likely reduce the applicability of our results.

Our scaling tries to overcome this problem in the simplest possible way. Namely, we have scaled the vehicle size by the minimum amount necessary to obtain a positive limit for it in terms of scaled inventory units (i.e. $V^{(n)}/\sqrt{n} = C^{(n)} = O(1)$). We cannot, however scale only the vehicle size. If we do not scale the arrival rate of the demand process, we must make sure that the ratio $V^{(n)}/\theta_T^{(n)}$ remains $O(1)$. This explains the scaling of the tour completion times. Finally, we also need to normalize the travel time variance. The proper scaling to use is of the same order of magnitude as for the truck size (i.e. $O(\sqrt{n})$). At first glance it might seem that this scaling reduces the variability in the system, since the the squared coefficient of variation for the tour completion times σ_T^2/θ_T^2 will be reduced by $\sqrt{n}$. Notice however that, in terms of the total inventory, increasing the truck size also increases the variability of the service process (i.e. deliveries are equal to tour completions times the vehicle size). Increasing the truck size and preserving the travel time coefficient of variation would yield a limiting system with infinite service variance.

Use of this scaling allows us to preserve the same level of variability in the system and obtain a positive truck size (in terms of normalized inventory units) in the limit, which corresponds to the actual behavior of the VRIP system. However, this scaling also preserves the discontinuous jumps in the inventory evolution at the delivery epochs. The introduction of the partial service completion process and its related netput process (another non-standard aspect of our scaling), allow us to work around these discontinuities. In particular, we can still invoke the Functional Central Limit Theorem (FCLT) and the Random Time Change Theorem (as in the standard heavy traffic analysis in Wein (1992) and Harrison (1988)) for equation (2.19) to obtain the following:

Proposition 1. *$X^{(n)}(t) \Rightarrow X(t)$, where $\Rightarrow$ denotes weak convergence and $X(t)$ is a Brownian motion process with drift $\mu_T > 0$ and variance*

$$\sigma_T^2 = \lambda\left(c_d^2 + C\frac{\varsigma_{TS}^2}{\vartheta_T^2}\right).$$

Notice that, both these quantities are $O(1)$ by assumption.

In terms of our original system, Proposition 1 is not enough to characterize the limiting behavior of the total inventory at the retailers. By equation (2.20) we still need to characterize the limiting control processes $Y(t)$ and $\hat{\xi}(t)$. Furthermore, in order to solve the problem at hand, we need to determine the behavior of the m individual inventory levels since the holding and backordering costs are driven by them. As it turns out, it is not possible to obtain a limit process for $(W_1, W_2, \ldots, W_m)$ in the usual sense because in the heavy traffic limit these levels move infinitely fast along a path that depends on the load allocation control. Similarly, the adjustment term $\hat{\xi}(t)$ does not converge to a limit in the usual sense, since it captures the discontinuous jumps from the actual deliveries over the tour. In order to overcome these difficulties we shall exploit a recent result in heavy traffic theory, the heavy traffic averaging principle (HTAP).

2.2.2 Limiting Behavior of The Inventory Processes: The Heavy Traffic Averaging Principle

The main aspect of the HTAP is a *time scale decomposition.* Recall that in the normalization used to obtain the Brownian motion (BM) limit for the total netput process time is sped up by a factor of n. We will henceforth refer to this as the *diffusion time scale.* According to the HTAP, at the diffusion time scale the m-dimensional retailer inventory process moves (asymptotically) infinitely fast. If instead of the diffusion scale one speeds up time only by $\sqrt{n}$ the m-dimensional inventory process moves at a positive and finite rate, while the netput remains fixed for an (asymptotically) infinite amount of time. Under this normalization the movement of the m-dimensional process is deterministic. Notice that this normalization essentially slows down time by a factor of $\sqrt{n}$ with respect to the diffusion time scale. We will henceforth refer to this case as the *deterministic time scale.* In both cases the state space is scaled down by $\sqrt{n}$, so that we may refer to normalized inventory levels without specifying whether the diffusion or the deterministic scaling is being used.

The HTAP was first established in Coffman, Puhalskii and Reiman (1993) (CPR) for polling systems with no switch-over times. Its far reaching implications for the control of systems in heavy traffic have been exploited in Reiman and Wein (1994) for the control of a polling system and in Markowitz, Reiman and Wein (1994) for the economic lot-sizing problem (ELSP). Our use of the HTAP in the VRIP follows the spirit of their work.

Providing a rigorous proof of the HTAP for our problem would be extremely demanding and take us far afield from our purpose of gaining insights into the VRIP. We will thus present only a heuristic argument of why an averaging principle should hold for the VRIP, and establish how it can be used for a remarkable simplification of our control problem.

Consider what happens in the unscaled VRIP system while the truck does a tour. The amount of time spent traveling along the route is $O(\sqrt{n})$ — by condition (2.22) — and the total amount delivered to the sites over the tour is $O(\sqrt{n})$ — by condition (2.21). From Proposition 1 of § 2.2.1, the total normalized netput behaves asymptotically like $X(t)$, a (μ_T, σ_T^2) Brownian motion. All three of these system characteristics are the same as in the polling system of the CPR paper, and hence we expect a similar time scale decomposition to apply to the VRIP.

The time scale decomposition alluded to in the previous paragraph becomes clearer if one considers the behavior of the normalized system over one tour in the diffusion time scale (i.e. speed up time by a factor of n and compress the inventory state space by $\sqrt{n}$). In terms of these normalized units, the amount of time spent doing a tour is $O(1/\sqrt{n})$ and the total amount delivered to the sites is $O(1)$. Meanwhile, the total normalized netput under this scaling behaves asymptotically like $X(t)$, a (μ_T, σ_T^2) Brownian motion. It is well known that the net change of a (μ_T, σ_T^2) Brownian motion over a time interval of length t is a normal random variable with mean $\mu_T t$ and variance $\sigma_T^2 t$. Therefore, over an interval of (normalized) time of length $O(1/\sqrt{n})$, the net change in the total netput will be $O(1/\sqrt{n})$ (in normalized inventory units). Furthermore, by definition, the remainder term $\hat{\xi}(t)$ vanishes at tour completion epochs, and so the change in total inventory itself (over successive completion epochs) will be $O(1/\sqrt{n})$.

This order-of-magnitude difference in the rate at which the total inventory and the individual inventories vary lies at the heart of the HTAP. By exercising the load allocation control, we can effectively make an $O(1)$ change in the individual inventories (by delivering more than the nominal allocation to one and less to the others), in an interval of normalized time that is infinitely small in the diffusion time scale. While this is accomplished, the total netput level will effectively remain the same (and so will the total inventory at tour completion epochs).

We may now characterize the limiting behavior of the inventory processes in the VRIP system. Recall from equation (2.20) that the total inventory in the system is equal to a Brownian motion (the netput process) regulated by an idleness control and superimposed with the tour delivery process $\hat{\xi}(t)$. As mentioned before, this last term is equal to zero at tour completion epochs and makes $O(1)$ jumps (in terms of scaled inventory units) as deliveries are made. Therefore, if we consider what happens in terms of normalized inventory units at tour completion epochs, total retailer inventory is (asymptotically) unchanged over successive cycles. However, we have seen that over this same time period we may produce an $O(1)$ change in the distribution of inventory among the retailers. It stands to reason that, since the total inventory is the same, the increase in the level of inventory at one of the retailers must (deterministically) correspond to a decrease in the level at some other retailer. This deterministic behavior of the individual inventory levels is one of the consequences of the HTAP for the VRIP. If the individual inventories behave deterministically over a tour, then the delivery process $\hat{\xi}(t)$ will also behave deterministically, the exact evolution being determined by the allocation control exercised over the tour.

In order to understand the implications of these consequences of HTAP for our control problem let us consider an example. To allow for a graphical representation of the inventory vector, consider the case when we have only two retailers (i.e. $m = 2$). To further simplify the discussion, let the load allocation control take the simple form where, on each particular tour, the manager of the system has only three options: give priority to retailer 1, give priority to retailer 2, or give them both equal importance. In this simplified scenario, retailer i is given preference during a tour by setting its delivery size to $V_i + \kappa V$, for

$$\kappa \in \left(0, 1 - \frac{1}{\lambda} \max\{\lambda_1, \lambda 2\}\right)$$

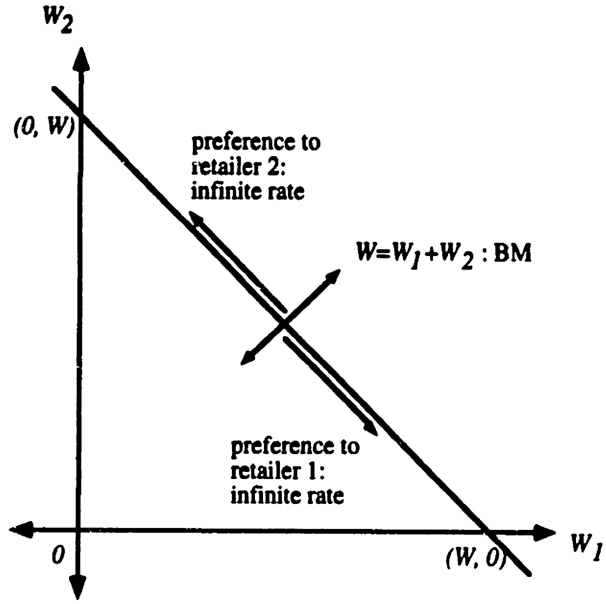

Figure 2-3: Normalized Inventory Evolution at Tour Completion Epochs under the HTAP.

a fixed parameter. Since a full load is delivered on every tour, the amount left at the other site (whose index, in this 2-retailer world, will be given by $3 - i$) must be $V_{3-i} - \kappa V$. Alternatively, when no retailer is given priority the manager simply sends amounts equal to the nominal allocation to both retailers.

In terms of units in the scaled state space, the additional amount delivered per cycle at the retailer with high priority will equal $\kappa V/\sqrt{n} = \kappa C$, which is $O(1)$. Considering the total inventory level at tour completion epochs, the HTAP implies that the rate at which the inventory at this retailer increases is infinite compared to the rate of change for the total inventory. Figure 2-3 illustrates the behavior of the normalized inventory level vector implied by this time scale decomposition at the diffusion time scale. As shown in the graph, at tour completion epochs this vector 'lives' in the constant total inventory line $W = w$, and it moves back and forth across this diagonal at an infinite rate, the direction being determined by which retailer is given preference (i.e. which site is getting κC units above its nominal allocation per cycle). Meanwhile, the total inventory line makes up and down parallel movements as a Brownian motion at a finite rate.

Consider now the evolution of the inventory levels at the deterministic time scale during the time the truck takes to complete one tour. Recall that, at this scale, time evolves $\sqrt{n}$ times slower than at the diffusion time scale, which makes the total netput appear to be fixed. The inventory at both retailers decreases deterministically (at the demand rate for each retailer) while the truck is traveling. At the epoch at which the retailer i is visited, its inventory level increases instantaneously by an amount equal to the normalized delivery size. Figure 2-4 plots the inventory evolution for a cycle in which retailer 2 is given preference in our 2-retailer example system. Notice that average travel times in this slower time scale are given by $\vartheta_{ij} = \theta_{ij}/\sqrt{n}$, and not by θ_{ij}/n as in the scaling giving rise to the BM limit for X. If we thus let $t = 0$ represent the time when the truck leaves the depot, then the inventory level at retailer 1 will instantaneously increase at time ϑ_{01} (by an amount κC units smaller than its nominal allocation), and the inventory at retailer 2 will increase instantaneously at the epoch $\vartheta_{01} + \vartheta_{12}$ (by an amount equal to $C_2 + \kappa C$). These are the only two epochs at which the inventories increase, since the truck spends the rest of the time traveling around the cycle. Notice that, since the HTAP holds under the assumption that $\rho_T^{(n)} \to 1$, we must take the traffic intensity to be one at the deterministic time scale to obtain an adequate trade-off between the inventory and transportation costs. Therefore, when the truck comes back to the depot (at time ϑ_T) the total normalized inventory in the system is unchanged but κC units have been 'transferred' from the low priority to the high priority retailer.

If a series of cycles giving preference to some retailers is implemented, then inventory is shifted from the least favored to the preferred sites. If, on the other hand, the nominal allocation sizes are used then, at the deterministic time scale, the inventory evolves in a fixed cycle in $\mathbf{R^m}$. Figure 2-5 illustrates the corresponding inventory evolution over a series of tours for the three possible allocation policies in our $m = 2$ example.

The implications for our problem in terms of the load allocation control are clear: no retailer subset should be given priority permanently, as this would immediately lead to ever increasing backorders at the rest of the sites; rather, for any given total

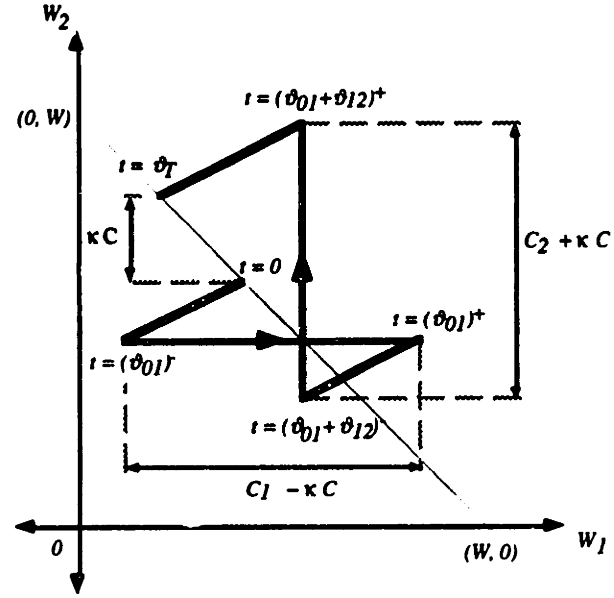

Figure 2-4: Inventory Evolution over a Cycle in the Slow Time Scale: Retailer 2 Given Preference.

inventory level at tour completion epochs, we should find the nominal-allocation cycle (i.e. the fixed cycle) that minimizes the inventory cost per unit time, and use load allocation policies that deviate from the nominal delivery size only as much as necessary to move there. Since movements along the constant total inventory hyperplane happen instantaneously at the diffusion time scale, the transient effects of shifting towards the desired fixed cycle and other temporary adjustments will not affect the long-run average inventory cost.

The HTAP also characterizes the evolution of the process $\hat{\xi}(t)$ that represents the adjustment for deliveries and partial tour completions necessary to relate the total inventory to the netput in (2.20). As we have argued, the HTAP implies that $\hat{\xi}(t)$ converges to a deterministic function of the time elapsed in the current tour. If we again consider the deterministic time scale and take $t = 0$ to represent the time at which the truck leaves the warehouse to start a new tour then, for the nominal allocation cycle, the behavior of $\hat{\xi}(t)$ in a two-retailer model corresponds to the path illustrated in Figure 2-6. While the truck is traveling, the value of $\hat{\xi}(t)$ decreases deterministically at the total demand rate λ. This process then increases instanta-

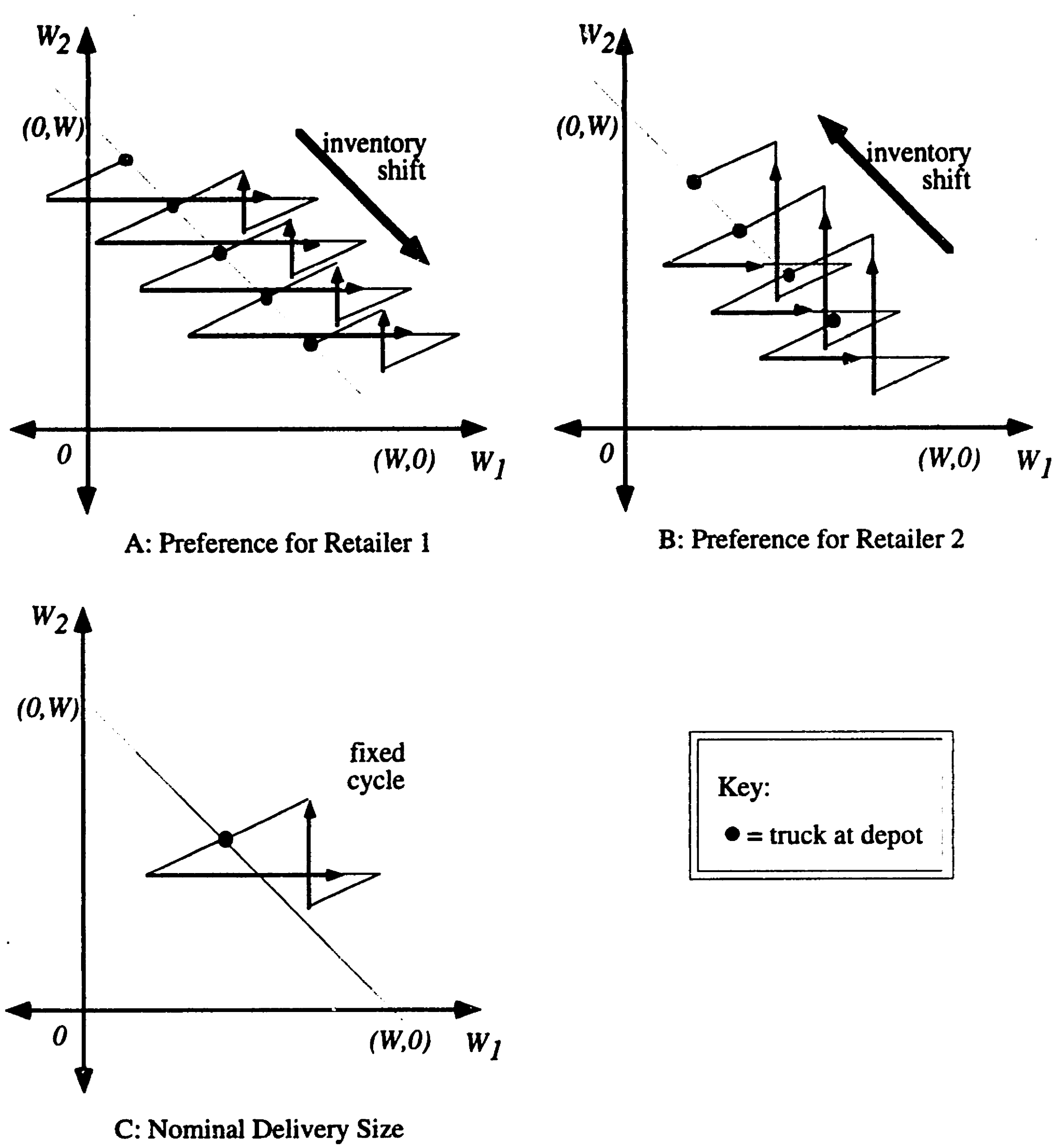

Figure 2-5: Effect on Inventory Evolution from Repetition of Allocation Schemes

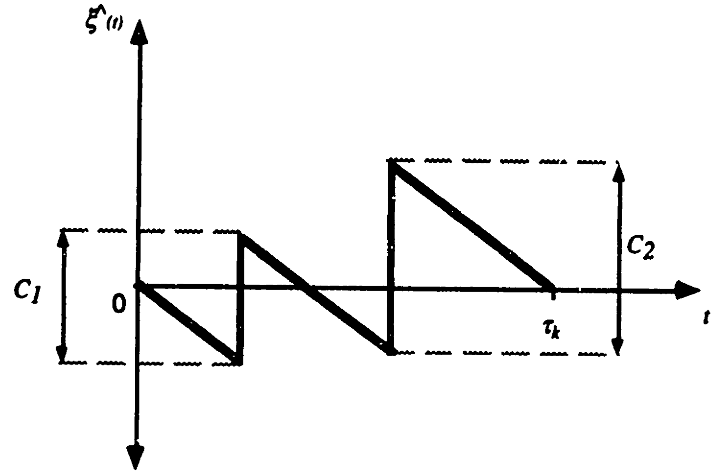

Figure 2-6: Evolution of $\hat{\xi}(t)$ Over a Nominal Allocation Cycle at the Deterministic Time Scale

neously by the nominal load allocation C_i (in terms of scaled inventory units) at the epoch where that retailer is visited. At the end of the cycle, $\hat{\xi}(\tau_k) = 0$. Recall that the system will use the nominal allocation most of the time, and hence we need only consider the evolution of $\hat{\xi}(t)$ under this delivery policy.

We will next use these implications to obtain a limiting control version for our problem. Again, we do this even though we have not rigorously proved that the HTAP holds for the VRIP. While we have chosen not to undertake this, we can exploit our understanding of the VRIP and the heuristic time scaling arguments of this section to provide an outline of such a proof. The analysis in Coffman, Puhalskii and Reiman (1993) is done under a specific queueing discipline: that of exhaustive service. The proof (for the 2 queues case) is based on the use of upper and lower bounds derived from a threshold queue. As we have argued heuristically, the VRIP seems to exhibit the same time scale decomposition as the polling system. The main difference between the systems is the inventory paths at the individual retailers. In the polling problem, the sample paths (on the deterministic time scale) look like those for the economic production quantity (EMQ): they go up and down at a finite rate. In the VRIP, the sample paths look like those from the economic order quantity model (EOQ): they go down at finite rate but go up at infinite rate (a vertical line). Whereas the HTAP for the polling system is derived with the aid of a threshold queue, it would appear that an averaging principle for the make-to-order version of the VRIP could be derived with the help of a threshold queue with instantaneous batch service, like a clearing system.

2.2.3 The Limiting Control Problem

Before we formally state the limiting control problem we need to exercise some care to make sure that we account for any distortions in the relative magnitudes of the transportation and inventory costs that may result from the scaling used to obtain the diffusion limit. This is necessary since, as implied by equation (2.12), our purpose is to find the control that minimizes the total system cost: that is, the one that best balances the inventory and transportation expenses. However, the time and space scaling involved in the limiting process essentially reduce the inventory holding and backordering costs by a factor of $n^{3/2}$. This is the case since one normalized unit of inventory held over one normalized time unit at the diffusion time scale corresponds to $\sqrt{n}$ items held over n time units in the original system. On the other hand, these same scalings only reduce the idling cost by $\sqrt{n}$. This follows from the observation that, in terms of the original system, an increase of 1 normalized time unit in the idling process $Y^{(n)}(t)$ corresponds at the diffusion time scale to $\sqrt{n}$ units of original system time. Since, by leaving b_i, h_i unscaled, we obtain holding costs that are incurred at a rate $O(n^{3/2})$ and idleness accrues at a rate $O(\sqrt{n})$, the travel cost $f^{(n)}$ must be increased by $O(n)$ for both cost rates to be of the same order. Consequently, let $\hat{f}^{(n)} = f^{(n)}/n$ denote the normalized travel cost. This implies that, in order to get a non-trivial problem, the heavy traffic approach requires the transportation cost rate f to be quite large relative to the inventory cost (roughly two orders of magnitude larger).

Substituting the limiting processes into the total inventory evolution in equation (2.20), we have that the limiting total inventory evolves as

$$W(t) = X(t) - \frac{V}{\theta_T} Y(t) + \hat{\xi}(t)$$

where X is a BM, Y is the idleness control (which we still need to optimize) and $\hat{\xi}$ is a deterministic function of the truck position along the tour. In this context, define the intrinsic total inventory process for the system as

$$Z(t) \equiv W(t) - \hat{\xi}(t).$$

This is the process that would be obtained if one were to observe the total inventory only at tour completion epochs. Notice that, since the evolution of $\hat{\xi}$ is independent of the inventory level, $W(t)$ can be recovered from $Z(t)$ and the current position along the tour. Also, recall from §2.2.2 that the system will almost always use the nominal allocation (other allocations are used only over infinitesimal amounts of time) so that the m-dimensional inventory process evolves around a fixed cycle for any given value of inventory at the beginning of a tour. Hence, knowing $Z(t)$ is sufficient to obtain the average inventory cost over a tour for a given nominal allocation cycle placement. Furthermore, the idling decision may also be based only on $Z(t)$ since it is related to $W(t)$ by a deterministic function. So that, if it is optimal for the system to idle when $Z(t) = z$ it is also optimal to idle when $W(t) = x + \hat{\xi}(t)$.

We can hence state the limiting stochastic control problem for the m-retailer TSP VRIP as follows: Find the optimal cycle placement under the nominal allocation for a given total intrinsic inventory level $Z(t) = x$, and its corresponding inventory cost rate $g(x)$; then, choose the nondecreasing RCLL process Y to minimize

$$\limsup_{T \to \infty} \frac{1}{T} E \left[\int_0^T g(Z(t)) dt - \hat{f} Y(T) \right] \tag{2.27}$$

subject to

$$Z(t) = X(t) - \frac{V}{\theta_T} Y(t) \tag{2.28}$$

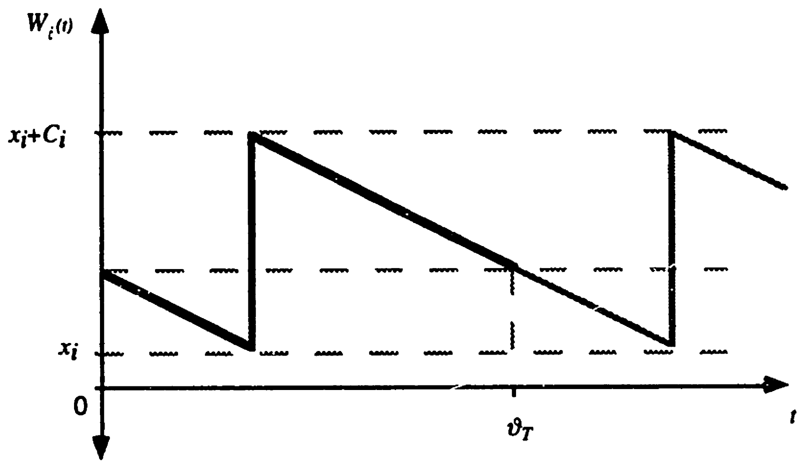

Figure 2-7: Inventory Evolution at Retailer i during a Nominal Allocation Cycle.

We now turn our attention to the solution of these two problems.

2.3 Optimal Cycle Placement

We address this problem in a similar fashion to the one Reiman and Wein (1994) use for the cycle placement in the polling problem and Markowitz, Reiman and Wein (1994) in the ELSP. Since backordering is allowed, the VRIP is closer to the ELSP than to the polling problem, in that there is no natural boundary at the origin; and hence, we can take the liberty of optimally placing the limit cycles in $\mathbf{R^m}$. 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 $W_i(t)$ (i.e. the amount of scaled inventory at retailer i just before delivery). Again, we assume that $\rho_T = 1$ (i.e., $\lambda_i \vartheta_T = C_i$) on the deterministic time scale to obtain an adequate trade-off among the different cost components. In this case, under a policy of nominal allocation delivery sizes, W_i will oscillate in deterministic fashion between x_i and $x_i + C_i$ at an infinite rate. Figure 2-7 illustrates the inventory level evolution at one retailer over a nominal load delivery cycle.

Of course, there are many alternative and equivalent characterizations of the cycle placement besides our proposed x_i — fixing any point along the deterministic path would do. However, this one seems to be as amenable to analysis as any other. In fact all one needs is to establish the relationship between the cycle placement variables x_i and the total intrinsic inventory level $Z(t) = x$. In order to do this we need some new notation. Recall that the warehouse index is 0 and that the retailer indices correspond to the order in which they are visited in the TSP tour. Denote the average scaled travel time between any two sites $i, j \in 0, 1, \ldots, m$ along the TSP path *in the deterministic time scale* by ϑ_{ij}^{TSP}. In terms of the original data, these quantities are defined as follows

$$\vartheta_{ij}^{TSP} = \frac{1}{\sqrt{n}} \sum_{k=i}^{j-1} \theta_{k,k+1} \quad \text{for} \quad j > i$$

and

$$\vartheta_{ii}^{TSP} = 0.$$

Notice that, while the notation is similar, the travel times ϑ_{ij}^{TSP} and ϑ_{ij} are different. The latter are parameters of the original system corresponding to the mean travel time at the deterministic time scale from site i to site j over the route with minimum expected delay, while the former corresponds to the travel time through all the legs of the TSP tour that lie between i and j at that same scale.

