# A Stochastic Vendor Managed Inventory Problem and Its Variations

A Thesis
Presented to
The Academic Faculty

by

**Pairote Balun**

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

School of Industrial and Systems Engineering
Georgia Institute of Technology
April 23, 2004

# A Stochastic Vendor Managed Inventory Problem and Its Variations

Approved by:

Chelsea C. White III, Advisor

Martin W. P. Savelsbergh

Hayriye Ayhan

Richard F. Serfozo

Mark E. Ferguson
(Dupree College of Management)

Date Approved

# ACKNOWLEDGEMENTS

I would like to thank Professor Chip White, my advisor, for his guidance and support throughout the course of this research. Several friends and colleagues have provided their inputs. Two of them are Tiravat Assavapokee, at Geogia Tech, and Hyun-soo Ahn, formerly at the University of Michigan, Ann Arbor.

I am grateful for the contributions of all members of the dissertation committee. Their meaningful questions and comments have made this dissertation significantly better.

PAIROTE BALUN

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# SUMMARY

We analyze the problem of distributing units of a product from one storage location (depot) to multiple retailers, which face stochastic demand. A capacitated vehicle is employed to transport the product from the depot to the retailers. We assume that the decision maker has available current inventory levels of the retailers and that of the vehicle. Moreover, the vehicle can travel freely between inventory locations. Whenever a retailer is visited, the decision maker must decide how many units of the product to leave and which location to visit next. Because of the stochastic demand and centralized inventory control, we refer to this problem as the stochastic vendor managed inventory (SVMI) problem.

We formulate the SVMI problem as a finite horizon non-homogeneous Markov decision process. Based on the single-crossing property, we show how a particular retailer continues to be a vehicle's optimal destination as inventory levels of the retailers vary. Furthermore, based on the super-additive and sub-additive properties, an optimal number of units of the product to deposit at the current retailer is shown to have monotone relations with the inventory levels. We provide conditions on the problem parameters that are sufficient for these structural results. The multiperiod SVMI problem and the infinite horizon SVMI problem with periodic reward and transition structures are analyzed. For computational efficiency, we develop three suboptimal solution procedures, including one that takes advantage of the structural results. Additionally, we present a numerical study of the performance measures in the SVMI problem and a case study, which involves a distribution problem at the Coca-Cola Enterprises, Inc.

We consider four variations of the SVMI problem. These variations differ in either one or both of the available state information and the vehicle routing procedure. Analytically,

we compare the optimal expected total rewards for the SVMI problem and its variations. Our computational experience suggests a complementary relationship between the quality of state information and the size of the set of retailers that the vehicle can visit.

# CHAPTER I

# INTRODUCTION

We consider a distribution system that has a depot and multiple retailers and is managed by a single vendor. We assume that there is only one vehicle used to distribute units of one product from the depot to the retailers. The demand processes for the product at the retailers are assumed to be independent of each other. At each retailer, the demand process is stochastic, time-dependent, and history-independent. The vehicle is allowed to travel freely between inventory locations. Once the vehicle arrives at an inventory location (the depot or one of the retailers), the decision maker is informed of the current inventory levels of the retailers and that of the vehicle. The decision maker then decides how many units of the product to deposit at the current retailer, or pick up at the depot, and which location to visit next. We refer to this problem as the stochastic vendor managed inventory (SVMI) problem.

We formulate the SVMI problem as a finite horizon non-homogeneous Markov decision process. Subsequently, we establish monotone relations between the optimal vehicle routing and inventory actions and inventory levels of the retailers. These results are extended to the multiperiod SVMI problem and to the infinite horizon SVMI problem with periodic reward and transition structures. To make the SVMI problem less computationally demanding, we develop a heuristic solution procedure based on the structural results for inventory control. There is also a heuristic procedure where the inventory decisions are based on a base-stock inventory policy. Additionally, myopic solutions of the infinite horizon SVMI problem are discussed. We present a numerical analysis of the performance measures in the SVMI problem. There is also a case study, which involves a distribution problem at the Coca-Cola Enterprises, Inc. For the case study, we establish structural results similar to those for the SVMI problem.

It is interesting to analyze how the quality of state information and the flexibility in vehicle routing procedure can affect the operating performance of the distribution system. To that goal, we consider four variations of the SVMI problem and present an analytical comparison of their optimal expected total rewards. Based on the theoretical results, the optimal expected total reward increases as the state information improves and/or the vehicle routing procedure becomes more flexible (there are more retailers that the vehicle can visit next). Furthermore, sample numerical results suggest a complementary relationship between the quality of state information and the size of the set of locations that may be visited next.

The contributions presented in this dissertation are as follows. We determine the existence of monotone optimal policies for the SVMI problem based on the single-crossing, super-additive, and sub-additive properties. An analytical comparison of the optimal expected total rewards for different Markov decision process (MDP) models is included. The SVMI problem has several potential applications, which we shall mention later. Furthermore, computationally useful procedures are presented. These include a heuristic solution procedure based on the structural results and alternative value iteration algorithms for multiperiod and infinite horizon periodic Markov decision problems. Finally, further insights for the SVMI problem are gained from the numerical study of some performance measures in the SVMI problem.

The sections of this chapter begin with some preliminary remarks and related applications. Then there is a preview of subsequent chapters, which are the SVMI problem, infinite horizon periodic Markov decision processes, suboptimal solutions and then variations of the SVMI problem, the case study, and the numerical study. Finally, we present the literature review.

## 1.1 Preliminary Remarks

Integrating transportation and inventory problems in distribution logistics is not a new research topic. Over twenty years ago, the two problems were treated mostly separately. Since then the focus has shifted toward the integration of the two problems. However, much of the obtained results are based on deterministic models. Heuristic algorithms are a common path to the solutions, such as optimal lot sizes and shipping frequencies. On the other hand, we study the problem in a relatively high level of detail. Our solutions specify optimal inventory and vehicle routing actions at each decision epoch, which may occur frequently.

The MDP model of the SVMI problem becomes much more computationally demanding as the problem size increases. In this case, the growth in size of the problem comes primarily from the increases in the number of retailers and in the capacities of the retailers and the vehicle. Meanwhile, a longer planning horizon also induces more computational requirements, though not as significantly as the aforementioned factors. Despite these computational implications, we believe the results we present here are widely applicable. A vendor may have a significant number of retailers to manage; however, it likely has also several depots or distribution centers. In that case, we may decompose the original problem into subproblems, in which each depot serves only a limited number of retailers. Furthermore, the vehicle routing procedure in the SVMI problem is very flexible. In many instances, this may imply less need for the retailers to have high capacities.

In the next chapter, we show that monotone backward induction algorithms based on the structural results reduce the run time in solving the SVMI problem by almost $50\%$. Furthermore, in Chapter 4, heuristic solution procedures are shown to further improve the computational efficiency while, especially for the first heuristic algorithm, maintaining great quality of the resulting solutions. Specifically, the expected total reward of the best heuristic solution is only $0.3\%$ less than that of the optimal one. This suggests that more aggressive use of the theoretical basis for that solution approach is still a viable option.

For a distribution problem with multiple vehicles, it is reasonable to assume that we can assign the vehicles to non-intersecting groups of retailers. The result is multiple SVMI problems. In fact, the finite horizon in the SVMI problem is intended to represent a working day. In some instances, two or more vehicles are not needed to service a retailer in such a short period of time. One way to decompose the distribution problem with multiple vehicles into multiple one-vehicle problems is to solve an instance of the assignment problem in combinatorial optimization. This approach may not be optimal for the expected total reward criterion. However, it allows the decision maker to consider other factors, such as travel distances and drivers' hours, when he or she assigns the retailers to each vehicle.

In the SVMI problem and most of its variations, we assume that current inventory information is available. Given the current state of technology, we believe that this is a reasonable assumption. Researchers of supply chains made this assumption as far back as several years ago. Among other tools, electronic data interchange (EDI) has been mentioned frequently in research papers as the provider of up-to-date demand information. References include Hammond (1993) and Srinivasan et al. (1994), which discuss the benefits of EDI in the apparel industry and manufacturing sector, respectively.

## *1.2 Related Applications*

Possible applications of the SVMI problem include deliveries of liquids or gases to multiple tanks at various locations. It is undesirable or even dangerous to store too much of these products at a single location. This is especially true for a residential area or university campus, where such storage tanks are often used. The vendor can use sophisticated gauges to remotely monitor inventory levels of these tanks. As a result, timely deliveries of the products to these tanks are possible.

A soft-drink company may also be interested in applying the results that we have obtained for the SVMI problem. We are aware of available technology that helps the company monitor inventory levels of vending machines and soft-drink retailers. According to Gillies et al. (1997), Coca-Cola Enterprises, Inc., considered implementing such technology from Harvest Electronics (U.S.A.) in 1996. Clearly, it is desirable for the company to avoid experiencing stock-outs or having expired soda cans. With such monitoring system in place, the vendor can profitably manage the inventory levels of soft-dink products at the retail locations.

The SVMI problem can also be applied on a factory floor. A certain part may be needed in multiple manufacturing cells. Because of its confined space, only a limited number of units of the part can be stored in each cell. As a result, deliveries of the part to these cells are necessary. It is reasonable to assume that current inventory information is available. In this case, the uncertainty in productivity levels in the manufacturing cells gives rise to the randomness in demand rates for the part.

Another application of the SVMI problem is in the distribution of money to banks and automated teller machines. In this setting, current inventory levels can be made readily available. Because of security concerns, a limited amount of money should be kept at these retail locations. On the other hand, customers must be able to access their money on demand. Therefore, timely deliveries of the money to these retail locations are necessary. This distribution problem can be formulated as an instance of the SVMI problem.

## *1.3 Preview of the Chapters*

In this section, we briefly introduce each of the upcoming chapters. The primary problem is introduced first. It is followed by the infinite horizon periodic Markov decision processes, suboptimal solutions, and then variations of the SVMI problem. We then preview our case study and, finally, introduce the numerical analysis.

### 1.3.1 The Stochastic Vendor Managed Inventory Problem

The SVMI problem presented and analyzed in Chapter 2 is essentially a stochastic optimization problem that integrates the vehicle routing and inventory control. Physical components of the problem include a depot and multiple retailers. A single vehicle transports units of a product from the depot to the retailers. We assume that the depot holds a countably infinite number of units of the product. The retailers and the vehicle have finite capacities. Demands at the retailers are independent and, at each retailer, demand is stochastic and time-dependent. We formulate the SVMI problem as a finite horizon non-homogeneous Markov decision process. At each decision epoch, the information available to the decision maker includes current inventory levels of the retailers, the vehicle's location, and its inventory level. The decision maker decides how many units of the product to drop off at the current retailer, or pick up at the depot, and which inventory location (the depot or one of the retailers) the vehicle will travel to next. The vehicle can stay at its current location. After the vehicle has arrived at its destination and current state information becomes known, another decision epoch occurs. The problem objective is to maximize the expected total reward, which takes both inventory and transportation costs into account.

Theoretical results for the SVMI problem include structural results for both the vehicle routing and inventory control. Precisely, assuming equal travel times and a relatively weak condition on the demand distributions, a particular retailer continues to be an optimal vehicle's destination (from the depot) as inventory level of that retailer decreases and/or inventory levels of the other retailers increase. This result is based on the single-crossing property. With regard to inventory control, we establish monotone relations between an optimal number of units of the product to deposit at the current retailer and inventory levels of all retailers. The assumption on travel times is not required for the results on inventory control. Our numerical examples show that the structural results for vehicle routing and inventory control help reduce the run time in solving the SVMI problem by almost $50\%$.

### 1.3.2 Infinite Horizon Periodic Markov Decision Processes

In Chapter 2, the SVMI problem is formulated as a finite horizon non-homogeneous Markov decision process. By assuming that this stochastic process repeats itself infinitely, we obtain the infinite horizon periodic Markov decision process. In Chapter 3, we describe this class of Markov decision processes and study the existence and convergence of their solutions. Results for the infinite horizon SVMI problem are included. In particular, we extend the algorithms and structural results for the SVMI problem to the infinite horizon version of the problem. Here only the expected total discounted reward criterion is considered. It is relatively straightforward to show that the structural results for the SVMI problem are applicable in the multiperiod SVMI problem.

### 1.3.3 Suboptimal Solutions of the SVMI Problem

In Chapter 4, we present three heuristic solution procedures for the SVMI problem. The first is a solution approach based on the structural results for inventory control that we have previously established. It gives us near-optimal solutions and reasonable computational efficiency in solving our sample problems. In the second heuristic, inventory actions are chosen according to a base-stock policy. The base-stock inventory levels are determined according to a formula equivalent to that of the Newsvender's problem. This approach provides greater computational efficiency but lower expected total rewards of the solutions than the first heuristic. Finally, we study the performance of myopic policies in the infinite horizon SVMI problem.

### 1.3.4 Variations of the SVMI Problem

Four variations of the SVMI problem are investigated in Chapter 5. The distinguishing features between these variations are the available state information and how the vehicle route is selected. In particular, the first variation is the case in which there is a delay in obtaining state information and the vehicle visits the retailers in a fixed order. The second variation is similar to the first one but without the delay in state information. Meanwhile, in the third variation, the order of the retailers may be varied but only before the vehicle departs the depot at the start of each round of service. In all three variations, at each non-final retailer, the vehicle has the option of making a stop at the depot for replenishment before travelling to the next retailer in the order. Finally, we study a variation of the SVMI problem featuring an intersection between each pair of inventory locations. At each intersection, the decision maker receives current state information and determines which of the two inventory locations accessible from that intersection to visit next.

Analytically, we compare the optimal expected total rewards for the SVMI problem and its four variations. The results are intuitive and sample numerical results confirm our findings. Altogether, the theoretical results imply that the optimal expected total reward increases as the state information improves and/or the vehicle routing procedure becomes more flexible (more choices of retailers for the vehicle to visit next). Based on our computational experience, we propose a hypothesis that there is a complementary relationship between the quality of state information and the flexibility in vehicle routing procedure towards improving the operating performance of the distribution system.

### 1.3.5 The Coca-Cola Distribution Problem: A Case Study

Based on our telephone conversations with the logistics team at the Coca-Cola Enterprises, Inc., we decided to study a distribution problem facing the company. This problem shares some characteristics with the SVMI problem and its variations. In particular, it involves the production and packaging of soda cans at the cannery and transporting the products to the distribution centers, which face stochastic demand. We simplify the vehicle routing by considering only one distribution center. For the inventory control, we establish monotone relations between the optimal delivery actions and inventory levels of the soft-drink products at the distribution center.

### 1.3.6 A Numerical Study of the Performance Measures in the SVMI Problem

In Chapter 7, we study how certain parameters in the SVMI problem affects the customer service level and the optimal expected total reward. Both of these performance measures are important in many businesses in the retail industry. The problem parameters that we study include reward parameters, such as revenue and costs per unit, demand variance, and distance from the depot. In order to better represent real problems, we solve almost one thousand instances of the SVMI problem in that chapter. The results obtained are generally intuitive. In particular, we observe a positive correlation between profit margin and customer service level. Meanwhile, demand variance has both positive and negative correlations with the optimal expected total reward. Additionally, as expected, our numerical results show that longer distance from the depot reduces both customer service level and the optimal expected total reward.

## *1.4 Literature Review*

Review of related literature is as follows. The structural results that we establish in Chapter 2 are based primarily on results found in Section 4.7 of Puterman (1994). In Puterman's book, it is shown that the optimality of monotone policies follows from the super-additive property of the reward function. The paper by Serfozo (1976) provides the basis for this result. Meanwhile, the single-crossing property of multivariate functions, also essential in our structural results, belongs to the book by Topkis (1998). White and White (1989) present a helpful review of Markov decision processes.