Again, measure time over a cycle so that the truck leaves the warehouse at $t = 0$. The state of the system is given by the scaled inventory level at the retailers, $(W_1(0), W_2(0), \ldots, W_m(0))$. Recall that we take $\rho_T = 1$ for the deterministic inventory evolution cycle and that this implies $\lambda_i \vartheta_T = C_i$. Hence each component of the inventory level vector $W_i(0)$ is related to its corresponding cycle placement value x_i by (see Figure 2-7):

$$W_i(0) = x_i + \lambda_i \vartheta_{0i}^{TSP} \quad \text{for} \quad i = 1, \ldots, m.$$

Summing these inventory levels over all retailers, we obtain the required relationship between the cycle placement vector $(x_1, \ldots, x_m)$ and the intrinsic system inventory $Z(0) = W(0) = x$ as:

$$\sum_i x_i = x - \sum_i \lambda_i \vartheta_{0i}^{TSP}. \tag{2.29}$$

Recall that, by the HTAP, the movement towards a particular cycle is instantaneous, and so we can ignore the inventory cost incurred while shifting in the calculations for the cost that corresponds to this particular placement. The average inventory cost per unit time while this stable cycle is used (i.e. while the total intrinsic inventory $Z(t) = x$) is therefore equal to the average cost over a cycle divided by the 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 times the (absolute value of) the average inventory. As illustrated in Figure 2-7, at retailer i the inventory jumps from x_i to $x_i + C_i$ once per cycle and decreases at a constant rate the rest of the time. Therefore the average inventory level over a cycle is simply $x_i + C_i/2$. 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 times h_i (b_i). To obtain the time average inventory cost we sum the areas of these triangles and divide by the cycle length $\vartheta_T = C_i/\lambda_i$. In summary, we have the following expression for retailer i:

$$g_i(x_i) = \begin{cases} h_i(x_i + \frac{C_i}{2}) & \text{if } x_i \geq 0 \\ \frac{h_i+b_i}{2C_i}x_i^2 + h_i x_i + \frac{h_i C_i}{2} & \text{if } 0 > x_i > -C_i \\ -b_i(x_i + \frac{C_i}{2}) & \text{if } -C_i \geq x_i. \end{cases} \tag{2.30}$$

Notice that $g_i(x_i)$ is a convex function of x_i. With equation (2.30) in hand, the optimal cycle placement problem can be expressed as

$$g(x) = \text{Min} \sum_i g_i(x_i) \quad \text{subject to (2.29).} \tag{2.31}$$

This mathematical programming problem may be solved by standard optimization procedures since the objective is a convex function. However, in order to make precise statements about the solution, we need to make some assumptions with regard to the cost structure. In particular, we shall assume that:

$$\begin{aligned} b_i &\geq h_i, \text{ for all } i \\ h_\ell &= h = \min_i h_i \\ b_p &= b = \min_i b_i \end{aligned} \tag{2.32}$$

The first assumption in (2.32) is made since backorders are in general considered something to be avoided. The second and third assumptions are just a labeling convention and so can be made without further loss of generality. Notice in particular that we are not assuming that the minimum inventory holding and backordering cost rates (h and b) occur at the same retailer (though $\ell = p$ is certainly allowed). A straightforward modification of the results allows one to handle other cost structures.

Under the assumptions in (2.32), we can find a closed-form solution to the optimal cycle placement problem. The basic idea is to use constraint (2.29), to turn the problem into one of unconstrained optimization over $m-1$ variables. After that, getting the optimal solution involves little else than standard non-linear optimization and cumbersome algebraic manipulations, and so we shall not go into the details. Interested readers are referred to the discussion of an analogous optimization in Markowitz, Reiman and Wein (1994) with regard to their analysis of the set-up cost ELSP. The solution to problem (2.29)-(2.32) yields the vector of optimal placements x_i^* and $g(x)$, the inventory cost as a function of the total total inventory at the warehouse epochs. Not surprisingly, $g(x)$ turns out to be quadratic with linear edges in the inventory level x. The solution is as follows:

Region 1. $\quad x < \hat{\alpha}_T = \sum_i \lambda_i \vartheta_{0i}^{TSP} - \sum_i \frac{b+h_i}{b_i+h_i} C_i$

$$\begin{aligned} x_i^* &= -\frac{b+h_i}{b_i+h_i} C_i \text{ for } i \neq p \\ x_p^* &= x - \sum_i \lambda_i \vartheta_{0i}^{TSP} + \sum_{i \neq p} \frac{b+h_i}{b_i+h_i} C_i \\ g(x) &= -bx + \hat{a}_1 \end{aligned}$$

$$\hat{a}_1 = b\sum_i \lambda_i \vartheta_{0i}^{TSP} + \frac{1}{2}\sum_i h_i C_i - \frac{1}{2}\sum_i \frac{(b+h_i)^2}{b_i+h_i} C_i$$

Region 2. $\hat{\alpha}_T \leq x \leq \hat{\beta}_T = \sum_i \lambda_i \vartheta_{0i}^{TSP} - \sum_i \frac{h_i - h}{b_i + h_i} C_i$

$$x_i^* = \frac{2\hat{a}_2 C_i}{h_i + b_i}\left(x - \sum_k \lambda_k \vartheta_{0k}^{TSP} - \sum_k \frac{(h_i - h_k)C_k}{b_k + h_k}\right)$$

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

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

$$\hat{a}_3 = 2\hat{a}_2\left(\sum_i \frac{h_i C_i}{b_i + h_i} - \sum_i \lambda_i \vartheta_{0i}^{TSP}\right)$$

$$\hat{a}_4 = \hat{a}_2\left(\sum_i \frac{h_i C_i}{b_i + h_i} - \sum_i \lambda_i \vartheta_{0i}^{TSP}\right)^2 + \frac{1}{2}\sum_i \frac{b_i h_i C_i}{b_i + h_i}$$

Region 3. $\hat{\beta}_T < x$

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

$$x_\ell^* = x - \sum_i \lambda_i \vartheta_{0i}^{TSP} + \sum_i \frac{h_i - h}{b_i + h_i} C_i$$

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

$$\hat{a}_5 = -h\sum_i \lambda_i \vartheta_{0i}^{TSP} + \frac{1}{2}\sum_i h_i C_i - \frac{1}{2}\sum_i \frac{(h_i - h)^2}{b_i + h_i} C_i$$

As can be seen from the equations above, the optimal cycle placement also has different functional representation in terms of the total intrinsic inventory x, along the same three regions that characterize the cost function $g(x)$. 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 cheaper 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 and 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 intrinsic inventory in the system.

It is worth noting that, for the symmetric cost case (i.e. when $h_i = h, \quad b_i = b$ for all $i = 1, \dots, m$), the more natural solution of

$$x_i^* = \begin{cases} \frac{x - \sum_k \lambda_k \vartheta_{0k}^{TSP}}{m} & \text{if } x < \hat{\alpha}_T \\ \frac{\lambda_i}{\lambda}\left(x - \sum_k \lambda_k \vartheta_{0k}^{TSP}\right) & \text{if } \hat{\alpha}_T \le x \le \hat{\beta}_T \\ \frac{x - \sum_k \lambda_k \vartheta_{0k}^{TSP}}{m} & \text{if } \hat{\beta}_T < x \end{cases}$$

is also optimal. This is so since the total inventory cost for this alternate cycle placement is the same as $g(x)$ given before.

We thus have a closed-form expression for the optimal nominal allocation cycle placement x_i and its corresponding average inventory cost $g(x)$ for any given total intrinsic inventory level $Z(t) = x$. This was the first part of the limiting control problem for the VRIP. We will next use $g(x)$ to find an optimal idling policy.

2.4 Optimal Base Stock Level

Once $g(x)$ is known, we can proceed with the solution to the one-dimensional stochastic control problem. However, we will not solve the general problem in (2.27) and (2.28), but the special case of base stock policies. The reason to specialize our analysis to this class of policies is threefold. First of all, by the problem decomposition obtained from the HTAP, the busy/idle decision will be based solely on the one-dimensional total intrinsic inventory process. A single parameter policy is the most natural in this context. Second, this assumption makes the problem quite easy to solve, since we can draw on well established results for its analysis. Specifically, under this kind of policy, the limiting idleness process corresponds to a single-sided regulator for the Brownian motion $X(t)$, and hence $Z(t)$ is a regulated brownian motion (RBM) on $(-\infty, z]$ (see §2.2 of Harrison (1985) for a definition). Thus, we can use known results for the steady state behavior of RBM to solve the control problem. Thirdly, single barrier policies often turn out to be asymptotically optimal under the heavy traffic regime (see e.g. Wein (1992)). This is in fact the case for our problem, and interested readers are referred to Appendix A of this thesis for a proof of this claim.

To find the optimal base stock level we must establish the steady state distribution for the total inventory level and the expected idleness rate. These are both well studied values, and both are characterized in Chapter 5 of Harrison (1985). For the expected idleness rate, a simple translation and sign change from his steady state analysis of the single barrier RBM (see formula 5.6.14, op. cit.) gives

$$\lim_{t\to\infty} \frac{1}{t} E_x[Y(t)] = \frac{\theta_T}{V}\mu_T \quad \text{for} \quad \mu_T \geq 0. \tag{2.33}$$

Substitution of the definition for the drift of $X(t)$ into (2.33) gives

$$\lim_{t\to\infty} \frac{1}{t} E_x[Y(t)] = \sqrt{n}\left(1 - \frac{\lambda\theta_T}{V}\right). \tag{2.34}$$

From this result, it is evident that the local time at the barrier depends only on the drift of the RBM, and is therefore independent of the base stock level. For our problem, this means that the transportation cost will not be affected by the selection of z, and so the optimal base stock level can be found by solving the equivalent problem:

$$\text{minimize} \ \limsup_{T\to\infty} \frac{1}{T} E\left[\int_0^T g(Z(t))dt\right] \tag{2.35}$$

subject to

$$Z(t) = X(t) - \frac{V}{\theta_T} Y(t) \tag{2.36}$$

independently of the transportation cost f. Notice that this is a direct consequence of the fact that we have constrained the operation of the truck to a single fixed route, so that the control has no effect on the drift of the diffusion approximation.

Turning now to the steady state density for $Z(t)$, it is a well known fact (see equation 5.6.11, op.cit.) that it is given by

$$p_Z(x) = \begin{cases} \hat{\nu}_T e^{\hat{\nu}_T(x-z)} & \text{if } x \leq z \\ \\ 0 & \text{if } x > z \end{cases} \tag{2.37}$$

where $\hat{\nu}_T = 2\mu_T/\sigma_T^2$, and $\hat{\nu}_T \geq 0$ by assumption. Using (2.37), the singular control problem becomes equivalent to finding z to optimize $\hat{F}_T(z)$, where, for $z \geq \hat{\beta}_T$

$$\begin{aligned}\hat{F}_T(z) &= \int_{-\infty}^{\hat{\alpha}_T}(-bx+\hat{a}_1)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}dx + \int_{\hat{\alpha}_T}^{\hat{\beta}_T}(\hat{a}_2x^2+\hat{a}_3x+\hat{a}_4)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}dx \\ &\quad + \int_{\hat{\beta}_T}^{z}(hx+\hat{a}_5)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}dx \qquad (2.38)\end{aligned}$$

and for $\hat{\alpha}_T < z < \hat{\beta}_T$

$$\hat{F}_T(z) = \int_{-\infty}^{\hat{\alpha}_T}(-bx+\hat{a}_1)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}dx + \int_{\hat{\alpha}_T}^{z}(\hat{a}_2x^2+\hat{a}_3x+\hat{a}_4)\hat{\nu}_T e^{\hat{\nu}_T(x-z)}dx \qquad (2.39)$$

The constants in (2.38) and (2.39) have the same definitions as in §2.3. Note that we are defining $\hat{F}_T(z)$ differently depending on whether $z \geq \hat{\beta}_T$ or not. This is necessary since, while the optimal base stock level z_T^* will always satisfy $z_T^* > \hat{\alpha}_T$[1], it need not be larger than $\hat{\beta}_T$.

Using integration by parts on (2.38) and(2.39), and then taking the first two derivatives of $\hat{F}_T(z)$ with respect to z we obtain the following:

Proposition 2. *The value that minimizes $\hat{F}_T(z)$ is z_T^*, where*

$$z_T^* = -\frac{1}{\hat{\nu}_T}\ln\left[\left(\frac{h}{b+h}\right)\left(\frac{\hat{\nu}_T(\hat{\beta}_T-\hat{\alpha}_T)}{e^{\hat{\nu}_T(\hat{\beta}_T-\hat{\alpha}_T)}-1}\right)\right]+\hat{\alpha}_T \qquad (2.40)$$

if $z_T^* \geq \hat{\beta}_T$, otherwise z_T^* is the solution to the following equation

$$\frac{2\hat{a}_2}{\hat{\nu}_T}e^{-\hat{\nu}_T(z_T^*-\hat{\alpha}_T)} + 2\hat{a}_2 z_T^* + \hat{a}_3 - \frac{2\hat{a}_2}{\hat{\nu}_T} = 0. \qquad (2.41)$$

Furthermore, the optimal cost is given by

$$\hat{F}_T(z_T^*) = hz_T^* + \hat{a}_5$$

if $z_T^* \geq \hat{\beta}_T$, and by

$$\hat{F}_T(z_T^*) = \hat{a}_2(z_T^*)^2 + \hat{a}_3 z_T^* + \hat{a}_4$$

[1]This is easily seen by the fact that $g(x)$ is linear and has a negative slope for $x < \hat{\alpha}_T$.

otherwise.

The proof is simple and is omitted. Optimality comes from the fact that, under the assumptions in (2.32) $\hat{F}_T''(z) \geq 0$ for both (2.38) and (2.39), and hence $\hat{F}_T(z)$ is convex. One can also show (from the fact that $\hat{F}_T(z)$ is convex and, hence continuously differentiable) that there is a unique optimum base stock level. That is, either there exists a solution z_T^* to (2.40) that satisfies $z_T^* \geq \hat{\beta}_T$ or a solution z_T^* to (2.41) that satisfies $z_T^* < \hat{\beta}_T$, but not both.

The base stock level z_T^* thus obtained, together with the cycle placement results of §2.3, completely characterize the optimal dynamic control for the fixed route VRIP under the heavy traffic regime. We are still left with the task of obtaining a control policy for the original (unscaled) system based on these results. This is a very interesting (and often quite challenging) issue which will be deferred until Chapter 4, where the implementation of results for alternative fixed-route schemes will also be discussed.

Chapter 3

Heavy Traffic Analysis of a Fixed Route VRIP: The Full-Load Direct Shipping Case

3.1 Problem Formulation

Since the setting for the full-load direct shipping (DS) VRIP is quite similar to the TSP case developed in Chapter 2, we will rely heavily in the material there, avoiding many details of the analysis except for those occasions where an important difference exists between the two cases. In terms of the problem set-up the DS case differs from the TSP in two main areas: the routing scheme used when the truck is operating, and the form taken by the delivery allocation control. All other characteristics of the problem (sources of uncertainty, operating costs, truck capacity, etc.) remain exactly as described in §2.1.

In terms of the routing scheme for this case (see Figure 3-1), the truck makes direct trips from the depot to each retailer. As before, the truck always leaves the warehouse with a full load and returns empty, so that every time a retail site is visited its inventory level increases by a full vehicle capacity V. Note that, in order to satisfy the expected demand in the long run, the system must, on the average, make a full

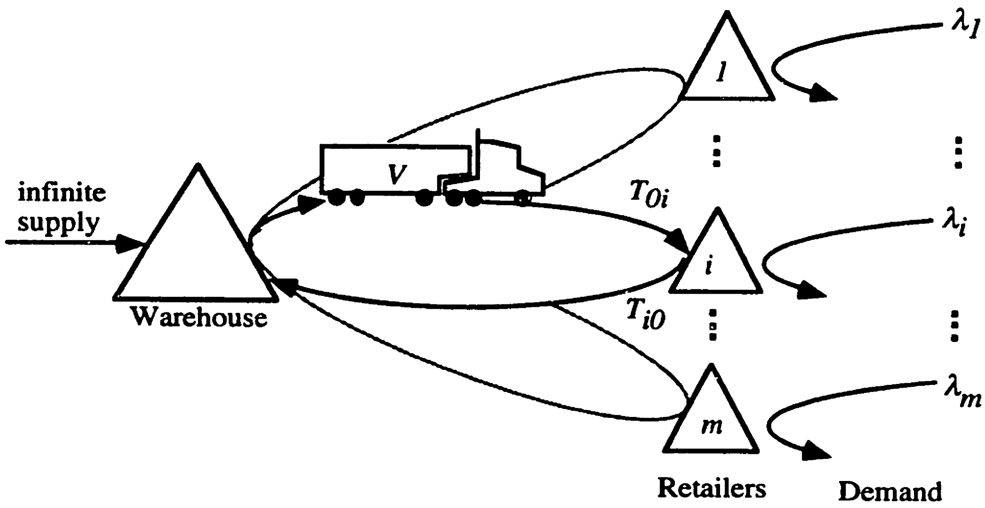

Figure 3-1: The Fixed-Route Full-Load DS VRIP.

load delivery to retailer i every $\lambda_i^{-1}V$ time units. In fact we take the nominal delivery allocation to correspond to a fixed sequence of visits to all the retailers (i.e. a polling table) which would meet this requirement exactly if travel times were deterministically equal to their means. It should be noted that, even for the deterministic travel time setting, it is not trivial (and in fact it may not be possible) to construct such a sequence. Nevertheless, we proceed as if such a polling table were available, since we shall again be invoking the machinery of heavy traffic approximations and, under the heavy traffic scaling, the diffusion limit for the total inventory is the same for any policy having the same average delivery rate under the nominal allocation.

Before we proceed any further, we must define the cycle completions for our proposed policy in order to identify the primitive service completion process. To this order, define the nominal relative visit frequency for retailer i as $\omega_i = \lambda_i/\lambda_\ell$, where $\lambda_\ell = \min_i \lambda_i$. Note that, while it is convenient to use ℓ to denote the index of the retailer with smallest demand and of the retailer with smallest holding cost, in reality these need not be the same. By the definition of ω_i, retailer i must receive ω_i deliveries for every time retailer ℓ (the one with the smallest demand rate) is visited. Besides this condition, in any feasible cycle each retailer must be visited an integer number of times. Within any desired accuracy, one can find a constant K such that

$K\omega_i$ is integer for all i. We therefore define a service cycle for the DS system as a sequence of $K\sum_i \omega_i = \lambda K/\lambda_\ell$ visits to the retailers. Notice that, with this definition, service completions need not correspond to the nominal allocation described in the previous paragraph (i.e. the truck might not visit retailer i exactly $K\omega_i$ times evenly spread over the course of the cycle). In fact the ability of the manager to deviate from the predefined polling table constitutes the dynamic load allocation control in the DS context. However, anticipating the use of the HTAP, delivery schemes that differ from the nominal allocation are only used for an infinitesimal instant. We therefore characterize the service completions based solely on the nominal polling table. Keeping the assumption that travel times over different tour legs are mutually independent, we have that the average cycle completion time is given by

$$\theta_D = \frac{2K}{\lambda_\ell}\sum_i \lambda_i \theta_{0i},$$

and that the cycle completion variance is

$$\sigma^2_{DS} = \frac{2K}{\lambda_\ell}\sum_i \lambda_i \sigma^2_{0i}.$$

These moments characterize the interarrival times for the process $S_D(t)$, which represents the cumulative service completions assuming the truck is continuously active in $[0,t]$. The manager affects the service completion process by deciding to let the truck idle. Correspondingly, denote the time allocation control by $T_D(t)$, the cumulative amount of the time that the truck is busy over $[0,t]$. We may hence define the service completion epochs in terms of these basic processes as:

$$\tau_k = \inf\{t \mid S_D(T_D(t)) \geq k\}.$$

As mentioned before, the manager may also affect the system through a dynamic delivery control that allows for deviations from the nominal allocation. Denote this control by the vector process $\varepsilon_i^D(t)$. The i-th component of this vector represents the allocation control in the following manner (see Figure 3-2): $\varepsilon_i^D(t)$ increases by a full load V every time a delivery at retailer i is made; the nominal delivery per cycle for retailer i (which equals $K\omega_i V$)) is then subtracted at every sequence completion epoch τ_k, $k = 1, 2, \ldots$. In this way, as in the fixed route TSP case, the control is expressed in terms of cumulative deviations from the nominal allocation over prior services, plus the amount delivered during the current cycle. Notice that, since a cycle completion occurs every time $K\lambda/\lambda_\ell$ shipments are made and a full load is delivered on every visit, the allocation control process must satisfy

$$\varepsilon_i(0) = 0, \text{ for all } i \tag{3.1}$$

$$\varepsilon_i(t) \geq \varepsilon_i(\tau_{k-1}), \text{ for } t \in (\tau_{k-1}, \tau_k) \text{ and all } i \tag{3.2}$$

$$\sum_i \varepsilon_i(\tau_k^-) = K\frac{\lambda}{\lambda_\ell}V, \tag{3.3}$$

$$\sum_i \varepsilon_i(\tau_k) = 0, \tag{3.4}$$

where τ_k is the time of the k-th polling cycle completion, and τ_k^- is the epoch an infinitesimal amount of time before this completion. As a consequence of (3.4), individual deviations from the nominal delivery allocation cancel out across the retailers, and $\varepsilon_D(t) = \sum_i \varepsilon_i^D(t)$ equals the amount delivered over the current cycle.

Finally, denote the demand arrival process by $D(t)$, and the state of the system by the inventory vector $Q_i(t)$, with its corresponding total inventory $Q(t) = \sum_i Q_i(t)$. The dynamics for the evolution of the state of the DS system, assuming that $Q_i(0) = 0$, for all i, are given by

$$Q_i(t) = K\omega_i V S_D(T_D(t)) - D_i(t) + \varepsilon_i^D(t) \text{ for } i = 1, \ldots, m \text{ and } t \geq 0. \tag{3.5}$$

Express the cumulative idle time process $I(t)$ as

$$I(t) = t - T_D(t) \text{ for } t \geq 0 \tag{3.6}$$

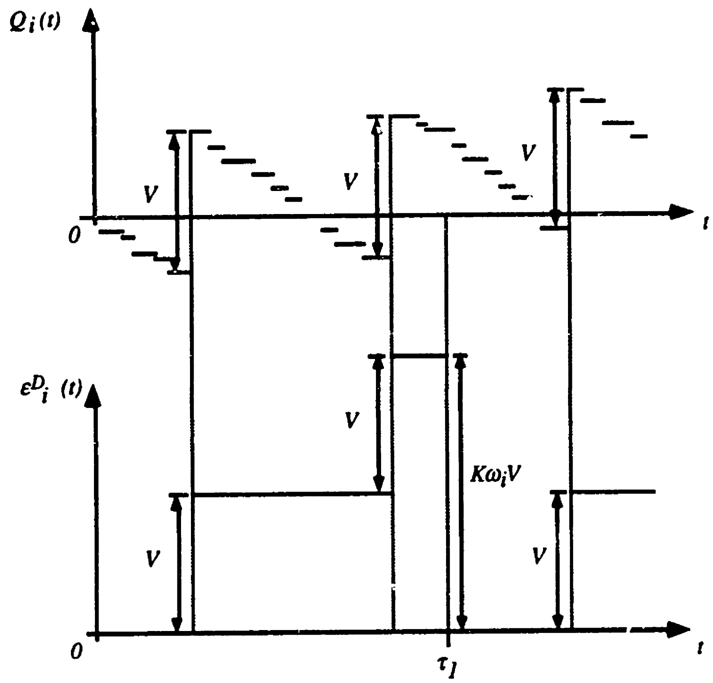

Figure 3-2: Joint Evolution of Inventory and Load Allocation Control: DS Case.

so that the control policy $\left(T_D(t), \varepsilon_i^D(t)\right)$ must satisfy

$$T_D \quad \text{is nondecreasing and continuous with } T_D(0) = 0 \tag{3.7}$$

$$T_D, \varepsilon_i^D \quad \text{are nonanticipating with respect to } Q \tag{3.8}$$

$$I \quad \text{is nondecreasing with } I(0) = 0 \tag{3.9}$$

The DS version of the fixed-route VRIP may thus be posed as finding a control policy $\left(T(t), \varepsilon_i^D(t)\right)$ to

$$\min \limsup_{T\to\infty} \frac{1}{T} E\left[\int_0^T \sum_i (h_i\{Q_i(t)\}^+ + b_i\{Q_i(t)\}^-)\, dt - fI(T)\right] \tag{3.10}$$

subject to (3.4) - (3.9). Again, the problem in its exact formulation does not appear to be tractable, and hence we resort to a heavy traffic approximation.

3.2 Heavy Traffic Normalizations and Averaging Principle

Since we follow the same procedure as that used in Chapter 2 for the TSP case, many details are omitted. Define the centered demand and partial service processes $\bar{S}_D(t), \bar{D}(t)$ in the same fashion that was used for the TSP case. The netput process for the n-th DS system is the given by

$$\chi^{(n)}(t) = \left(\frac{\lambda^{(n)}C^{(n)}}{2\sum_i \lambda_i^{(n)}\vartheta_{0i}^{(n)}} - \lambda^{(n)}\right)t + \frac{K^{(n)}\lambda^{(n)}C^{(n)}}{\lambda_\ell^{(n)}}\bar{S}_D^{(n)}(T_D^{(n)}(t)) - \bar{D}^{(n)}(t). \qquad (3.11)$$

Summing the inventory dynamics equations (3.5) over all retailers, and substituting the suitable definitions gives the following expression for the dynamics of the total inventory in the n-th system:

$$Q^{(n)}(t) = \chi^{(n)}(t) - \frac{\lambda^{(n)}V^{(n)}}{2\sum_i \lambda_i^{(n)}\theta_{0i}^{(n)}}I^{(n)}(t) + \xi_D^{(n)}(t), \qquad (3.12)$$

where

$$\xi_D^{(n)}(t) = \varepsilon_D^{(n)}(t) + \frac{K^{(n)}\lambda^{(n)}V^{(n)}}{\lambda_\ell^{(n)}}S_D^{(n)}(T_D^{(n)}(t)) - \frac{K^{(n)}\lambda^{(n)}C^{(n)}}{\lambda_\ell^{(n)}}\mathcal{S}_D^{(n)}(T_D^{(n)}(t))$$

Starting from (3.11)-(3.12), we can follow exactly the same path as for the TSP tour to obtain a diffusion limit for the system netput process. Again, we generate a sequence of normalized systems, and the corresponding normalized inventory, idleness, and netput processes $(W, Y, X)^{(n)}$, using the same scaling as in §2.2.1 to obtain:

Proposition 3. *$X^{(n)}(t) \Rightarrow X(t)$, where $\Rightarrow$ denotes weak convergence and $X(t)$ is a Brownian motion process with drift*

$$\mu_D = \sqrt{n}\left(\frac{\lambda C}{2\sum_i \lambda_i \vartheta_{0i}} - \lambda\right)$$

and variance

$$\sigma_D^2 = \lambda\left(c_d^2 + \frac{\lambda C \sum_i \lambda_i \varsigma_{0i}^2}{2(\sum_i \lambda_i \vartheta_{0i})^2}\right),$$

where c_d^2 represents the squared coefficient of variation of the demand interarrival times.