The paper by Yang et al. (2000) is a recent study on the stochastic vehicle routing problem that has some characteristics similar to our SVMI problem. In the distribution problem investigated by the authors, which is referred to as the stochastic vehicle routing problem (SVRP) with restocking, physical components include a depot and multiple retailers. The problem objective is to determine an optimal route for the vehicle to visit the retailers and, in that route, the time(s) at which the vehicle goes back to the depot for replenishment before proceeding to the next retailer. Costs under consideration are the transportation cost, the replenishment cost, and the cost of emergency replenishment when the vehicle inventory is exhausted. Heuristics algorithms to identify superior routes are developed for the single-vehicle and the multiple-vehicle cases.

The differences between the SVRP with restocking and our SVMI problem are in the available state information and the way vehicle routes are selected. In the SVRP, the vehicle route, which includes trip(s) to the depot for replenishment, is specified before the vehicle's first departure from the depot. On the other hand, in our SVMI problem, from its current location, the vehicle is allowed to travel to any inventory location. Furthermore, current inventory levels of all retailers are available at each decision epoch, as suppose to only that of the current retailer in the SVRP with restocking.

The SVRP is a special case of the well-studied vehicle routing problem (VRP). Unlike our SVMI problem, the VRP generally focuses on designing optimal delivery or collection routes, where vehicles originate from a single depot and visit multiple geographically scattered locations. The inventory problem is usually either ignored or simplified. Deterministic models of the VRP are investigated in Christofides (1985), Golden and Assad (1988), and Laporte (1992). In Bertsimas (1992), the stochastic VRP is considered. In the paper, demands are assumed to be random and the objective is to find the vehicle route that minimizes the total distance traveled. The solution technique involves updating the vehicle route as demand information becomes available, without re-optimization. Computational approaches for the stochastic VRP are discussed in Bertsimas et al. (1995).

Gendreau et al. (1995) solve the stochastic VRP, in which each customer is randomly present and has random demand. Before the set of customers is known, planned collection routes are chosen. Then, in each collection route, absent customers are simply skipped. The stochastic VRP is formulated as a mixed integer program and solved by an integer L-shaped method. A tabu-search heuristic algorithm is developed for the problem in the follow-up paper, Gendreau et al. (1996). In a recent paper, Kleywegt et al. (2002), a variation of the VRP, which is referred to as the inventory routing problem (IRP), is formulated as a Markov decision process. Approximation methods for solving the stochastic IRP are proposed. For the case that only one customer is visited on a vehicle route, namely the stochastic IRP with direct deliveries, computational results are presented.

Previous studies of integrated inventory and transportation problems focus on deterministic models. We shall refer interested readers to the survey by Bertazzi and Speranza (1999). In the survey, the authors classify these problems into continuous-time and discrete-time models. Further classifications are based on the numbers of origins and destinations in the logistics problems. For instance, the SVMI problem would be regarded as a discrete time, one origin-multiple destinations case.

The paper by Federgruen and Zipkin (1984a) was among the first studies on stochastic models of the integrated inventory and transportation problems. In that setting, demand is random. Results presented in the paper include an algorithm that solves the inventory and routing problems separately and later combines the two solutions. The inventory allocation problem is formulated as a constrained non-linear optimization problem. This problem is solved repeatedly as part of the algorithm that solves the original integrated inventory and transportation problem.

Chien et al. (1989) study a combined inventory allocation and vehicle routing problem as a mixed integer program. This is the problem of delivering a limited amount of inventory from a single depot to multiple customers using a fleet of vehicles. A Lagrangian-relaxation-based procedure is developed to solve the mixed integer program. According to the authors, the heuristic algorithm performs well in several cases.

In the paper by Cetinkaya and Lee (2000), a renewal model is used in the study of stock replenishment and shipment scheduling for vendor managed inventory systems. In that setting, customer demand is random. The objective is to determine the replenishment quantity and shipment-release policy. The vehicle routing is not considered. Therefore, the distribution problem studied in the paper is simpler than our SVMI problem.

An application of integrated inventory and transportation problems is presented in the paper by Dror and Ball (1987). In the paper, the authors investigate the problem of distributing heating oil among customers using a fleet of vehicles. The problem objective is to minimize the annual delivery and stock-out costs. Both the deterministic and stochastic demand cases are studied. A heuristic algorithm was developed based on an interchange procedure and an LP-based assignment algorithm. According to the authors, the algorithm provides an increase in performance of 25 percent over a previous one.

Review of relatively recent literature on the effect of information sharing in supply chains is as follows. Gavirneni et al. (1999) consider a two-echelon capacitated supply chain model with three levels of information sharing between the supplier and retailer. Numerical results are used to study the relationships between the value of information and the supplier's capacity, inventory, and information, and the retailer's order quantity and demand distribution. In Lee et al. (2000), analytical results for a two-level supply chain show that the benefits of information sharing can be significantly high when demand is correlated over time. Cachon and Fisher (2000) study the inventory model with one supplier and multiple retailers. Each retailer faces stationary stochastic demand. Their numerical results suggest that information is significantly more valuable when it is used to reduce lead time and batch size than when it is simply shared with the supplier. There are numerous other papers on supply chains with one depot and multiple retailers. However, none that we know of has studied the relationship between state information quality and vehicle routing strategy in improving the operating performance of the distribution system.

# CHAPTER II

# THE STOCHASTIC VENDOR MANAGED INVENTORY PROBLEM

The integrated vehicle routing and inventory control problem that we study involves a depot, multiple retailers, and a single vehicle. The vehicle is used to transport units of a product from the depot to the retailers. We assume that the demand processes for the product at the retailers are independent of each other. At each retailer, the demand process is history-independent, stochastic, and time-dependent. We shall study five different models of this distribution problem. These models differ in the available state information and/or how the vehicle route is selected. The primary model is the topic of this chapter. In this setting, we assume that, at each decision epoch, the decision maker has available current inventory levels of the retailers and that of the vehicle. Furthermore, the vehicle can travel freely between all inventory locations (the depot and the retailers). We refer to this problem as the stochastic vendor managed inventory (SVMI) problem.

In the next two sections, we describe and formulate the SVMI problem. Preliminary results are presented in Section 2.3. In the two sections that follow, we establish structural results for the vehicle routing and inventory control. Sufficient conditions on the problem parameters are provided for both sets of structural results. In Section 2.6, we mention some applications of the structural results. Numerical examples are presented in Section 2.7. Subsequently, we briefly discuss distribution problems with multiple vehicles.

## *2.1 Problem Description*

We assume that the depot holds a countably infinite number of units of the product. The vehicle and the retailers have finite capacities. A decision epoch occurs when the vehicle arrives at an inventory location (the depot or one of the retailers). The time between the current and the next decision epochs equals the time required for the vehicle to travel from its current location to its chosen destination. If the vehicle's destination is its current location, then the next decision epoch occurs one time unit from the current time. We assume that the travel time between any two inventory locations is deterministic.

At each decision epoch, the decision maker is assumed to have available the demand distribution for each retailer. These demand distributions depend only on the current time and the time until the next decision epoch. The sufficient conditions for the structural results to be established later in this chapter are based on this assumption. To satisfy this assumption, the number of orders that arrive at each retailer between any two successive decision epoches must be history-independent.

By definition, the SVMI problem has history-independent demand processes. In particular, the demand at each retailer is such that the demands in disjoint time intervals are independent (the demands have independent increments that need not be stationary). This allows the demand in an interval to depend on the location of the interval in time as well as the length of the interval. For instance, the demand might be a time-dependent Poisson process or time-dependent compound Poisson process. Our assumption rules out renewal processes since they do not have independent increments. In Appendix B, we claim that any demand processes that have independent increments are history-independent. We remark that the history-independent property of the demand processes also implies that the SVMI problem is Markovian.

We assume that over the finite problem horizon, the (capacitated) vehicle picks up units of the product at the depot and deposits them at the retailers. The vehicle is not allowed to pick up units of the product at the retailers. At each decision epoch, the decision maker selects two actions:

1. the amount of inventory to leave (if the vehicle is at a retailer) or pick up (if the vehicle is at the depot);

2. the location to visit next.

We refer to the two actions above as the inventory action and the vehicle routing action, respectively. The decision maker selects these actions based on the current location of the vehicle, the vehicle's current inventory level, and the current inventory levels of the retailers. Before the end of the problem horizon, at each decision epoch, we assume that the vehicle can stay at its current location or travel to any one of the retailers or to the depot. Thus, a vehicle routing action prior to the end of the horizon can be any one of the locations. After the end of the horizon, the vehicle must return to the depot.

The net reward accrued between two successive decision epochs is the sum of the net rewards for the retailers minus the transportation cost for the vehicle. The costs and revenue generated at each retailer include a holding cost for current inventory, a revenue from filling orders, and a penalty cost for lost orders. There is no backlogging. In Appendix A, we discuss how our theoretical results would change if unfilled orders are backlogged. The retailer also incurs a per-unit procurement cost when the vehicle drops off non-zero units of the product. The net reward for each retailer is defined as the revenue minus the sum of the penalty cost, the holding cost, and, where applicable, the procurement cost. There is no discounting in the finite horizon problem.

## 2.2 Problem Formulation

We formulate the SVMI problem as follows. We assume that the depot holds a countably infinite number of units of the product. There are $N$ retailers, each of which has finite capacity, and a vehicle, also with finite capacity. We define $K = \{0, 1, 2, ..., N\}$ as the set of all inventory locations, where location $0$ represents the depot and location $i > 0$ represents retailer $i$. Let $q_i$, for $i = 1, 2, ..., N$, be the capacity of retailer $i$. Also, let $q_v$ denote the capacity of the vehicle.

We assume that there is a route between each pair of the locations. Assume that the problem horizon is finite and of length $T$. Let $t_j$ be the time of the $j^{th}$ decision epoch, where $t_1 = 1$. Let $J$ be the random integer such that $t_J \leq T$ and $t_{J+1} > T$. We note that if the vehicle is at location $l$ at time $t_j$ and, from there, it travels to location $k$, then $t_{j+1} = t_j + d_{lk}$, where $d_{lk}$ is the time required for the vehicle to travel from $l$ to $k$. We assume that $0 < d_{lk} < \infty$, for all $l, k \in K$.

We define the state at time $t$ as $s_t = (x, x_v, l)$, where $x$ is the vector of current inventory levels of the retailers, $x_v$ is the current inventory level of the vehicle, and $l$ is the current location of the vehicle. Thus, $x = (x_1, x_2, ..., x_N)$, in which $x_i$, for $i = 1, 2, ..., N$, is the inventory level of retailer $i$. Also, $l \in K = \{0, 1, 2, ..., N\}$. We have that $x_v \in X_v = \{0, 1, 2, ..., q_v\}$ and $x \in X = X_1 \times X_2 \times ... \times X_N$, where, for $i = 1, 2, ..., N$, $X_i = \{0, 1, 2, ..., q_i\}$.

Suppose the vehicle arrives at location $l$ at time $t$. The state information is immediately revealed and the inventory and vehicle routing decisions are then made. Let integer $a$ represent the inventory action. If $a > 0$, then $a$ units of inventory are removed from the vehicle and dropped off at the current location, which must be one of the retailers. No inventory action is taken if $a = 0$. On the other hand, if $a < 0$, then $-a$ units of inventory are picked up by the vehicle at the current location, which must be the depot. We assume that inventory cannot be removed from the retailers. We define $A_t(x, x_v, l)$ as the set of inventory actions available at a decision epoch at time $t$, given that the state of the system is $s_t = (x, x_v, l)$. It follows that, for $l = 0$,

$$A_t(x, x_v, l) = \{a : -(q_v - x_v) \leq a \leq 0\}$$

and, for $l > 0$,

$$A_t(x, x_v, l) = \{a : 0 \leq a \leq min\{q_l - x_l, x_v\}\}.$$

The corresponding set of vehicle routing actions is $K_t(x, x_v, l)$. Since the vehicle can travel to any location, $K_t(x, x_v, l) = K$, for $t = 1, 2, ..., T$. At a decision epoch after time $T$, the vehicle must return directly to the depot.

Let $t$ be a decision epoch. We define $r_t((x, x_v, l), a, k)$ as the reward accrued between time $t$ and time $t + d_{lk}$, where $a$ is the inventory action taken at time $t$, and $k$ is the next location to visit. Let $r_t^i(x_i, l, k)$ and $\tilde{r}_t^l(x_l, l, a, k)$ be the net rewards for the non-current retailer $i$ and the current retailer $l$, respectively. For $l = 0$,

$$r_t((x, x_v, l), a, k) = \sum_{1 \leq i \leq N} r_t^i(x_i, l, k) - c_{lk},$$

where $c_{lk}$ represents the cost of travelling from $l$ to $k$, for $l, k \in K$. If $l > 0$,

$$r_t((x, x_v, l), a, k) = \tilde{r}_t^l(x_l, l, a, k) + \sum_{i \in K \setminus \{0, l\}} r_t^i(x_i, l, k) - c_{lk}.$$

Additional parameters are defined as follows. For $i = 1, 2, ..., N$, $h_i$ is the inventory holding cost per unit per unit time, $b_i^1$ is the revenue per unit sold, $b_i^2$ is the penalty cost per unfilled order, and $b_i^3$ is the procurement cost per unit. We assume that $h_i, b_i^1, b_i^2, b_i^3 > 0$. Let $D_{t,i}^{l,k}$, a random variable with known distribution, represent the number of units demanded at retailer $i$ between time $t$ and time $t + d_{lk}$. Then,

$$r_t^i(x_i, l, k) = -h_i d_{lk} x_i + b_i^1 E[min\{D_{t,i}^{l,k}, x_i\}] - b_i^2 E[max\{0, D_{t,i}^{l,k} - x_i\}]$$

and, where $\tilde{x}_l = x_l + a$,

$$\tilde{r}_t^l(x_l, l, a, k) = -h_l d_{lk} \tilde{x}_l + b_l^1 E[min\{D_{t,l}^{l,k}, \tilde{x}_l\}] - b_l^2 E[max\{0, D_{t,l}^{l,k} - \tilde{x}_l\}] - b_l^3 a.$$

By convention, we denote any time $t > T$ by $T + 1$. The time-invariant terminal reward accrued at time $T + 1$ is

$$\bar{r}_{T+1}(x, x_v, l) = \sum_{1 \leq j \leq N} (e_j - h_j \tau) x_j - c_{l0},$$

where $e_j$ is the per unit salvage value at retailer $j$, $\tau$ is the holding period before salvage values are realized, and $c_{l0}$ is the the transportation cost for the vehicle to return to the depot from its current location. We assume that $e_j \geq h_j \tau$, for $j = 1, 2, ..., N$.

The transition structure is as follows. For $x = (x_1, x_2, ..., x_N)$, $y = (y_1, y_2, ..., y_N)$ and a decision epoch at time $t < T$, we define $p_t(y|(x, x_v, l), a, k)$ as the probability that the vector of inventory levels of the retailers at time $t + d_{lk}$ is $y$, given that the state at time $t$ is $s_t = (x, x_v, l)$ and actions $a$ and $k$ are taken. By independence, for $l = 0$,

$$p_t(y|(x, x_v, l), a, k) = \prod_{1 \leq i \leq N} p_t^i(y_i|x_i, l, k).$$

For $l > 0$,

$$p_t(y|(x, x_v, l), a, k) = p_t^l(y_l|x_l + a, l, k) \prod_{i \in K \setminus \{0,l\}} p_t^i(y_i|x_i, l, k).$$

In these two equations, for $j = 1, 2, ..., N$, $p_t^j(y'_j|x'_j, l, k)$ is the probability that the inventory level at retailer $j$ at time $t + d_{lk}$ is $y'_j$, given that the inventory level at this retailer after inventory action is taken at time $t$ is $x'_j$. This probability can be determined from the distribution of $D_{t,j}^{l,k}$. In particular, given that actions $a$ and $k$ are taken at time $t$,

$$x_l(t + d_{lk}) = max\{x_l(t) + a - D_{t,l}^{l,k}, 0\},$$

and for $i \in K \setminus \{0, l\}$,

$$x_i(t + d_{lk}) = max\{x_i(t) - D_{t,i}^{l,k}, 0\}.$$

We note that $D_{t,i}^{l,k}$ is history-independent. It follows that this model of the SVMI problem is Markovian.

We call $\delta_t : X \times X_v \times K \longrightarrow \tilde{A}_t \times \tilde{K}$, a decision rule, where $\tilde{A}_t = \bigcup_{x,x_v,l} A_t(x, x_v, l)$ and $\tilde{K}_t = \bigcup_{x,x_v,l} K_t(x, x_v, l)$. A policy $\pi$ is defined as a sequence of decision rules, $\pi = \{\delta_1, \delta_2, ..., \delta_T\}$. Any policy $\pi \in \Pi$, where $\Pi$ is the set of all deterministic Markov policies. Because the state and action sets are finite and the model is Markovian, we can restrict our attention to only this set of policies. For a proof, see Section 4.4 of Puterman (1994). We define the problem criterion as follows:

$$v_T^\pi(s_1) = E_{s_1}^\pi\{\sum\nolimits_{1 \leq j \leq J} r_{t_j}(s_{t_j}, a_{t_j}, k_{t_j}) + \bar{r}_{T+1}(s_{T+1})\}$$

That is, $v_T^\pi(s_1)$ is the expected total reward over the finite time horizon, given that policy $\pi$ is followed and that the state at time $t = 1$ is $s_1$. We say that a policy $\pi^*$ is optimal if $v_T^{\pi^*}(s_1) \geq v_T^\pi(s_1)$, for all $\pi \in \Pi$.

## 2.3 Preliminaries

By convention, we denote any time $t > T$ by $T + 1$. We write the optimality equations, including the boundary condition, for the SVMI problem as follows:

for $t \leq T$,

$$u_t((x, x_v, l), k) = max_{a \in A_t(x, x_v, l)} \{ r_t((x, x_v, l), a, k) + \sum_y p_t(y|(x, x_v, l), a, k) u^*_{t+d_{lk}}(y, x_v - a, k) \}$$

$$u^*_t(x, x_v, l) = max_{k \in K} \{ u_t((x, x_v, l), k) \}$$

for any $t > T$,

$$u^*_{T+1}(x, x_v, l) = \bar{r}_{T+1}(x, x_v, l) = \sum_{1 \leq j \leq N} (e_j - h_j \tau) x_j - c_{l0}.$$

From above, we may conclude that the value function of the problem satisfies the optimality equations. We shall refer to an optimal inventory action in the first optimality equation as $a^*(k)$ for the state $(x, x_v, l)$. For $i \in \{1, 2, ..., N\}$, let $x_i$ be the inventory level of interest. We may write $x$ as $x = (x_i, x^c_i)$, where $x^c_i$ is the row vector of inventory levels at retailers $j \in \{1, 2, ..., N\}$, $j \neq i$. As a result, the first two equations above may be written as follows:

$$u_t(((x_i, x^c_i), x_v, l), k) = max_{a \in A_t((x_i, x^c_i), x_v, l)} \{ r_t(((x_i, x^c_i), x_v, l), a, k) + \sum_{(y_i, y^c_i)} p_t((y_i, y^c_i)|((x_i, x^c_i), x_v, l), a, k) u^*_{t+d_{lk}}((y_i, y^c_i), x_v - a, k) \} \tag{2.3.1}$$

$$u^*_t((x_i, x^c_i), x_v, l) = max_{k \in K} \{ u_t(((x_i, x^c_i), x_v, l), k) \} \tag{2.3.2}$$

Since the ranges of $y_i$ and $y^c_i$ are independent, we may write equation 2.3.1 as

$$u_t(((x_i, x^c_i), x_v, l), k) = max_{a \in A_t((x_i, x^c_i), x_v, l)} \{ r_t(((x_i, x^c_i), x_v, l), a, k) + \sum_{y^c_i} \sum_{y_i} p_t((y_i, y^c_i)|((x_i, x^c_i), x_v, l), a, k) u^*_{t+d_{lk}}((y_i, y^c_i), x_v - a, k) \}. \tag{2.3.3}$$

Let $w_t(((x_i, x_i^c), x_v, l), a, k)$ be defined as follows:

$$w_t(((x_i, x_i^c),x_v, l), a, k) = r_t(((x_i, x_i^c), x_v, l), a, k) + \sum_{y_i^c} \sum_{y_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k) u^*_{t+d_{lk}}((y_i, y_i^c), x_v - a, k)$$

Consequently, equation 2.3.3 is the same as

$$u_t(((x_i, x_i^c), x_v, l), k) = max_{a \in A_t((x_i, x_i^c), x_v, l)} \{w_t(((x_i, x_i^c), x_v, l), a, k)\}. \quad (2.3.4)$$

For the rest of this section, we state definitions, properties, and concepts that will be important in establishing the structural results for the SVMI problem. We say that $\preceq$ is a binary relation on a set $V$ if, for all $v', v'' \in V$, the statement $v' \preceq v''$ is either true or false. If $v' \preceq v''$ and $v' \neq v''$, we write $v' \prec v''$. A partially ordered set is defined next.

**Definition 2.3.1.** *A partially ordered set is a set $V$ on which there is a binary relation $\preceq$ that has the following properties:*

1. $v \preceq v$*, for all* $v \in V$*,*

2. $v' \preceq v''$ *and* $v'' \preceq v'$ *imply that* $v' = v''$*, for all* $v', v'' \in V$*, and*

3. $v' \preceq v''$ *and* $v'' \preceq v'''$ *imply that* $v' \preceq v'''$*, for all* $v', v'', v''' \in V$*.*

The single-crossing property is defined next. This definition is derived from the one in Topkis (1998). Let $R$ denote the set of real numbers.

**Definition 2.3.2.** *(Topkis) Assume that $V$,$T$ are partially ordered sets with binary relations $\preceq_1$ and $\preceq_2$, respectively. Let $f(v, t)$ be a function from subset $S$ of $V \times T$ into $R$. Then $f(v, t)$ satisfies the single-crossing property in $(v, t)$ on $S$ if and only if, for all $v', v'' \in V$ and $t', t'' \in T$ with $v' \prec_1 v''$ and $t' \prec_2 t''$, and $\{v', v''\} \times \{t', t''\}$ being a subset of $S$, the following two conditions hold:*

$$f(v', t') \leq f(v'', t') \Rightarrow f(v', t'') \leq f(v'', t'')$$

$$f(v', t') < f(v'', t') \Rightarrow f(v', t'') < f(v'', t'')$$

We may write the two conditions above as follows:

$$f(v',t') \leq (<) f(v'',t') \Rightarrow f(v',t'') \leq (<) f(v'',t''),$$

which is equivalent to

$$f(v',t'') \geq (>) f(v'',t'') \Rightarrow f(v',t') \geq (>) f(v'',t').$$

Throughout the next section and beyond, we refer to the super-additive property, which is defined below. This definition is the same as the one given in Section 4.7 of Puterman (1994).

**Definition 2.3.3.** *(Puterman) Assume that $V$,$T$ are partially ordered sets with binary relations $\preceq_1$ and $\preceq_2$, respectively. Let $g(v,t)$ be a function from $V \times T$ into $R$. We say that $g(v,t)$ is <u>super-additive</u> in $(v,t)$ on $V \times T$ if and only if, for all $v', v'' \in V$ and $t', t'' \in T$ with $v' \prec_1 v''$ and $t' \prec_2 t''$,*

$$g(v',t') + g(v'',t'') \geq g(v',t'') + g(v'',t'). \tag{2.3.5}$$

When the inequality 2.3.5 is reversed, we say that $g(v,t)$ is <u>sub-additive</u> in $(v,t)$ on $V \times T$. Clearly, the inequality 2.3.5 is the same as

$$g(v'',t'') - g(v',t'') \geq g(v'',t') - g(v',t').$$

This proves the following lemma.

**Lemma 2.3.4.** *If $g(v,t)$ is super-additive in $(v,t)$ on $V \times T$, then $g(v,t)$ satisfies the single-crossing property in $(v,t)$ on $V \times T$.*

The following lemma establishes useful results on the single-crossing property and the super-additive property.

**Lemma 2.3.5.** *Assume that $V$ and $T$ are partially ordered sets with binary relations $\preceq_1$ and $\preceq_2$, respectively. Let $A$ be a finite set. Also, assume that the function $f : (V \times T) \to R$ has the single-crossing property in $(v,t)$ on $V \times T$. Finally, assume that the function $g : (V \times T) \to R$ is super-additive in $(v,t)$ on $V \times T$. Then,*

1. $\alpha f(v,t)$, *with* $\alpha > 0$, *and*

2. $f(v,t) + g(v,t)$,

*have the single-crossing property in* $(v,t)$ *on* $V \times T$.

*Proof.* First, $\alpha f(v,t)$, with $\alpha > 0$, has the single-crossing property. This follows from the definition of single-crossing property.

We now show that $f(v,t) + g(v,t)$ has the single-crossing property. Because $f(v,t)$ has the single-crossing property, the following conditions hold for $v', v'' \in V$ and $t', t'' \in T$ such that $v' \prec_1 v''$ and $t' \prec_2 t''$:

$$f(v',t') \leq f(v'',t') \Rightarrow f(v',t'') \leq f(v'',t'') \tag{2.3.6}$$

$$f(v',t') < f(v'',t') \Rightarrow f(v',t'') < f(v'',t'') \tag{2.3.7}$$

Also, $g(v,t)$ is super-additive in $(v,t)$ on $V \times T$ means that the following inequality holds:

$$g(v',t') + g(v'',t'') \geq g(v',t'') + g(v'',t') \tag{2.3.8}$$

This is equivalent to

$$g(v'',t') - g(v',t') \geq g(v'',t'') - g(v',t''). \tag{2.3.9}$$

Conditions (2.3.6) and (2.3.9) imply the following result:

$$f(v',t') + g(v',t') \leq f(v'',t') + g(v'',t') \Rightarrow f(v',t'') + g(v',t'') \leq f(v'',t'') + g(v'',t'') \tag{2.3.10}$$

Similarly, (2.3.7) and (2.3.9) imply the next result:

$$f(v',t') + g(v',t') < f(v'',t') + g(v'',t') \Rightarrow f(v',t'') + g(v',t'') < f(v'',t'') + g(v'',t'') \tag{2.3.11}$$

From the last two results, it follows that $f(v,t) + g(v,t)$ has the single-crossing property.

$\square$

Lemma 2.3.6 and Lemma 2.3.7 state the key results of super-additive property and sub-additive property, respectively. The following lemma is a variation of Lemma 4.7.1 in Puterman (1994).

**Lemma 2.3.6.** *Let $V$ and $T$ be partially ordered sets with binary relations $\preceq_1$ and $\preceq_2$, respectively. Furthermore, the set $T$ is finite. Assume that $g(v,t)$ is a function from $V \times T$ into $R$ and that $g(v,t)$ has super-additive property in $(v,t)$ on $V \times T$. For each $v \in V$, assume that $max_{t \in T} g(v,t)$ exists. Then*

$$max\{argmax_{t\in T} g(v'',t)\} \geq max\{argmax_{t\in T} g(v',t)\},$$

*for all $v' \preceq_1 v''$, $v', v'' \in V$.*

*Proof.* Since the set $T$ is finite, $max\{argmax_{t\in T} g(v,t)\}$ exists for all $v \in V$. Let $f(v) = max\{argmax_{t\in T} g(v,t)\}$. For $v' \preceq_1 v''$ and $t' \preceq_2 f(v')$, by the super-additive property of $g(v,t)$, we have that

$$g(v', f(v')) - g(v', t') \geq 0$$

and that

$$g(v', t') + g(v'', f(v')) \geq g(v', f(v')) + g(v'', t').$$

The second inequality is the same as

$$g(v'', f(v')) \geq [g(v', f(v')) - g(v', t')] + g(v'', t').$$

From the first inequality above, it follows that

$$g(v'', f(v')) \geq g(v'', t'),$$

for all $t' \preceq_2 f(v')$. As a result, $f(v'') \geq f(v')$. This is the same as the assertion of the lemma. □

**Lemma 2.3.7.** *Let $V$ and $T$ be partially ordered sets with binary relations $\preceq_1$ and $\preceq_2$, respectively. Furthermore, the set $T$ is finite. Assume that $\tilde{g}(v,t)$ is a function from $V \times T$*

*into $R$ and that $\tilde{g}(v,t)$ has sub-additive property in $(v,t)$ on $V \times T$. For each $v \in V$, assume that $max_{t \in T} \tilde{g}(v,t)$ exists. Then*

$$min\{argmax_{t \in T} \tilde{g}(v',t)\} \geq min\{argmax_{t \in T} \tilde{g}(v'',t)\},$$

*for all $v' \preceq_1 v''$, $v', v'' \in V$.*

*Proof.* Without loss of generality, let $\tilde{g}(v,t) = -g(v,t)$, where $g(v,t)$ is defined in the previous lemma. By an analogous proof to that of the previous lemma, the desired result holds. $\square$

As shown below, the ordering of vehicle destinations, excluding the depot, is defined with respect to a particular retailer.

**Definition 2.3.8.** *For each $i \in \{1, 2, ..., N\}$ and $k', k'' \in K \backslash \{0\}$, we say that $k' \prec_i k''$ if and only if $k' \in K_i'$ and $k'' \in K_i''$, where $K_i' = \{i\}$ and $K_i'' = K \backslash \{0, i\}$.*

The following definition of the single-crossing property shall be used in the structural results for vehicle routing.

**Definition 2.3.9.** *For each $i \in \{1, 2, ..., N\}$, we say that a function $f : K \times X_i \to R$ has single-crossing property w.r.t. retailer $i$ in $(k, x_i)$ on $K \times X_i$ if and only if, for each pair of $k', k'' \in K$ such that $k' \prec_i k''$ and for each pair of $x_i', x_i'' \in X_i$ such that $x_i' \leq x_i''$, the following condition holds:*

$$f(k', x_i') \leq (<) f(k'', x_i') \Rightarrow f(k', x_i'') \leq (<) f(k'', x_i'').$$

The following well-known lemma, from Section 4.7 of Puterman (1994) among others, will be referred to frequently later in this chapter.

**Lemma 2.3.10.** *Assume that $v_{i+1} \geq v_i$, for $i = \{1, 2, ..., M-1\}$ and that, for real-valued and non-negative sequences $\alpha_i$ and $\alpha_i'$, $\sum_{k \leq i \leq M} \alpha_i \geq \sum_{k \leq i \leq M} \alpha_i'$, for all $k$, with equality holding for $k = 0$. Then $\sum_{1 \leq i \leq M} \alpha_i v_i \geq \sum_{1 \leq i \leq M} \alpha_i' v_i$.*

Finally, we define a property that will be part of the sufficient conditions for a structure in the optimal inventory actions.