Notice that the parameters of the Brownian motion are independent of the multiplicative factor K which was used to guarantee that each retailer was visited an integer number of times during a cycle. This limiting result holds subject to the heavy traffic condition that

$$\sqrt{n}\left(\frac{\lambda^{(n)}C^{(n)}}{2\sum_i \lambda_i^{(n)}\vartheta_{0i}^{(n)}} - \lambda^{(n)}\right) = \mu_D = O(1), \quad \text{for all } n.$$

In other words, the result holds subject to the traffic intensity, given by

$$\rho_D^{(n)} = \frac{2\sum_i \lambda_i^{(n)}\theta_{0i}^{(n)}}{V^{(n)}},$$

approaching unity at the right rate. As in the TSP case, we also make the parameter scalings necessary for the delivery size to remain $O(1)$ in terms of normalized inventory units without affecting the load level nor the variability of the system. The appropriate definitions for the DS case are

$$C^{(n)} = \frac{V^{(n)}}{\sqrt{n}},$$

$$\vartheta_D^{(n)} = \frac{\theta_D^{(n)}}{\sqrt{n}},$$

and

$$\varsigma_{DS}^{2\,(n)} = \frac{\sigma_{DS}^{2\,(n)}}{\sqrt{n}}.$$

The parameters are scaled in order to satisfy the conditions (analogous to the TSP case) that

$$\lim_{n\to\infty} C^{(n)} = C = O(1) \tag{3.13}$$

$$\lim_{n\to\infty} \vartheta_D^{(n)} = \vartheta_D = O(1) \tag{3.14}$$

$$\lim_{n\to\infty} \varsigma_{DS}^{2\,(n)} = \varsigma_{DS}^2 = O(1) \tag{3.15}$$

The only difference between the limiting process for the system netput of Proposition 3 and the analogous result for the TSP case (c.f. Proposition 1 of §2.2.1) are the formulae for the parameters of the Brownian motion in terms of the data of the original system. This does not affect the behavior of the system with regard to the time scale decomposition that gives rise to the HTAP. In fact, the same argument put forth in §2.2.2 for the TSP case goes through for the DS policy. Hence, we have that the HTAP should hold for this case as well. As a consequence, the delivery allocation control may be used to alter individual inventories at a rate that is an order of magnitude faster than the one at which the netput varies. Define the total intrinsic inventory process as

$$Z(t) \equiv W(t) - \hat{\xi}_D(t).$$

By the HTAP, the retailer inventory levels $W_i(t)$ will move deterministically along a fixed path for any given value of $Z(t) = x$; as long as the nominal delivery allocation is used. This leads to the following decomposition for the limiting stochastic control problem for the DS VRIP: first solve for the optimal cycle placement given $Z(t) = x$, and obtain the corresponding optimal inventory cost rate $g_D(x)$; then choose the nondecreasing RCLL process Y to minimize

$$\limsup_{T\to\infty} \frac{1}{T} E[\int_0^T g_D(Z(t))dt - \hat{f}Y(T)] \tag{3.16}$$

subject to

$$Z(t) = X(t) - \frac{\lambda V}{2\sum_i \lambda_i \theta_{0i}} Y(t). \tag{3.17}$$

3.3 Optimal Cycle Placement

Again we face the problem of optimally placing the limit cycles for the deterministic evolution of the retailer inventories in $\mathbf{R}^{\mathbf{m}}$. We shall 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 $W_i(t)$ (i.e. the amount of scaled inventory at retailer i just before a delivery). The HTAP implies that, while the nominal delivery allocation is used, W_i varies

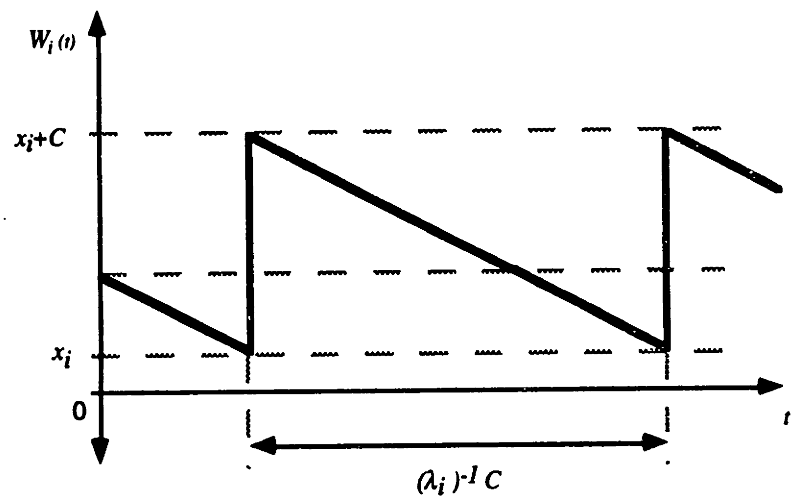

Figure 3-3: Inventory at Retailer i for the DS Case with Nominal Allocation.

deterministically in $[x_i, x_i+C]$. The evolution of the normalized inventory at retailer i over time is illustrated in Figure 3-3. The only differences with the TSP case are the delivery size (in this case we deliver full loads on each visit to a retailer) and the visit frequency. Again, since the HTAP is obtained under the assumption that $\rho_D = 1$, the inter-visit time at retailer i for the nominal load allocation must equal $\lambda_i^{-1}C$.

The next step is to establish the relationship between the cycle placement variables x_i and the total intrinsic inventory level $Z(t) = x$. Here we encounter a difference with the TSP case in that we would need to specify the polling table for the DS policy in order to establish this relationship precisely. However, the optimal cycle placement may be found based on an arbitrary epoch along the deterministic inventory evolution. To this order, and seeking to avoid the otherwise unnecessary task of defining the polling table, we shall assume that the total intrinsic inventory equals the average inventory over the cycle. This assumption will simplify the exposition significantly and will not reduce the applicability of our results. We shall revisit this issue in Chapter 4 when the solutions from our heavy traffic analysis are interpreted in terms of the original system. There we will develop a simple rule to recover the true intrinsic inventory level from the system evolution.

Based on the evolution of the inventory at each retailer (remember, it is the same as in the TSP case except that the (scaled) delivery size is always C), we obtain the following expression for $\bar{x}_i$ the average scaled inventory level, at retailer i over a cycle:

$$\bar{x}_i = x_i + \frac{C}{2}$$

Now, we can sum these average inventories at the retailers to obtain the average total inventory level over a cycle. We now make use of the assumption that the intrinsic inventory equals the average total inventory over a cycle to obtain the following relationship between the cycle placement parameters x_i and $Z(t) = x$:

$$\sum_i x_i = x - \frac{mC}{2}. \tag{3.18}$$

Since we have already observed that the inventory evolution for the DS case differs from the TSP case only in the delivery size, we obtain the inventory cost at retailer i as a function of the cycle placement parameter x_i as

$$g_{Di}(x_i) = \begin{cases} h_i(x_i + \frac{C}{2}) & \text{if } x_i \geq 0 \\ \frac{h_i+b_i}{2C} x_i^2 + h_i x_i + \frac{h_i C}{2} & \text{if } 0 > x_i > -C \\ -b_i(x_i + \frac{C}{2}) & \text{if } -C \geq x_i \end{cases} \tag{3.19}$$

and the optimal cycle placement problem for the DS case can be expressed as

$$\begin{aligned} g_D(x) \quad = \quad & \text{Min} \sum_i g_{Di}(x_i) \\ & \text{subject to (3.18).} \end{aligned} \tag{3.20}$$

As in the TSP case, we are left with a convex-objective mathematical program, and hence the cycle placement problem is easy to solve. We still make the assumptions on the cost structure specified in equation (2.32) of §2.3 and apply standard optimization techniques to obtain the optimal cycle placements and their corresponding cost function $g_D(x)$. Not surprisingly, this function will again have three different functional forms, according to the value of the total scaled inventory x. The results are as follows:

$$\text{Region 1.} \qquad x < \hat{\alpha}_D = -C\sum_i \frac{b+h_i}{b_i+h_i} + \frac{mC}{2}$$

$$\begin{aligned}
x_i^* &= -\frac{b+h_i}{b_i+h_i}C \text{ for } i \neq p \\
x_p^* &= x - \frac{mC}{2} + C\sum_{i\neq p} \frac{b+h_i}{b_i+h_i} \\
g_D(x) &= -bx + \hat{a}_6 \\
\hat{a}_6 &= b\frac{mC}{2} + \frac{C}{2}\sum_i h_i - \frac{C}{2}\sum_i \frac{(b+h_i)^2}{b_i+h_i}
\end{aligned}$$

$$\text{Region 2.} \qquad \hat{\alpha}_D \leq x \leq \hat{\beta}_D = C\sum_i \frac{h-h_i}{b_i+h_i} + \frac{mC}{2}$$

$$\begin{aligned}
x_i^* &= \frac{2\hat{a}_7 C}{h_i+b_i}\left(x - \frac{mC}{2}\right) \\
g_D(x) &= \hat{a}_7 x^2 + \hat{a}_8 x + \hat{a}_9 \\
\hat{a}_7 &= \frac{1}{2C}\left(\sum_i \frac{1}{b_i+h_i}\right)^{-1} \\
\hat{a}_8 &= 2C\hat{a}_7\left(\sum_i \frac{h_i}{b_i+h_i} - \frac{m}{2}\right) \\
\hat{a}_9 &= \hat{a}_7\left(C\sum_i \frac{h_i}{b_i+h_i} - \frac{mC}{2}\right)^2 + \frac{C}{2}\sum_i \frac{b_i h_i}{b_i+h_i}
\end{aligned}$$

$$\text{Region 3.} \qquad \hat{\beta}_D < x$$

$$\begin{aligned}
x_i^* &= -\frac{h_i-h}{b_i+h_i}C \text{ for } i \neq \ell \\
x_\ell^* &= x - \frac{mC}{2} + C\sum_i \frac{h_i-h}{b_i+h_i} \\
g(x) &= hx + \hat{a}_{10} \\
\hat{a}_{10} &= -h\frac{mC}{2} + \frac{C}{2}\sum_i h_i - \frac{C}{2}\sum_i \frac{(h_i-h)^2}{b_i+h_i}
\end{aligned}$$

Observe that the overall qualitative behavior of the cycle placement is the same as in the TSP case: when the total inventory is far away from the origin use the cheapest site to accumulate the excess inventory or backorders; if the total inventory level is close to the origin, then the cycle placement varies linearly with the intrinsic inventory level.

3.4 Optimal Base Stock Level

We now proceed with the solution to the stochastic control problem in (3.16) and (3.17), again specialized to the case of base stock policies. The reasoning behind this decision is the same as for the TSP case: not only is this a natural policy to consider, but it is also asymptotically optimal. A proof of the optimality of base stock policies for the TSP case is included in Appendix A. While no proof is provided for the DS case, the argument would proceed in essentially the same manner.

As in the TSP case, the total intrinsic inventory process $Z(t)$ is a regulated Brownian motion on $(-\infty, z]$ when a base stock policy is used. Therefore, in order to establish the optimal base stock level for the DS-VRIP all we need are standard results in the theory of diffusion processes. By arguments identical to those in §2.4 the optimal base stock level will be the value of z that minimizes $\hat{F}_D(z)$, where, for $z \geq \hat{\beta}_D$

$$\begin{aligned}\hat{F}_D(z) &= \int_{-\infty}^{\hat{\alpha}_D}(-bx+\hat{a}_6)\hat{\nu}_D e^{\hat{\nu}_D(x-z)}dx + \int_{\hat{\alpha}_D}^{\hat{\beta}_D}(\hat{a}_7x^2+\hat{a}_8x+\hat{a}_9)\hat{\nu}_D e^{\hat{\nu}_D(x-z)}dx \\ &\quad + \int_{\hat{\beta}_D}^{z}(hx+\hat{a}_{10})\hat{\nu}_D e^{\hat{\nu}_D(x-z)}dx \qquad (3.21)\end{aligned}$$

and for $\hat{\alpha}_D < z < \hat{\beta}_D$

$$\begin{aligned}\hat{F}_D(z) &= \int_{-\infty}^{\hat{\alpha}_D}(-bx+\hat{a}_6)\hat{\nu}_D e^{\hat{\nu}_D(x-z)}\,dx \\ &\quad + \int_{\hat{\alpha}_D}^{z}(\hat{a}_7x^2+\hat{a}_8x+\hat{a}_9)\hat{\nu}_D e^{\hat{\nu}_D(x-z)}dx \qquad (3.22)\end{aligned}$$

where $\hat{\nu}_D = 2\mu_D/\sigma_D^2$ is the parameter for the exponential distribution that characterizes the steady state of $Z(t)$. The rest of the constants in (3.21) and (3.22) have the same definitions as in §3.3. As in the TSP case, the optimal value of the base stock level need not be larger than $\hat{\beta}_D$, and hence the two definitions for $\hat{F}_D(z)$ given above.

We now proceed with the minimization. As before, we use integration by parts, and then take the first two derivatives of $\hat{F}_D(z)$ with respect to z to obtain:

Proposition 4. *The value of z that minimizes $\hat{F}_D(z)$ is given by*

z_D^*, where

$$z_D^* = -\frac{1}{\hat{\nu}_D} \ln \left[\left(\frac{h}{b+h} \right) \left(\frac{\hat{\nu}_D(\hat{\beta}_D - \hat{\alpha}_D)}{e^{\hat{\nu}_D(\hat{\beta}_D - \hat{\alpha}_D)} - 1} \right) \right] + \hat{\alpha}_D \tag{3.23}$$

if $z_D^* \geq \hat{\beta}_D$, otherwise z_D^* is the solution to the following equation

$$\frac{1}{\hat{\nu}_D} e^{-\hat{\nu}_D(z_D^* - \hat{\alpha}_D)} + z_D^* - \frac{1}{\hat{\nu}_D} - \frac{mC}{2} + C \sum_i \frac{h_i}{b_i + h_i} = 0. \tag{3.24}$$

Furthermore, the optimal cost $\hat{F}_D(z_D^*)$ is given by

$$\hat{F}_D(z_D^*) = h z_D^* + \hat{a}_{10}$$

if $z_D^* \geq \hat{\beta}_D$, and by

$$\hat{F}_D(z_D^*) \;\; = \;\; \hat{a}_7 (z_D^*)^2 + \hat{a}_8 z_D^* + \hat{a}_9$$

otherwise.

The proof is simple and is omitted. The analysis follows essentially the same outline described after Proposition 2.

The results in Proposition 4, together with the optimal cycle placement equations of §3.3 provide a complete characterization for the optimal dynamic control of the fixed route DS-VRIP under a heavy traffic regime. Notice that, since we have used the cycle placement formulae of §3.3 to obtain z_D^*, we have implicitly use the assumption that $Z(t)$ equals the average inventory over a cycle. In Chapter 4 we will consider the implications of these results for the control of the original system, and show that this assumption is really innocuous.

Chapter 4

Fixed Route VRIP Results

4.1 Interpretation of Limiting Optimal Control

4.1.1 The Fixed Route TSP Case

The analysis in Chapter 2 allowed us to characterize the control policy that is optimal for the fixed-route TSP VRIP under heavy traffic. We will now interpret those results to extract a policy that may be implemented in the original system. Again, there are two main control decisions available for this system: whether the truck should be busy or idle, and how to assign the load among the retailers during a particular tour. We will address the load allocation decision first, and then consider the busy/idle control.

Since the state of the system evolves dynamically in time, the decision of how much of the load to leave at each retailer i is best delayed until the truck arrives at the site. To this end, let t corresponds to the epoch at which the truck leaves the warehouse with a full load, and consider then $t_i^- > t$, the point in time just before the truck arrives at retailer i. Using the unscaled version for the travel times *along the TSP tour* that we defined in §2.3, the average travel time between any two sites $i, j \in 0, 1, \ldots, m$ is $\theta_{ij}^{TSP} = \sqrt{n}\,\vartheta_{ij}^{TSP}$. Under the deterministic inventory evolution, if the truck leaves the depot at time t it would be arriving at retailer i at time $t_i^- = t + \theta_{0i}^{TSP-} \equiv \lim_{\epsilon \to 0} t + \theta_{0i}^{TSP} - \epsilon$, for $\epsilon > 0$. At time t_i^-, the state of the system is given by the vector of inventory levels at the retailers, $\left(Q_1(t_i^-), \ldots, Q_m(t_i^-)\right)$, and the amount of goods left in the truck, $L(t_i^-)$. Given this information, how many units should the truck deliver at retailer i? The answer, according to the results in Chapter 2, is to set the delivery size d_i so as to shift the vector of retailer inventories towards the optimal cycle placement.

Recall that the optimal cycle placement problem of Chapter 2 (In particular, equation (2.29)) was obtained by means of a relationship between the (scaled) cycle placement vector x_i and the total (scaled) intrinsic inventory $Z(t)$. Since it seems advantageous to delay the load allocation decision until each retailer is reached, the first necessary step to establish the delivery size is to find a relation between the total system inventory at the epochs where the truck is at retailer i and the corresponding intrinsic inventory level.

To this order, consider the value of $Q_j(t_i)$, the inventory level at retailer j, at the epoch t_i when the truck arrives at retailer i. In keeping with the behavior predicted by the CPR results we characterize these values under a deterministic evolution for the retailer inventories over the course of a cycle. The relationship between $Q_j(t_i)$ and the unscaled cycle placement $q_j = \sqrt{n}\, x_j$ at the time where the truck visits retailer i depends on whether site j lies after or before i in the TSP tour (see Figure 4-1). Again under the deterministic evolution, assuming the truck leaves the warehouse at time t, it arrives at retailer i at time $t_i^- = t + \theta_{0i}^{TSP-}$. Therefore, the retailer inventories relate to the cycle placement parameters by:

$$Q_j(t_i^-) = q_j + \lambda_j \theta_{ij}^{TSP} \text{ for } j \geq i$$

and

$$Q_j(t_i^-) = q_j + V_j - \lambda_j \theta_{ji}^{TSP} \text{ for } j < i.$$

Also, the total inventory at time $t_i^- = t + \theta_{0i}^{TSP-}$, relates to the cycle placement parameters through

$$Q(t_i^-) = \sum_j Q_j(t_i^-) = \eta_i + \sum_j q_j, \tag{4.1}$$

where

$$\eta_i = \sum_{j<i} V_j - \sum_{j<i} \lambda_j \theta_{ji}^{TSP} + \sum_{j\geq i} \lambda_j \theta_{ij}^{TSP}$$

is an epoch locator constant for each retailer i.

According to the unscaled version of constraint (2.29) the cycle placement vector q_i must satisfy

$$\sum_i q_i = q - \sum_i \lambda_i \theta_{0i}^{TSP} \tag{4.2}$$

when $Q(t) = q$ (i.e. when the intrinsic inventory equals q). Equation (4.2) is obtained by substituting $q_i/\sqrt{n} = x_i$, $q/\sqrt{n} = x$, and $\theta_{ij}^{TSP}/\sqrt{n} = \vartheta_{ij}^{TSP}$ into (2.29). Readers may verify that, with these substitutions, the scaling factor n cancels out. Using equations (4.1) and (4.2), we can express the total inventory at time t (i.e. when the truck was at the warehouse) as a translation of the inventory vector at time t_i^-. Namely,

$$Q(t) = q(t_i^-) = \sum_j Q_j(t_i^-) - \eta_i + \sum_j \lambda_j \theta_{0j}^{TSP}.$$

The value $q(t_i)$ thus defined is the intrinsic inventory that we need to use to find the optimal cycle placement. Notice that we have essentially used the deterministic inventory evolution to define a constant (for each retailer) which we should use to translate the total inventory at the time when the retailer is visited into the total intrinsic inventory for the cycle.

To reduce the notational burden, and to make the connection with the heavy traffic results of Chapter 2, we will henceforth denote the intrinsic inventory level simply as q. The reader should note, however, that this quantity will change in a dynamic fashion as the total inventory changes. It plays the role of the system state in the load allocation part of the VRIP. This being said, for a given value of q we can find $q_i^* = \sqrt{n}\, x_i^*$, the corresponding optimal (unscaled) cycle placement parameters from the unscaled version of the formulae in §2.3. As before, this is done by replacing the parameters of the original system with their appropriate scalings. Namely we let $q_i/\sqrt{n} = x_i$, $q/\sqrt{n} = x$, $V_i/\sqrt{n} = C_i$, and $\theta_{ij}^{TSP}/\sqrt{n} = \vartheta_{ij}^{TSP}$ to obtain the following

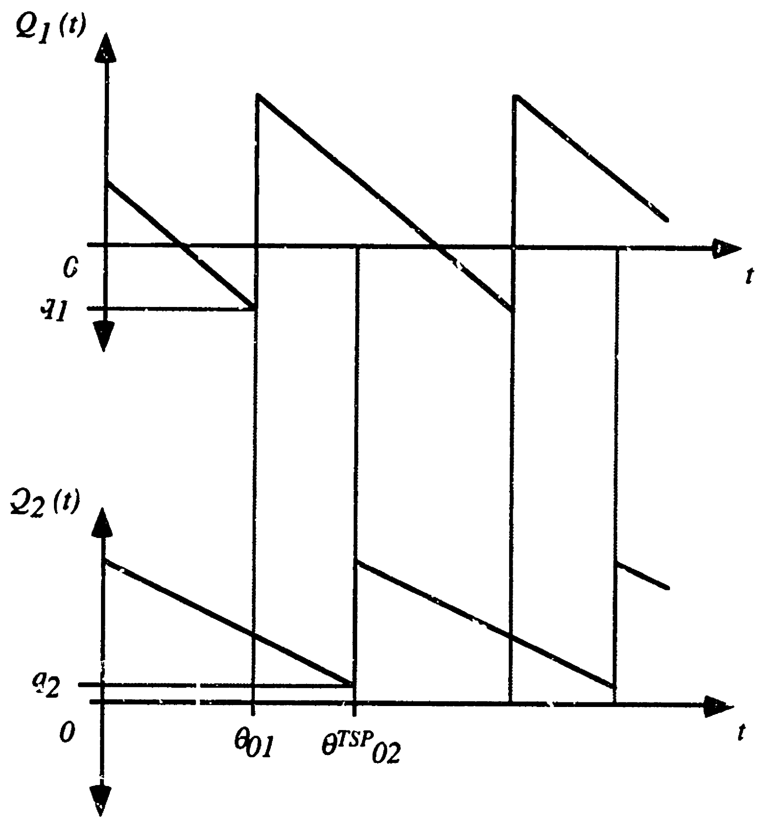

Figure 4-1: Deterministic Inventory Evolution: TSP tours.

expressions for q_i^* in terms of the state q:

Region 1. $\quad q < \alpha_T = \sum_i \lambda_i \theta_{0i}^{TSP} - \sum_i \frac{b+h_i}{b_i+h_i} V_i$

$$q_i^* = -\frac{b+h_i}{b_i+h_i} V_i \text{ for } i \neq p$$

$$q_p^* = q - \sum_i \lambda_i \theta_{0i}^{TSP} + \sum_{i \neq p} \frac{b+h_i}{b_i+h_i} V_i$$

Region 2. $\quad \alpha_T \le q \le \beta_T = \sum_i \lambda_i \theta_{0i}^{TSP} - \sum_i \frac{h_i - h}{b_i+h_i} V_i$

$$q_i^* = \frac{2a_2 V_i}{h_i + b_i} \left(q - \sum_k \lambda_k \theta_{0k}^{TSP} - \sum_k \frac{(h_i - h_k) V_k}{b_k + h_k} \right)$$

$$a_2 = \frac{1}{2} \left(\sum_i \frac{V_i}{b_i + h_i} \right)^{-1}$$

Region 3. $\quad \beta_T < q$

$$q_i^* = -\frac{h_i - h}{b_i+h_i} V_i \text{ for } i \neq \ell$$

$$q_\ell^* = q - \sum_i \lambda_i \theta_{0i}^{TSP} + \sum_i \frac{h_i - h}{b_i+h_i} V_i$$

which again are independent of the scaling parameter n.

We may now use these results to decide the delivery size at retailer i. Observe that, 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

$$Q_i(t_i^+) = q_i^* + V_i. \tag{4.3}$$

Therefore, we set d_i^*, the ideal delivery size for this cycle, so as to come as close as possible to this state. The actual inventory in the system after the delivery is made will be $Q_i(t_i^-) + d_i$, and equating this to (4.3), its equivalent point in the optimal cycle, we obtain

$$d_i^* = q_i^* + V_i - Q_i(t_i^-).$$

for the ideal dynamic load allocation. Note that these delivery sizes are adjusted dynamically to the state of the system, through the dependence of q_i^* on q.

We may not use this delivery size directly since we cannot allocate more than the available load $L(t_i^-)$, nor desire to make negative deliveries. We shall therefore deliver an amount that is as close as possible to the ideal value d_i^*, but does not violate either constraint. Denoting the delivery size by d_i, this amount is given by

$$d_i = \max[d_i^*, 0] + \min[0, L(t_i^-) - d_i^*] \quad \text{for} \quad i = 1, 2, \ldots, m-1. \tag{4.4}$$

The delivery size at the last retailer will simply equal the load on the truck at that time, so that V units are always delivered on a cycle. That is we set

$$d_m = L(t_m^-). \tag{4.5}$$

Notice that we could in principle allow negative deliveries, as long as there is inventory available at the retailer i and the total amount of load the truck carries as it leaves this retailer is kept under its total capacity. However, this is not necessarily an improvement over the delivery sizes found above since under our policy the truck returns empty, and so the items picked in a negative delivery would more likely end up being shifted to the last few retailers of the tour. This will not necessarily bring the state of the system closer to the optimal cycle. Hence, we disallow negative deliveries and will use the delivery sizes as defined in (4.4)-(4.5).

We have thus obtained a dynamic load allocation rule for the VRIP from the solution to the optimal cycle placement for the heavy traffic limit problem. The rule was developed based on a recalculation of the optimal cycle placement as the truck visits every retailer. We expect this to do better than the case where the optimal cycle placement (and the corresponding delivery sizes) is found only once per cycle; for instance when the truck is ready to leave the depot. Notice however that, according to the HTAP, the evolution of the retailer inventories should behave increasingly in a deterministic fashion as $\rho_T \to 1$. Therefore we expect that, in systems where the utilization is high, the advantage obtained by recalculation of the cycle placement at the retailers should be quite small (and asymptotically disappear).

We now turn our attention to second aspect of the dynamic control policy for the TSP-VRIP: the busy/idle decision. In this case, the interpretation of the heavy traffic results is quite straightforward. For this aspect of the control we consider the decision epochs to occur when the truck is at the warehouse, so that the truck always completes its tour and returns to the warehouse empty. At these points in time, the truck idles if the total inventory level is above $w_T^* = \sqrt{n}\, z_T^*$, otherwise it starts a new tour. Notice that if we substitute w_T* as just defined, and replace the parameters of the scaled system by their appropriately normalized counterparts from the original system (as done above for the cycle placement formulae) into the formulae of Proposition 2, we obtain that the optimal unscaled base stock level equals

$$w_T^* = -\frac{1}{\nu_T} \ln \left[\left(\frac{h}{b+h} \right) \left(\frac{\nu_T(\beta_T - \alpha_T)}{e^{\nu_T}(\beta_T - \alpha_T) - 1} \right) \right] + \alpha_T \tag{4.6}$$

if $w_T^* \geq \beta_T$, otherwise w_T^* is the solution to the following equation

$$\frac{2a_2}{\nu_T} e^{-\nu_T(w_T^* - \alpha_T)} + 2a_2 w_T^* + a_3 - \frac{2a_2}{\nu_T} = 0. \tag{4.7}$$

Furthermore, the optimal cost is given by

$$F_T(w_T^*) = h w_T^* + a_5$$

if $w_T^* \geq \beta_T$, and by

$$F_T(w_T^*) = a_2 (w_T^*)^2 + a_3 w_T^* + a_4$$

otherwise. The constants in these formulae are the unscaled counterparts of the ones in Chapter 2. Their values are given in terms of the original system parameters as

$$\nu_T = \frac{2(1-\rho_T)V}{\lambda \theta_T (c_d^2 + V c_T^2)},$$

$$a_3 = 2a_2 \left(\sum_i \frac{h_i V_i}{b_i + h_i} - \sum_i \lambda_i \theta_{0i}^{TSP} \right),$$

$$a_4 = a_2 \left(\sum_i \frac{h_i V_i}{b_i + h_i} - \sum_i \lambda_i \theta_{0i}^{TSP} \right)^2 + \frac{1}{2} \sum_i \frac{b_i h_i V_i}{b_i + h_i},$$

and,

$$a_5 = -h\sum_i \lambda_i \theta_{0i}^{TSP} + \frac{1}{2}\sum_i h_i V_i - \frac{1}{2}\sum_i \frac{(h_i - h)^2}{b_i + h_i} V_i.$$

Hence, the optimal base stock level is also independent of the scaling factor n.

We have thus completely characterized a dynamic control policy for the fixed Route TSP that depends exclusively on the original system parameters. The policy specifies the two controllable aspects of the system: (1) the allocation of load to retailers, through the delivery sizes d_i defined in (4.4)-(4.5); (2) the busy idle control for the truck, through the system-wide base stock level defined in (4.6)-(4.7).

As mentioned before, due to the underlying heavy traffic assumptions, we expect this policy to perform quite well for systems with high utilizations. We shall shortly use some computational experiments to gage how robust these results are when the system is not the heavily loaded. For now we turn our attention to the extraction of an operating policy for the DS system from the analytic results in Chapter 3.

4.1.2 The Fixed Route DS Case

Again we need to address the two aspects of the control separately. In this case the two decisions involve: deciding whether the truck should depart immediately or remain in the warehouse and idle; and — if the truck is to remain busy — choosing the next retailer to visit (to deliver a full load to). As for the TSP case, we address the delivery allocation among the retailers first.

The interpretation of the heavy traffic results with regard to the retailer selection follows the same principle as in the TSP case: the next delivery should be made in order to bring the inventory vector as close as possible to the optimal cycle. We shall therefore attempt to obtain an unscaled optimal cycle placement based on the state of the system. Recall from §3.3 that the heavy traffic cycle placement results were obtained assuming that the intrinsic inventory in the system was equal to the average inventory over a cycle. The first step toward the dynamic delivery allocation control will therefore be to establish a relationship between the total inventory in the system at the decision epoch and the cycle placement parameters.