**Definition 2.3.11.** *For a discrete random variable $Z$ having $\{0, 1, 2, ...\}$ as its domain space, we say that $Z$ has non-increasing probability mass function if and only if, for any non-negative integers $z_1, z_2$ such that $z_1 \leq z_2$, $Pr(Z = z_1) \geq Pr(Z = z_2)$.*

## *2.4 Structural Results for Vehicle Routing*

We now examine how the optimal destination of the vehicle varies with inventory levels of the retailers, assuming all travel times are identical. We begin by making the following assumption.

**Assumption 2.4.1.** *For all $l, k \in K$, $d_{lk}$ = 1.*

We remark that there is no loss of generality between this assumption and the assumption that, for all $l, k \in K$, $d_{lk} = \bar{\tau}$, where $\bar{\tau}$ is a positive real number. Assumption 2.4.1 applies to this section only. Numerical examples that suggest the importance of this assumption in the structural results for vehicle routing are presented in Appendix C.

Next we establish an interesting relation between the optimal destination of the vehicle (from the depot) and inventory levels of the retailers. First, in Subsection 2.4.1, the structural results and their sufficient conditions are presented. Then, in Subsection 2.4.2, we show how these conditions hold in terms of the parameters of the SVMI problem. Finally, we summarize the results of the first two subsections.

### 2.4.1 Sufficient Conditions for the Structural Results

Proposition 2.4.2 specifies the property of the optimal value function which implies the desired structure in the optimal vehicle routing actions when the vehicle is at the depot.

**Proposition 2.4.2.** *Let $i \in \{1, 2, ..., N\}$ and $l = 0$, assume that $u_t(((x_i, x_i^c), x_v, l), k)$ has the single-crossing property w.r.t. retailer $i$ in $(k, x_i)$ on $K \times X_i$. At the depot, if an* *optimal (vehicle routing) action for the state $s''_t = ((x''_i, x^c_i), x_v, l)$ is to go to retailer $i$, then an optimal action for state $s'_t = ((x'_i, x^c_i), x_v, l)$, in which $x'_i \leq x''_i$, is to go to retailer $i$ or the depot.*

*Proof.* By assumption, for each pair of $x'_i, x''_i \in X_i$ such that $x'_i \leq x''_i$ and for each pair of $k', k'' \in K\backslash\{0\}$ such that $k' \prec_i k''$, the following condition holds:

$$u_t(((x''_i, x^c_i), x_v, l), k') \geq (>) u_t(((x''_i, x^c_i), x_v, l), k'')$$
$$\Rightarrow u_t(((x'_i, x^c_i), x_v, l), k') \geq (>) u_t(((x'_i, x^c_i), x_v, l), k''),$$

where $k' = i$ and $k'' \in K\backslash\{0, i\}$. That is, retailer $i$ is still preferred to other retailers, as the vehicle's destination, when its inventory level decreases. The depot is not part of this argument. So it could still be an optimal destination. □

The next proposition specifies the structural result for vehicle routing when the optimal value function has the single-crossing property w.r.t. all retailers.

**Proposition 2.4.3.** *For $l = 0$ and for all $n \in \{1, 2, ..., N\}$, assume that $u_t(((x_n, x^c_n), x_v, l), k)$ has the single-crossing property w.r.t. retailer $n$ in $(k, x_n)$ on $K \times X_n$. At the depot, if, for an $i \in \{1, 2, ..., N\}$, an optimal (vehicle routing) action for the state $\tilde{s}_t = ((\tilde{x}_i, \tilde{x}^c_i), x_v, l)$ is to go to retailer $i$, then an optimal action for state $s_t = ((x_i, x^c_i), x_v, l)$, in which $x_i \leq \tilde{x}_i$ and $x_j \geq \tilde{x}_j$, for all $j \in \{1, 2, ..., N\}$, $j \neq i$, is to go to retailer $i$ or the depot.*

*Proof.* By assumption, for all $n \in \{1, 2, ..., N\}$, for each pair of $x'_n, x''_n \in X_n$ such that $x'_n \leq x''_n$ and for each pair of $k', k'' \in K\backslash\{0\}$ such that $k' \prec_n k''$, the following condition holds:

$$u_t(((x'_n, x^c_n), x_v, l), k') \leq (<) u_t(((x'_n, x^c_n), x_v, l), k'')$$
$$\Rightarrow u_t(((x''_n, x^c_n), x_v, l), k') \leq (<) u_t(((x''_n, x^c_n), x_v, l), k'')$$

Equivalently, we may write the above condition as

$$u_t(((x''_n, x^c_n), x_v, l), k') \geq (>) u_t(((x''_n, x^c_n), x_v, l), k'')$$
$$\Rightarrow u_t(((x'_n, x^c_n), x_v, l), k') \geq (>) u_t(((x'_n, x^c_n), x_v, l), k'').$$

Let $n = j$, for all $j \in \{1, 2, ..., N\}$, $j \neq i$, in the first condition above, and let $n = i$ in the second condition. Let us recall that $k' \prec_n k''$ if and only if $k' = n$ and $k'' \in K''_n = K \backslash \{0, n\}$. This implies that, for $i, j \in \{1, 2, ..., N\}$, $\bigcap_{j \neq i} K''_j = \{i\}$. Precisely, the first condition states that as inventory levels of retailers $j \in \{1, 2, ..., N\}$, $j \neq i$ increase, retailer $i$ is still preferred to other retailers as the vehicle's destination. From the second condition, in which $n = i$, it follows that retailer $i$ continues to be the preferred destination as inventory level of retailer $i$ decreases. The depot is not part of these arguments. So it could still be an optimal destination. $\square$

Proposition 2.4.4 specifies sufficient conditions for the optimal value function to be non-decreasing in $x_i$. This result is equivalent to Proposition 4.7.3 in Puterman's book.

**Proposition 2.4.4.** *Let $i \in \{1, 2, ..., N\}$, assume that the following conditions are satisfied:*

1. *$r_t(((x_i, x_i^c), x_v, l), k)$ is non-decreasing in $x_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$, for all $k \in K$, and for $t = 1, 2, ..., T$,*

2. *$\bar{r}_{T+1}((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$, and*

3. *$\sum_{m \leq y_i \leq q_i} p_t((y_i, y_i^c) | ((x_i, x_i^c), x_v, l), a, k)$ is non-decreasing in $x_i$, for all $m \in X_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$, for all $k \in K$, for all $y_i^c$, and for $t = 1, 2, ..., T$.*

*Then $u_t^*((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$, for $t = 1, 2, ..., T$.*

*Proof.* We use induction to prove the proposition. At the end of the horizon,

$$u_{T+1}^*((x_i, x_i^c), x_v, l) = \bar{r}_{T+1}((x_i, x_i^c), x_v, l).$$

By (2), $u_{T+1}^*((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$. Assume that $u_n^*((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$, for $n = T, T-1, ..., t+1$. We shall show that $u_t^*((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$. This will complete the proof of the proposition.

Let us recall the optimality equations 2.3.3 and 2.3.4. If we can show that $u_t(((x_i, x_i^c), x_v, l), k)$ is non-decreasing in $x_i$, for all $k \in K$, then $u_t^*((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$. Let

$a^*(k) \in A_t((x_i, x_i^c), x_v, l)$ be such that the following equation holds:

$$u_t(((x_i, x_i^c), x_v, l), k) = r_t(((x_i, x_i^c), x_v, l), a^*(k), k) \\ + \sum_{y_i^c} \sum_{y_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a^*(k), k) u_{t+1}^*((y_i, y_i^c), x_v - a^*(k), k)$$

Condition (3) states that $\sum_{m \le y_i \le q_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k)$ is non-decreasing in $x_i$, for all $m \in X_i$ and $a \in A_t((x_i, x_i^c), x_v, l)$. From the induction hypothesis, $u_{t+1}^*((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$, for all $k \in K$. Lemma 2.3.10 then implies that

$$\sum_{y_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a^*(k), k) u_{t+1}^*((y_i, y_i^c), x_v - a^*(k), k)$$

is non-decreasing in $x_i$. Since this result holds for all $y_i^c$, it follows that

$$\sum_{y_i^c} \sum_{y_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a^*(k), k) u_{t+1}^*((y_i, y_i^c), x_v - a^*(k), k)$$

is non-decreasing in $x_i$. Because the sum of two non-decreasing functions is non-decreasing, this result and condition (1) above imply that $u_t(((x_i, x_i^c), x_v, l), k)$ is non-decreasing in $x_i$, for all $k \in K$. This completes the proof of the proposition. □

We now show why the vehicle should be replenished to its full capacity. This result is important in establishing the single-crossing property (w.r.t. retailer $i$ in $(k, x_i)$ on $K \times X_i$) of $u_t(((x_i, x_i^c), x_v, l), k)$, for $l = 0$. We remark that this result holds without Assumption 2.4.1.

**Proposition 2.4.5.** *At the depot, it is always optimal to refill the vehicle to its full capacity.*

*Proof.* We note that, for $i \in \{1, 2, ..., N\}$,

1. $r_t(((x_i, x_i^c), x_v, l), a, k)$,
2. $\bar{r}_{T+1}((x_i, x_i^c), x_v, l)$, and
3. $\sum_{m \le y_i \le q_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k)$, for all $y_i^c$,

are independent in $x_v$. These results can be verified by inspecting the definitions of $r_t(((x_i, x_i^c), x_v, l), a, k)$, $\bar{r}_{T+1}((x_i, x_i^c), x_v, l)$, and $p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k)$ in Section 2.2. Furthermore, for $l > 0$, the set $A_t((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_v$. In particular, for $x_v'' \geq x_v'$,

$$A_t((x_i, x_i^c), x_v', l) \subseteq A_t((x_i, x_i^c), x_v'', l).$$

At the depot, the quantities in (1), (2), and (3) are independent of $a$. For these reasons, it is relatively straight-forward to show, by induction, that $u_t^*((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_v$, for $t = 1, 2, ..., T$.

Consider the first optimality equation for $l = 0$. At the depot, the quantities in (1), (2), and (3) are independent of $a$. Furthermore, $u_{t+d_{lk}}^*(y, x_v - a, k)$ is non-decreasing in $x_v - a$. The assertion of the proposition follows. □

The next theorem specifies the sufficient conditions that imply the desired structural result in the vehicle routing problem.

**Theorem 2.4.6.** *Let $i \in \{1, 2, ..., N\}$ and $l = 0$, assume that the conditions of Proposition 2.4.4 hold. Furthermore, assume that $r_t(((x_i, x_i^c), x_v, l), a, k)$ has the single-crossing property w.r.t. retailer $i$ in $(k, x_i)$ on $K \times X_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$. If an optimal (vehicle routing) action for the state $s_t'' = ((x_i'', x_i^c), x_v, l)$ is to go retailer $i$, then an optimal action for state $s_t' = ((x_i', x_i^c), x_v, l)$, in which $x_i' \leq x_i''$, is to go to retailer $i$ or the depot.*

*Proof.* Consider the optimality equation 2.3.4. If we can show that $w_t(((x_i, x_i^c), x_v, l), a, k)$ has the single-crossing property w.r.t. retailer $i$ in $(k, x_i)$ on $K \times X_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$, then $u_t(((x_i, x_i^c), x_v, l), k)$ has the same property. This is because, by Proposition 2.4.5, the only optimal inventory action at the depot is $-(q_v - x_v)$, which is independent of $k$ and $x_i$. The assertion of the theorem then follows from Proposition 2.4.2.

When its conditions are satisfied, Proposition 2.4.4 implies that $u_t^*((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$, for $t = 1, 2, ..., T$. Furthermore, by Assumption 2.4.1,

$$p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k)$$

is independent in $k$. It can be shown that these two results and Lemma 2.3.10 imply that

$$\sum_{y_i^c} \sum_{y_i} p_t((y_i, y_i^c)|(x_i, x_i^c), x_v, l), a, k) u_{t+1}^*((y_i, y_i^c), x_v - a, k)$$

is super-additive in $(k, x_i)$ on $K \times X_i$.

By assumption, $r_t((x_i, x_i^c), x_v, l), a, k)$ has the single-crossing property w.r.t. retailer $i$ in $(k, x_i)$ on $K \times X_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$. By Lemma 2.3.4 and Lemma 2.3.5(2), we have that $w_t(((x_i, x_i^c), x_v, l), a, k)$ has the single-crossing property w.r.t. retailer $i$ in $(k, x_i)$ on $K \times X_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$. This completes the proof of the theorem. $\square$

When the conditions of Theorem 2.4.6 are satisfied for all $i \in \{1, 2, ..., N\}$, the conditions of Proposition 2.4.3 hold and its result follows. Theorem 2.4.7, stated here without proof, formalizes this statement.

**Theorem 2.4.7.** *For $l = 0$, assume that the conditions of Proposition 2.4.4 hold for all $i \in \{1, 2, ..., N\}$. Furthermore, assume that, for all $i \in \{1, 2, ..., N\}$, $r_t(((x_i, x_i^c), x_v, l), a, k)$ has the single-crossing property w.r.t. retailer $i$ in $(k, x_i)$ on $K \times X_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$. If an optimal (vehicle routing) action for the state $\tilde{s}_t = ((\tilde{x}_i, \tilde{x}_i^c), x_v, l)$ is to go retailer $i$, then an optimal action for state $s_t = ((x_i, x_i^c), x_v, l)$, in which $x_i \leq \tilde{x}_i$ and $x_j \geq \tilde{x}_j$, for all $j \in \{1, 2, ..., N\}$, $j \neq i$, is to go to retailer $i$ or the depot.*

### 2.4.2 Sufficient Conditions on the Problem Parameters

We have presented sufficient conditions for the structural results for vehicle routing. It is relatively straightforward to show that, with few additional assumptions, the parameters of the SVMI problem satisfy these conditions. The following corollary shows that the reward structure of the SVMI problem has the super-additivity property.

**Corollary 2.4.8.** *For all $i \in \{1, 2, ..., N\}$, $r_t(((x_i, x_i^c), x_v, l), a, k)$ is super-additive in $(k, x_i)$ on $K \times X_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$, for all $k \in K$, and for $t = 1, 2, ..., T$.*

*Proof.* For $l = 0$,

$$r_t(((x_i, x_i^c), x_v, l), a, k) = \sum_{1 \le j \le N} r_t^j(x_j, l, k) - c_{lk},$$

and for $l > 0$,

$$r_t(((x_i, x_i^c), x_v, l), a, k) = \tilde{r}_t^l(x_l, l, a, k) + \sum_{j \in K \setminus \{0,l\}} r_t^j(x_j, l, k) - c_{lk}.$$

By Assumption 2.4.1, $r_t^j(x_j, l, k)$ and $\tilde{r}_t^l(x_l, l, a, k)$ are independent in $k$. Moreover, $c_{lk}$ is independent of $x_i$. Therefore, $r_t^j(x_j, l, k)$, $\tilde{r}_t^l(x_l, l, a, k)$, and $-c_{lk}$ have the super-additive property and, as a result, so does their sum. The assertion of the proposition follows. □

Note that, by Lemma 2.3.4, the above result implies that $r_t(((x_i, x_i^c), x_v, l), a, k)$ has the single-crossing property in $(k, x_i)$ on $K \times X_i$. The next three results, i.e., Proposition 2.4.9, Corollary 2.4.10, and Corollary 2.4.11, hold without Assumption 2.4.1. Let $\bar{F}_{t,i}^{l,k}(x_i) = 1 - F_{t,i}^{l,k}(x_i)$, for $x_i \in X_i$, where $F_{t,i}^{l,k}$ denotes the cumulative probability distribution function of $D_{t,i}^{l,k}$. The following proposition specifies a sufficient condition for $r_t(((x_i, x_i^c), x_v, l), a, k)$ to be non-decreasing in $x_i$.