Unlike the TSP case, in the DS model assuming that the decision epochs are when the truck is at the depot is not enough to obtain a single such relationship since retailers are visited with different frequency. Therefore the relationship between the inventory at the retailers and the cycle placement value x_i when the truck is at the warehouse evolves dynamically as the truck follows the polling table. This is not a major problem since, according to the HTAP, we may use a deterministic model for the inventory evolution over a cycle. Under this deterministic model, there are several possible ways in which one can define a dynamic epoch location function to relate the inventory at the current epoch to the cycle placement parameters. One way to do it is to keep a vector process $(r_1(t), r_2(t), \ldots, r_m(t))$ of the time epoch for the latest visit to each retailer. If we then denote the current time by t, the value $t - r_i(t)$ represents the time elapsed since the truck last visited retailer i. Using these definitions and the deterministic evolution over a cycle of the inventory at retailer i (see Figure 3-3), we obtain the following expression for the inventory at retailer i and decision epoch t:

$$Q_i(t) = q_i + V - \lambda_i(t - r_i(t)) \equiv q_i + u_{io}(t).$$

where $q_i = \sqrt{n}\, x_i$ are the unscaled cycle placement parameters. Summing these expressions over all retailers we obtain a candidate relationship between the total inventory at the decision epoch t and the cycle placement parameters as:

$$\sum_i q_i = Q(t) - mV + \sum_i \lambda_i(t - r_i(t)) \equiv Q(t) - u_o(t).$$

While in the deterministic evolution model each retailer would always be visited every V/λ_i time units, it is certainly possible for the inter-visit times in the actual system to exceed this nominal value (since demand and travel times are stochastic). We need to account for this possibility in the definition of $u(t)$. If the current inter-visit period is extremely long, we may have $V - \lambda_i(t - r_i(t)) < 0$. So that $u_{io}(t)$ as defined above would make the cycle placement of the retailer higher than the current inventory level — contradicting the definition of q_i as the lowest level reached by the inventory at retailer i over a cycle. In particular, by this definition, $q_i \geq Q_i(t) - \lambda_i \theta_{0i}$

if the truck is at the warehouse at time t. Therefore, define the epoch locator function as

$$u(t) = \sum_i u_i(t) = \sum_i \max\left[V - \lambda_i(t - r_i(t)), \lambda_i\theta_{0i}\right]. \tag{4.8}$$

Using (4.8) (and the appropriate scalings) in equation (3.18) we find that $q(t)$, the intrinsic inventory level as defined for our heavy traffic analysis relates to the inventory vector at time t by

$$q(t) = \sum_i Q_i(t) - u(t) + \frac{mV}{2}. \tag{4.9}$$

Again, though we shall henceforth drop its functional dependency on time, notice that q is a dynamic system state which depends on the current total inventory and the cycle evolution. We next use this value to find the corresponding optimal cycle placement.

Recall that for the heavy traffic analysis we assumed that the intrinsic inventory was equal to the average inventory over a cycle. Observe that, by using (3.18) to find q, we have related the total current inventory to the average inventory over a cycle, so that the we can use the unscaled formulae form the heavy traffic analysis to find the cycle placement directly. Therefore, at least in terms of the load allocation control, the assumption that $Z(t) = mC/2$ was innocuous. To this order, we use the appropriate scaling relations for the system parameters to obtain the formulae for the unscaled optimal cycle placement vector $q_i^* = \sqrt{n}\, x_i^*$ given q, as

Region 1. $\quad q < \alpha_D = -V \sum_i \frac{b + h_i}{b_i + h_i} + \frac{mV}{2}$

$$\begin{aligned} q_i^* &= -\frac{b + h_i}{b_i + h_i} V \text{ for } i \neq p \\ q_p^* &= q - \frac{mV}{2} + V \sum_{i \neq p} \frac{b + h_i}{b_i + h_i} \end{aligned}$$

Region 2. $\quad \alpha_D \leq q \leq \beta_D = V \sum_i \frac{h - h_i}{b_i + h_i} + \frac{mV}{2}$

$$q_i^* = \frac{2a_7 V}{h_i + b_i}\left(q - \frac{mV}{2}\right)$$

$$
\begin{aligned}
a_7 &= \frac{1}{2V}\left(\sum_i \frac{1}{b_i + h_i}\right)^{-1} \\
\text{Region 3.} \qquad & \beta_D < q \\
q_i^* &= -\frac{h_i - h}{b_i + h_i} V \ \text{ for } i \neq \ell \\
q_\ell^* &= x - \frac{mV}{2} + V \sum_i \frac{h_i - h}{b_i + h_i},
\end{aligned}
$$

which are independent of the scaling factor n.

As mentioned before, we will use this cycle placement as a reference or 'ideal' m-dimensional inventory state for a given total inventory level and epoch locator (i.e. for a give q), and choose the retailer to visit so as to bring the current inventory vector as close as possible to this reference cycle. To state this precisely, let the epoch at which the truck is ready to depart from the warehouse be t and consider the deterministic inventory evolution over the next delivery trip. In the deterministic inventory evolution, the truck will reach retailer i (if it chooses to go there next) at time $t_i = t + \theta_{0i}$. In the deterministic nominal allocation tour corresponding to the optimal cycle placement, $Q_j(t_i^+)$, the inventory at retailer j right after a delivery is made to retailer i, is given by:

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

and

$$
Q_i^*(t_i^+) = q_i^* + V,
$$

where the maximization used for the case when $j \neq i$ makes the necessary adjustment for long inter-visit times as discussed above. If we now take the actual inventory levels $Q_i(t)$ as a starting point, under the deterministic inventory evolution the inventory vector after a delivery to retailer i is given by

$$
Q_j(t_i^+) = Q_j(t) - \lambda_j \theta_{0i}, \quad \text{for } j \neq i
$$

and

$$Q_i(t_i^+) = Q_i(t) + V - \lambda_i \theta_{0i}.$$

Therefore, the resulting euclidean distance between the ideal and actual inventory level vectors after a delivery to retailer i is

$$\Delta(i) = \sqrt{\sum_j \left(Q_j(t_i^+) - Q_j^*(t_i^+)\right)^2}\,. \tag{4.10}$$

The proposed control is to send the truck to deliver at retailer k, where $k = \arg\min_i \Delta(i)$.

Finally, as in the TSP case, the busy/idle control is a direct unscaling of the heavy traffic results. The only added complexity is that for the DS case the truck visits the warehouse several times during the course of a cycle, and so has many possible idling decision epochs. Exploiting again the deterministic inventory evolution we may allow this possibility by letting the truck idle whenever $Q(t) - u(t) + \frac{mV}{2} > w_D^*$, where $w_D^* = \sqrt{n}\, z_D^*$ is the unscaled idling threshold. Notice that, using the deterministic evolution model to relate the total inventory at time t to the implied average inventory over a cycle, we have again found a solution that cancels the effects of the assumption that $Z(t) = mC/2$, which we made on Chapter 3. Making the suitable scaling substitutions into the formulae of Proposition 4, we have that

$$w_D^* = -\frac{1}{\nu_D} \ln \left[\left(\frac{h}{b+h} \right) \left(\frac{\nu_D(\beta_D - \alpha_D)}{e^{\nu_D(\beta_D - \alpha_D)} - 1} \right) \right] + \alpha_D \tag{4.11}$$

if $w_D^* \geq \beta_D$, otherwise w_D^* is the solution to the following equation

$$\frac{1}{\nu_D} e^{-\nu_D(w_D^* - \alpha_D)} + w_D^* - \frac{1}{\nu_D} - \frac{mV}{2} + V \sum_i \frac{h_i}{b_i + h_i} = 0. \tag{4.12}$$

Furthermore, the predicted optimal cost $F(w_D^*)$ is given by

$$F_D(w_D^*) = h w_D^* - h\frac{mV}{2} + \frac{V}{2}\sum_i h_i - \frac{V}{2}\sum_i \frac{(h_i - h)^2}{b_i + h_i}$$

if $w_D^* \geq \beta_D$, and by

$$F_D(w_D^*) = a_7(w_D^*)^2 + a_8 w_D^* + a_9$$

otherwise. The constants in these equations are the unscaled counterparts of the ones defined in Chapter 4. Namely,

$$\nu_D = \frac{(1-\rho_D)\lambda V}{\sigma_D^2 \sum_i \lambda_i \theta_{0i}},$$

$$a_8 = 2V a_7 \left(\sum_i \frac{h_i}{b_i + h_i} - \frac{m}{2} \right),$$

$$a_9 = a_7 \left(V \sum_i \frac{h_i}{b_i + h_i} - \frac{mV}{2} \right)^2 + \frac{V}{2} \sum_i \frac{b_i h_i}{b_i + h_i}$$

and (α_D, β_D, a_7) are as defined in the unscaled cycle placement formulae above.

We have thus obtained a dynamic control policy that determines the delivery allocation and the busy/idle decisions for the DS VRIP system in terms of the original system parameters. As in the TSP case, we expect this policy to perform quite well for systems with relatively high values of ρ_D, since the limiting control problem solved corresponds to the heavy traffic regime.

4.2 Performance Analysis of Fixed-Route VRIP Models

Besides an asymptotically optimal control policy for the fixed route VRIP, the results of §4.1 also provide a prediction for the system cost when this policy is used. This will in general be only an approximation for the true cost performance since it was obtained under a regime where individual inventories behave deterministically. Our results predict that this estimate will approach the true value as the traffic intensity approaches 1 (i.e. as the proposed policy becomes optimal). While we have not yet established how accurate these results are when the system experiences lower utilization rates, we shall nevertheless use these cost predictions to compare the performance of the two routing modes considered. Based on previous uses of heavy traffic approximations for the stochastic control of queueing systems we expect these predictions to be fairly accurate in terms of the *relative* costs of the two policies. In a later section of this chapter we shall use some numerical experiments to confirm that this is also true in our case.

We will analyze the operating cost predictions obtained from the heavy traffic analysis in order to address two issues: firstly, to underscore the importance of taking into account the stochastic nature of the VRIP when choosing a control policy; and secondly to investigate the relative performance of the two fixed route schemes considered in this thesis (TSP and DS).

Recall from the problem formulations in (2.6)-(2.12) and (3.4)-(3.10), for the TSP and DS case respectively, that the total cost for the VRIP system has two components: an inventory holding/backordering component, and a transportation component. The analysis of each of the two fixed route cases later revealed that the control could not affect the transportation cost (since it could not affect the rate at which the product could be shipped to the retailers). As a consequence, the predicted cost functions F_T, F_D in §4.1 represent only the inventory component of the total system cost. Denote a generic fixed routing scheme by $\Re$, and by $C(\Re)$ the total cost for the system under this scheme. To obtain this cost we need to add the transportation cost (or equivalently subtract the idleness reward) to the corresponding inventory cost expressions in §4.1. That is, we set

$$C(\Re) = F_{\Re}(w^*_{\Re}) - f(1 - \rho_{\Re}). \tag{4.13}$$

Note that this is true for a general fixed route system and not just for the $\Re =$ TSP and $\Re =$ DS cases which we have considered in this thesis.

Careful examination of (4.13) and of the inventory cost functions in §4.1 helps illustrate why stochastic modeling is of great importance for the analysis of the VRIP, and why the deterministic models which have been so successful in pure routing problems may fail in this context. Notice in particular that the second term in the RHS of equation (4.13), which represents the transportation cost, does not depend at all on the stochastic nature of the system. The steady state transportation cost depends only on the utilization rate $\rho_{\Re}$ for the given route, which in turn is a function only of the first moments of the tour completion and demand arrival processes. Hence, if transportation costs are all important, a deterministic model is indeed the right approach to the problem.

The situation is quite different for the first term in the RHS of (4.13), which represents the inventory related cost. As should be clear form the analyses in Chapters 2 and 3, the optimal base stock level depends on the variance of both the demand arrival processes (which we assumed are independent of the route) *and* the tour completion times. Therefore, for any instance of the VRIP in which the inventory costs are at least of the same order of magnitude as the transportation costs, a deterministic model will fail to set the correct inventory levels and might give a solution far away from the true optimal.

Concentrate for now on the transportation cost for the TSP and DS cases. It is interesting to observe that the DS policy will have an advantage over TSP in terms of this cost component. In particular, if we denote by $R_{\Re}$ the steady state expected idleness reward per unit time for routing scheme $\Re$, the additional transportation cost incurred by the TSP over the DS policy is

$$R_T - R_D = -f\left[(1-\rho_T) - (1-\rho_D)\right] = -f(\rho_D - \rho_T).$$

We have the following

Proposition 5. *As long as* $\theta_{ij} = \theta_{ji}$, $R_T - R_D \geq 0$.

Proof: We need to show that $\rho_D \leq \rho_T$. By definition of the traffic intensities this is equivalent to showing that

$$2\sum_i \lambda_i \theta_{0i} \leq \lambda\left(\sum_{j=0}^{m-1} \theta_{j,j+1} + \theta_{m0}\right). \tag{4.14}$$

It is well known that the length of the shortest tour that starts at node 0, visits all other sites and returns there is bounded from bellow by the distance of the longest direct route to the retailers, that is

$$\sum_{j=0}^{m-1} \theta_{j,j+1} + \theta_{m0} \geq 2 \max_k \theta_{0k} \tag{4.15}$$

(this is a direct consequence of the triangle inequality, which requires $\theta_{ij} = \theta_{ji}$). Relaxing the maximization on the RHS of (4.15), it is seen that

$$\sum_{j=0}^{m-1} \theta_{j,j+1} + \theta_{m0} \geq 2\theta_{0i}, \quad \text{for all} \quad i = 1, \ldots, m. \tag{4.16}$$

Multiplying both sides of (4.16) by λ_i and then adding up over all the retailers proves (4.14). ■

The result in Proposition 5 is quite interesting since, at first glance, one might expect the TSP route to provide the best transportation cost. However, as the previous analysis reveals, minimization of the route length is not equivalent to minimization of the steady state transportation cost in the VRIP context. Rather, what one needs to do is increase the distribution efficiency of the system (i.e. maximize the amount of items delivered per unit time traveled). In this context, full load direct shipping provides the highest transportation efficiency of *any* fixed route scheme.

One consequence of the transportation dominance of the DS policy is that as the demand rate increases the DS policy will also dominate the TSP tours in terms of the inventory cost. This is the case since, for any system where the retailers are not all in the same geographic location, $\rho_D < 1$ when $\rho_T = 1$. That is, there exist some levels of traffic intensity where the DS would be stable while the TSP would not. Based on this observation, note that as $\rho_T \to 1$, $\nu_T \to 0$. This has two implications: first in the limit as $\rho_T \to 1$, w_T^* as given by the closed form expression (4.6) grows without bound, and so it is the correct value to use as a base stock; second, $F_T(w_T^*)$ also grows without bound as utilization approaches 1, and hence will eventually become higher than the $F_D(w_D^*)$ which is finite when $\rho_T = 1$.

We should note that $\rho_{\Re} < 1$ is a necessary condition for stability of any fixed route scheme $\Re$ but that 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 our statement about DS trips dominating TSP tours as $\rho_T \to \infty$ holds as long as some form of stable dynamic allocation is used in the DS case.

Another consequence of the higher transportation efficiency of the DS scheme is that it will be preferred to the TSP policy if the transportation cost is high enough. In particular if we let

$$f_c = \frac{F_T(w_T^*) - F_D(w_D^*)}{\rho_T - \rho_D}$$

then for any $f > f_c$ the DS policy provides the best overall cost. While the value of f_c may be found numerically for any particular problem instance, a more precise characterization of f_c 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. Part of the reason is that the base stock levels, and hence the predicted inventory costs in §4.1.1 and 4.1.2 are not in closed form (see equations (4.6)-(4.7) and (4.11)-(4.12) respectively). Besides, even if one assumes that $w_{\Re}^* > \beta_{\Re}$ holds, the closed form expression for $F_{\Re}$ will still depend on $\sigma_{\Re}^2$ the variance of the BM, which in turn depends on the variance of the cycle time under routing scheme $\Re$. While the expected travel times will observe the triangle inequality and so the drift of the DS system will always be higher than the TSP case, any ordering is possible between the variance of these systems. Nevertheless, something ca be said about the relative performance of the TSP and DS policies as the inventory cost parameters are taken to their limits.

In order 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. Namely we assume that $h_i = h$ and $b_i = b$ for all $i = 1, \ldots, m$. We shall analyze the behavior of the inventory cost for the TSP and DS cases as b becomes large. First observe that the value of w_T^* and w_D^* in (4.6) and (4.11) is increasing in b/h. One therefore 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, they themselves have closed form expressions. Namely,

$$b_T = h\left[\frac{\nu_T V e^{\nu_T V}}{e^{\nu_T V} - 1} - 1\right], \tag{4.17}$$

and

$$b_D = h\left[\frac{\nu_D m V e^{\nu_D m V}}{e^{\nu_D m V} - 1} - 1\right]. \tag{4.18}$$

Consider now the difference between the closed-form inventory costs for the TSP and DS policies which, for the symmetric cost case, can be expressed as

$$\begin{aligned} F_T(w_T^*) - F_D(w_D^*) &= h\left(\frac{\nu_D - \nu_T}{\nu_D \nu_T}\ln\left[1+\frac{b}{h}\right] + \frac{1}{\nu_T}\ln\left[\frac{e^{\nu_T V}-1}{\nu_T V}\right]\right. \\ &\left. - \frac{1}{\nu_D}\ln\left[\frac{e^{\nu_D m V}-1}{\nu_D m V}\right] + \frac{(m-1)V}{2}\right). \end{aligned} \tag{4.19}$$

Clearly, when $F_T(w_T^*) - F_D(w_D^*) > 0$ the DS policy provides a better inventory cost than the TSP, and vice versa. Notice that as $b \to \infty$ the value of (4.19) is dominated by the term

$$\frac{\nu_D - \nu_T}{\nu_D \nu_T}\ln\left[1+\frac{b}{h}\right]$$

whose sign will be the same as the sign of $\nu_D - \nu_T$. Define the critical value b_c as

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

where $\exp[x] = e^x$. Then, for $b > \max\{b_T, b_D, b_c\}$, the DS policy provides the best inventory cost when $\nu_D - \nu_T > 0$, while the TSP routing will dominate the inventory cost for this range of b in the cases where $\nu_D - \nu_T < 0$. Recall that $\nu_{\Re}$ is the attenuation constant for the exponential steady state distribution of the RBM limit for the total intrinsic inventory under the routing scheme $\Re$. Hence, in the limit, the policy that concentrates more of its density close to the idling threshold dominates. Notice that $\nu_D - \nu_T < 0$ requires $\sigma_D^2 > \sigma_T^2$, since the TSP policy always has a slower drift than DS.

This is a rather unexpected result that again highlights the influence of the stochastic nature of the system in the inventory costs. Note in particular that, since it makes smaller and more frequent deliveries to each retailer, the TSP policy might be expected to outperform the DS scheme in terms of the inventory cost at low utilization levels. Yet, as the preceding analysis showed, the policy that provides the best inventory cost for large backorder penalties is determined by the effect of the routing schemes on the variance of the system. We should note that, for the case of deterministic travel times $\sigma_D^2 = \sigma_T^2$, and so DS dominates in the limit.

4.3 Simulation Experiments

In order to validate the solutions obtained for the VRIP by our heavy traffic analyses, we resort to a Monte Carlo simulation. Several experiments were performed in order to test the results under a variety of conditions. The main purpose of the experiments reported here will thus be to confirm that the policies obtained from the analyses in Chapters 2 and 3 perform well over a reasonable range of utilization levels (and not only when the traffic intensity is very close to 1).

The following assumptions were used in all the simulation experiments:

1. There are 5 retailers in the system.

2. Demand arrivals follow a Poisson process. Different values of the total arrival rate λ are used in the experiments to obtain higher utilization rates. However, the fraction of demand represented by retailer i is fixed. In particular, we set

$$\lambda_1 = \frac{\lambda}{5}, \ \lambda_2 = \frac{\lambda}{10}, \ \lambda_3 = \frac{\lambda}{10}, \ \lambda_4 = \frac{\lambda}{5}, \ \text{and} \ \lambda_5 = \frac{2\lambda}{5}.$$

3. The travel time random variables T_{ij} are distributed like second order Erlang,

independent of each other and of past travel times. The travel times are adjusted in order to consider several truck sizes without affecting the delivery rate of the system. To achieve this, the mean TSP tour time θ_T is set so that $10\theta_T = V$.

4. Since for a fixed route scheme the control does not affect the transportation cost, we leave the transportation cost rate unspecified (equivalently, we set $f = 0$) and concentrate on the inventory cost component.

The first set of simulation runs was performed in order to test the cost improvement obtained under the TSP policy from recalculation of the cycle placement at the retailers as opposed to determining these values only once per cycle; when the truck is at the warehouse. Recall from §2.3 that we expect this improvement to be small for systems that have high utilization rates (since the higher the utilization, the closer we are to the time scale decomposition implied by the HTAP). One might also expect this improvement to be more dramatic for big truck sizes and long inter-retailer travel times. With this in mind, we consider the following travel time structure (which corresponds to the pentagon topology illustrated in Figure 4-2):

$$\theta_{01} = \theta_{50} = \frac{\theta_T}{10}, \quad \theta_{12} = \theta_{23} = \theta_{34} = \theta_{45} = \frac{\theta_T}{5}.$$

To complete the characterization of the system, we take the cost structure to be symmetric over the retailers. Namely, we let $b_i = b = 5$, $h_i = h = 1$ for all $i = 1, \ldots, 5$.

We considered a total of nine cases under this setup, corresponding to three different vehicle sizes $(100, 10, 5)$ and three demand arrival rates $(5, 7, 9)$. For the cases with $\lambda \leq 8$, we simulated three replications of 36,000 time units each with cycle placement recalculation at the retailers and three more with calculation only at the warehouse. For those instances where $\lambda > 8$, the length of each replication was increased to 240,000 time units in order to reduce the variance of the results. This allowed us to keep the standard deviation of the cost estimate under 1% of its mean. Hence, roughly speaking, differences of 2% or more may be taken as significant. In all cases we used the base stock level and cycle placement formulae from §4.1.1.

Table 4.1 summarizes the results of the experiment. The entries in the table

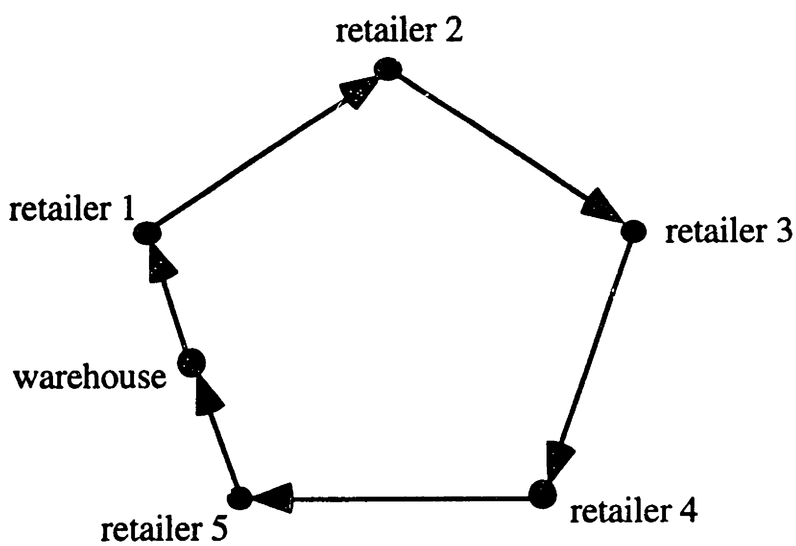

Figure 4-2: TSP Tour on Pentagon Topology For Simulation Experiments.

	$\lambda = 5.0$ $\rho_T = 0.5$	$\lambda = 7.0$ $\rho_T = 0.7$	$\lambda = 9.0$ $\rho_T = 0.9$
$V = 100$	4.0%	5.0%	1.1%
$V = 10$	7.4%	18.6%	0.7%
$V = 5$	7.3%	18.1%	1.1%

Table 4.1: Experiment 1. Increase in Cost for Placement Calculation only at Depot

represent the increase in the long-run average inventory cost when calculation of the delivery sizes is done only at the warehouse, and not adjusted over the course of the tour. Notice that, as expected, the advantage obtained by recalculation at the retailers becomes quite small when the system is subject to a high traffic intensity (on the order of 1% when $\rho_T = 0.9$). On the other hand, this effect does not appear to be sensitive to the truck size. These results certainly support our conjecture that the HTAP applies for the VRIP. Even though the advantage from cycle placement recalculation at the retailers is small at high values of ρ_T, we shall still use this policy in all subsequent TSP simulations.

For the second simulation experiment we keep the same set-up as in the first with the exception that we shall also consider the case of asymmetric costs. In particular, we consider the following inventory cost rates:

$$h_1 = h_2 = 1, \quad h_3 = h_4 = h_5 = 2,$$

and

$$b_1 = 5, \quad b_2 = 10, \quad b_3 = 5, \quad b_4 = 10, \quad b_5 = 5.$$

For each of these eighteen cases (3 Truck sizes, 3 arrival rates and 2 cost structures) we performed an exhaustive search in a series of simulations (each consisting of three replications with the length described in Experiment 1) in order to find the base stock level that provides the best cost for the system.

Table 4.2 summarizes the results of the second simulation experiment. The entries in the table represent the suboptimality (within the class of base stock policies) incurred by using the base stock level obtained from the formulae in §4.1.1 instead of the optimal base stock level found by exhaustive search. Notice that the base stock level found by the heavy traffic analysis performs remarkably well even at low traffic intensities, and that, as expected, the suboptimality of the results decreases dramatically as $\rho_T \rightarrow 1$. This holds true for both the symmetric and asymmetric cost cases. The average suboptimality is 5.1% for the symmetric cost cases, and it is 3.8% for th uns with asymmetric cost. Even more telling is the fact that, out of the eighteen combinations considered, only four cases have a suboptimality that exceeds 5%. Furthermore, though it is not evident from the data as reported here, in those cases where the suboptimality does exceed 5%, the error in absolute terms is not so substantial (since the inventory cost of systems with small utilizations is quite small).

In the third simulation experiment we study the performance of the control policies of §4.1.2 for the DS case, and compare it to the performance of the TSP policy on the same system. As before, this is done by comparing the long run average cost obtained when using the base stock levels given by (4.11)-(4.12) with that of the optimal base stock level found by exhaustive search. As it turns out, the DS policy has a huge drift advantage over TSP in the pentagon topology used for experiments 1 and 2. Therefore, in order to obtain a meaningful cost comparison, the third experiment was

		$\lambda = 5.0$ $\rho_T = 0.5$	$\lambda = 7.0$ $\rho_T = 0.7$	$\lambda = 9.0$ $\rho_T = 0.9$
$V = 100$	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%

Table 4.2: Suboptimality of H.T. Base Stock for Pentagon Topology TSP VRIP

done over a different topology. Specifically, we assume that the travel times satisfy

$$\theta_{0i} = \frac{9\theta_T}{20}, \quad \text{for } i = 1, \ldots, 5, \quad \text{and} \quad \theta_{12} = \theta_{23} = \theta_{34} = \theta_{45} = \frac{\theta_T}{40}.$$

These travel times correspond to the wedge structure illustrated in Figure 4-3. Notice that the traffic intensity advantage of the DS policy is proportional to the ratio of the inter-retailer to the warenouse-retailer travel times which, for the simulation data, is given by:

$$\frac{\rho_D}{\rho_T} = \frac{2\theta_{01}}{2\theta_{01} + (m-1)\theta_{12}} = 0.9.$$

The rest of the simulation parameters are left as in the symmetric cost case 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 since it corresponds to $\rho_T = 1$, and so the system is not stable under this scheme. Hence we consider a dozen cases (4 traffic intensities and 3 vehicle sizes) for the DS policy and 9 cases for the TSP.

Tables 4.3 and 4.4 summarize the results of this experiment. The entries of Table 4.3 represent the cost degradation incurred by using the base stock level obtained from the formulae in §4.1.1 instead of the optimal base stock level found by exhaustive search for *the same routing scheme.* In other words, the data in this table compare the cost obtained under the best base stock level for each policy (found by exhaustive search) with the performance of the proposed base stock level of the same policy. For the TSP case the figures show an average suboptimality (within the class of base

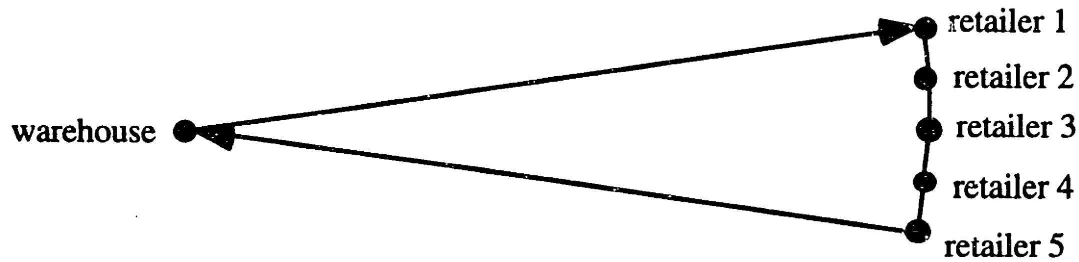

Figure 4-3: TSP Tour on Wedge Topology For Simulation Experiments.

		$\lambda = 5.0$, $\rho_T = 0.50$, $\rho_D = 0.45$	$\lambda = 7.0$, $\rho_T = 0.70$, $\rho_D = 0.63$	$\lambda = 9.0$, $\rho_T = 0.90$, $\rho_D = 0.81$	$\lambda = 10.0$, $\rho_T = 1.00$, $\rho_D = 0.90$
$V = 100$	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%

Table 4.3: Same-class Suboptimality of Proposed DS and TSP Policies on Wedge Topology

stock-controlled TSP policies) for the proposed base stock level of 5.2%, and only three of the nine cases have a suboptimality of more than 6%. The average suboptimality for the proposed DS policy (within the class of base stock-controlled DS policies) is 2.1%, and it is never higher than 5%. These results confirm again that the heavy traffic results are quite accurate over a broad range of parameters, for both the DS and TSP policies.

Table 4.4 presents a comparison of the inventory cost for the DS and TSP policies. All entries in the table represent the percentage difference between the TSP cost and the DS cost. That is,

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

		$\lambda = 5.0$ $\rho_T = 0.50$ $\rho_D = 0.45$	$\lambda = 7.0$ $\rho_T = 0.70$ $\rho_D = 0.63$	$\lambda = 9.0$ $\rho_T = 0.90$ $\rho_D = 0.81$	$\lambda = 10.0$ $\rho_T = 1.00$ $\rho_D = 0.90$
$V = 100$	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%	∞

Table 4.4: Inventory Cost Comparison $\frac{\text{TSP-DS}}{\text{DS}}$: Wedge Topology

The entries labeled 'Sim.' represent the percentage difference between the best TSP inventory cost and the corresponding DS case from the simulation experiment (found by exhaustive search over base stock levels in both cases). Notice that for low utilization levels and large truck sizes the TSP policy enjoys a considerable advantage over the DS scheme in terms of the inventory cost. This advantage erodes as the traffic intensity increases. Of course, for the cases where $\rho_T = 1$ (see last column of Table 4.4), the TSP cost becomes unbounded and the DS policy is obviously preferred. For their part, the entries labeled 'Pred.' represent the difference between the inventory costs predicted by the optimal base stock formulae in §4.1.1 and §4.1.2 for the TSP and DS cases respectively. These values are quite good estimates for the relative inventory cost performance found by simulation. In all but one case the heavy traffic model would identify the policy that dominates the inventory cost correctly. In the case where the prediction errs in the sign of the percentage difference (when $\lambda = 9$ and $V = 10$), it still correctly predicts that the costs for both policies are very close.

Finally, we address the question of whether the performance of the system is much improved by using the proposed control, as opposed to some direct extension of the deterministic models. Since the proposed policy has two components: the delivery allocation control and the busy/idle threshold control; we shall consider the system's sensitivity to each of this aspects separately.

In terms of the deliverv allocation control, a deterministic model for the TSP-

VRIP would provide a fixed delivery size for each site while the solution for the delivery allocation control in the DS-VRIP case would take the form of a fixed polling table. However, fixed delivery allocations perform very badly in a stochastic environment. In fact, if deliveries are set without regard to the system state then, as time goes on, some sites will experience high backordering levels while other sites will be holding large inventories. This is so because, even though the time average demand at retailer i over a period of n time units will converge to λ_i as $n \to \infty$ by the law of large numbers, the deviation from the mean $D_i(n) - n\lambda_i$ grows as $\sqrt{n}$ by the FCLT. To verify that these policies will do poorly we simulated a few instances of the TSP and DS VRIP with fixed delivery allocations. Again, we searched over different values of the base stock to find the one that provided the best cost performance. In all cases, the best cost for the fixed allocation was at least 10 times higher than the cost for the proposed dynamic allocation policy. Hence we conclude that a closed loop (i.e. state dependent) delivery allocation control is essential to achieve a good performance in the stochastic VRIP.

We next address the question of the sensitivity of the system performance to the busy/idle control, by looking at the increase in cost incurred by using a base stock level different from the one proposed in the formulae of §4.1.1 and §4.1.2. We already have the required data for this analysis from the exhaustive search performed in the simulation experiments above.

Figure 4-4 plots three examples of the increase in cost with respect to the proposed policy, as a function of the base stock level (expressed in vehicle sizes). 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. Three characteristics worth noting are: (1) the inventory cost is convex in the base stock level; (2) the cost performance remains relatively constant over a range of approximately one vehicle size around the optimal base stock level; (3) once the base stock level moves beyond this range in either direction the cost performance deteriorates rapidly.

These results illustrate the relevance of our model for the VRIP and should help

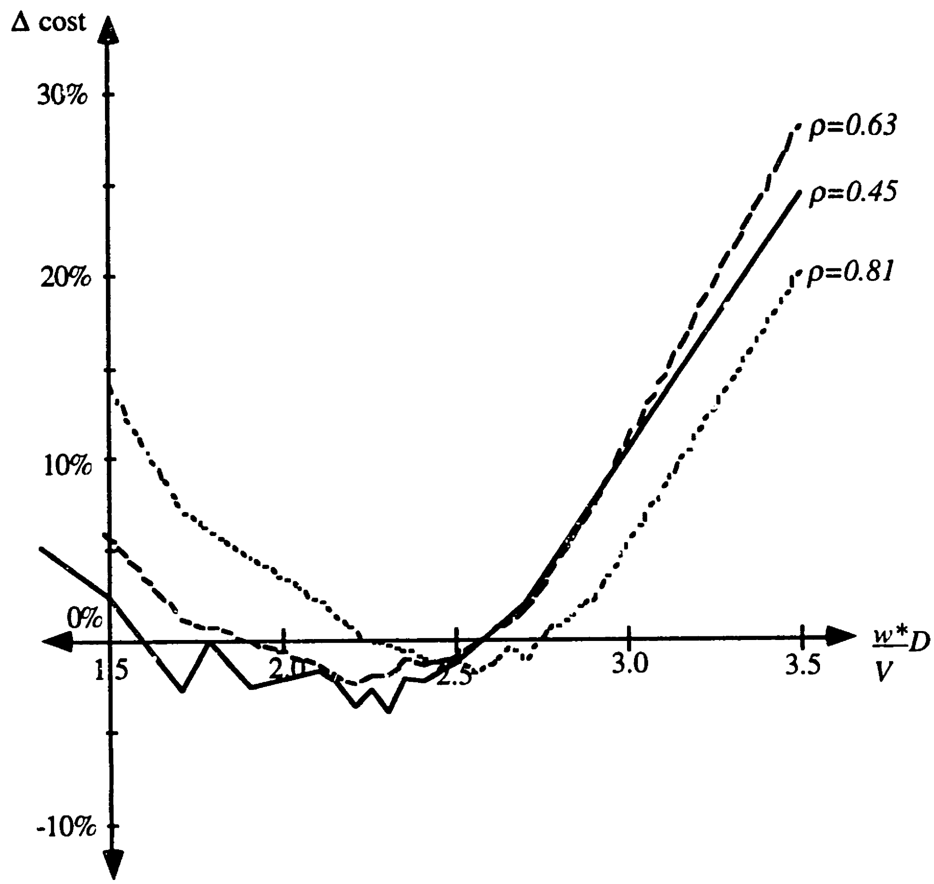

Figure 4-4: Example of Sensitivity of Inventory Cost To Base Stock Level.

justify the rather impressive machinery invoked to solve it. In particular, the system performance is much deteriorated in the absence of an adequate dynamic delivery allocation or base stock level. As mentioned before, stochastic modeling is necessary to obtain an adequate base stock level for the system. Also, the optimal base stock level will depend on the load allocation rule. The heavy traffic analyses of Chapters 2 and 3 allowed us to obtain a tractable approximation for the stochastic VRIP which yields a complete characterization of these two controls.

Chapter 5

Heavy Traffic Analysis of the VRIP with Dynamic Routing

5.1 Problem Formulation

Consider now a situation where the manager of the distribution system has the option to select, at every point in time when the vehicle is about to leave the warehouse, among the two routing schemes analyzed in Chapters 2 and 3. That is, once the truck is loaded at the depot, it can either be sent to do a full load TSP tour, or one direct shipment of a complete vehicle capacity (V) to some retailer. All other aspects of the problem (e.g. sources of uncertainty, cost structure) remain the same as in the fixed route cases. We shall take advantage of this and retain many of the definitions and notation from the corresponding static cases.

The service to the retailers in this system will be characterized by two counting processes $(S_T(t), S_D(t))$. The process $S_T(t)$ represents the cumulative number of TSP tours completed by the truck, assuming it has been continuously active and using the TSP route over the period $[0, t]$. Similarly, $S_D(t)$ counts the DS cycle completions, assuming the truck has been performing direct deliveries without idling in $[0, t]$. Recall from §3.1 that a DS service completion occurs after $K \sum_i \omega_i$ visits to the retailers. Here, as in the fixed route case, $\omega_i = \lambda_i / \lambda_\ell$ is the relative frequency of visits to retailer i, $\lambda_\ell = \min_i \lambda_i$, and K is a scaling parameter chosen so that all $K\omega_i$ are integer.

The routing control will be expressed as the cumulative time-allocation processes $(T_T(t), T_D(t))$, which respectively represent the amount of time over the interval $[0, t]$ that the delivery truck has spent operating under the TSP and DS policy. The cycle completion epochs for the TSP and DS routing scheme will depend on these controls, and are given by

$$\tau_k^T = \inf\{t \mid S_T(T_T(t)) \geq k\},$$

and

$$\tau_k^D = \inf\{t \mid S_D(T_D(t)) \geq k\}$$

respectively.

As before, the manager of this system is also allowed to exercise a delivery allocation control. This control takes different forms depending on the routing scheme being used. Let $\varepsilon_i^T(t)$ represent the load allocation control for the TSP policy. This process is defined in the same way as in the corresponding fixed route model. Namely, $\varepsilon_i^T(t)$ increases by the delivery amount when retailer i is visited *during the course of a TSP tour*, and the decreases by the nominal delivery amount V_i at the TSP tour completion epochs τ_k^T. We continue to assume that the truck always delivers a full load over the course of a tour, and therefore ε_i^T must satisfy

$$\begin{aligned}
\varepsilon_i^T(0) &= 0, \text{ for all } i && (5.1)\\
\varepsilon_i^T(t) &\geq \varepsilon_i^T(\tau_{k-1}^T), \text{ for } t \in (\tau_{k-1}^T, \tau_k^T) \text{ and all } i && (5.2)\\
\sum_i \varepsilon_i^T(\tau_k^{T-}) &= V, && (5.3)\\
\sum_i \varepsilon_i^T(\tau_k^T) &= 0, && (5.4)
\end{aligned}$$

where τ_k^T is the time of the k-th TSP tour completion, and τ_k^{T-} is the epoch an infinitesimal amount of time before this completion.

With regard to the DS policy, the delivery allocation control is denoted by $\varepsilon_i^D(t)$. As in the fixed route DS model, $\varepsilon_i^D(t)$ increases by V every time a direct shipping is made to retailer i, and decreases by $K\omega_i V$ whenever a DS cycle is completed. Hence, by definition of the DS cycle, the delivery allocations must satisfy

$$\varepsilon_i^D(0) = 0, \text{ for all } i \tag{5.5}$$

$$\varepsilon_i^D(t) \geq \varepsilon_i(\tau_{k-1}^D), \text{ for } t \in (\tau_{k-1}^D, \tau_k^D) \text{ and all } i \tag{5.6}$$

$$\sum_i \varepsilon_i(\tau_k^{D-}) = K\frac{\lambda}{\lambda_\ell}V, \tag{5.7}$$

$$\sum_i \varepsilon_i(\tau_k^D) = 0, \tag{5.8}$$

where τ_k^D is the time of the k-th polling cycle completion, and τ_k^{D-} is the epoch an infinitesimal amount of time before this completion. Notice that, with these definitions for the delivery allocation controls, $\varepsilon_i^T(t) + \varepsilon_i^D(t)$ represents the cumulative deviation from the nominal allocation over past service cycles plus the amount delivered over the current cycle at retailer i.

Under this set-up, and assuming that $Q_i(0) = 0$, the dynamics of individual inventory levels in this system are given by

$$Q_i(t) = V_iS_T(T_T(t)) + VS_D(T_D(t)) - D_i(t) + \varepsilon_i^D(t) + \varepsilon_i^T(t) \text{ for } t \geq 0. \tag{5.9}$$

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

$$I(t) = t - T_T(t) - T_D(t) \text{ for } t \geq 0. \tag{5.10}$$

We then have that the control $(T_T(t), T_D(t), \varepsilon_i^T(t), \varepsilon_i^D(t))$ must satisfy

$$T_T, T_D \text{ are nondecreasing continuous with } T_T(0) = T_D(0) = 0 \tag{5.11}$$

$$T_T, T_D, \varepsilon_i^T, \varepsilon_i^D \text{ are nonanticipating with respect to } Q \tag{5.12}$$

$$I \text{ is nondecreasing with } I(0) = 0 \tag{5.13}$$

The dynamic routing VRIP may thus be formulated as finding the routing control

$(T_T(t), T_D(t))$ and its accompanying delivery allocation process $(\varepsilon_i^T(t), \varepsilon_i^D(t))$ to

$$\min \limsup_{T\to\infty} \frac{1}{T} E\left[\int_0^T \sum_i \left(h_i\{Q_i(t)\}^+ + b_i\{Q_i(t)\}^-\right) dt - fI(T)\right] \tag{5.14}$$

subject to (5.1) - (5.13). As each of the fixed route cases analyzed before corresponds to a special case of this dynamic routing formulation, it should come as no surprise that the resulting control problem (5.1) - (5.14) does not seem to be tractable in its exact form. And so, we shall use the same heavy traffic machinery called upon for the fixed route cases to solve a limiting problem and gain insight into this instance of the VRIP.

5.2 Heavy Traffic Normalizations and an Averaging Principle

5.2.1 Normalizations and Diffusion Limit for the System Netput

The first step in our development of a heavy traffic limit version for the dynamic VRIP is to define a sequence of systems indexed by n. We make the same definitions as in the static cases for the n-th system parameters that we use to define the partial service completion process. Namely, we let

$$C^{(n)} = \frac{V^{(n)}}{\sqrt{n}}, \tag{5.15}$$

$$\vartheta_T^{(n)} = \frac{\theta_T^{(n)}}{\sqrt{n}}, \tag{5.16}$$

$$\vartheta_D^{(n)} = \frac{\theta_D^{(n)}}{\sqrt{n}}, \tag{5.17}$$

$$\varsigma_{TS}^{2\,(n)} = \frac{\sigma_{TS}^{2\,(n)}}{\sqrt{n}}, \tag{5.18}$$

$$\varsigma_{DS}^{2\,(n)} = \frac{\sigma_{DS}^{2\,(n)}}{\sqrt{n}}. \tag{5.19}$$

As in the fixed route cases, we next define the centered partial service and demand processes for the n-th system as

$$\bar{S}_T^{(n)}(t) \equiv S_T^{(n)}(t) - \frac{t}{\vartheta_T^{(n)}},$$

$$\bar{S}_D^{(n)}(t) \equiv S_D^{(n)}(t) - \frac{t}{\vartheta_D^{(n)}},$$

and

$$\bar{D}^{(n)}(t) \equiv D^{(n)}(t) - \lambda^{(n)} t.$$

In order to characterize the influence of the routing control on the total netput, define the cumulative fraction of busy time that the DS service has been used as

$$\delta^{(n)}(t) = \frac{T_D^{(n)}(t)}{T_T^{(n)}(t) + T_D^{(n)}(t)}.$$

We are now in a position to define the netput for the n-th dynamic routing VRIP system as

$$\begin{aligned}\chi^{(n)}(t) &= \sqrt{n}\left(\frac{C^{(n)}}{\vartheta_T^{(n)}}(1-\delta^{(n)}(t)) + \frac{\lambda^{(n)} C^{(n)}}{2\sum_i \lambda_i^{(n)} \vartheta_{0i}^{(n)}}\delta^{(n)}(t) - \lambda^{(n)}\right) t \\ &\quad + C^{(n)}\left[\bar{S}_T^{(n)}(T_T^{(n)}(t)) + \frac{K^{(n)}\lambda^{(n)}}{\lambda_\ell^{(n)}}\bar{S}_D^{(n)}(T_D^{(n)}(t))\right] - \bar{D}^{(n)}(t). \qquad (5.20)\end{aligned}$$

Summing the equations in (5.9) over the retailers, and substituting the relevant definitions into (5.20) we obtain the following expression for the total inventory in terms of the netput and controls:

$$Q^{(n)}(t) = \chi^{(n)}(t) - \left(\frac{C^{(n)}}{\vartheta_T^{(n)}}(1-\delta^{(n)}(t)) + \frac{\lambda^{(n)} C^{(n)}}{2\sum_i \lambda_i^{(n)} \vartheta_{0i}^{(n)}}\delta^{(n)}(t)\right) I^{(n)}(t) + \xi^{(n)}(t), \quad (5.21)$$

where

$$\begin{aligned}\xi^{(n)}(t) &= \xi_T^{(n)}(t) + \xi_D^{(n)}(t) \\ \xi_T^{(n)}(t) &= \varepsilon_T^{(n)}(t) + V^{(n)} S^{(n)}(T^{(n)}(t)) - C^{(n)} S^{(n)}(T^{(n)}(t))\end{aligned}$$

$$\xi_D^{(n)}(t) \;=\; \varepsilon_D^{(n)}(t) + \frac{K^{(n)}\lambda^{(n)}V^{(n)}}{\lambda_\ell^{(n)}} S_D^{(n)}(T_D^{(n)}(t)) - \frac{K^{(n)}\lambda^{(n)}C^{(n)}}{\lambda_\ell^{(n)}} \mathcal{S}_D^{(n)}(T_D^{(n)}(t)).$$

We next define the normalized processes for the n-th system according to the following expressions:

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

$$W^{(n)}(t) = \sum_i W_i^{(n)}(t) = \frac{Q^{(n)}(nt)}{\sqrt{n}},$$

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

$$X^{(n)}(t) = \frac{\chi^{(n)}(nt)}{\sqrt{n}},$$

$$\hat{D}^{(n)}(t) = \frac{\bar{D}^{(n)}(nt)}{\sqrt{n}},$$

$$\hat{\xi}^{(n)}(t) = \frac{\xi^{(n)}(nt)}{\sqrt{n}},$$

$$\hat{\mathcal{S}}_T^{(n)}(t) = \frac{\bar{\mathcal{S}}_T^{(n)}(nt)}{\sqrt{n}}, \quad \hat{\mathcal{S}}_D^{(n)}(t) = \frac{\bar{\mathcal{S}}_D^{(n)}(nt)}{\sqrt{n}},$$

$$\tau_T^{(n)}(t) = \frac{T_T^{(n)}(nt)}{n}, \quad \tau_D^{(n)}(t) = \frac{T_D^{(n)}(nt)}{n},$$

and

$$\hat{\delta}^{(n)}(t) = \delta(nt).$$

The expressions for the dynamic behavior of the n-th normalized system are found by applying these scalings into equations (5.20) and (5.21). From this procedure we obtain an expression for the normalized system netput process $X^{(n)}(t)$ as

$$\begin{aligned} X^{(n)}(t) \;=\; & \sqrt{n}\left(\frac{C^{(n)}}{\vartheta_T^{(n)}}(1-\hat{\delta}^{(n)}(t)) + \frac{\lambda^{(n)}C^{(n)}}{2\sum_i \lambda_i^{(n)}\vartheta_{0i}^{(n)}}\hat{\delta}^{(n)}(t) - \lambda^{(n)}\right)t \\ & + C^{(n)}\left[\hat{\mathcal{S}}_T^{(n)}(\tau_T^{(n)}(t)) + \frac{K^{(n)}\lambda^{(n)}}{\lambda_\ell^{(n)}}\hat{\mathcal{S}}_D^{(n)}(\tau_D^{(n)}(t))\right] - \hat{D}^{(n)}(t), \quad (5.22) \end{aligned}$$

while the normalized inventory can be expressed as

$$W^{(n)}(t) = X^{(n)}(t) - \left(\frac{C^{(n)}}{\vartheta_T^{(n)}}(1 - \hat{\delta}^{(n)}(t)) + \frac{\lambda^{(n)} C^{(n)}}{2\sum_i \lambda_i^{(n)} \vartheta_{0i}^{(n)}} \hat{\delta}^{(n)}(t) \right) Y^{(n)}(t) + \hat{\xi}^{(n)}(t). \tag{5.23}$$

We find the required limit by letting the scaling index $n \to \infty$, subject to the following heavy traffic conditions on the parameters of the system:

$$\lim_{n\to\infty} C^{(n)} = C = O(1) \tag{5.24}$$
$$\lim_{n\to\infty} \vartheta_T^{(n)} = \vartheta_T = O(1) \tag{5.25}$$
$$\lim_{n\to\infty} \vartheta_D^{(n)} = \vartheta_D = O(1) \tag{5.26}$$
$$\lim_{n\to\infty} \varsigma_{TS}^{(n)} = \varsigma_{TS} = O(1) \tag{5.27}$$
$$\lim_{n\to\infty} \varsigma_{DS}^{(n)} = \varsigma_{DS} = O(1) \tag{5.28}$$
$$\lim_{n\to\infty} \lambda^{(n)} = \lambda \tag{5.29}$$
$$\lambda \vartheta_T = C \tag{5.30}$$
$$\sqrt{n}\left(\frac{C^{(n)}}{\vartheta_T^{(n)}} - \lambda^{(n)} \right) = \mu_T = O(1), \quad \text{for all } n. \tag{5.31}$$
$$\sqrt{n}\left(\frac{\lambda^{(n)} C^{(n)}}{2\sum_i \lambda_i^{(n)} \vartheta_{0i}^{(n)}} - \lambda^{(n)} \right) = \mu_D = O(1), \quad \text{for all } n. \tag{5.32}$$

Taken together, conditions (5.31) and (5.32) require the traffic intensity under both the TSP and DS policies to approach 1 in the limit. In terms of the problem data, this requires the retailers to be located fairly close together relative to their distance to the warehouse. More precisely, the average travel times should satisfy

$$\frac{\vartheta_T - 2\vartheta_{0i}}{\vartheta_T} = O\left(\frac{1}{\sqrt{n}} \right), \quad \text{for all } i = 1, \ldots, m.$$

As a canonical example consider the case where $\rho_D = 0.9$, and the average direct travel time to all retailers is the same and $\theta_{0i} = 1$. In this case the average demand requires the truck to be busy 90% of the time under the DS policy. If we now take

$n = 100$, the heavy traffic conditions are satisfied when

$$\sum_{i=1}^{m-1} \theta_{i,i+1} \leq \frac{1}{10}.$$

This requirement is not too surprising given the performance analysis results in §4.2. There we saw was that as $\rho_D \to 1$ the DS policy dominates both cost components in systems where the distance between the retailers is not zero.

As in the fixed route cases, the next step towards the limiting control problem consists in characterizing

$$X(t) = \lim_{n\to\infty} X^{(n)}(t).$$

The main difference between this undertaking and its fixed route counterpart is that the diffusion limit for the netput process depends on the routing control. Let us therefore consider the behavior of the systems during an interval of scaled time $[t_o, t_f]$ where the TSP routing scheme is used exclusively. Let $X^*(t) = X(t_o + t) - X(t_o)$ for $t \in [0, t_f - t_o]$ be the limiting in-period netput process, and define $\hat{\delta}^{*(n)}(t)$, the relative routing allocation over this time interval as

$$\hat{\delta}^{*(n)}(t) = \frac{\tau_D^{(n)}(t_o + t) - \tau_D^{(n)}(t_o)}{\tau_D^{(n)}(t_o + t) - \tau_D^{(n)}(t_o) + \tau_T^{(n)}(t_o + t) - \tau_T^{(n)}(t_o)}.$$

By assumption, we use only the TSP routing during this interval, hence $\tau_D^{(n)}(t_o + t) - \tau_D^{(n)}(t_o) = 0$, and $\hat{\delta}^{*(n)}(t) = 0$ must hold for any $t \in [0, t_f - t_o]$.

With these observations, it is a matter of simple algebra to obtain the following:

$$\begin{aligned} X^*(t) \;&=\; \lim_{n\to\infty} \Bigg\{ \sqrt{n}\left(\frac{C^{(n)}}{\vartheta_T^{(n)}} - \lambda^{(n)}\right) t + C^{(n)} \left[\hat{S}_T^{(n)}(\tau_T^{(n)}(t_o + t)) - \hat{S}_T^{(n)}(\tau_T^{(n)}(t_o))\right] \\ &\quad - \hat{D}^{(n)}(t_o + t) + \hat{D}^{(n)}(t_o) \Bigg\} \quad \text{for } t \in [0, t_f - t_o] \qquad (5.33) \end{aligned}$$

for the in-period netput process.

In order to characterize the RHS of (5.33), we need the limiting in-period demand and partial service completion processes. Consider the partial service completions first. The corresponding definitions for the scaled processes give

$$\hat{S}_T^{(n)}(\tau_T^{(n)}(t_o+t)) - \hat{S}_T^{(n)}(\tau_T^{(n)}(t_o)) =$$

$$\frac{1}{\sqrt{n}}\left[S_T^{(n)}(T_T^{(n)}(nt_o+nt)) - S_T^{(n)}(T_T^{(n)}(nt_o)) - \frac{T_T^{(n)}(nt_o+nt) - T_T^{(n)}(nt_o)}{\vartheta_T^{(n)}}\right].$$

Consider now the in-period control $T_T^{(n)*}(t)$, given by

$$T_T^{(n)*}(nt) = T_T^{(n)}(nt_o+nt) - T_T^{(n)}(nt_o) \text{ for } t \in [0, t_f - t_o].$$

Notice that, by construction, $T_T^{(n)*}(t)$ is an increasing, continuous process with $T_T^{(n)*}(0) = 0$, that is nonanticipating with respect to the inventory process. It therefore constitutes a feasible control for the dynamic routing VRIP system. According to the relevant definitions, the scaled and centered partial service completion process that corresponds to this control is

$$\hat{S}_T^{(n)*}(\tau_T^{(n)*}(t)) = \frac{1}{\sqrt{n}}\left[S_T^{(n)}(T_T^{(n)*}(nt)) - \frac{T_T^{(n)*}(nt)}{\vartheta_T^{(n)}}\right] \text{ for } t \in [0, t_f - t_o], \tag{5.34}$$

where

$$\tau_T^{(n)*}(t) = \frac{T_T^{(n)*}(nt)}{n} \text{ for } t \in [0, t_f - t_o].$$

In general, the process defined in (5.34) will be different from the desired in-period partial service completion process. This is the case since the sequence of partial service completions in $[t_o, t_f]$ could include the residual partial tour at time $t = t_o$ followed by the subsequent iid partial tour times. Also, there is the elapsed time of the (possibly unfinished) partial tour at $t = t_f$. Therefore, the two boundary partial service completion times do not have the same distributions as the rest of the service times. However, Iglehart and Witt (1970) have shown that, under heavy traffic, this boundary times vanish and the same heavy traffic limit is obtained in both cases. We thus have that:

$$\lim_{n\to\infty} \hat{\mathcal{S}}_T^{(n)}(\tau_T^{(n)}(t_o+t)) - \hat{\mathcal{S}}_T^{(n)}(\tau_T^{(n)}(t_o)) \sim \lim_{n\to\infty} \hat{\mathcal{S}}_T^{(n)*}(\tau_T^{(n)*}(t)),$$

where $\sim$ denotes equality in distribution. The treatment for the in-period demand process follows the same line of argument in simpler terms since no control is involved. This leads to

$$\lim_{n\to\infty} \hat{D}^{(n)}(t_o+t) - \hat{D}^{(n)}(t_o) \sim \lim_{n\to\infty} \hat{D}^{(n)*}(t).$$

Since demand arrivals are independent of service completions, these previous two equalities in distribution give

$$X^*(t) \sim \lim_{n\to\infty} \sqrt{n}\left(\frac{C^{(n)}}{\vartheta_T^{(n)}} - \lambda^{(n)}\right)t + C^{(n)}\hat{\mathcal{S}}_T^{(n)*}(\tau_T^{(n)*}(t)) - \hat{D}^{(n)*}(t). \tag{5.35}$$

The diffusion limit for $X^*(t),\ \ t \in [0, t_f - t_o]$ may now be found by the same basic results used in the fixed route cases. We thus obtain that, under the heavy traffic conditions (5.24)-(5.32), $X^*(t)$ is a (μ_T, σ_T^2) Brownian motion, the same diffusion that characterizes the system netput in the fixed route TSP VRIP.