**Proposition 2.4.9.** *Let $i \in \{1, 2, ..., N\}$, assume that $(b_i^1 + b_i^2)\bar{F}_{t,i}^{l,k}(q_i) \ge h_i d_{lk}$. Then, $r_t(((x_i, x_i^c), x_v, l), a, k)$ is non-decreasing in $x_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$.*

*Proof.* We shall show that if $(b_i^1 + b_i^2)\bar{F}_{t,k}^i(q_i) \ge h_i d_{lk}$, then $r_t^i(x_i, l, k)$ and $\tilde{r}_t^i(x_i, l, a, k)$ are non-decreasing in $x_i$. From this and the definition of $r_t(((x_i, x_i^c), x_v, l), a, k)$, the assertion of the proposition then follows. Let $\rho(d)$ be the probability that the number of orders that arrive at retailer $i$, from time $t$ to time $t + d_{lk}$, is $d$. Then,

$$\begin{aligned}
r_t^i(x_i, l, k) &= -h_i d_{lk} x_i + b_i^1 E[min\{D_{t,i}^{l,k}, x_i\}] - b_i^2 E[max\{0, D_{t,i}^{l,k} - x_i\}] \\
&= -h_i d_{lk} x_i + b_i^1 \sum_{0 \le d < \infty} \rho(d)(min\{d, x_i\}) - b_i^2 \sum_{0 \le d < \infty} \rho(d)(max\{0, d - x_i\}) \\
&= -h_i d_{lk} x_i + b_i^1 \sum_{0 \le d < \infty} \rho(d)[min\{d, x_i\} - (b_i^2/b_i^1) max\{0, d - x_i\}] \\
&= -h_i d_{lk} x_i + b_i^1 \sum_{0 \le d < \infty} \rho(d)[min\{d, x_i\} - max\{0, (b_i^2/b_i^1)(d - x_i)\}],
\end{aligned}$$

and

$$r_t^i(x_i+1,l,k) - r_t^i(x_i,l,k) = b_i^1 \sum_{x_i \le d < \infty} \rho(d)[1+(b_i^2/b_i^1)] - h_i d_{lk}$$
$$= (b_i^1 + b_i^2) \sum_{x_i \le d < \infty} \rho(d) - h_i d_{lk}$$
$$= (b_i^1 + b_i^2) \bar{F}_{t,i}^{l,k}(x_i) - h_i d_{lk}.$$

Note that if $(b_i^1 + b_i^2)\bar{F}_{t,i}^{l,k}(q_i) \ge h_i d_{lk}$, then $(b_i^1 + b_i^2)\bar{F}_{t,i}^{l,k}(x_i) \ge h_i d_{lk}$, for all $x_i \in X_i$. Therefore, $r_t^i(x_i+1,l,k) - r_t^i(x_i,l,k) \ge 0$, for all $x_i \in X_i$. That is, $(b_i^1+b_i^2)\bar{F}_{t,i}^{l,k}(q_i) \ge h_i d_{lk}$ implies that $r_t^i(x_i,l,k)$ is non-decreasing in $x_i$. Similarly, it can be shown that

$$\tilde{r}_t^i(x_i+1,l,a,k) - \tilde{r}_t^i(x_i,l,a,k) = (b_i^1 + b_i^2)\bar{F}_{t,i}^{l,k}(x_i) - h_i d_{lk}.$$

So the same condition applies here. This completes the proof of the proposition. □

To interpret the condition of Proposition 2.4.9, we rewrite it as

$$\bar{F}_{t,i}^{l,k}(q_i) \ge h_i d_{lk}/(b_i^1 + b_i^2).$$

We observe that the left hand side of the above inequality is the probability that the demand at retailer $i$ from time $t$ to time $t + d_{lk}$ exceeds the capacity of the retailer. On the right hand side, the numerator and denominator represents the potential cost and reward, respectively, of having one more unit of inventory at retailer $i$. In general, the quantity on the right hand side is very small.

Corollary 2.4.10 shows that the transition structure of the SVMI problem has the desired property.

**Corollary 2.4.10.** *For all* $i \in \{1, 2, ..., N\}$, $\sum_{m \le y_i \le q_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k)$ *is non-decreasing in* $x_i$*, for all* $a \in A_t((x_i, x_i^c), x_v, l)$, $k \in K$, $y_i^c$, $m$*, and for* $t = 1, 2, ..., T$.

*Proof.* For $i \in K \backslash \{0, l\}$, it can be shown that

$$\sum_{m \le y_i \le q_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k) = F_{t,i}^{l,k}(x_i - m).$$

Similarly, for the case $i = l$,

$$\sum_{m \le y_i \le q_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k) = F_{t,i}^{l,k}(x_i + a - m).$$

Clearly, in both cases, the result follows. $\square$

The following corollary proves the non-decreasing property of the terminal reward.

**Corollary 2.4.11.** *$\bar{r}_{T+1}((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$.*

*Proof.* By definition, $\bar{r}_{T+1}((x_i, x_i^c), x_v, l) = \sum_{1 \le j \le N}(e_j - h_j\tau)x_j - c_{l0}$. The result follows from the assumption $e_j \ge h_j\tau$, for all $j \in \{1, 2, ..., N\}$. $\square$

Given that its condition is satisfied, Proposition 2.4.9, along with Corollary 2.4.10, and Corollary 2.4.11 imply the conditions of Proposition 2.4.4. This result and Corollary 2.4.8 then imply the conditions of Theorem 2.4.6. We may conclude that, with few additional assumptions, the parameters of the SVMI problem satisfy the sufficient conditions for the structural results for vehicle routing.

### 2.4.3 Summary of the Structural Results

Theorem 2.4.12 summarizes the structural results for vehicle routing. This result is based on Assumption 2.4.1, and the results in Subsection 2.4.1 and Subsection 2.4.2.

**Theorem 2.4.12.** *For all $l, k \in K$, let $d_{lk} = 1$. Furthermore, for all $n \in \{1, 2, ..., N\}$, assume that*

$$(b_n^1 + b_n^2)\bar{F}_{t,n}^{l,k}(q_n) \ge h_n d_{lk},$$

*for all $l, k \in K$, and for $t = 1, 2, ..., T$. At the depot, if, for an $i \in \{1, 2, ..., N\}$, an optimal (vehicle routing) action for the state $\tilde{s}_t = ((\tilde{x}_i, \tilde{x}_i^c), x_v, l)$ is to go to retailer $i$, then an optimal action for state $s_t = ((x_i, x_i^c), x_v, l)$, in which $x_i \le \tilde{x}_i$ and $x_j \ge \tilde{x}_j$, for all $j \in \{1, 2, ..., N\}$, $j \ne i$, is to go to retailer $i$ or the depot.*

In summary, the main structural result for vehicle routing states (informally) that, when the vehicle is at the depot, a particular retailer continues to be preferred to other retailers as the vehicle's destination if the inventory level of that retailer decreases and/or inventory levels of the other retailers increase.

## 2.5 Structural Results for Inventory Control

We now establish monotone relations between the optimal inventory action and inventory levels of the retailers. In particular, we show that the optimal number of units to deposit at the current retailer is non-decreasing in inventory levels of the other retailers. Additionally, with a stronger assumption on the demand distribution at the current retailer, the optimal inventory action is shown to be non-increasing in inventory level of the current retailer. Note that Assumption 2.4.1 is not applied here, and hence, for $l, k \in K$, we allow the travel time from location $l$ to location $k$, $d_{lk}$, to be dependent on $l$ and $k$.

We first present sufficient conditions for the structural results for inventory control. In the subsection that follows, we show how the parameters of the SVMI problem satisfy these conditions. Finally, in Subsection 2.5.3, we summarize the structural results for inventory control.

### 2.5.1 Sufficient Conditions for the Structural Results

The following proposition establishes the first structural result for inventory control.

**Proposition 2.5.1.** *Let $l \in \{1, 2, ..., N\}$, assume that $w_t(((x_i, x_i^c), x_v, l), a, k)$ is super-additive in $(x_i, a)$ on $X_i \times A_t((x_i, x_i^c), x_v, l)$, for an $i \in K \backslash \{0, l\}$. Then there exists $a^*(k)$ for the state $s_t = ((x_i, x_i^c), x_v, l)$ which is non-decreasing in $x_i$.*

*Proof.* By assumption, for each pair of $x_i', x_i'' \in X_i$ such that $x_i' \leq x_i''$ and for each pair of $\tilde{a}, a \in \{A_t((x_i', x_i^c), x_v, l) \cap A_t((x_i'', x_i^c), x_v, l)\}$ such that $\tilde{a} \geq a$, the following inequality holds:

$$w_t(((x_i'', x_i^c), x_v, l), \tilde{a}, k) - w_t(((x_i'', x_i^c), x_v, l), a, k)$$
$$\geq w_t(((x_i', x_i^c), x_v, l), \tilde{a}, k) - w_t(((x_i', x_i^c), x_v, l), a, k)$$

By Lemma 2.3.6, for all $x_i'' \geq x_i'$,

$$max\{argmax_{a \in A_t((x_i'', x_i^c), x_v, l)} w_t(((x_i'', x_i^c), x_v, l), a, k)\}$$
$$\geq max\{argmax_{a \in A_t((x_i', x_i^c), x_v, l)} w_t(((x_i', x_i^c), x_v, l), a, k)\}.$$

Note that $A_t((x_i', x_i^c), x_v, l) = A_t((x_i'', x_i^c), x_v, l)$. The assertion of the proposition follows.

□

Corollary 2.5.2, stated here without proof, follows from Proposition 2.5.1 when its condition holds for all $i \in K \backslash \{0, l\}$.

**Corollary 2.5.2.** *Let $l \in \{1, 2, ..., N\}$, assume that $w_t(((x_i, x_i^c), x_v, l), a, k)$ is super-additive in $(x_i, a)$, for all $i \in K \backslash \{0, l\}$. Then there exists $a^*(k)$ for the state $s_t = ((x_i, x_i^c), x_v, l)$ which is non-decreasing in $x_i^c$.*

Proposition 2.5.3 specifies the sufficient conditions for the optimal value function, $u_t^*((x_i, x_i^c), x_v, l)$, to be non-decreasing in $x_i$. This result is analogous to Proposition 4.7.3 in Puterman (1994).

**Proposition 2.5.3.** *Let $i \in \{1, 2, ..., N\}$, assume that the following conditions are satisfied:*

1. *$r_t(((x_i, x_i^c), x_v, l), k)$ is non-decreasing in $x_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$, for all $k \in K$, and for $t = 1, 2, ..., T$,*

2. *$\bar{r}_{T+1}((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$, and*

3. *$\sum_{m \leq y_i \leq q_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k)$ is non-decreasing in $x_i$, for all $m \in X_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$, for all $k \in K$, for all $y_i^c$, and for $t = 1, 2, ..., T$.*

*Then $u_t^*((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$, for $t = 1, 2, ..., T$.*

*Proof.* The proof is analogous to that of Proposition 2.4.4. In particular, Assumption 2.4.1 is not required in the proof of Proposition 2.4.4. $\square$

The above result is needed in the following theorem, whose result follows from Proposition 2.5.1. Theorem 2.5.4 is similar to Theorem 4.7.4 in Puterman's book.

**Theorem 2.5.4.** *Let $l \in \{1, 2, ..., N\}$, assume that the conditions of Proposition 2.5.3 hold for an $i \in K \backslash \{0, l\}$. Furthermore, assume that, for all $a \in A_t((x_i, x_i^c), x_v, l)$, for all $k \in K$, and for $t = 1, 2, ..., T$,*

1. *$r_t(((x_i, x_i^c), x_v, l), a, k)$ is super-additive in $(x_i, a)$,*

2. *$\sum_{y_i^c} \sum_{m \leq y_i \leq q_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k)$, for all $m$, is super-additive in $(x_i, a)$.*

*Then there exists $a^*(k)$ for the state $s_t = ((x_i, x_i^c), x_v, l)$ which is non-decreasing in $x_i$.*

*Proof.* By definition,

$$\begin{aligned} w_t(((x_i, x_i^c), x_v, l), a, k) &= r_t(((x_i, x_i^c), x_v, l), a, k) \\ &\quad + \sum_{y_i^c} \sum_{y_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k) u^*_{t+d_{lk}}((y_i, y_i^c), x_v - a, k). \end{aligned}$$

Given that its conditions hold for $i$, Proposition 2.5.3 implies that $u_t^*((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$, for $t = 1, 2, ..., T$. This result along with condition (2) and Lemma 2.3.10 then imply that

$$\sum_{y_i^c} \sum_{y_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k) u^*_{t+d_{lk}}((y_i, y_i^c), x_v - a, k)$$

is super-additive in $(x_i, a)$. Because the sum of super-additive functions is super-additive, condition (1) then imply that $w_t(((x_i, x_i^c), x_v, l), a, k)$ is super-additive in $(x_i, a)$. The result of the theorem follows from Proposition 2.5.1. $\square$

When the conditions of Theorem 2.5.4 hold for all $i \in K \backslash \{0, l\}$, the result of the following theorem, stated here without proof, follows from Corollary 2.5.2.

**Theorem 2.5.5.** *Let $l \in \{1, 2, ..., N\}$, assume that the conditions of Proposition 2.5.3 hold for all $i \in K \backslash \{0, l\}$. Furthermore, assume that, for all $a \in A_t((x_i, x_i^c), x_v, l)$, for all $k \in K$, and for $t = 1, 2, ..., T$,*

1. *$r_t(((x_i, x_i^c), x_v, l), a, k)$ is super-additive in $(x_i, a)$,*

2. *$\sum_{y_i^c} \sum_{m \leq y_i \leq q_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k)$, for all $m$, is super-additive in $(x_i, a)$.*

*Then there exits $a^*(k)$ for the state $s_t = ((x_i, x_i^c), x_v, l)$ which is non-decreasing in $x_i^c$.*

In the next proposition, we show how the optimal number of units to deposit at the current retailer can be non-increasing in inventory level of the retailer. This is the second structural result for inventory control.

**Proposition 2.5.6.** *Let $l \in \{1, 2, ..., N\}$, assume that $w_t(((x_l, x_l^c), x_v, l), a, k)$ is sub-additive in $(x_l, a)$. Then there exists $a^*(k)$ for the state $s_t = ((x_l, x_l^c), x_v, l)$ which is non-increasing in $x_l$.*

*Proof.* By assumption, $w_t(((x_l, x_l^c), x_v, l), a, k)$ is sub-additive in $(x_l, a)$. That is, for each pair of $x_l', x_l'' \in X_i$ such that $x_l' \leq x_l''$ and for each pair of $\tilde{a}, a \in \{A_t((x_l', x_l^c), x_v, l) \cap A_t((x_l'', x_l^c), x_v, l)\}$ such that $\tilde{a} \geq a$, the following inequality holds:

$$\begin{aligned} w_t(((x_l', x_l^c), x_v, l), \tilde{a}, k) - w_t(((x_l', x_l^c), x_v, l), a, k) \\ \geq w_t(((x_l'', x_l^c), x_v, l), \tilde{a}, k) - w_t(((x_l'', x_l^c), x_v, l), a, k) \end{aligned}$$

By Lemma 2.3.7,

$$\begin{aligned} min\{argmax_{a \in A_t((x_l', x_l^c), x_v, l)} w_t(((x_l', x_l^c), x_v, l), a, k)\} \\ \geq min\{argmax_{a \in A_t((x_l'', x_l^c), x_v, l)} w_t(((x_l'', x_l^c), x_v, l), a, k)\}. \end{aligned}$$

Note that $A_t((x_l'', x_l^c), x_v, l) \subseteq A_t(x_l', x_l^c), x_v, l)$. The result of the proposition follows. $\square$

Theorem 2.5.7 specifies sufficient conditions that imply the conditions and, therefore, result of Proposition 2.5.6.