Consider now the case when during the period $[t_o, t_f]$ the system uses exclusively the DS policy. Following similar lines of argument as for the TSP case, we obtain

$$X^*(t) \sim \lim_{n\to\infty} \sqrt{n}\left(\frac{\lambda^{(n)}C^{(n)}}{2\sum_i \lambda_i^{(n)}\vartheta_{0i}^{(n)}} - \lambda^{(n)}\right)t + \frac{K^{(n)}C^{(n)}\lambda^{(n)}}{\lambda_\ell^{(n)}}\hat{\mathcal{S}}_D^{(n)*}(\tau_D^{(n)*}(t)) - \hat{D}^{(n)*}(t). \tag{5.36}$$

Therefore, under the heavy traffic conditions (5.24)-(5.32), $X^*(t)$ is a (μ_D, σ_D^2) Brownian motion, the same diffusion that characterizes the netput in the DS case.

Notice that neither of these results require that $[t_o, t_f]$ be large in any sense. We have thus established the following

Proposition 6. *$X^{(n)}(t) \Rightarrow X(t, \Re(t))$, where $\Rightarrow$ denotes weak convergence and $X(t, \Re(t))$ is a diffusion processes with control-dependent drift and variance given* *by*

$$\mu(\Re(t)) = \begin{cases} \mu_D & \text{if} \quad \Re(t) = DS \\ \\ \mu_T & \text{if} \quad \Re(t) = TSP \end{cases} \qquad \text{and} \quad \sigma^2(\Re(t)) = \begin{cases} \sigma_D^2 & \text{if} \quad \Re(t) = DS \\ \\ \sigma_T^2 & \text{if} \quad \Re(t) = TSP \end{cases}$$

respectively.

In other words, the routing control switches the netput of the system from one Brownian motion to another. Furthermore, these two Brownian motions have the same parameters as the diffusion limits of the corresponding fixed route cases.

5.2.2 Time Scale Decomposition and Limiting Control Problem

We must now establish the behavior of the (scaled) retailer inventory processes $(W_1, \ldots, W_m)$ to obtain a limiting control problem. To this order, we shall again invoke the HTAP. Notice that the individual inventories will still vary an order of magnitude faster than $X(t)$ since, by Proposition 6, the system netput still changes like a Brownian motion and the ability to quickly shift inventory among the retailers is retained (by means of delivery size changes in the TSP mode, and by retailer selection in the DS mode). Hence, a HTAP should also hold for this dynamic routing system.

As in the fixed route cases, we will use the implications of the HTAP to obtain a dramatic simplification for the limiting control problem. First, let us define the intrinsic total inventory level for the system as

$$Z(t) \equiv W(t) - \hat{\xi}(t).$$

According to the HTAP, over the course of a cycle $Z(t)$ varies like a diffusion with control-dependent drift and variance, while the individual inventories move deterministically at a rate that is an order of magnitude faster. The exact evolution of

$(W_1(t), \ldots, W_m(t))$ over a service completion will be determined by the intrinsic inventory level, and the cycle placement and routing scheme being used at time t. Hence, the total inventory in the system $W(t)$ will, in the limit, look like the intrinsic inventory diffusion superimposed with a deterministic process $\hat{\xi}(t)$.

These results will again allow for a decomposition of the dynamic routing VRIP: (1) for any given total intrinsic inventory level $Z(t) = x$ and routing mode $\Re(t) \in \{\text{TSP}, \text{DS}\}$ use the results in Chapters 2 and 3 to determine the optimal cycle placement and the corresponding inventory cost function $g_{\Re}(x)$; (2) choose the non-anticipating control $(Y(t), \Re(t))$ to minimize

$$\limsup_{T \to \infty} \frac{1}{T} E \left[\int_0^T g_{\Re}(Z(t, \Re(t)))dt - \hat{f} Y(T) \right]. \tag{5.37}$$

Subject to

$$Z(t, \Re(t)) = X(t, \Re(t)) - \left(\frac{C}{\vartheta_T}(1 - \hat{\delta}(t)) + \frac{\lambda C}{2 \sum_i \lambda_i \vartheta_{0i}} \hat{\delta}(t) \right) Y(t) \tag{5.38}$$

$$Y(t) \qquad \text{a non-decreasing RCLL process} \tag{5.39}$$

If we define the control-dependent infinitesimal generator of the diffusion $Z(t, \Re(t))$ as

$$\Gamma^{\Re} = \mu(\Re) \frac{\partial}{\partial x} + \frac{\sigma^2(\Re)}{2} \frac{\partial^2}{\partial x^2},$$

we can use Ito's Lemma (see e.g. Kushner (1971) and Taksar (1985)) to obtain the optimality conditions for this problem as:

$$\min \left\{ \hat{f} - V'(x), \min_{\Re} \left\{ \Gamma^{\Re} V(x) + g_{\Re}(x) - \gamma \right\} \right\} = 0 \tag{5.40}$$

$$V(0) = 0. \tag{5.41}$$

So, if one can find a constant γ, which is referred to as the gain, and a potential function $V(x)$ (not to be confused with the vehicle capacity V) that solve (5.40) and (5.41), then the controls $(\Re(t), Y(t))$ which minimize the expression in (5.40) are optimal and

γ is the minimal average cost per unit time.

5.3 Optimization of a Double Threshold Policy

As it turns out, a direct analytic approach to the solution of the partial differential equations (PDE) in (5.40) and (5.41) is a rather formidable task. This is caused by the complicated nature of the control-dependent inventory cost $g_{\Re}(x)$. In fact, the PDE system (5.40) and (5.41) could be solved if $g_{\Re}(x)$ were linear. Therefore, in order to gain some insight into the optimal control of the VRIP, we specialize our analysis to a certain class of policies. In particular, we shall consider a two-parameter policy that can be described as follows: the truck is busy whenever the total intrinsic inventory $Z(t) < z$, and correspondingly it idles when $Z(t) \geq z$; while busy, the truck uses the TSP routing scheme whenever $Z(t) \in (s, z)$, and it switches to DS mode whenever $Z(t) \leq s$. We will attempt to find the optimal values for the parameters (s, z).

We choose to analyze this class of policies for similar reasons than we used to select the base stock policy in the fixed route cases. First of all, since in heavy traffic the busy/idle control depends only on the total amount of inventory at the retailers, it makes sense to consider policies in which the idling decision is based only on a single threshold. In this context, it is natural to consider a second threshold at which the server should turn to the fast drift DS policy, and avoid running more costly backorders. Secondly, it is often the case in this sort of problem that double threshold policies are asymptotically optimal (see e.g. Taksar (1985), and Harrison and Taksar (1983)). In particular, if the drift for the DS policy were infinite (i.e. if the inter-retailer travel times are relatively big) then a double threshold policy would be optimal as long as $g_T(x)$ is convex. Finally, this assumption makes the problem tractable.

As in the fixed route cases, to obtain the optimal parameters (s, z) we must establish the steady state distribution for the total inventory level and the expected idleness rate. Let us begin by considering the steady state distribution of $Z(t)$ under this type of policy.

We established in the previous section that, at the arbitrary epochs t, $Z(t)$ is a diffusion process with control-dependent drift and variance parameters. Under the proposed double-threshold policy, $Z(t)$ will be reflected at the idling threshold z and the parameters for the brownian motion component $X(t)$ will depend only on the total inventory level. Fortunately, the subject of steady state distributions of single dimensional diffusion processes is old and well understood (see e.g. Karlin and Taylor, 1981). For our particular case, the steady state density function of the total inventory $\pi(x)$ must satisfy the following system of differential equations

$$\frac{1}{2}\frac{d^2}{dx^2}[\sigma^2(x)\pi(x)] - \frac{d}{dx}[\mu(x)\pi(x)] = 0 \qquad x < z \tag{5.42}$$

$$\frac{1}{2}\frac{d}{dx}[\sigma^2(x)\pi(x)] - \mu(x)\pi(x) = 0 \qquad x = z \tag{5.43}$$

where, according to our characterization of $Z(t)$ and the shape of the proposed policy, the diffusion parameters are given by

$$\mu(x) = \begin{cases} \mu_T & \text{if } x > s \\ \mu_D & \text{if } x \le s \end{cases} \qquad \text{and} \qquad \sigma^2(x) = \begin{cases} \sigma_T^2 & \text{if } x > s \\ \sigma_D^2 & \text{if } x \le s \end{cases}$$

Any regular diffusion with smooth infinitesimal coefficients has a continuous transitory density function, and hence, if it exists, its steady state density will be continuous. The diffusion limit for the total inventory in our case is not regular because of the jump at s in the infinitesimal coefficients of the diffusion process $Z(t)$. However, it is possible to approximate this process by a regular diffusion using a suitable interpolation for $\mu(x)$ and $\sigma^2(x)$ for $x \in [s-\epsilon, s+\epsilon]$, a small interval. By letting $\epsilon \downarrow 0$ the inaccuracy of the approximation can be reduced to arbitrarily small levels. Hence we proceed as if the process $Z(t)$ were regular in $(-\infty, z]$, and will attempt to find a continuous density function $\pi(x)$ that satisfies (5.42) and (5.43).

The solution to these differential equations has a rather intuitive form. Recall from the analyses of the fixed route cases in Chapters 2 and 3, that an RBM on $(-\infty, z]$ will be exponentially distributed in steady state, *iff the drift is positive.* Also recall that the requirement that the drift be positive is equivalent to the requiring the utilization of the system to be less than 1. In this dynamic routing setting we have a diffusion that behaves like an RBM inside the interval $(s, z]$, and then takes a different drift and variance over the interval $(-\infty, s]$. We therefore expect that the steady state distribution for this process will be given by suitably scaled exponential distributions; the parameter of the density in each interval being $\nu_T = 2\mu_T/\sigma_T^2$ and $\nu_D = 2\mu_D/\sigma_D^2$ (i.e. the parameters of the corresponding fixed route cases). In other words, we expect $\pi(x)$ to have the form:

$$\pi(x) = \begin{cases} k_1 \nu_T e^{\nu_T(x-z)} & \text{if } s < x \leq z \\ \\ k_2 \nu_D e^{\nu_D(x-s)} & \text{if } x \leq s \end{cases} \tag{5.44}$$

where the scaling constants k_1, k_2 should satisfy

$$\int_{-\infty}^{z} \pi(x)dx = 1$$

$$\lim_{x \uparrow s} \pi(x) = \pi(s).$$

That is, the total probability over the relevant interval should be equal to one, and the density should be continuous at s. These conditions uniquely determine the value of the scaling factors k_1, k_2 for the candidate density in (5.44). Straightforward calculations give:

$$\pi(x) = \begin{cases} \left[\frac{\nu_D e^{\nu_T(z-s)}}{\nu_T + \nu_D(e^{\nu_T(z-s)} - 1)}\right] \nu_T e^{\nu_T(x-z)} & \text{if } s < x \leq z \\ \\ \left[\frac{\nu_T}{\nu_T + \nu_D(e^{\nu_T(z-s)} - 1)}\right] \nu_D e^{\nu_D(x-s)} & \text{if } x \leq s \end{cases} \tag{5.45}$$

The reader may verify that the density in (5.45) satisfies the system (5.42) and (5.43).

Furthermore, it is the only continuous density that satisfies this set of equations and hence it is the required steady state density for the total inventory process $Z(t)$.

In order to fully characterize the steady state cost expression in (5.37) we need only further determine the expected idleness rate. Taking expectations on both sides of (5.38), rearranging terms, dividing through by t and taking the limit as $t \to \infty$ we have that

$$\lim_{t\to\infty} \frac{1}{t} E\left[\left((1-\hat{\delta}(t))\frac{C}{\vartheta_T} + \hat{\delta}(t)\frac{\lambda C}{2\sum_i \lambda_i \vartheta_{0i}}\right) Y(t)\right] =$$

$$\lim_{t\to\infty} \frac{1}{t} E[X(t, \Re(t))] - \lim_{t\to\infty} \frac{1}{t} E[Z(t, \Re(t))] \tag{5.46}$$

The first term on the RHS of (5.46) corresponds to the asymptotic growth rate of the netput process under the proposed policy. Let $\delta = lim_{t\to\infty}\hat{\delta}(t)$ denote the long run fraction of time that the truck uses the DS routing policy, so that $(1-\delta)$ represent the fraction of time that corresponds to the TSP policy. Then, using standard results for diffusion processes, and the characteristics of $X(t, \Re(t))$, the desired long-run growth rate is given by,

$$\lim_{t\to\infty} \frac{1}{t} E[X(t, \Re(t))] = (1-\delta)\mu_T + \delta\mu_D. \tag{5.47}$$

Since the DS control will be exercised whenever the inventory process falls below the switching threshold s, it is a straightforward matter to obtain δ form the steady state distribution for $Z(t)$. Specifically,

$$\delta = \int_{-\infty}^{s} \pi(x)dx = \frac{\nu_T}{\nu_T + \nu_D(e^{\nu_T(z-s)} - 1)}.$$

Notice that the limit for $\hat{\delta}(t)$ as $t \to \infty$ exists in the strong law of large numbers sense. We can thus rewrite the LHS of (5.46) as

$$\lim_{t\to\infty} \frac{1}{t} E\left[\left((1-\hat{\delta}(t))\frac{C}{\vartheta_T} + \hat{\delta}(t)\frac{\lambda C}{2\sum_i \lambda_i \vartheta_{0i}}\right) Y(t)\right] =$$

$$\left((1-\delta)\frac{C}{\vartheta_T}+\delta\frac{\lambda C}{2\sum_i \lambda_i\vartheta_{0i}}\right)\lim_{t\to\infty}\frac{1}{t}E\left[Y(t)\right]. \tag{5.48}$$

Finally, according to the steady state distribution for $Z(t)$ in (5.45), $\bar{z} = \lim_{t\to\infty} E[Z(t)]$ exists and is finite. Hence, the second term on the RHS of (5.46) will vanish in the limit. Canceling this term and substituting (5.47) and (5.48) into (5.46) we obtain the following

$$\bar{y} = \lim_{t\to\infty}\frac{1}{t}E[Y(t)] = \sqrt{n}\left(1-\frac{\rho_T(1-\delta)2\sum_i \lambda_i\vartheta_{0i}+\rho_D\delta\lambda\vartheta_T}{(1-\delta)2\sum_i \lambda_i\vartheta_{0i}+\delta\lambda\vartheta_T}\right) \tag{5.49}$$

We may now write an expression for the steady state cost of the VRIP under the double threshold dynamic routing policy as a function of the control parameters (s, z). The problem is hence reduced to finding (s^*, z^*) that satisfy

$$(s^*, z^*) = \arg\min_{s\leq z} F(s, z) \tag{5.50}$$

where

$$F(s,z) = \int_{-\infty}^{s} g_D(x)\pi(x)dx + \int_{s}^{z} g_T(x)\pi(x)dx - \hat{f}\bar{y} \tag{5.51}$$

Unfortunately, $F(s, z)$ is rather complicated and a closed form solution for the optimal control parameters does not seem possible. Furthermore, an expansion for $F(s, z)$ is not possible in general since its exact form will depend on the relationship between $(\hat{\alpha}_T, \hat{\beta}_T)$ and $(\hat{\alpha}_D, \hat{\beta}_D)$ the parameters that define the three characteristic regions for the cycle placement and inventory cost solutions. By considering different problem parameters, one may get either $\hat{\beta}_D > \hat{\beta}_T$ or $\hat{\beta}_D < \hat{\beta}_T$. This is most unfortunate since without a general explicit characterization of (z^*, s^*) and $F(z^*, s^*)$ we cannot follow an approach analogous to Appendix A to verify that the double threshold solution satisfies the optimality conditions (5.40) and (5.41). Still, in an attempt to gain further insight into the VRIP we shall turn our attention to the numerical solution of specific instances.

5.4 Numerical Experiments

5.4.1 Proposed Double Threshold Policy

In order to keep our exposition simple we specialize our numerical results to the wedge topology cases with balanced costs described in Chapter 4. Other cases are handled in exactly the same fashion, the only difference being that the constants in the cycle placement formulae are generally more complicated expressions of the original system parameters. For this particular family of cases, the inventory cost for a given intrinsic inventory level $Z(t) = x$ is given by

$$\text{Region 1 TSP.} \qquad x \leq \hat{\alpha}_T = \sum_i \lambda_i \vartheta_{0i}^{TSP} - C \cdot$$

$$g_T(x) = -bx + b\sum_i \lambda_i \vartheta_{0i}^{TSP} - \frac{bC}{2} \tag{5.52}$$

$$\text{Region 2 TSP.} \qquad \hat{\alpha}_T < x \leq \hat{\beta}_T = \sum_i \lambda_i \vartheta_{0i}^{TSP}$$

$$g_T(x) = \hat{a}_{11}x^2 + \hat{a}_{12}x + \hat{a}_{13} \tag{5.53}$$

$$\hat{a}_{11} = \frac{b+h}{2C}$$

$$\hat{a}_{12} = h - \frac{b+h}{C}\sum_i \lambda_i \vartheta_{0i}^{TSP}$$

$$\hat{a}_{13} = \frac{C}{2(b+h)}\left(\hat{a}_{12}^2 + bh\right)$$

$$\text{Region 3 TSP.} \qquad \hat{\beta}_T < x$$

$$g_T(x) = hx - h\sum_i \lambda_i \vartheta_{0i}^{TSP} + \frac{hC}{2} \tag{5.54}$$

or by

$$\text{Region 1 DS.} \qquad x \leq \hat{\alpha}_D = -\frac{mC}{2}$$

$$g_D(x) = -bx \tag{5.55}$$

$$\text{Region 2 DS.} \qquad \hat{\alpha}_D < x \leq \hat{\beta}_D = \frac{mC}{2}$$

$$g_D(x) = \hat{a}_{14}x^2 + \hat{a}_{15}x + \hat{a}_{16} \tag{5.56}$$

$$\hat{a}_{14} = \frac{b+h}{2mC}$$

$$\hat{a}_{15} = \frac{h-b}{2}$$

$$\hat{a}_{16} = \frac{(b+h)mC}{8}$$

Region 3 DS. $\hat{\beta}_D < x$

$$g_D(x) = hx \tag{5.57}$$

depending on which routing scheme is being used at time t. Making the observation that $\sum_i \lambda_i \vartheta_{0i}^{TSP} < C$ is true as long as $\rho_T < 1$, we have that for $m \geq 2$, the following relationship holds among the parameters that characterize the regions along which the inventory cost behavior changes:

$$\hat{\alpha}_D < \hat{\alpha}_T < \hat{\beta}_T < \hat{\beta}_D.$$

Also, as long as $b > h$ it is clear that $z^* > 0$. With the cost expressions in (5.52)-(5.57) and the previous observations we may write an explicit expansion for the cost of the double threshold policy given in (5.51). As it turns out, $F(s,z)$ has a different functional form along eight regions of the intersection of the half-spaces $\{s \leq z\}$ and $\{z \geq 0\}$. While the qualitative nature of the cost function is similar for all regions, the detailed expressions are unfortunately rather long. Thus, in the interest of brevity, we present here only two of these regions as examples and refer the interested reader to Appendix B for the complete details. The $\hat{a}_k$ terms are constants for any given problem instance. Their detailed definitions in terms of the primitive system parameters are also included in Appendix B. The long-run average cost for the double threshold policy (s,z) is given by

Case 1. $s \leq \hat{\alpha}_D$, and $z > \hat{\beta}_T$

$$F(s,z) = \int_{-\infty}^{s} g_D(x)\pi(x)\,dx + \int_{s}^{z} g_T(x)\pi(x)\,dx$$

$$= \delta\left[\hat{a}_{20}s + \hat{a}_{21} + (\hat{a}_{22} + \hat{a}_{23})e^{-\hat{\nu}_T s} + (\hat{a}_{24}z + \hat{a}_{25})e^{\hat{\nu}_T(z-s)}\right] + \hat{a}_{26} + \frac{\hat{a}_{27}\delta}{(1-\delta)\vartheta_{01} + \delta\vartheta_T} \quad (5.58)$$

Case 2. $\quad s \leq \hat{\alpha}_D, \quad \text{and} \quad z \in [0, \hat{\beta}_T]$

$$F(s,z) = \delta\left[\hat{a}_{20}s + \hat{a}_{21} + \hat{a}_{23}e^{-\hat{\nu}_T s} + (\hat{a}_{28}z^2 + \hat{a}_{29}z + \hat{a}_{30})e^{\hat{\nu}_T(z-s)}\right] + \hat{a}_{26} + \frac{\hat{a}_{27}\delta}{(1-\delta)\vartheta_{01} + \delta\vartheta_T} \quad (5.59)$$

and similar expressions for the other six possible cases. As mentioned before, even if we ignore the constraint $s \leq z$, there is no closed form expression for the values (s^*, z^*) that minimize $F(s,z)$ as given by these formulae. We should note that, if we define the unscaled idling and switching thresholds as $w = \sqrt{n}\,z$ and $c = \sqrt{n}\,s$ respectively, and we make the parameter scaling substitutions as in the fixed route cases, the scaling factor n cancels out of these expressions for $F(c,w)$, and what remains is a function only of the original system parameters. Hence we can proceed to find the optimal control values numerically without having to specify a scaling factor.

We used a simple steepest descent method to find the optimal control (c^*, w^*), for several different values of the system parameters. In the course of these numerical experiments we found that $F(c,w)$ will frequently have two local minima (one of which is a global minimum). When there are two local minima, each of them corresponds to a perturbation of the optimal value of the fixed route extremes. In all cases we found the global minimum by starting the steepest descent algorithm first at the point (w_D^*, w_D^*) that corresponds to the optimal DS base stock level, and then from the point $(-10mV, w_T^*)$, which approximates the optimal fixed route TSP case. In the cases when the algorithm converged to a different answer for each starting point, we selected the one with the smallest cost. We later verified that this was indeed the global minimum by a search over the whole (c,w) plane.

The results for a sample of the case considered are presented in Table 5.1. These represent 36 different cases (3 vehicle sizes, 4 traffic intensities and 3 transportation cost values) of the wedge topology as described in Chapter 4. The entries in the table give the idling (w^*) and switching (c^*) thresholds proposed by the heavy traffic analysis, as well as the limiting fraction of the active time that the truck is expected to use the DS mode (δ^*) when operating this policy. As can be seen, in many instances δ^* is either 100.0% or zero. Hence, the dynamic policy reverts to one of the fixed route modes in many instances of the problem. While results are not reported here, we found this to still be the case for different values of the backordering to holding cost ratio b/h, as well as for wider or narrower wedges (i.e. for cases where θ_{01}/θ_{12} was bigger or smaller than the value of 18 used here and in Chapter 4). Since the implementation of a dynamic routing policy is more complex, this suggests that in many instances the manager of a distribution system may be better off by just choosing the right fixed route policy.

5.4.2 An Algorithmic Solution

An alternative numerical approach to the solution of (5.40) and (5.41) is to approximate the limiting diffusion by a discrete time and space Markov chain, and then solve the control problem by dynamic programming. The main advantage to this approach is that we may use it to consider general switching policies, and not just the proposed double threshold case. Weak convergence methods have been developed to verify that the controlled Markov chain and its optimal cost approximate arbitrarily closely (at an increased computational expense) the controlled diffusion process and its optimal cost. Interested readers are referred to Kushner and Dupuis (1992) for an up-to-date account of this research area.

Based on our heavy traffic analysis, we consider the total intrinsic inventory as the state of the system and allow for the arbitrary choice of either the TSP or DS routing scheme for any give value of the state. In this context, it is natural to assume a control that idles whenever the intrinsic inventory reaches a base stock level z. Notice that the numerical complexity of this approach is independent of m, the number of retailers in the system, due to the state space collapse allowed by the HTAP. Even so, it remains more computationally intensive than the steepest descent minimization

			$\lambda = 5.0$	$\lambda = 7.0$	$\lambda = 8.0$	$\lambda = 9.0$
$V = 100$	$f = 500$	w^*	20	45	75	138
		c^*	-312	-151	-100	-51
		δ^*	0.0%	0.0%	1.3%	8.5%
	$f = 100$	w^*	20	45	76	141
		c^*	-385	-168	-111	-58
		δ^*	0.0%	0.0%	1.0%	7.7%
	$f = 50$	w^*	20	45	76	142
		c^*	-404	-171	-112	-59
		δ^*	0.0%	0.0%	1.0%	7.5%
$V = 50$	$f = 500$	w^*	10	23	39	112
		c^*	-10^6	-64	-43	112
		δ^*	0.0%	0.0%	2.0%	100.0%
	$f = 100$	w^*	10	24	39	73
		c^*	-202	-80	-53	-28
		δ^*	0.0%	0.0%	1.3%	8.2%
	$f = 50$	w^*	10	24	39	73
		c^*	-202	-82	-54	-29
		δ^*	0.0%	0.0%	1.2%	8.0%
$V = 10$	$f = 500$	w^*	18	20	21	24
		c^*	18	20	21	24
		δ^*	100.0%	100.0%	100.0%	100.0%
	$f = 100$	w^*	2	6	10	24
		c^*	-24	-11	-8	24
		δ^*	0.0%	0.0%	3.9%	100.0%
	$f = 50$	w^*	2	6	10	18
		c^*	-31	-13	-9	-4
		δ^*	0.0%	0.0%	3.3%	11.9%

Table 5.1: Proposed Double Threshold Control

done for the double threshold policy.

Let $\epsilon > 0$ denote the finite difference interval, which dictates how finely time and space are discretized. One can think of a sequence of discrete time and space Markov chains that become better approximations for the diffusion problem at hand as $\epsilon \to 0$.

A computer implementation requires that, besides discretizing the state space, we confine the evolution of the Markov chain to a bounded region. Since, under a base stock idling policy, the total intrinsic inventory process resides on a halfline, we set the state space of the controlled Markov chain to be $\{-N, -N+\epsilon, \ldots, N-\epsilon, N\}$ for N a positive integer multiple of ϵ. Fix, for the time being, the idling threshold at z, an integer multiple of ϵ, that satisfies $z < N$ (this parameter will be optimized later on). Then, the state space for the Markov chain approximation is $\{-N, -N+\epsilon, \ldots, z-\epsilon, z\}$. Assuming that the the two diffusion drifts μ_T, μ_D are positive (i.e. that $\rho_T < 1$ and $\rho_D < 1$), the approximating Markov chain has non-zero control-dependent transition probabilities given by

$$P^\epsilon(x, x+\epsilon) = \frac{\sigma^2(\Re(x)) + 2\epsilon\mu(\Re(x))}{2\sigma^2(\Re(x)) + 2\epsilon\mu(\Re(x))},$$

and

$$P^\epsilon(x, x-\epsilon) = \frac{\sigma^2(\Re(x))}{2\sigma^2(\Re(x)) + 2\epsilon\mu(\Re(x))}$$

for states in the interior of $[-N, z]$, and time intervals of length

$$\Delta t^\epsilon(x) = \frac{\epsilon^2}{\sigma^2(\Re(x)) + \epsilon\mu(\Re(x))}.$$

In order to obtain a controlled Markov chain approximation to the VRIP two issues must be addressed: (1) for an ergodic cost problem the interpolation interval $\Delta t^\epsilon(x)$ must be independent of the state x; (2) we need to characterize the probabilistic behavior of the chain at the boundary states $x = -N$ and $x = z$. To obtain a state-independent discretization interval, define $\eta^\epsilon = max_{\Re \in \{TSP, DS\}}\{\sigma^2_\Re + \epsilon\mu_\Re\}$. We next use η^ϵ to define modified interior transition probabilities for the Markov chain as

$$\bar{P}^\epsilon(x, x+\epsilon) = \frac{\sigma^2(\Re(x)) + 2\epsilon\mu(\Re(x))}{2\eta^\epsilon}, \tag{5.60}$$

$$\bar{P}^\epsilon(x, x-\epsilon) = \frac{\sigma^2(\Re(x))}{2\eta^\epsilon}, \tag{5.61}$$

$$\bar{P}^\epsilon(x, x) = 1 - \frac{\sigma^2(\Re(x)) + \epsilon\mu(\Re(x))}{\eta^\epsilon}, \tag{5.62}$$

as well as the new (state-independent) interpolation interval

$$\Delta t^\epsilon = \frac{\epsilon^2}{\eta^\epsilon}. \tag{5.63}$$

We next turn our attention to the definition of appropriate transition probabilities for the boundary states. Our choice of idling control specifies a reflecting barrier at $x = z$. However, the Markov chain approximation method assumes that the reflection at the boundary is instantaneous. Since this would require an interpolation interval at this state that is different from η^ϵ (the chain would not spend any time there), we eliminate the transitions into this state. We thus adjust the transition probabilities at the state $x = z - \epsilon$ to

$$\tilde{P}^\epsilon(z-\epsilon, z-\epsilon) = 1 - \bar{P}^\epsilon(z-\epsilon, z-2\epsilon), \tag{5.64}$$

and

$$\tilde{P}^\epsilon(z-\epsilon, z) = 0. \tag{5.65}$$

In order to keep the Markov chain inside the state space (i.e. in order to remain within the memory allocation in the computer) we also impose a reflecting barrier at $x = -N$. That is, we adjust the transition probabilities at state $x = -N + \epsilon$ to

$$\tilde{P}^\epsilon(-N+\epsilon, -N+\epsilon) = 1 - \bar{P}^\epsilon(-N+\epsilon, -N+2\epsilon), \tag{5.66}$$

and

$$\tilde{P}^{\epsilon}(-N+\epsilon, -N) = 0. \tag{5.67}$$

This barrier is artificial in the sense that $P^{\epsilon}(-N+\epsilon, -N)$ is positive in the real system. However, if the value of N is sufficiently large, this probability becomes very small and the effect of this artifice on the control problem should be negligible. In the implementation of this algorithm, we adjust the size of the Markov chain until a further increase in N produces no change in the optimal solution $(\Re^*(x), z^*)$.

In summary the Markov chain approximation to the diffusion control problem (5.40) and (5.41) has state space $\{-N+\epsilon, -N+2\epsilon, \ldots, z-2\epsilon, z-\epsilon\}$, interpolation interval Δt^{ϵ} defined by (5.63), and nonzero transition probabilities $\tilde{P}^{\epsilon}(x,y)$ defined by (5.64)-(5.67) and $\tilde{P}^{\epsilon}(x,y) = \bar{P}^{\epsilon}(x,y)$ otherwise, where $\bar{P}^{\epsilon}(x,y)$ are defined in equations (5.60)-(5.62).

The dynamic programming optimality equation for the controlled Markov chain is given by

$$\begin{aligned} V^{\epsilon}(x) &= \min_{\Re\in\{TSP,DS\}}\left\{\sum_{y}\tilde{P}^{\epsilon}(x,y)V^{\epsilon}(y) + (g_{\Re}(x)-\gamma^{\epsilon})\Delta t^{\epsilon}\right\} \quad \text{for } -N < x \le z-\epsilon \\ V^{\epsilon}(z) &= 0 \end{aligned} \tag{5.68}$$

where $g_{\Re}(x)$ is the inventory cost function given by the fixed route optimal cycle placement results for the corresponding routing scheme $\Re$.

We can now solve the Markov chain control problem by means of a policy improvement algorithm. We start by choosing an arbitrary initial routing policy $\Re_0(x)$ and idling threshold z_0. The algorithm next iterates over two steps: (1) an *evaluation* step where the potential function and gain $(V_k^{\epsilon}(x), \gamma_k^{\epsilon})$ are found recursively from the given controls $(\Re_k(x), z_k)$; (2) an *improvement* stage where $\Re_{k+1}(x)$ is found by minimization of (5.68) for the given $(V_k^{\epsilon}(x), \gamma_k^{\epsilon})$, and z_{k+1} is found by choosing the gain minimizing value given $(\Re_{k+1}(x), V_k^{\epsilon})$. The iteration is stopped when there are no gain improvements.

At the evaluation stage the main tasks are to compute the gain and the potential function for the current control. In order to compute the gain, which corresponds to the long-run time average cost, we need to characterize the steady state distribution for the system. For this purpose, we take advantage of the birth-death structure of the Markov chain approximation. In particular, for a Markov chain with this kind of structure, we may obtain the steady state distribution $\pi^\epsilon(x)$ by

$$\pi^\epsilon(x) = \begin{cases} 0 & \text{for } x > z-\epsilon \\ \pi^\epsilon(z-\epsilon)\prod_{k=x+\epsilon}^{z-\epsilon} \frac{\tilde{P}^\epsilon(k,k-\epsilon)}{\tilde{P}^\epsilon(k-\epsilon,k)} & \text{for } x < z-\epsilon \\ \left(1+\sum_{j=-N+\epsilon}^{z-\epsilon}\prod_{k=j+\epsilon}^{z-\epsilon}\frac{\tilde{P}^\epsilon(k,k-\epsilon)}{\tilde{P}^\epsilon(k-\epsilon,k)}\right)^{-1} & \text{for } x = z-\epsilon \end{cases} \tag{5.69}$$

where the summations and products are done over the appropriately discretized space (i.e. the step size is ϵ). The time average cost for the VRIP may now be found as the sum of the time average transportation cost (or idleness reward) and the time average inventory cost. The transportation cost component depends on $\bar{y}$, the steady state fraction of time that the system idles, which in turn depends on δ, the steady state fraction of its busy time that the truck uses the DS routing scheme. The fraction of time DS is used in the Markov approximation may be found by

$$\delta^\epsilon = \sum_{x\in(N,z)} \pi^\epsilon(x)\mathbf{1}_{\{\Re(x)=DS\}}$$

where

$$\mathbf{1}_{\{A\}} = \begin{cases} 1 & \text{if } A \\ 0 & \text{otherwise} \end{cases}$$

is the indicator function. Once δ^ϵ is available, we can get the steady state fraction of time that the system idles by

$$\bar{y}^\epsilon = \sqrt{n}\left(1-\frac{\rho_T(1-\delta^\epsilon)2\sum_i \lambda_i\vartheta_{0i}+\rho_D\delta^\epsilon\lambda\vartheta_T}{(1-\delta^\epsilon)2\sum_i\lambda_i\vartheta_{0i}+\delta^\epsilon\lambda\vartheta_T}\right). \tag{5.70}$$

In terms of these definitions, the gain is given by

$$\gamma^\epsilon = -\hat{f}\bar{y}^\epsilon + \sum_{x \in (N,z)} \pi^\epsilon(x) g_{\Re}(x). \tag{5.71}$$

Using (5.68) and $V^\cdot(x) = 0$ for $x \geq z$, $V^\epsilon(x)$ can be calculated recursively by

$$V^\epsilon(x-\epsilon) = \frac{[\gamma^\epsilon - g_{\Re}(x)]\Delta t^\epsilon + (1 - \tilde{P}^\epsilon(x,x)V^\epsilon(x) - \tilde{P}^\epsilon(x,x+\epsilon)V^\epsilon(x+\epsilon)}{\tilde{P}^\epsilon(x,x-\epsilon)}. \tag{5.72}$$

In the policy improvement step, we first solve for the routing policy $\Re(x)$ and then for the idling threshold z. The routing decision is straightforward: for each state x we choose the scheme $\Re$ that minimizes the RHS of (5.68). Then, keeping $\Re(x)$ constant, we determine an improved idling threshold z by evaluating (via equation (5.71)) the gain γ^ϵ for all values of $z \in (-N, N)$ and selecting the gain minimizing value.

These two steps are repeated until the improvement in gain becomes sufficiently small. The output of the algorithm includes an idling threshold z and a state-dependent routing policy $\Re(x)$ for $x \in (-N, z)$. However, the mapping from this numerical solution to a proposed policy is not as straightforward as in our analytical results. Specifically, there is no way to develop a proposed scheduling policy that is independent of the heavy traffic scaling factor n. For instance, the drifts of the underlying Brownian motions are

$$\mu_T = \sqrt{n}\,(1-\rho_T)\,\frac{V}{\theta_T}, \quad \text{and} \quad \mu_D = \sqrt{n}\,(1-\rho_D)\,\frac{\lambda V}{2\sum_i \lambda_i \theta_{0i}}$$

and the expression for $\bar{y}$ in (5.70) also depends on $\sqrt{n}$. Therefore, a value of n must be chosen in order to compute a numerical solution to the Markov chain control problem. We deal with this quandary in the most natural way: we set $\sqrt{n}(1-\rho_T) = 1$ and hence take $n = (1-\rho_T)^{-2}$. This value is used to scale the parameters for the computation and to unscale the solution. Because of this, it is possible for the (unscaled) solution to be quite insensitive to the choice of n.

We implemented this algorithm and have so far tested it on the $V = 100$ and $V = 50$ cases of Table (5.1) when $f = 500$. Though results are still preliminary at the point of this writing, it is interesting to point out that the routing policy we obtained from this algorithm was a double threshold policy for all these cases. Furthermore, the unscaled thresholds (c, w) obtained with this approach were very close to the corresponding ones in Table (5.1). This suggests that a double threshold policy is likely to do reasonably well in most instances. However, more work will be necessary before we can strengthen this conclusion.

5.4.3 Simulation Experiment

We performed a series of simulation experiments in order to gage the accuracy of our heavy traffic approximations over a range of values for the problem parameters. To this order, we compared the average total cost obtained by the use of our proposed policy with that of the best values of the control parameters found by an exhaustive search over the (c, w) plane. The addition of a second control parameter will of course have the undesirable effect of making the number of runs required to find a reasonable approximation for the optimal threshold levels roughly equal to the square of those needed in the static cases. We therefore set $f = 500$ and considered a total of six cases (3 traffic intensities and 2 vehicle sizes). The basic test case is the same 5-retailer wedge system presented in Chapter 4.

Table 5.2 summarizes the results of these simulation experiments. The entries in the table represent the cost increase incurred by using the proposed dynamic policy, the best fixed-route TSP system or the best fixed-route DS system instead of the best double threshold policy found by exhaustive search. In all cases, the delivery allocation uses the dynamic rule derived from the heavy traffic optimal cycle placement. The average suboptimality (within the class of double threshold policies) for the double threshold policy found by the heavy traffic analysis is 2.1%, and in only one case it is higher than 3%. Once again the proposed policy, which was obtained from heavy traffic results, performs very well. Furthermore, its performance does not seem to deteriorate much at lower traffic intensities or smaller truck sizes.

A glance at the appropriate entries in Table 5.1, shows that in 5 out of the 6 simulation cases the value of δ is close to either 0 or 1. The exception is the $(\lambda = 9, V = 100)$

		$\lambda = 7.0$	$\lambda = 8.0$	$\lambda = 9.0$
$V = 100$	Prop.	2.3%	2.2%	2.2%
	TSP*	1.4%	2.0%	6.1%
	DS*	42.5%	27.6%	10.6%
$V = 50$	Prop.	0.8%	1.3%	3.8%
	TSP*	0.3%	0.8%	0.9%
	DS*	18.4%	11.8%	3.3%

Table 5.2: Suboptimality of Proposed Double Threshold and Fixed Route Policies

case with a $\delta = 8.5\%$. In this case we expected the optimal dynamic policy to outperform either of the static routing schemes. The results in Table 5.2 confirm this. We made an attempt to find a case where the advantage for the dynamic policy would be even more dramatic by searching over different values for λ (leaving everything else fixed). As it turns out, the $(\lambda = 9, V = 100)$ case turned out to be the best one. Further increasing the traffic intensity $(\lambda > 9)$ improves the DS performance, while reducing the traffic intensity will favor TSP. In the end, the dynamic policy found from our heavy traffic analysis was at most 5% better than the best fixed-route policy. Furthermore in the cases where the proposed double threshold policy coincides with either of the fixed route schemes, the cost increase incurred by choosing the wrong fixed route scheme is quite significant (higher than 10% in all 5 cases, and up to 43% for $\lambda = 7, V = 100$). This suggests that, while finding the best fixed route scheme is very important, the advantage obtained from dynamic routing selection is quite small in most problem instances. Note that this is true even when the transportation cost f is several orders of magnitude higher than the holding rate.

The dynamic routing model is quite accurate in its characterization of the best routing policy for the system. In fact only in the $(\lambda = 9, V = 50)$ case the best policy found by exhaustive search is different from the one obtained from the minimization of $F(c, w)$. Our model predicts that DS will dominate in this problem instance but TSP turns out to be better. The predicted cost advantage for DS in this case was slightly higher than 10%, but this relative measure might be misleading since at this (relatively) moderate truck size the total cost itself is small,and that the predicted absolute cost difference is also small. Therefore, our heavy traffic model was still correct in predicting that both policies would have similar total costs.

Chapter 6

Conclusions

6.1 Main Results and Extensions

We believe one of the main contributions in this thesis is that we have formally modeled the dynamic-stochastic nature of the VRIP. In most practical cases demand arrivals are subject to considerable stochastic variation. In this context, the stochastic nature of the system must be taken into account if one is to properly address the inventory component of the problem.

This modeling approach, and the use of some powerful heavy traffic results, allowed us to fully characterize an asymptotically optimal operating policy for VRIP systems under two fixed routes. We should note that the approach used in Chapters 2 and 3 for the TSP and DS schemes respectively can easily be extended to any other case where a single truck cycles through a fixed sequence of trips.

In our treatment of the fixed route cases we ignored loading and unloading times (as is common in the vehicle routing literature) and concentrated on the travel time component of the service. It should be noted that this assumption is not critical. In fact, a straightforward extension of our results would allow us to handle load/unload times. The diffusion limit for the system netput and HTAP are not affected (as long as these times are appropriately scaled). The inclusion of the unload times at the retailers would change the deterministic inventory evolution over a cycle, so that inventories do not grow instantaneously but increase at a finite rate while the truck is being unloaded. We can still solve the optimal cycle placement for this case. Finally the steady state average driver and truck cost components are simply found from the average fraction of the cycle time that the truck is running (as opposed to being loaded/unloaded).

The results obtained from our analysis of the fixed-route cases provided valuable insights about the VRIP. In particular

- The heavy traffic analysis is quite accurate and provides a control policy that is close to optimal even in systems operating far away from the assumed regime.

- The allocation of load among the retailers is dictated by the desire to concentrate most of the total inventory (backorders) at the site where it is cheaper to hold (backorder).

- Closed-loop delivery allocations greatly outperform their open-loop counterparts in a stochastic environment.

- As utilization increases the relative advantage of recalculating the load allocation within the cycle as opposed to setting it at the beginning of each cycle decreases, as predicted by the HTAP.

- The inventory component of the total long-run average cost depends on the stochastic characteristics of the system, while the long-run average transportation cost for a fixed routing scheme is determined from first moments of the processes involved.

- The performance of the system is rather sensitive to the base stock level, so that the ability of the proposed policy to yield a near-optimal value for this parameter is very important.

- The policy that provides the best transportation cost in this context is not the one with the shortest cyclic route, but the one with largest amount delivered per unit time traveled. Therefore, if one ignores the inventory cost, DS is always more efficient than any other fixed route scheme (as long as the triangle inequality holds for the travel times).

- Since DS has a higher transportation efficiency, it is stable over a bigger range of arrival rates than the TSP (as long as individual retailers sites have different locations). This implies that, at higher utilization rates, DS will also dominate the inventory cost component, since the TSP policy would yield very high inventory levels.

We also applied our modeling approach to the case where a some dynamic route selection is allowed. While we are able to obtain a proposed policy for the system on the basis of our heavy traffic analysis, we do not do this with the same level of precision as in the fixed-route cases. Still, we gain some valuable additional insights. Namely

- The range over which the best dynamic routing policy differs from the static route cases (at least in this limited 2-choice version) is rather narrow (in system parameter space) even at high transportation cost rates. This range corresponds to systems in which the best TSP and DS policies have fairly similar performance. Therefore, it appears that finding the best fixed route policy is very important while allowing for dynamic routing provides a much less substantial benefit.

- The heavy traffic analysis is also quite accurate for this setting, and will be able to identify good control policies even when system utilization is moderate.

6.2 Further Research

As we hope the results in this thesis have illustrated, the dynamic/stochastic modeling of distribution systems, aided by the tools of heavy traffic analysis, allows for the development of simple good solutions for, as well as furthers our insight into very complex problems. We are confident that this area of research will continue to attract attention and provide valuable results.

There are several areas of the problem considered in this thesis that merit future development. First to come to mind is the extension of the dynamic routing case to allow for more general route choices. In particular, one might allow for the choice among K different cyclic routes. Assuming (without loss of generality) that these K schemes are indexed in increasing order for the corresponding drift parameters (i.e. assuming $\mu_1 < \mu_2 < \cdots < \mu_K$), one might then try to find the optimal K-threshold policy that switches among these schemes. The slowest drift policy to be used when the total cost is small, and then increasingly faster drift schemes as inventory levels fall and the risk of backorders increases.

Another area for future research would be to further develop the necessary steps for a hierarchical approach to the general (multi-truck, multi-depot) VRIP. The results in this thesis provide estimates for the operating cost for any such system given a particular assignment of retailers and trucks to depots and a fixed sequence of trips among the assigned retailers for each truck. Motivated by our insight that the best fixed route policy is close to the best dynamic policy over a broad range of parameters, the first level up in the hierarchy could implement some interexchange optimization algorithm (similar to those 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 or the total number of truck to have in the system. At an even higher level, these results could be used to decide on the number and location of depots.

Appendix A

Proof of the Optimality of The Proposed Idling Policy for the Fixed-Route TSP VRIP

We want to prove that a reflective barrier at z_T^*, or equivalently

$$Y^*(t) = \sup_{0 \le s \le t} \{X(s) - z_T^*\}^+ , \tag{A.1}$$

is the optimal solution for the diffusion control problem:

(P1) Choose the nondecreasing RCLL process Y to minimize

$$F_Y(x) = \limsup_{T \to \infty} \frac{1}{T} E_x \left[\int_0^T g(Z(t))\, dt - \hat{f} Y(T) \right] \tag{A.2}$$

subject to

$$Z(t) = X(t) - \frac{V}{\theta_T} Y(t) \tag{A.3}$$

$$X(0) = x \tag{A.4}$$

We proceed in three steps: (1) show that the transportation cost component may be ignored for this case; (2) find a lower bound for the long-run expected cost in (A.2);

(3) show that a base stock policy achieves this lower bound.

Step 1. Take expectations on both sides of (A.3), divide by t, take $\lim_{t\to\infty}$ and rearrange terms to obtain:

$$\lim_{t\to\infty}\frac{1}{t}Y(t) = \frac{\theta_T}{V}\lim_{t\to\infty}\frac{1}{t}X(t) - \frac{\theta_T}{V}\lim_{t\to\infty}\frac{1}{t}Z(t) \tag{A.5}$$

Since $X(t)$ is a (μ_T, σ_T^2) BM the first term in the RHS of (A.5) equals

$$\frac{\theta_T}{V}\lim_{t\to\infty}\frac{1}{t}X(t) = \frac{\theta_T}{V}\mu_T.$$

In order to characterize the second term in the RHS of (A.5) we need the following

Claim A.1: For any policy Y that satisfies $F_Y(x) < \infty$, $\lim_{t\to\infty} E_x[Z(t)] < \infty$

Proof of Claim A.1 We will show that if $\lim_{t\to\infty} E_x[Z(t)] = \infty$ then $F_Y(x) = \infty$.

Since $g(x) \geq 0$ for all x, we can use Tonelli's Theorem to get

$$E_x\left[\int_O^\infty g(Z(t))\,dt\right] = \int_O^\infty E_x\left[g(Z(t))\right]\,dt.$$

Also, since $g(x)$ is convex, Jensen's Inequality gives

$$E_x\left[g(Z(t))\right] \geq g(E_x\left[Z(t)\right]), \quad \text{for all } t.$$

Therefore, if $E_x\left[Z(t)\right]$ becomes unbounded as $t \to \infty$ the integral term in $F_Y(x)$ has an unbounded growth rate, and hence $F_Y(x) = \infty$. ∎

Since $Z(t)$ has a finite steady state expected value,

$$\lim_{t\to\infty}\frac{1}{t}Z(t) = 0$$

holds for any policy Y with finite long-run average cost. As we are interested in the policy that minimizes $F_Y(x)$, we may consider only the class of policies for which this value is finite. Within this class of policies the time average idleness rate is

$$\lim_{t\to\infty} \frac{1}{t} Y(t) = \frac{\theta_T}{V} \mu_T,$$

a constant. Therefore the idleness reward for all finite cost policies is the same and (P1) is equivalent to:

(P2) Choose the nondecreasing RCLL process Y to minimize

$$F_Y(x) = \limsup_{T\to\infty} \frac{1}{T} E_x \left[\int_0^T g(Z(t))\, dt \right] \tag{A.6}$$

subject to (A.3)-(A.4)

Step 2. Let,

$$\Gamma = \frac{1}{2}\sigma_T^2 \frac{\partial^2}{\partial x^2} + \mu_T \frac{\partial}{\partial x}$$

denote the infinitesimal generator for the BM $X(t)$, let γ represent the minimal average cost of problem (P2), and let $V(x)$ represent the cost incurred under the optimal policy when the initial state of the BM is x minus the cost incurred under the optimal policy when the initial state is a reference state z.

Proposition A.2 *Suppose* $(\gamma, V(x))$ *satisfy*

$$\min\{\Gamma V(x) + g(x) - \gamma, -V'(x)\} = 0 \tag{A.7}$$

and there exist constants K_0, K_1, K_2 *such that*

$$0 \le V(x) \le K_0 + K_1 x + K_2 x^2 \text{ for all } x. \tag{A.8}$$

Then

$$\gamma \le F_Y(x) \text{ for all } Y.$$

Proof: This result follows from Ito's formula applied on $V(Z(t))$. ■

We now propose a particular solution to (A.7) from which we obtain the desired lower bound on the optimal cost for problem (P2). The proposed solution $(\gamma^*, V^*(x), z^*)$ takes two forms, depending on the values of the system parameters. It is given by:

Case 1

$$\text{Condition:} \quad -\frac{1}{\hat{\nu}_T} \ln\left[\left(\frac{h}{b+h}\right)\left(\frac{\hat{\nu}(\hat{\beta}_T - \hat{\alpha}_T)}{e^{\hat{\nu}_T(\hat{\beta}_T - \hat{\alpha}_T)} - 1}\right)\right] + \hat{\alpha}_T \geq \hat{\beta}_T$$

$$z^* = -\frac{1}{\hat{\nu}_T} \ln\left[\left(\frac{h}{b+h}\right)\left(\frac{\hat{\nu}(\hat{\beta}_T - \hat{\alpha}_T)}{e^{\hat{\nu}_T(\hat{\beta}_T - \hat{\alpha}_T)} - 1}\right)\right] + \hat{\alpha}_T \tag{A.9}$$

$$\gamma^* = hz^* + \hat{a}_5 \tag{A.10}$$

$$V^*(x) = \begin{cases} 0 & \text{if } x \geq z^* \\ -\int_x^{z^*} V'(y)\,dy & \text{if } x < z^* \end{cases} \tag{A.11}$$

$$V'^*(x) = \begin{cases} \frac{1}{\mu_T}[\gamma^* + bx - \frac{b}{\hat{\nu}_T} - \hat{a}_1] & \text{if } x \leq \hat{\alpha}_T \\ \frac{1}{\mu_T}\left[\gamma^* + \frac{(h+b)e^{-\hat{\nu}_T(x - \hat{\alpha}_T)}}{\hat{\nu}_T^2(\hat{\beta}_T - \hat{\alpha}_T)} - \hat{a}_2 x^2 + \hat{a}_{17} x + \hat{a}_{18}\right] & \text{if } x \in (\hat{\alpha}_T, \hat{\beta}_T) \\ \frac{1}{\mu_T}\left[\gamma^* - \frac{(h+b)(e^{\hat{\nu}_T \hat{\beta}_T} - e^{\hat{\nu}_T \hat{\alpha}_T})e^{-\hat{\nu}_T x}}{\hat{\nu}_T^2(\hat{\beta}_T - \hat{\alpha}_T)} - hx + \frac{h}{\hat{\nu}_T} - \hat{a}_5\right] & \text{if } x \in [\hat{\beta}_T, z^*] \end{cases} \tag{A.12}$$

Case 2

$$\text{Condition:} \quad -\frac{1}{\hat{\nu}_T} \ln\left[\left(\frac{h}{b+h}\right)\left(\frac{\hat{\nu}(\hat{\beta}_T - \hat{\alpha}_T)}{e^{\hat{\nu}_T(\hat{\beta}_T - \hat{\alpha}_T)} - 1}\right)\right] + \hat{\alpha}_T < \hat{\beta}_T$$

z^* is the solution to

$$0 = \frac{2\hat{a}_2}{\hat{\nu}_T} e^{-\hat{\nu}_T(z^* - \hat{\alpha}_T)} + 2\hat{a}_2 z^* + \hat{a}_3 - \frac{2\hat{a}_2}{\hat{\nu}_T} \tag{A.13}$$

$$\gamma^* = \hat{a}_2 (z^*)^2 + \hat{a}_3 z^* + \hat{a}_4 \tag{A.14}$$

$$V^*(x) = \begin{cases} 0 & \text{if } x \geq z^* \\ \\ -\int_x^{z^*} V'(y)\,dy & \text{if } x < z^* \end{cases} \tag{A.15}$$

$$V'^*(x) = \begin{cases} \frac{1}{\mu_T}[\gamma^* + bx - \frac{b}{\hat{\nu}_T} - \hat{a}_1] & \text{if } x \leq \hat{\alpha}_T \\ \\ \frac{1}{\mu_T}\left[\gamma^* + \frac{(h+b)e^{-\hat{\nu}_T(x-\hat{\alpha}_T)}}{\hat{\nu}_T^2(\hat{\beta}_T-\hat{\alpha}_T)} - \hat{a}_2x^2 + \hat{a}_{17}x + \hat{a}_{18}\right] & \text{if } x \in (\hat{\alpha}_T, z^*] \end{cases} \tag{A.16}$$

where, in both cases

$$\hat{a}_{17} = \frac{2\hat{a}_2}{\hat{\nu}_T} - \hat{a}_3, \text{ and } \hat{a}_{18} = -\frac{2\hat{a}_2}{\hat{\nu}_T^2} + \frac{\hat{a}_3}{\hat{\nu}_T} - \hat{a}_4.$$

Proposition A.3 (γ^*, V^*, z^*) *satisfy (A.8) and*

$$\Gamma V(x) + g(x) - \gamma \geq 0 \text{ for } x \geq z \tag{A.17}$$

$$\Gamma V(x) + g(x) - \gamma = 0 \text{ for } x < z \tag{A.18}$$

$$V'(x) \leq 0 \text{ for } x < z \tag{A.19}$$

$$V'(x) = 0 \text{ for } x \geq z \tag{A.20}$$

Proof:

Case 1: We start by substituting γ^* and z^* into (A.18) and solving

$$\frac{1}{2}\sigma_T^2 V''(x) + \mu_T V'(x) + g(x) - \gamma^* = 0 \tag{A.21}$$

for $V'(x)$ which yields

$$V'(x) = Ke^{-\hat{\nu}_T x} + \frac{\gamma^*}{\mu_T} - \frac{2e^{-\hat{\nu}_T x}}{\sigma_T^2}\int_{-\infty}^{x} g(y)e^{\hat{\nu}_T y}\,dy \text{ for } x \leq z^* \tag{A.22}$$

where K is a constant. Setting $V'^*(z^*) = 0$ in (A.22) and integrating by parts gives

$$0 = \frac{he^{\hat{\nu}_T z^*}}{\mu_T \hat{\nu}_T} + K\frac{(h+b)(e^{\hat{\nu}_T\hat{\beta}_T} - e^{\hat{\nu}_T\hat{\alpha}_T})}{\hat{\nu}_T^2\mu_T(\hat{\beta}_T - \hat{\alpha}_T)}.$$

From the definition of z^* in the candidate solution we have that

$$\frac{he^{\hat{\nu}_T z^*}}{\mu_T \hat{\nu}_T} = \frac{(b+h)}{\mu_T\hat{\nu}_T^2}\left(\frac{e^{\hat{\nu}_T\hat{\beta}_T} - e^{\hat{\nu}_T\hat{\alpha}_T}}{\hat{\beta}_T - \hat{\alpha}_T}\right)$$

and hence $K = 0$. Using $K = 0$ and the definition of $g(x)$ from the optimal cycle placement solution in §2.3 into (A.22) gives (A.12) for $x \leq z^*$. Defining $V^*(x)$ as in (A.11) gives $V'^*(x) = 0$ for $x > z^*$ and hence the proposed solution satisfies conditions (A.18) and (A.20)

From (A.11) it follows that $\Gamma V^*(x) = 0$ for $x \geq z^*$, so that

$$\Gamma V^*(x) + g(x) - \gamma^* = h(x - z^*) \geq 0 \text{ for } x > z^*$$

and condition (A.17) is verified for the proposed solution.

We next verify that the proposed solution satisfies (A.19). To this order we proceed in two steps: (1) show that $V'^*(x)$ is increasing when $x < z^*$; (2) show that $V'^*(x) = 0$. These two results together imply the required condition.