**Theorem 2.5.7.** *Let $l \in \{1, 2, ..., N\}$, assume that the conditions of Proposition 2.5.3 hold for $i = l$. Furthermore, assume that, for all $k \in K$,*

1. $r_t(((x_l, x_l^c), x_v, l), a, k)$*, and*

2. $\sum_{y_l^c} \sum_{m \leq y_l \leq q_l} p_t((y_l, y_l^c)|((x_l, x_l^c), x_v, l), a, k)$*, for all* $m$*,*

*are sub-additive in $(x_l, a)$. Then there exists $a^*(k)$ for the state $s_t = ((x_l, x_l^c), x_v, l)$ which is non-increasing in $x_l$.*

*Proof.* By assumption, $r_t(((x_l, x_l^c), x_v, l), a, k)$ is sub-additive in $(x_l, a)$. Proposition 2.5.3 implies that $u^*_{t+d_{lk}}((y_l, y_l^c), x_v - a, k)$ is non-decreasing in $y_l$, for all $y_l^c$, for $t = 1, 2, ..., T$. By Lemma 2.3.10 and assumption (2) above, we have that

$$\sum_{y_i^c} \sum_{y_l} p_t((y_l, y_l^c)|((x_l, x_l^c), x_v, l), a, k) u^*_{t+d_{lk}}((y_l, y_l^c), x_v - a, k)$$

is sub-additive in $(x_l, a)$. Since the sum of sub-additive functions is sub-additive, it follows that $w_t(((x_l, x_l^c), x_v, l), a, k)$ is sub-additive in $(x_l, a)$. The result of the theorem follows from Proposition 2.5.6. □

Combining Theorem 2.5.5 and Theorem 2.5.7 gives us the following result.

**Theorem 2.5.8.** *Let $l \in \{1, 2, ..., N\}$, assume that the conditions of Proposition 2.5.3 hold for all $i \in K \backslash \{0, l\}$. Furthermore, assume that, for all $a \in A_t((x_i, x_i^c), x_v, l)$, for all $k \in K$, and for $t = 1, 2, ..., T+1$,*

1. $r_t(((x_i, x_i^c), x_v, l), a, k)$ *is super-additive in* $(x_i, a)$*,*

2. $\sum_{y_i^c} \sum_{m \leq y_i \leq q_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k)$*, for all* $m$*, is super-additive in* $(x_i, a)$*,*

3. $r_t(((x_l, x_l^c), x_v, l), a, k)$ *is sub-additive in* $(x_l, a)$*, and*

4. $\sum_{y_l^c} \sum_{m \leq y_l \leq q_l} p_t((y_l, y_l^c)|((x_l, x_l^c), x_v, l), a, k)$*, for all* $m$*, is sub-additive in* $(x_l, a)$*.*

*Then there exists $a^*(k)$ for the state $s_t = ((x_l, x_l^c), x_v, l)$ which is non-decreasing in $x_i^c$ and non-increasing in $x_l$.*

### 2.5.2 Sufficient Conditions on the Problem Parameters

One of the sufficient conditions for the first structural result for inventory control is that $r_t(((x_i, x_i^c), x_v, l), a, k)$ is super-additive in $(x_i, a)$, in which $i \in K \backslash \{0, l\}$. Corollary 2.5.9 shows that this is the case.

**Corollary 2.5.9.** *For all* $l \in \{1, 2, ..., N\}$*,* $r_t(((x_i, x_i^c), x_v, l), a, k)$ *is super-additive in* $(x_i, a)$*, for all* $i \in K \backslash \{0, l\}$*, for all* $k \in K$*, and for* $t = 1, 2, ..., T$*.*

*Proof.* For $l \in \{1, 2, ..., N\}$,

$$r_t(((x_i, x_i^c), x_v, l), a, k) = \tilde{r}_t^l(x_l, l, a, k) + r_t^i(x_i, l, k) + \sum_{j \in K \backslash \{0, l, i\}} r_t^j(x_j, l, k) - c_{lk}.$$

Since $r_t^i(x_i, l, k)$ is independent of $a$, it is super-additive in $(x_i, a)$. Similarly, $\tilde{r}_t^l(x_l, l, a, k)$ is independent of $x_i$ and, thus, super-additive in $(x_i, a)$. The last two terms on the right hand side of the above equation need not be considered because both of them are independent of $x_i$ and $a$. Since the sum of super-additive functions is super-additive, the result follows. □

The next result shows that the transition structure has the super-additive property.

**Corollary 2.5.10.** *For all* $l \in \{1, 2, ..., N\}$*,*

$$\sum_{y_i^c} \sum_{m \leq y_i \leq q_i} p_t((y_i, y_i^c) | ((x_i, x_i^c), x_v, l), a, k)$$

*is super-additive in* $(x_i, a)$*, for all* $i \in K \backslash \{0, l\}$*, for all* $m$*, for all* $k \in K$*, and for* $t = 1, 2, ..., T$*.*

*Proof.* It can be shown that

$$\begin{aligned} \sum_{y_i^c} \sum_{m \leq y_i \leq q_i} p_t((y_i, y_i^c) | ((x_i, x_i^c), x_v, l), a, k) &= \sum_{m \leq y_i \leq q_i} p_t^i(y_i | x_i, l, k) \\ &= F_{t,i}^{l,k}(x_i - m). \end{aligned}$$

$F_{t,i}^{l,k}(x_i - m)$ is independent of $a$ and, thus, super-additive in $(x_i, a)$. The result follows. □

Assumption 2.4.1 is not required in the proofs of Proposition 2.4.9, Corollary 2.4.10, and Corollary 2.4.11. These results are as follows:

1. let $i \in \{1, 2, ..., N\}$, assume that $(b_i^1+b_i^2)\bar{F}_{t,i}^{l,k}(q_i) \geq h_i d_{lk}$. Then, $r_t(((x_i, x_i^c), x_v, l), a, k)$ is non-decreasing in $x_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$;

2. for all $i \in \{1, 2, ..., N\}$, $\sum_{m \leq y_i \leq q_i} p_t((y_i, y_i^c)|((x_i, x_i^c), x_v, l), a, k)$ is non-decreasing in $x_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$, $k \in K$, $y_i^c$, $m$, and for $t = 1, 2, ..., T$;

3. $\bar{r}_{T+1}((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$, for all $i \in \{1, 2, ..., N\}$.

For the second structural result for vehicle routing, the next corollary establishes the sub-additive property of the reward function.

**Corollary 2.5.11.** *For all $k \in K$, $r_t(((x_l, x_l^c), x_v, l), a, k)$ is sub-additive in $(x_l, a)$.*

*Proof.* For $l \in \{1, 2, ..., N\}$,

$$r_t(((x_l, x_l^c), x_v, l), a, k) = \tilde{r}_t^l(x_l, l, a, k) + \sum_{j \in K \setminus \{0, l\}} r_t^j(x_j, l, k) - c_{lk}.$$

It can be shown that

$$\tilde{r}_t^l(x_l + 1, l, a, k) - \tilde{r}_t^l(x_l, l, a, k) = -h_l d_{lk} + (b_l^1 + b_l^2)\bar{F}_{t,l}^{l,k}(x_l).$$

Because this quantity is independent of $a$, we have that $r_t(((x_l, x_l^c), x_v, l), a, k)$ is sub-additive in $(x_l, a)$. □

In Proposition 2.5.12, we show how the sub-additive property of the transition probability at the current retailer follows from the non-increasing property of the demand distribution. The latter property was first introduced in Definition 2.3.11.

**Proposition 2.5.12.** *Assume that, for all $k \in K$, the demand at retailer $l$ between time $t$ and $t + d_{lk}$, that is $D_{t,l}^{l,k}$, has non-increasing probability mass function. Then*

$$\sum_{y_l^c} \sum_{m \leq y_l \leq q_l} p_t((y_l, y_l^c)|((x_l, x_l^c), x_v, l), a, k)$$

*is sub-additive in $(x_l, a)$, for all $m$.*

*Proof.* It can be shown that

$$\sum_{m \leq y_l \leq q_l} p_t((y_l, y_l^c)|((x_l, x_l^c), x_v, l), a, k) = F_{t,l}^{l,k}(x_l + a - m),$$

where $F_{t,l}^{l,k}$ is the cumulative probability distribution function of $D_{t,l}^{l,k}$. By assumption, $D_{t,l}^{l,k}$ has non-increasing probability mass function. As a result, it can be shown that

$$F_{t,l}^{l,k}(x_l + a + 1 - m) - F_{t,l}^{l,k}(x_l + a - m)$$

is non-increasing in $x_l$. Thus,

$$\sum_{y_l} p_t((y_l, y_l^c)|((x_l, x_l^c), x_v, l), a, k)$$

is sub-additive in $(x_l, a)$, for all $y_l^c$. The assertion of the proposition follows. □

### 2.5.3 Summary of the Structural Results

Based on the results of Subsections 2.5.1 and 2.5.2, we summarize the two main structural results for inventory control in Theorem 2.5.13 and Theorem 2.5.14. In Theorem 2.5.13, an optimal inventory action at a particular retailer is shown to be non-decreasing in inventory levels of the other retailers.

**Theorem 2.5.13.** *For all $n \in \{1, 2, ..., N\}$, assume that*

$$(b_n^1 + b_n^2)\bar{F}_{t,n}^{l,k}(q_n) \geq h_n d_{lk},$$

*for all $l, k \in K$, and for $t = 1, 2, ..., T$. Then there exits $a^*(k)$ for the state $s_t = ((x_i, x_i^c), x_v, l)$ which is non-decreasing in $x_i^c$.*

The next theorem summarizes the sufficient conditions for an optimal inventory action to be non-increasing in inventory level of the current retailer.

**Theorem 2.5.14.** *For all $n \in \{1, 2, ..., N\}$, assume that*

$$(b_n^1 + b_n^2)\bar{F}_{t,n}^{l,k}(q_n) \geq h_n d_{lk},$$

*for all $l, k \in K$, and for $t = 1, 2, ..., T$. Furthermore, assume that, for all $k \in K$, the demand at retailer $l$ between time $t$ and $t + d_{lk}$, that is $D_{t,l}^{l,k}$, has non-increasing probability mass function. Then there exists $a^*(k)$ for the state $s_t = ((x_l, x_l^c), x_v, l)$ which is non-decreasing in $x_i^c$ and non-increasing in $x_l$.*

## 2.6 Applications of the Structural Results

We have presented the structural results for vehicle routing and inventory control in the SVMI problem. It is often the case that structural results lead to computational simplifications. We can now develop monotone backward induction algorithms to solve several instances of the SVMI problem. In the next section, we show how effective the algorithms are and how the effectiveness varies with the problem size and its specifications.

For larger problems, heuristic solution procedures based on the structural results that we have obtained can be developed. One such method assumes that the monotone relations in the optimal inventory actions are piecewise linear. We investigate this solution technique and present our findings in Chapter 4. The heuristic solution procedure helps reduce the computational requirement in solving the SVMI problem noticeably.

Other potential benefits of the structural results include the insight and intuition for management. This would help the decision makers improve the operating performance of their distribution systems. Practically, structured policies are relatively easy to implement. This, by itself, may have important implications in reducing operating costs for the vendor.

## 2.7 Numerical Results

In this section, we first present a set of optimal vehicle routing and inventory actions for a sample SVMI problem with two retailers. The parameters for this sample problem satisfy the sufficient conditions for both sets of structural results. In this case, each of the two retailers has ten-unit capacity. Demands for the product at the two retailers have the same distribution (a discrete version of an exponential distribution). We define the revenue and cost parameters such that the second retailer is much more profitable than the first one.

Table 1 presents optimal vehicle routing actions for different inventory levels of the retailers ($x_1$ and $x_2$), when the vehicle is at the depot and its capacity level is 20 units. The structural result for vehicle routing holds here. In particular, each retailer continues to be an optimal destination for the vehicle as its inventory level decreases and/or inventory level of the other retailer increases. Here we also observe the effect of each retailer's profitability on whether or not the retailer is an optimal destination. If the revenue and cost parameters for both retailers are the same, then, as we have verified, the results in Table 1 will be symmetric.

**Table 1:** The vehicle's optimal destinations from the depot for the two-retailer SVMI problem

| $x_1 \backslash x_2$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 9 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 8 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 7 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 |
| 6 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 |
| 5 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 |
| 4 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 |
| 3 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |

Optimal inventory actions are presented in Table 2. In this case, the vehicle, with 14 units of inventory, is currently at the first retailer and the decision maker is considering the second retailer as a destination. It is easy to verify that the first and second structural results for inventory control hold in Table 2. Specifically, an optimal inventory action is non-decreasing in inventory level of the second retailer and non-increasing in the inventory level of the first retailer.

**Table 2:** Optimal inventory actions for the two-retailer SVMI problem

| $x_1 \backslash x_2$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 8 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 7 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 6 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 4 | 4 | 5 | 5 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| 3 | 4 | 5 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| 2 | 4 | 5 | 6 | 6 | 7 | 8 | 8 | 8 | 8 | 8 | 8 |
| 1 | 5 | 5 | 6 | 7 | 7 | 8 | 9 | 9 | 9 | 9 | 9 |
| 0 | 5 | 6 | 6 | 7 | 8 | 8 | 9 | 9 | 10 | 10 | 10 |

Next we examine the direct computational advantage of the structural results in backward induction (BI) algorithms. The computational measures of interest include the number of vehicle routing actions evaluated, the number of inventory actions evaluated, and the run time (CPU time). In addition, we study how the computational advantage of structured solutions varies with certain parameters of the problem. These parameters include the number of retailers, the capacities of the retailers, and the capacity of the vehicle.

In Table 3, we show how each structural result improves the computational measures of interest. The numerical examples were done on a SUN Ultra 60 workstation. Let us recall that Theorem 2.4.7 establishes the structure in the optimal vehicle routing actions (SV). Meanwhile, Theorem 2.5.5 and Theorem 2.5.7, specifies the first and second structural results for inventory control (SIA and SIB), respectively. Theorem 2.5.8 combines these two results (SI). Based on the numerical results, we may conclude that each of these structures reduces run time by about a third. When they are applied simultaneously, these structures reduce run time by about one half.

Table 4 illustrates how the computational advantage of the structural result for vehicle routing varies with the number of retailers. We present numerical results for the SVMI problems with two, three, and four retailers. In this case, the run time reduction varies from

**Table 3:** Computational measures for the SVMI problem with three retailers

| | Inventory actions (x1000) | Vehicle routing actions (x1000) | Actions (x1000) | Run time (seconds) | Run time reduction |
|---|---|---|---|---|---|
| BI | 4,125 | 598 | 4,723 | 817 | - |
| BI-SV | 2,659 | 345 | 3,004 | 540 | 33.90% |
| BI-SIA | 2,789 | 598 | 3,387 | 551 | 32.56% |
| BI-SIB | 2,794 | 598 | 3,393 | 561 | 31.33% |
| BI-SI | 2,595 | 598 | 3,198 | 524 | 35.86% |
| BI-SV-SI | 2,096 | 336 | 2,432 | 433 | 47.00% |

28 percent to 33 percent and, finally, to 48 percent. We may conclude that the computational advantage increases with the number of retailers.