To show that $V'^*(x)$ is increasing for $x \in (-\infty, z^*]$, notice first that (A.22) and the fact that $g(x)$ is continuous in x imply that $V''(x)$ is continuous. Now, for $x < \hat{\alpha}_T$:

$$\begin{aligned} V'^*(x) &= \frac{1}{\mu_T}\left[\gamma^* + bx - \frac{b}{\hat{\nu}_T} - \hat{a}_1\right] \\ V''^*(x) &= \frac{b}{\mu_T} > 0\text{rule}0.0\text{cm}1.0\text{cm} \end{aligned}$$

and hence $V'^*(x)$ is increasing over $x < \hat{\alpha}_T$. Consider now the range $x \in [\hat{\alpha}_T, \hat{\beta}_T]$:

$$\begin{aligned} V''^*(x) &= \frac{1}{\mu_T}\left[-2\hat{a}_2 x + \hat{a}_{17} - \frac{h+b}{\hat{\nu}_T(\hat{\beta}_T - \hat{\alpha}_T)}e^{\hat{\nu}_T(x-\hat{\alpha}_T)}\right] \\ V'''^*(x) &= \frac{h+b}{\mu_T(\hat{\beta}_T - \hat{\alpha}_T)}(e^{-\hat{\nu}_T(x-\hat{\alpha}_T)} - 1) < 0\text{rule}0.0\text{cm}1.0\text{cm} \end{aligned}$$

and hence $V''^*(x)$ is decreasing in $x \in [\hat{\alpha}_T, \hat{\beta}_T]$. Now, for $x \in [\hat{\beta}_T, z^*]$

$$\begin{aligned} V''^*(x) &= \frac{1}{\mu_T}\left[-h + \left(\frac{h+b}{\hat{\nu}_T(\hat{\beta}_T - \hat{\alpha}_T)}\right)(e^{\hat{\nu}_T\hat{\beta}_T} - e^{\hat{\nu}_T\hat{\alpha}_T})e^{-\hat{\nu}_T x}\right] \\ V'''^*(x) &= -\frac{h+b}{\mu_T(\hat{\beta}_T - \hat{\alpha}_T)}(e^{\hat{\nu}_T\hat{\beta}_T} - e^{\hat{\nu}_T\hat{\alpha}_T})e^{-\hat{\nu}_T x} < 0\text{rule0.0cm1.0cm} \end{aligned}$$

so that $V''^*(x)$ is still decreasing in $x \in [\hat{\beta}_T, z^*]$. Also, by definition of the proposed z^*, we have that

$$V''^*(z^*) = \frac{1}{\mu_T}\left[-h + \left(\frac{(h+b)(e^{\hat{\nu}_T\hat{\beta}_T} - e^{\hat{\nu}_T\hat{\alpha}_T})}{\hat{\nu}_T(\hat{\beta}_T - \hat{\alpha}_T)}\right)\left(\frac{h\hat{\nu}_T(\hat{\beta}_T - \hat{\alpha}_T)}{(h+b)(e^{\hat{\nu}_T\hat{\beta}_T} - e^{\hat{\nu}_T\hat{\alpha}_T})}\right)\right] = 0.$$

Therefore, $V''^*(x) \geq 0$ for $x < z^*$, which implies that $V'^*(x)$ increases over this range. Since,

$$V'^*(z^*) = \frac{1}{\mu_T}\left[hz^* + \hat{a}_5 - hz^* + \frac{h}{\hat{\nu}_T} - \hat{a}_5 - \frac{h}{\hat{\nu}_T}\right] = 0$$

the condition (A.19) is verified.

All that remains for Case 1 is to show that the proposed solution satisfies (A.8). The non-negativity requirement is easily verified since, by definition, $V^*(x) = 0$ for $x \geq z^*$, and by condition (A.19) $V^*(x)$ is decreasing in $x < z^*$. Now, since $V'^*(x)$ is decreasing for $x \in (-\infty, z^*]$ it follows that $V^*(x) < V^*(\hat{\alpha}_T)$ for $x \in (\hat{\alpha}_T, z^*]$. By definition of the proposed solution

$$V^*(x) = V^*(\hat{\alpha}_T) - \int_x^{\hat{\alpha}_T}\left(by + \gamma^* - \frac{b}{\hat{\nu}_T} - \hat{a}_1\right)dy \quad \text{for } x \leq \hat{\alpha}_T. \tag{A.23}$$

Let

$$\begin{aligned} K_3 &= -\gamma^* + \frac{b}{\hat{\nu}_T} + \hat{a}_1 \\ K_4 &= -b, \end{aligned}$$

then equation (A.23) yields that

$$V^*(x) = V^*(\hat{\alpha}_T) + K_3 x + K_4 x^2 \quad \text{for } x \leq \hat{\alpha}_T.$$

Finally, defining

$$K_5 = V^*(\hat{\alpha}_T) + \sup_{s\in[\hat{\alpha}_T, z^*]} \left| V^*(s) - V^*(\hat{\alpha}_T) - K_3 s - K_4 s^2 \right|$$

we have that

$$V(x) \leq K_5 + K_3 x + K_4 x^2$$

and the proof of Proposition A.3 is complete for Case 1.

Case 2: Note that, as mentioned in Chapter 2 after Proposition 2, $z^* < \hat{\beta}_T$ for this case. We start by substituting γ^* and z^* into (A.18) and then solving for $V'(x)$ to obtain

$$V'(x) = Ke^{-\hat{\nu}_T x} + \frac{\gamma^*}{\mu_T} - \frac{2e^{-\hat{\nu}_T x}}{\sigma_T^2} \int_{-\infty}^{x} g(y) e^{\hat{\nu}_T y}\, dy \quad \text{for } x \leq z^*. \tag{A.24}$$

Setting $V'^*(z^*) = 0$ and integrating (A.24) by parts (using the fact that $z^* < \hat{\beta}_T$) yields

$$Ke^{-\hat{\nu}_T z^*} + \frac{\gamma^*}{\mu_T} - \frac{2}{\hat{\nu}_T \sigma_T^2} \left[\hat{a}_2 z^{*2} - \hat{a}_{17} z^* - \hat{a}_{18} \right] + \frac{2\hat{a}_2 e^{-\hat{\nu}_T (z^* - \hat{\alpha}_T)}}{\mu_T \hat{\nu}_T^2} = 0. \tag{A.25}$$

Using the definition of γ^*, z^* and the fact that $\hat{\nu}_T \sigma_T^2/2 = \mu_T$ in (A.25) gives

$$Ke^{-\hat{\nu}_T z^*} + \frac{1}{\hat{\nu}_T \mu_T} \left[2\hat{a}_2 z^* - \hat{a}_{17} - 2\hat{a}_2 z^* + \hat{a}_{17} \right] = 0.$$

Therefore, as in Case 1, we have $K = 0$. Using this and $g(x)$ in (A.24) gives (A.16). Defining $V^*(x)$ as in (A.15) implies $V'^*(x) = 0$ for $x > z^*$, and hence conditions (A.18) and (A.20) are verified for this case.

In order to verify (A.17), we start from the fact that, by definition of the proposed solution, $\Gamma V^*(x) = 0$ for $x \geq z^*$. Now, for $x \in [\hat{\beta}_T, z^*]$

$$\Gamma V^*(x) + g(x) - \gamma^* \;=\; \hat{a}_2 \left[(x^2 - z^{*2}) + \frac{\hat{a}_3}{\hat{a}_2}(x - z^*) \right] \tag{A.26}$$

where, as defined in the cycle placement solution of §2.3,

$$\hat{a}_2 = \frac{h+b}{2(\hat{\beta}_T - \hat{\alpha}_T)} > 0. \tag{A.27}$$

We need to show that the RHS of (A.26) is non-negative. In view of (A.27), all we need is to establish the non-negativity of the term inside the brackets. This will be done by showing that

$$z^* \geq \hat{a}_{19} = -\frac{\hat{a}_3}{2\hat{a}_2}.$$

Notice that this would imply

$$(x^2 - z^{*2}) + \frac{\hat{a}_3}{\hat{a}_2}(x - z^*) \geq (x - z^*)^2 \geq 0.$$

According to the definition of the proposed solution for Case 2, z^* solves

$$-\frac{1}{\hat{\nu}_T} e^{-\hat{\nu}_T(z^* - \hat{\alpha}_T)} = z^* + \frac{\hat{a}_3}{2\hat{a}_2} - \frac{1}{\hat{\nu}_T} \tag{A.28}$$

and, as discussed after Proposition 2 of §2.4, this solution satisfies $z^* > \hat{\alpha}_T$. Now, $z^* > \hat{\alpha}_T$ implies

$$-\frac{1}{\hat{\nu}_T} e^{-\hat{\nu}_T(z^* - \hat{\alpha}_T)} > \frac{1}{\hat{\nu}_T}. \tag{A.29}$$

Using (A.29) in (A.28) we obtain $z^* \geq \hat{a}_{19}$, which was the desired condition. Consider now the case where $x \geq \hat{\beta}_T$. Again, using the definition of $V^*(x)$ we have that $\Gamma V^*(x) = 0$ for x in this range. We thus obtain

$$\Gamma V^*(x) + g(x) - \gamma^* \quad = \quad hx + \hat{a}_5 - \hat{a}_2 z^{*2} - \hat{a}_3 z^* - \hat{a}_4 \ \text{ for } x > \hat{\beta}.$$

From the characterization of $g(x)$ in §2.3, we have that $g'(x) = h$ for $x > \hat{\beta}_T$ and that

$$h\hat{\beta}_T + \hat{a}_5 = \hat{a}_2\hat{\beta}_T^2 + \hat{a}_3\hat{\beta}_T + \hat{a}_4.$$

Therefore,

$$\begin{aligned} hx + \hat{a}_5 - \hat{a}_2 z^{*2} - \hat{a}_3 z^* - \hat{a}_4 &\geq \hat{a}_2 \hat{\beta}_T^2 + \hat{a}_3 \hat{\beta}_T - \hat{a}_2 z^{*2} - \hat{a}_3 z^* \text{ for } x > \hat{\beta} \\ &= g(\hat{\beta}) - g(z^*). \end{aligned}$$

The equation $\hat{a}_2 z^{*2} + \hat{a}_3 z^* + \hat{a}_4 = g(z^*)$ follows from the fact (as discussed after Proposition 2 of §2.4) that $z^* \in (\hat{\alpha}_T, \hat{\beta}_T)$ when the system parameters satisfy the condition stated for Case 2. For $x \in (\hat{\alpha}_T, \hat{\beta}_T)$ we have that $g''(x) = 2\hat{a}_2 > 0$. Also, by substitution of the relevant definitions, it is a simple matter to establish that $\hat{a}_{19} \in (\hat{\alpha}_T, \hat{\beta}_T)$ and $g'(\hat{a}_{19}) = 0$. Therefore, $g(x)$ is increasing in $[\hat{a}_{19}, \hat{\beta}_T]$, and hence

$$g(\hat{\beta}_T) \geq g(z^*).$$

Which completes the verification of condition (A.17) for Case 2.

In order to verify that the proposed solution satisfies condition (A.19) we again proceed in two steps: (1) show that $V'(x)$ is increasing in $x < z^*$; (2) show that $V'(z^*) = 0$.

As in Case 1, we get started by noticing that the continuity of $g(x)$ and (A.24) imply that $V'^*(x), V''^*(x)$ are continuous in x. Now, for $x \leq \hat{\alpha}_T$

$$V''(x) = \frac{b}{\mu_T} > 0 \tag{A.30}$$

so that $V'(x)$ increases in this range. Considering now the range $x \in (\hat{\alpha}_T, z^*]$, we have that

$$\begin{aligned} V''^*(x) &= \frac{1}{\mu_T}\left[-\frac{(h+b)e^{-\hat{\nu}_T(x-\hat{\alpha}_T)}}{\hat{\nu}_T(\hat{\beta}_T - \hat{\alpha}_T)} - 2\hat{a}_2 x + \hat{a}_{17}\right] && \text{(A.31)} \\ V'''^*(x) &= \frac{1}{\mu_T}\left[\frac{(h+b)e^{-\hat{\nu}_T(x-\hat{\alpha}_T)}}{\hat{\beta}_T - \hat{\alpha}_T} - 2\hat{a}_2\right] \end{aligned}$$

$$= \frac{(h+b)}{\mu_T(\hat{\beta}_T - \hat{\alpha}_T)} \left[e^{-\hat{\nu}_T(x-\hat{\alpha}_T)} - 1\right] < 0 \tag{A.32}$$

Therefore, $V'''^*(x)$ is decreasing for $x \in (\hat{\alpha}_T, z^*]$. Also, by (A.16) $V'''^*(z^*) = 0$, and hence $V'''^*(z^*) \geq 0$ for $x \in (\hat{\alpha}_T, z^*]$. We have then that $V'''^*(z^*) \geq 0$ for $x \in (-\infty, z^*]$ which implies that $V'^*(x)$ increases over this range. Using (A.13) in (A.16) gives $V'^*(z^*) = 0$, and therefore condition (A.19) is verified.

Finally, that the proposed solution for this case satisfies (A.8) follows from the same argument as in Case 1. ■

Step 3. We now prove the optimality of the candidate policy.

Proposition A.4 *The policy $Y(t)$ defined in (A.1) is an optimal solution to problem (A.2)-(A.4).*

Proof. Conditions (A.17)-(A.20) imply the variational inequality (A.7). Thus, by Propositions A.1 and A.3, the value of γ^* is a lower bound on the optimal long run average cost for any arbitrary policy in problem (P2). Comparing γ^* with $\hat{F}_T(z_T^*)$ of Proposition 2, we have that the cost of base stock policy proposed in Chapter 2 achieves this lower bound. Hence this policy is an optimal solution to (P2). Furthermore, (P2) and (A.2)-(A.4) are equivalent problems and so the proof is complete. ■

Appendix B

Complete Cost Expression for the Double Threshold Policy: Dynamic Routing VRIP

This appendix contains the detailed steady state cost function that corresponds to the double threshold policy proposed for the dynamic routing VRIP. This particular instance of the function is obtained from (5.51) under the assumptions that all retailers have the same holding and backordering costs (i.e. $h_i = h$ and $b_i = b$ for $i = 1, \ldots, m$), and that the system has the wedge topology described in the fixed-route simulations of Chapter 4 (i.e. $\theta_{0i} = \theta_{01}$ for all i and $\theta_{i,i+1} = \theta_{12}$ for $i = 1, \ldots, m-1$).

The cost takes a different functional form over eight sections of the region in the (s, z) plane over which it is defined. Namely:

$$
\begin{aligned}
&\text{Case 1.} && s \leq \hat{\alpha}_D, \ \text{ and } \ z > \hat{\beta}_T \\
&F(s,z) &=\ & \delta\left[\hat{a}_{20}s + \hat{a}_{21} + (\hat{a}_{22} + \hat{a}_{23})e^{-\hat{\nu}_T s} + (\hat{a}_{24}z + \hat{a}_{25})e^{\hat{\nu}_T (z-s)}\right] \\
& && + \hat{a}_{26} + \frac{\hat{a}_{27}\delta}{(1-\delta)\vartheta_{01} + \delta\vartheta_T} \qquad \text{(B.1)} \\
&\text{Case 2.} && s \leq \hat{\alpha}_D, \ \text{ and } \ z \in [0, \hat{\beta}_T] \\
&F(s,z) &=\ & \delta\left[\hat{a}_{20}s + \hat{a}_{21} + \hat{a}_{23}e^{-\hat{\nu}_T s} + (\hat{a}_{28}z^2 + \hat{a}_{29}z + \hat{a}_{30})e^{\hat{\nu}_T (z-s)}\right]
\end{aligned}
$$

$$+ \hat{a}_{26} + \frac{\hat{a}_{27}\delta}{(1-\delta)\vartheta_{01} + \delta\vartheta_T} \tag{B.2}$$

Case 3. $s \in (\hat{\alpha}_D, \hat{\alpha}_T]$, and $z > \hat{\beta}_T$

$$F(s,z) = \delta\left[\hat{a}_{31}s^2 + \hat{a}_{32}s + \hat{a}_{33} + \hat{a}_{34}e^{-\hat{\nu}_D s} + (\hat{a}_{22} + \hat{a}_{23})e^{-\hat{\nu}_T s} + (\hat{a}_{24}z + \hat{a}_{25})e^{\hat{\nu}_T(z-s)}\right] + \hat{a}_{26} + \frac{\hat{a}_{27}\delta}{(1-\delta)\vartheta_{01} + \delta\vartheta_T} \tag{B.3}$$

Case 4. $s \in (\hat{\alpha}_D, \hat{\alpha}_T]$, and $z \in [0, \hat{\beta}_T]$

$$F(s,z) = \delta\left[\hat{a}_{31}s^2 + \hat{a}_{32}s + \hat{a}_{33} + \hat{a}_{34}e^{-\hat{\nu}_D s} + \hat{a}_{23}e^{-\hat{\nu}_T s} + (\hat{a}_{28}z^2 + \hat{a}_{29}z + \hat{a}_{30})e^{\hat{\nu}_T(z-s)}\right] + \hat{a}_{26} + \frac{\hat{a}_{27}\delta}{(1-\delta)\vartheta_{01} + \delta\vartheta_T} \tag{B.4}$$

Case 5. $s \in (\hat{\alpha}_T, \hat{\beta}_T]$, and $z > \hat{\beta}_T$

$$F(s,z) = \delta\left[(\hat{a}_{31} + \hat{a}_{35})s^2 + \hat{a}_{36}s + \hat{a}_{37} + \hat{a}_{34}e^{-\hat{\nu}_D s} + \hat{a}_{22}e^{-\hat{\nu}_T s} + (\hat{a}_{24}z + \hat{a}_{25})e^{\hat{\nu}_T(z-s)}\right] + \hat{a}_{26} + \frac{\hat{a}_{27}\delta}{(1-\delta)\vartheta_{01} + \delta\vartheta_T} \tag{B.5}$$

Case 6. $s \in (\hat{\alpha}_T, \hat{\beta}_T]$, and $z \in [s, \hat{\beta}_T]$

$$F(s,z) = \delta\left[(\hat{a}_{31} + \hat{a}_{35})s^2 + \hat{a}_{36}s + \hat{a}_{37} + \hat{a}_{34}e^{-\hat{\nu}_D s} + (\hat{a}_{28}z^2 + \hat{a}_{29}z + \hat{a}_{30})e^{\hat{\nu}_T(z-s)}\right] + \hat{a}_{26} + \frac{\hat{a}_{27}\delta}{(1-\delta)\vartheta_{01} + \delta\vartheta_T} \tag{B.6}$$

Case 7. $s \in (\hat{\beta}_T, \hat{\beta}_D]$, and $z \in [s, \infty)$

$$F(s,z) = \delta\left[\hat{a}_{31}s^2 + \hat{a}_{38}s + \hat{a}_{39} + \hat{a}_{34}e^{-\hat{\nu}_D s} + (\hat{a}_{24}z + \hat{a}_{25})e^{\hat{\nu}_T(z-s)}\right] + \hat{a}_{26} + \frac{\hat{a}_{27}\delta}{(1-\delta)\vartheta_{01} + \delta\vartheta_T} \tag{B.7}$$

Case 8. $s > \hat{\beta}_D$, and $z \in [s, \infty)$

$$F(s,z) = \delta\left[\hat{a}_{40}s + \hat{a}_{41} + (\hat{a}_{34} + \hat{a}_{42})e^{-\hat{\nu}_D s} + (\hat{a}_{24}z + \hat{a}_{25})e^{\hat{\nu}_T(z-s)}\right] + \hat{a}_{26} + \frac{\hat{a}_{27}\delta}{(1-\delta)\vartheta_{01} + \delta\vartheta_T} \tag{B.8}$$

where

$$\delta = \frac{\nu_T}{\nu_T + \nu_D(e^{\nu_T(z-s)} - 1)}.$$

The parameters in equations (B.1)-(B.1) are defined as follows: $(\hat{\alpha}_T, \hat{\beta}_T, \hat{\nu}_T)$ and $(\hat{\alpha}_D, \hat{\beta}_D, \hat{\nu}_D)$ have the same definitions as in the fixed route cases of Chapters 2 and 3 respectively; the $\hat{a}_k$ constants are given by

$$
\begin{aligned}
\hat{a}_{20} &:= \left(\frac{\hat{\nu}_D}{\hat{\nu}_T} - 1\right) b \\
\hat{a}_{21} &= \left(\frac{1}{\hat{\nu}_D} - \frac{\hat{\nu}_D}{\hat{\nu}_T^2} + \frac{C}{2} - \sum_i \lambda_i \vartheta_{0i}^{TSP}\right) b \\
\hat{a}_{22} &= \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(\frac{b+h}{\hat{\nu}_T^2 C}\right) e^{\hat{\nu}_T \hat{\beta}_T} \\
\hat{a}_{23} &= -\frac{\hat{\nu}_D}{\hat{\nu}_T}\left(\frac{b+h}{\hat{\nu}_T^2 C}\right) e^{\hat{\nu}_T \hat{\alpha}_T} \\
\hat{a}_{24} &= \frac{\hat{\nu}_D}{\hat{\nu}_T} h \\
\hat{a}_{25} &= \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(\frac{C}{2} - \frac{1}{\hat{\nu}_T} - \sum_i \lambda_i \vartheta_{0i}^{TSP}\right) h \\
\hat{a}_{26} &= -\frac{f}{\sqrt{n}}(1 - \rho_T) \\
\hat{a}_{27} &= -\frac{f}{\sqrt{n}}\vartheta_T(\rho_T - \rho_D) \\
\hat{a}_{28} &= \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(\frac{b+h}{2C}\right) \\
\hat{a}_{29} &= \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(h - \frac{b+h}{C}\sum_i \lambda_i \vartheta_{0i}^{TSP} - \frac{b+h}{\hat{\nu}_T C}\right) \\
\hat{a}_{30} &= \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(\frac{b+h}{2C}\left(\frac{hC}{b+h} - \sum_i \lambda_i \vartheta_{0i}^{TSP}\right)^2 + \frac{bhC}{2(b+h)} + \frac{b+h}{\hat{\nu}_T^2 C} - \frac{h}{\hat{\nu}_T}\right. \\
&\qquad \left. + \frac{b+h}{\hat{\nu}_T C}\sum_i \lambda_i \vartheta_{0i}^{TSP}\right) \\
\hat{a}_{31} &= \frac{b+h}{2mC} \\
\hat{a}_{32} &= \frac{\hat{\nu}_D}{\hat{\nu}_T} b + \frac{h-b}{2} - \frac{b+h}{m\hat{\nu}_D C} \\
\hat{a}_{33} &= \frac{b+h}{m\hat{\nu}_D^2 C} + \frac{b-h}{2\hat{\nu}_D} + \frac{(b+h)mC}{8} + \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(\frac{C}{2} - \sum_i \lambda_i \vartheta_{0i}^{TSP} - \frac{1}{\hat{\nu}_T}\right) b \\
\hat{a}_{34} &= -\frac{b+h}{m\hat{\nu}_D^2 C} e^{-\hat{\nu}_D mC/2} \\
\hat{a}_{35} &= -\frac{(b+h)\hat{\nu}_D}{2\hat{\nu}_T C} \\
\hat{a}_{36} &= \frac{h-b}{2} - \frac{b+h}{m\hat{\nu}_D C} + \frac{(b+h)\hat{\nu}_D}{\hat{\nu}_T^2 C} - \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(h - \frac{b+h}{C}\sum_i \lambda_i \vartheta_{0i}^{TSP}\right)
\end{aligned}
$$

respectively; the $\hat{a}_k$ constants are given by

$$
\begin{aligned}
\hat{a}_{20} &= \left(\frac{\hat{\nu}_D}{\hat{\nu}_T} - 1\right) b \\
\hat{a}_{21} &= \left(\frac{1}{\hat{\nu}_D} - \frac{\hat{\nu}_D}{\hat{\nu}_T^2} + \frac{C}{2} - \sum_i \lambda_i \vartheta_{0i}^{TSP}\right) b \\
\hat{a}_{22} &= \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(\frac{b+h}{\hat{\nu}_T^2 C}\right) e^{\hat{\nu}_T \hat{\beta}_T} \\
\hat{a}_{23} &= -\frac{\hat{\nu}_D}{\hat{\nu}_T}\left(\frac{b+h}{\hat{\nu}_T^2 C}\right) e^{\hat{\nu}_T \hat{\alpha}_T} \\
\hat{a}_{24} &= \frac{\hat{\nu}_D}{\hat{\nu}_T} h \\
\hat{a}_{25} &= \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(\frac{C}{2} - \frac{1}{\hat{\nu}_T} - \sum_i \lambda_i \vartheta_{0i}^{TSP}\right) h \\
\hat{a}_{26} &= -\frac{f}{\sqrt{n}}(1-\rho_T) \\
\hat{a}_{27} &= -\frac{f}{\sqrt{n}}\vartheta_T(\rho_T - \rho_D) \\
\hat{a}_{28} &= \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(\frac{b+h}{2C}\right) \\
\hat{a}_{29} &= \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(h - \frac{b+h}{C}\sum_i \lambda_i \vartheta_{0i}^{TSP} - \frac{b+h}{\hat{\nu}_T C}\right) \\
\hat{a}_{30} &= \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(\frac{b+h}{2C}\left(\frac{hC}{b+h} - \sum_i \lambda_i \vartheta_{0i}^{TSP}\right)^2 + \frac{bhC}{2(b+h)} + \frac{b+h}{\hat{\nu}_T^2 C} - \frac{h}{\hat{\nu}_T}\right. \\
&\quad \left. + \frac{b+h}{\hat{\nu}_T C}\sum_i \lambda_i \vartheta_{0i}^{TSP}\right) \\
\hat{a}_{31} &= \frac{b+h}{2mC} \\
\hat{a}_{32} &= \frac{\hat{\nu}_D}{\hat{\nu}_T} b + \frac{h-b}{2} - \frac{b+h}{m\hat{\nu}_D C} \\
\hat{a}_{33} &= \frac{b+h}{m\hat{\nu}_D^2 C} + \frac{b-h}{2\hat{\nu}_D} + \frac{(b+h)mC}{8} + \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(\frac{C}{2} - \sum_i \lambda_i \vartheta_{0i}^{TSP} - \frac{1}{\hat{\nu}_T}\right) b \\
\hat{a}_{34} &= -\frac{b+h}{m\hat{\nu}_D^2 C} e^{-\hat{\nu}_D mC/2} \\
\hat{a}_{35} &= -\frac{(b+h)\hat{\nu}_D}{2\hat{\nu}_T C} \\
\hat{a}_{36} &= \frac{h-b}{2} - \frac{b+h}{m\hat{\nu}_D C} + \frac{(b+h)\hat{\nu}_D}{\hat{\nu}_T^2 C} - \frac{\hat{\nu}_D}{\hat{\nu}_T}\left(h - \frac{b+h}{C}\sum_i \lambda_i \vartheta_{0i}^{TSP}\right)
\end{aligned}
$$

References

Chan, L. M. A., A. Federgruen and D. Simchi-Levi. 1994. Probabilistic Analyses and Practical Algorithms for Inventory-Routing Models. Working Paper. Columbia University.

Coffman, E.G., A. A. Puhalskii and M. I. Reiman. 1993. Polling Systems with Zero Switchover Times: A Heavy-Traffic Averaging Principle. Submitted for publication.

Federgruen, A. and D. Simchi-Levi. 1992. Analytical Analysis of Vehicle Routing and Inventory-Routing Problems. *Handbooks in Operations Research an Management Science*, the volume on Networks and Distribution. M. Ball, T. Magnanti, C. Monma and G. Nemhauser, eds.

Federgruen, A. and P. Zipkin. 1984. A Combined Vehicle Routing and Inventory Allocation Problem. *Operations Research.* **32**, 1019-1037.

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

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

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

Harrison, J.M. and M.I. Taksar. 1983. Instantaneous Control of Brownian Motion. *Mathematics of Operations Research.* **8**, 439-453.

Iglehart, D. L. and W. Whitt. 1970. Multiple Channel Queues in Heavy Traffic I. *Advances in Applied Probability.* **2**, 150-177.

Kushner, H.J. 1971. *Introduction to Stochastic Control.* Holt, Rinehart and Winston, Inc. New York, NY.

Kushner, H.J. and P.G. Dupuis. 1992. *Numerical Methods for Stochastic Control Problems in Continuous Time.* Springer-Verlag. New York, NY.

Karlin S. and H. M. Taylor. 1981. *A Second Course in Stochastic Processes.* Academic Press, Inc. San Diego, CA.

Kumar, A., L. B. Schwarz and J.E. Ward . 1994. Risk-Pooling Along a Fixed Delivery Route Using a Dynamic Inventory Allocation Policy. Working Paper. Krannert Graduate School of Management, Purdue University.

Markowitz, D. M., M. I. Reiman and L. M. Wein. 1994. Heavy Traffic Analysis of the Stochastic Economic Lot Scheduling Problem. In preparation.

Reiman, M. I. and L. M.Wein. 1994. Dynamic Scheduling of a Two-Class Queue with Setups. Working Paper. Sloan School of Management.

Reiman, M. I. 1984. Open Queueing Networks in Heavy Traffic. *Mathematics of Operations Research* **9**, 441-458.

Simchi-Levi, D. 1992. Hierarchical Planning for Probabilistic Distribution Systems in Euclidean Spaces. *Management Sciences* **38**, 198-211.

Taksar, M. I. 1985. Average Optimal Singular Control and a Related Stopping Problem. *Mathematics of Operations Research.* **10**, 63-81.

Wein, L.M. 1992. Dynamic Scheduling of a Multiclass Make-to-Stock Queue. *Operations Research*. **40**, 724-735.