**Table 4:** Computational measures for the SVMI problems with two, three, and four retailers

| Number of retailers | Solution procedure | Inventory actions (x1000) | Vehicle routing actions (x1000) | Actions (x1000) | Run time (seconds) | Run time reduction |
|---|---|---|---|---|---|---|
| 2 | BI | 435 | 50 | 484 | 12.27 | - |
| 2 | BI-SV | 357 | 38 | 395 | 8.81 | 28.20% |
| 3 | BI | 4,125 | 598 | 4,723 | 817 | - |
| 3 | BI-SV | 2,659 | 345 | 3,004 | 540 | 33.90% |
| 4 | BI | 35,759 | 5,988 | 41,746 | 63,788 | - |
| 4 | BI-SV | 19,697 | 2,971 | 22,698 | 38,049 | 40.35% |

The capacities of the vehicle and the retailers directly affect the computational advantage of the structural results for inventory control. Table 5 presents supporting numerical examples. From the table, it is clear that the computational efficiency increases with these capacities. Specifically, as the capacities get larger, the run time reduction varies from 34 percent to 36 percent and, finally, to 38 percent.

**Table 5:** Computational measures for the SVMI problems with three retailers and different capacities

| Retailer capacity | Vehicle capacity | Solution procedure | Inventory actions (x1000) | Vehicle routing actions (x1000) | Actions (x1000) | Run time (seconds) | Run time reduction |
|---|---|---|---|---|---|---|---|
| 3 | 12 | BI | 510 | 110 | 620 | 209 | - |
| 3 | 12 | BI-SI | 329 | 110 | 439 | 137 | 34.45% |
| 5 | 20 | BI | 4,125 | 598 | 4,723 | 1,689 | - |
| 5 | 20 | BI-SI | 2,595 | 598 | 3,193 | 1,082 | 35.94% |
| 7 | 28 | BI | 17,910 | 1,960 | 19,870 | 7,436 | - |
| 7 | 28 | BI-SI | 11,106 | 1,960 | 13,066 | 4,585 | 38.34% |

## 2.8 Distribution Problems with Multiple Vehicles

Actual distribution problems normally involves multiple vehicles. It is reasonable to solve such problems in two stages. First, the vehicles are assigned to non-intersecting groups of retailers. In other words, each retailer is assigned to a vehicle and no retailer is assigned to more than one vehicle. This is an instance of the assignment problem in combinatorial optimization. After it is solved, we will have multiple SVMI problems and each can be solved by the approach that we have discussed in this chapter. Balinski (1986) and Goldfarb (1985) present polynomial-time dual network simplex algorithm for the assignment problem. This approach may not be optimal for the expected total reward criterion. However, it allows the decision maker to consider other factors, such as travel distances and drivers' hours, when he or she assigns the retailers to each vehicle.

## 2.9 Conclusions and Future Research

We have formulated and analyzed an MDP model of the SVMI problem. By assuming certain conditions on the demand distributions at the retailers, we established structural results for vehicle routing and inventory control. These results helped reduce the computational requirement in solving the problem noticeably. In the next chapter, we extend the structural results, plus the algorithms, to the multiperiod SVMI problem and the infinite horizon

SVMI problem.

An interesting extension of the SVMI problem is the case in which the vehicle is allowed to pick up the product at the retailers. This increases the replenishment flexibility. For geographical reasons, it might be less expensive to replenish inventory at a retailer with units of the product from a nearby retailer than by requiring the vehicle to return to the depot. This extension of the SVMI problem is suitable for a distribution system with multiple clusters of retailers.

It is interesting to see how suboptimal solution procedures perform for the SVMI problem. These solution techniques may include myopic policies, base-stock inventory policies, and heuristic solution procedures based on the structural results for inventory control that we have obtained. We investigate these alternative solution approaches and present our findings in Chapter 4.

We also study the operating performance of variations of the SVMI problem. These variations differ in the available state information and how the vehicle route is selected. We are particularly interested in how the quality of state information and the flexibility in the vehicle routing procedure (or the size of the set of inventory locations that the vehicle can visit next) help improve the operating performance of the distribution system. The results are presented in Chapter 5.

# CHAPTER III

# INFINITE HORIZON PERIODIC MARKOV DECISION PROCESSES

In the preceding chapter, we formulate the stochastic vendor managed inventory (SVMI) problem as a finite horizon non-homogeneous Markov decision process. By assuming periodic reward and transition structures in the infinite problem horizon, we obtain an instance of the infinite horizon periodic Markov decision process. This class of stochastic processes is the topic of this chapter. In particular, we formulate the infinite horizon periodic Markov decision problem and study the existence and convergence of its solutions. We focus on the expected total discounted reward criterion. Relevant results for the infinite horizon SVMI problem (with periodic reward and transition structures) are emphasized. Subsequently, we discuss how the structural results and algorithms for the SVMI problem can be extended to the multiperiod and infinite horizon versions of the problem.

The finite horizon non-homogeneous Markov decision problem and the infinite horizon periodic Markov decision problem are formulated next. For the latter, we show that there is an equivalent infinite horizon stationary Markov decision problem. In Section 3.3, we present theoretical results for the expected total discounted reward model of the infinite horizon periodic Markov decision problem. A discussion on the infinite horizon and multiperiod SVMI problems then follows in Section 3.4.

## *3.1 Finite Horizon Non-Homogeneous Markov Decision Problem*

We define $S$ as the set of states and $T$ as the length of the problem horizon. The process is $\{s(t), t = 1, 2, ..., T\}$, where $s(t) \in S$. We assume that $S$ is a finite set and $T$ is finite.

The transition structure of the Markov process is described by the following conditional transition probabilities:

$$p_t(j|i,a) = Pr[s(t+1) = j|s(t) = i, a(t) = a],$$

where $a \in A_t(i)$. We define $A_t(i)$ as the set of actions available at time $t$ if the state at time $t$ is $i$. For $t = 1, 2, ..., T$, $A_t(i) \subseteq A$, where $A$ is the finite set of actions. We define $r_t(i,a)$ as the reward accumulated from time $t$ to time $t+1$ if the state at time $t$ is $i$ and action $a$ is chosen at time $t$. Additionally, $\hat{r}_{T+1}(i)$ is defined as the terminal reward to be accumulated at time $T+1$ if the state at time $T+1$ is $i$.

A decision rule $d_t$ specifies an action to take at time $t$, given the state of the system. That is, $d_t$ is a mapping from $S$ to $A$. A policy $\delta$ is a sequence of decision rules from time $t = 1$ to time $t = T$, $\delta = (d_1, d_2, ..., d_T)$. We define $\Delta$ as the set of deterministic Markov policies for the finite horizon non-homogeneous Markov decision problem. Because of the finite state and action sets, we can restrict our attention to this type of policies. The optimality criterion for the problem is the expected total reward accumulated from time $t = 1$ to time $t = T+1$:

$$V_T^\delta(i) = E_i^\delta \sum_{1 \leq t \leq T} r_t(s(t), a(t)) + \hat{r}_{T+1}(s(T+1)).$$

That is, $V_T^\delta(i)$ is the expected total reward if the state at time $t = 1$ is $i$ and policy $\delta$ is followed. The problem objective is to find a policy from the set $\Delta$ that maximizes the optimality criterion.

## 3.2 Infinite Horizon Periodic Markov Decision Problem

Unless specified otherwise, the parameters that appear here are defined as in the previous section. The process is $\{s(t), t = 1, 2, ...\}$, where $s(t) \in S$. The reward function, $r_t(i,a)$, and the transition probability, $p_t(j|i,a)$, are defined as in the finite horizon non-homogeneous Markov decision problem. We assume that, for $n = 1, 2, ...$, $r_t(i,a) = r_{nT+t}(i,a)$, $p_t(j|i,a) = p_{nT+t}(j|i,a)$, and $A_t(i) = A_{nT+t}(i)$. Furthermore, after every $T$

time periods, there is a reward accrued as a function of the state of the system. We denote this reward by $\bar{r}_{nT+1}(i)$, for $n = 1, 2, ...$, and remark that it may be different from the terminal reward $\hat{r}_{T+1}(i)$ for the finite horizon non-homogeneous Markov decision problem.

Because of the periodicity in the reward and transition structures, we can formulate an equivalent infinite horizon stationary Markov decision problem. For $n = 1, 2, ...$, let $\hat{A}_n$ be the set of actions for the stationary problem. In this case, an action is $\hat{a}$, where $\hat{a} \in \hat{A}_n$. We note that this action is a policy for the finite horizon non-homogeneous Markov decision problem. The reward and transition structures for the stationary process are as follows:

$$R_n(i, \hat{a}) = \sum_{1 \leq t \leq T} r_{nT+t}(s(nT+t), a(nT+t)) + \bar{r}_{(n+1)T+1}(s((n+1)T+1))$$

$$P_n(j|i, \hat{a}) = Pr[s((n+1)T) = j | s(nT) = i, \hat{a}(nT) = \hat{a}]$$

Clearly, these two parameters can be determined from the relevant parameters of the finite horizon non-homogeneous Markov decision process. Because of the periodic reward and transition structures, for $n = 1, 2, ...$, we have that $\hat{A}_n(i) = \hat{A}$, $R_n(i, \hat{a}) = R(i, \hat{a})$, and $P_n(j|i, \hat{a}) = P(j|i, \hat{a})$. We assume that $|R(i, \hat{a})| < \infty$, for all $i \in S$ and $\hat{a} \in \hat{A}$.

A policy $\pi$ is defined as $\pi = (\hat{a}_1, \hat{a}_2, ...)$. We have that $\pi \in \Pi$, where $\Pi$ is the set of all deterministic Markov policies for this problem. Let $\beta$ be the discount factor, where $0 \leq \beta < 1$. Given that $s(1) = s$, the expected total discounted reward criterion is

$$V^{\pi}_{\beta,T}(s) = E^{\pi}_{s}\{\sum_{0 \leq n < \infty} \beta^n R(s(nT), \hat{a}(nT))\}.$$

Under this criterion, we say that a policy $\pi^*$ is optimal if, for all $s \in S$ and all $\pi \in \Pi$,

$$V^{\pi^*}_{\beta,T}(s) \geq V^{\pi}_{\beta,T}(s).$$

Given that $s(1) = s$, the average reward criterion is

$$g^{\pi}(s) = lim_{N \to \infty}(1/N)V^{\pi}_{N}(s),$$

where

$$V^{\pi}_{N}(s) = E^{\pi}_{s}\{\sum_{0 \leq n \leq N} R(s(nT), \hat{a}(nT))\}.$$

We say that a policy $\pi^*$ is optimal if, for all $s \in S$ and all $\pi \in \Pi$,

$$liminf_{N \to \infty}(1/N)V_N^{\pi^*}(s) \geq limsup_{N \to \infty}(1/N)V_N^{\pi}(s).$$

## 3.3 Theoretical Results

In this section, we present the theoretical results for the infinite horizon periodic Markov decision process, particularly for the expected total discounted reward criterion.

By assumption, we have that $0 \leq \beta < 1$ and $|R(i, \hat{a})| < \infty$, for all $i \in S$ and $a \in \hat{A}$. The sets $S$ and $\hat{A}$ are finite. The optimality equation for the infinite horizon stationary Markov decision problem is as follows:

$$v(i) = max_{\hat{a} \in \hat{A}}\{R(i, \hat{a}) + \beta \sum_j P(j|i, \hat{a})v(j)\}$$

Standard value iteration algorithm, such as Algorithm 3.2.1 below, is a common approach to solve the optimality equations in the above form. We define $V$ as the set of bounded real-valued functions on $S$. Additionally, $\|v\| = max_{i \in S}\{v(i)\}$ and $d_\epsilon$ is the $\epsilon$-optimal decision rule.

For this problem, the set $\hat{A}$ could be very large and, thus, render the algorithm computationally intractable. We propose an alternative value iteration algorithm, i.e., Algorithm 3.2.2. Here $d_\epsilon^t$ is the $\epsilon$-optimal decision rule at time $t$. It follows that $d_\epsilon = (d_\epsilon^1, d_\epsilon^2, ..., d_\epsilon^T)$.

We now show that Algorithm 3.2.1 and Algorithm 3.2.2 are equivalent.

**Theorem 3.3.1.** *The value iteration algorithm and the alternative value iteration algorithm for the infinite horizon stationary Markov decision problem are equivalent.*

*Proof.* To show that the two algorithms are equivalent is the same as showing that, for $n = 1, 2, ...,$ $v^n = u_1^n$. Theorem 4.5.1(b) in Puterman (1994) states that a policy consisting of the optimal decision rules, as determined by backward induction, for a finite horizon Markov decision process is optimal. By definition, $v^0 = u_1^0$. It follows from the theorem that $v^1 = u_1^1$, and then $v^2 = u_1^2$, and so on. $\square$

**Algorithm 3.3.1** Value iteration algorithm

0. Select $v^0 \in V$ and set $n = 0$. Also, specify $\epsilon > 0$.

1. For each $i \in S$, compute

$$v^{n+1}(i) = max_{\hat{a} \in \hat{A}}\{R(i, \hat{a}) + \beta \sum_j P(j|i, \hat{a})v^n(j)\}.$$

2. If $\|v^{n+1} - v^n\| < \epsilon(1 - \beta)/(2\beta)$, go to step 3.

Otherwise, set $n = n + 1$ and go to step 1.

3. For each $i \in S$, choose

$$d_\epsilon(i) = argmax_{\hat{a} \in \hat{A}}\{R(i, \hat{a}) + \beta \sum_j P(j|i, \hat{a})v^{n+1}(j)\}.$$

**Algorithm 3.3.2** Alternative value iteration algorithm

0. Set $n = 0$ and select $u_1^0 = v^0 \in V$. Specify $\epsilon > 0$.

1. Compute $u_{T+1}^{n+1}(i) = \bar{r}_{T+1}(i) + \beta u_1^n(i)$, for each $i \in S$.

2. For $t = T, T - 1, ..., 1$, and for each $i \in S$, compute

$$u_t^{n+1}(i) = max_{a \in A_t(i)}\{r_t(i, a) + \sum_{j \in S} p_t(j|i, a)u_{t+1}^{n+1}(j)\}.$$

3. If $\|u_1^{n+1} - u_1^n\| < \epsilon(1 - \beta)/(2\beta)$, go to step 4.

Otherwise, set $n = n + 1$ and return to step 1.

4. For $t = 1, 2, ..., T$, and for each $i \in S$, select

$$d_\epsilon^t(i) = argmax_{a \in A_t(i)}\{r_t(i, a) + \sum_{j \in S} p_t(j|i, a)u_{t+1}^{n+1}(j)\}.$$

## 3.4 Multiperiod and Infinite Horizon SVMI Problems

We intend the horizon of the SVMI problem to represent a working day. (There are other possibilities.) Let us refer to the length of the horizon as period. Then multiperiod SVMI problems have a horizon that spans two or more days. We may solve these problems by a value iteration algorithm that finds the optimal set of decision rules for one period at a time. In fact, the step in the value iteration algorithm that solves each period of the problem is essentially the same as the backward induction algorithm for the SVMI problem. It is relatively straightforward to show that the structural results that we have presented apply to each period of the multiperiod SVMI problem.

In Chapter 2, we formulate the SVMI problem as a finite horizon non-homogeneous Markov decision process. By assuming periodic reward and transition structures in the infinite problem horizon, we have that the infinite horizon SVMI problem is an instance of the infinite horizon periodic Markov decision problem that we have studied in this chapter. The expected total discounted reward criterion and, thus, the results in the previous section, are applicable to the infinite horizon SVMI problem.

We have some remarks on the expected total discounted reward criterion. We use the finite horizon to represent a working day. This implies that, when the expected total discounted reward criterion is used in the problem formulation of the infinite horizon SVMI problem, the discount factor should be very close to unity. In our model of the SVMI problem, the finite horizon represents a working day. Time-value of reward can be significant when the problem horizon in the infinite horizon SVMI problem spans over several months. In that case, the expected total discounted reward criterion should be preferred to the average reward criterion.

In the previous section, we use the backward induction algorithm for the finite horizon non-homogeneous Markov decision problem to develop the value iteration algorithm for the infinite horizon periodic Markov decision problem. We obtain the alternative value iteration algorithms for the expected total discounted reward model. These results directly apply to the MDP model of the infinite horizon SVMI problem.

Let us consider the alternative value iteration algorithms. Assume that the sufficient conditions for the structural results in the (finite horizon) SVMI problem are satisfied. From the periodicity in the reward and transition structures, it follows that these structural results also hold in the infinite horizon SVMI problem. This can be shown by direct reasoning or induction. Therefore, we can use the monotone value iteration algorithms based on the structural results in Chapter 2 to solve the MDP model of the infinite horizon SVMI problem.

## 3.5 Conclusions

In this chapter, we have described the infinite horizon periodic Markov decision processes and present relevant theoretical results. By assuming periodicity in the reward and transition structures, we have that the MDP model of the infinite horizon SVMI problem belongs to this class of stochastic processes. For the infinite horizon periodic Markov decision problem, we developed an alternative value iteration algorithm that is based on the backward induction algorithm for the finite horizon non-homogeneous Markov decision problem. This allows us to extend the structural results and algorithms for the MDP model of the (finite horizon) SVMI problem to the MDP model of the infinite horizon version of the problem.

# CHAPTER IV

# SUBOPTIMAL SOLUTIONS OF THE SVMI PROBLEM

In this chapter, we develop suboptimal solution procedures for the SVMI problem and study their computational advantages over the optimal one. Detailed description of the SVMI problem can be found in Chapter 2. Three suboptimal solution procedures (or heuristics) are considered. First, we develop a heuristic based on the structural results for inventory control. These previously established structural results specify monotone relations between the optimal inventory action and inventory levels of the retailers. The second heuristic has base-stock inventory policy. In this case, the base-stock inventory levels are determined via a formula equivalent to that of the Newsvendor's problem. Finally, we study how myopic policies perform in the infinite horizon SVMI problem. As the name implies, we define myopic policy as the result of solving the infinite horizon SVMI problem one finite horizon (or period) at a time.

The first and second heuristic solution procedures are investigated in Section 4.1 and Section 4.2, respectively. In Section 4.3, we study myopic solutions of the infinite horizon SVMI problem. We include sample numerical results in each of the three sections on suboptimal solution procedures.

## *4.1 Suboptimal Solutions Based on the Structural Results for Inventory Control*

In Chapter 2, we establish the structural results for vehicle routing and inventory control in the SVMI problem. The structural results for inventory control are restated below, with the sufficient conditions on the problem parameters included. Unless stated otherwise, the parameters in this chapter are defined as in Chapter 2. Let us recall that $\bar{F}_{t,i}^{l,k} = 1 - F_{t,i}^{l,k}$, where $F_{t,i}^{l,k}$ is the cumulative probability distribution of $D_{t,i}^{l,k}$. Also, note that we can write the vector $x$ as $x = (x_i, x_i^c)$, for $i = 1, 2, ..., N$. Theorem 4.1.1 specifies monotone relations between the optimal inventory action at the current retailer and inventory levels of the non-current retailers.

**Theorem 4.1.1.** *(Inventory control) For $t = 1, 2, ..., T$, an $l \in \{1, 2, ..., N\}$, an $i \in K \backslash \{0, l\}$, and all $k \in K$, assume that*

$$(b_i^1 + b_i^2)\bar{F}_{t,i}^{l,k}(q_i) \geq h_i d_{lk}.$$

*Then there exists $a^*(k)$ for the state $s_t((x_i, x_i^c), x_v, l)$ which is non-decreasing in $x_i$.*

In this section, we develop a heuristic solution procedure based on the monotone relations between the optimal inventory action and inventory levels of the retailers. The solution approach is described next. Subsequently, we present sample numerical results.

### 4.1.1 Solution Approach

Consider a structural result which states that the optimal inventory action, $a^*$, is non-decreasing in inventory level of retailer $i$, denoted by $x_i$. Based on this result, we can develop a heuristic in which optimal inventory actions are determined for states with certain values of $x_i$. Inventory actions for states with the remaining values of $x_i$ are selected by assuming piecewise-linear relationship between $a^*$ and $x_i$. We expect the resulting policy to have the expected total reward that is closer to that of the optimal one as we increase the number of inventory actions optimally determined.

For the SVMI problem, Theorem 4.1.1 provides us with the monotone relations between the optimal inventory action and inventory levels of the retailers that we can use in the heuristic solution procedure as described above. Computational results for some sample problems are presented next. For this heuristic, it is obvious that the more linear sections used in the solution the better the solution will be. On the other hand, increasing the number of linear sections in the heuristic solution requires longer computational time.

### 4.1.2 Numerical Results

Numerical results for the heuristic solution procedure are presented in Table 6 and Table 7. To measure the computational efficiency of the heuristic, we set the efficiency of the optimal solution (by direct backward induction) to be $0\%$ and that of the efficient solution to be $100\%$. The efficient solution gives a lower bound on the run time by randomly assigning inventory and vehicle routing actions at each decision epoch. However, as in the heuristic solution, when the vehicle is at the depot, inventory and vehicle routing actions are optimally computed.

**Table 6:** Solutions based on the structural results for Problem I

| | Expected total reward | Quality | Run time (min.) | Efficiency |
|---|---|---|---|---|
| Optimal solution | 9,220.42 | 100% | 1,571 | 0% |
| Efficient solution | - | - | 867 | 100% |
| Solution I-A | 9,190.78 | 99.68% | 1,120 | 64.06% |
| Solution I-B | 9,193.74 | 99.71% | 1,141 | 61.08% |

**Table 7:** Solutions based on the structural results for Problem II

| | Expected total reward | Quality | Run time (sec.) | Efficiency |
|---|---|---|---|---|
| Optimal solution | 571.14 | 100% | 5,443 | 0% |
| Efficient solution | - | - | 2,699 | 100% |
| Solution II-A | 564.80 | 98.89% | 3,786 | 60.38% |
| Solution II-B | 564.80 | 98.89% | 3,862 | 57.62% |

The results in Table 6 are for a four-retailer SVMI problem (Problem I). The two heuristic solutions for Problem I are referred to as solutions I-A and I-B. In solution I-A, one in ten inventory actions are optimal. In solution I-B, the ratio is two in ten. That is, solution I-B has twice as many linear sections as solution I-A. From the results, it is clear that both heuristic solutions are near optimal, with the first solution slightly better than the second. The efficiency measures for solutions I-A and I-B are about $64\%$ and $61\%$, respectively.

In Table 7, we present the corresponding results for Problem II, which is a three-retailer SVMI problem. It is worth noting that solution II-A gives the same expected total reward as solution II-B. That is, for this example, a change in the number of optimal inventory actions computed does not affect the quality of the heuristic solution.

## 4.2 Suboptimal Solutions with Base-Stock Inventory Policy

Base-stock inventory policy is attractive to practitioners of inventory control because it is relatively easy to implement and yet optimal in various situations. In this section, we first describe how a base-stock policy can be used in the SVMI problem. Computational examples then follow.

### 4.2.1 Determination of Base-Stock Inventory Levels

In the SVMI problem, it is not clear when the vehicle will return to the current retailer since the vehicle's next destination is optimally determined at each decision epoch. At the current decision epoch at time $t$, assume that we know the time until the next visit to the current retailer (retailer $l$). Furthermore, let us suppose that the optimal inventory action is base-stock in nature. That is, there is a target inventory level $S_t^l$ such that the optimal inventory action is

$$a^* = min\{x_v, max\{0, S_t^l - x_l\}\}.$$

Then the inventory control problem is a variation of the classic Newsvendor's problem. Consequently, the target inventory level $S_t^l$ can be determined based on the demand distribution at the current retailer.

We shall proceed to identify the target inventory level as follows. For $l \in \{1, 2, ..., N\}$, assume that the time until the next visit to this retailer is known to be $\delta$. Then, the expected reward for the current retailer $l$ from time $t$ until time $t + \delta$ is

$$g_t(\tilde{x}_l, x_l) = -h_l \delta \tilde{x}_l + b_l^1 E[min\{Q_t^l, \tilde{x}_l\}] - b_l^2 E[max\{0, Q_t^l - \tilde{x}_l\}] - b_l^3(\tilde{x} - x_l),$$

where $\tilde{x}_l = x_l + a$ and $Q_t^l$ is the demand at retailer $l$ from time $t$ until time $t + \delta$. It will later be shown that $g_t(\tilde{x}_l, x_l)$ is concave in $\tilde{x}_l$.

Solely for the purpose of determining the target inventory level, it is reasonable to estimate the time until the next visit to the current retailer. Let $\bar{H}_t^l = 1 - H_t^l$, where $H_t^l$ is the cumulative demand distribution at the current retailer from time $t$ until the next visit to this retailer. The following theorem provides us with the formula to compute the target inventory level, $S_t^l$.

**Theorem 4.2.1.** *Assume that $H_t^l$ and $g_t(\tilde{x}_l, x_l)$ are known. Then the optimal base-stock inventory level for this retailer, $S_t^l$, is such that*

$$\bar{H}_t^l(S_t^l) = (h_l \delta + b_l^3)/(b_l^1 + b_l^2).$$

*Proof.* Let $\rho(d)$ be the probability that $Q_t^l = d$. The expected reward for the current retailer $l$ from time $t$ until the next visit is

$$\begin{aligned}
g_t(\tilde{x}_l, x_l) &= -h_l \delta \tilde{x}_l + b_l^1 E[min\{Q_t^l, \tilde{x}_l\}] - b_l^2 E[max\{0, Q_t^l - \tilde{x}_l\}] - b_l^3(\tilde{x} - x_l) \\
&= -h_l \delta \tilde{x}_l + b_l^1 \sum_{0 \le d < \infty} \rho(d)(min\{d, \tilde{x}_l\}) - b_l^2 \sum_{0 \le d < \infty} \rho(d)(max\{0, d - \tilde{x}\}) - b_t^3(\tilde{x} - x_l) \\
&= -h_l \delta \tilde{x}_l + b_l^1 \sum_{0 \le d < \infty} \rho(d)[min\{d, \tilde{x}\} - (b_l^2/b_l^1) max\{0, d - \tilde{x}_l\}] - b_l^3(\tilde{x} - x_l) \\
&= -h_l \delta \tilde{x}_l + b_l^1 \sum_{0 \le d < \infty} \rho(d)[min\{d, \tilde{x}\} - max\{0, (b_l^2/b_l^1)(d - \tilde{x}_l)\}] - b_l^3(\tilde{x} - x_l).
\end{aligned}$$

It follows that

$$
\begin{aligned}
g_t(\tilde{x}_l + 1, x_l) - g_t(\tilde{x}_l, x_l) &= b_l^1 \sum_{\tilde{x}_l \le d < \infty} \rho(d)(1 + (b_l^2/b_l^1)) - h_l\delta - b_l^3 \\
&= (b_l^1 + b_l^2) \sum_{\tilde{x}_l \le d < \infty} \rho(d) - h_l\delta - b_l^3 \\
&= (b_l^1 + b_l^2)\bar{H}_t^l(\tilde{x}_l) - h_l\delta - b_l^3.
\end{aligned}
$$

This quantity is non-increasing in $\tilde{x}_l$. Thus, $g_t(\tilde{x}_l, x_l)$ is concave in $\tilde{x}_l$. As a result, the value of $\tilde{x}_l$ such that the above quantity is zero maximizes $g_t(\tilde{x}_l, x_l)$. Let $S_t^l$ be this value of $\tilde{x}_l$. The assertion of the proposition follows. □

Note that, since $Q_t^l$ is a discrete random variable, there is almost always no $S_t^l$ that satisfies the equation in the above theorem. Therefore, we may choose the value of $S_t^l$ such that $\bar{H}_t^l(S_t^l)$ is closest to $(h_l\delta + b_l^3)/(b_l^1 + b_l^2)$.

It can be shown that the result of Theorem 4.2.1 is equivalent to that of the Newsvendor's problem with the following underage cost, $c_u$, and overage cost, $c_o$:

$$c_u = b_l^1 + b_l^2 - h_l\delta - b_l^3,$$

and

$$c_o = h_l\delta + b_l^3.$$

**Algorithm 4.2.1** Base-stock inventory algorithm

1. At current time $t$, estimate the time until the next visit to current retailer $l$ and call it $\delta$.

2. Compute the cumulative distribution of $Q_t^l$, namely $H_t^l$.

3. Determine $S_t^l$ such that

$$\bar{H}_t^l(S_t^l) \approx (h_l\delta + b_l^3)/(b_l^1 + b_l^2).$$

4. The base-stock inventory action is $a(S_t^l) = min\{x_v, max\{0, S_t^l - x_l\}\}$.

Algorithm 4.2.1 summarizes how we apply the result of Theorem 4.2.1 to determine the base-stock inventory actions in the SVMI problem.

### 4.2.2 Numerical Results

We present numerical results for the heuristic solution procedure in Table 8 and Table 9. The results in Table 8 are for the same three-retailer SVMI problem as that in Table 7 (Problem II). This allows us to compare the performance of this heuristic with the previous one, which is based on the structural results for inventory control. In Table 8, the solution quality varies noticeably with the base-stock inventory level. The highest quality level is almost $94\%$. Meanwhile, the computational efficiency at around $81\%$ is good. Note that the efficiency levels for the three solutions are the same because they involve the same procedure.

Let us recall that, in the algorithm for this heuristic, we first estimate the time until the next visit to this retailer. In this three-retailer case, we assume that $d_{lk} = 1$, for all $l, k \in K$, where $d_{lk}$ is the travel time from location $l$ to location $k$. At current time $t$, we assume that the vehicle visits the current retailer again at time $t + 4$. This implicitly assumes that the vehicle visits the depot once during a round of service in which all retailers are visited. From the numerical results, the quality level increases with the base-stock inventory level. This implies that we may have underestimated the time until the next visit to the current retailer. That is, the vehicle may visit the depot more than once during each round of service.

When we compare Table 8 with Table 7, it is clear that the solution quality of this heuristic is less than that of the first one. On the other hand, the computational efficiency of this heuristic is better than that of the first one. These observations are intuitive. In particular, once based-stock inventory levels are known, inventory actions are easily determined. But this convenience comes with a loss in the quality of the solution.

Table 9 presents the corresponding results for a two-retailer SVMI problem (Problem III). Clearly, the quality and efficiency measures are quite similar to those in the previous table. This may imply that the performance of this heuristic is independent of the number of retailers.

**Table 8:** Solutions with base-stock inventory policy for Problem II

| | Expected total reward | Quality | Run time (sec.) | Efficiency | Base-Stock level |
|---|---|---|---|---|---|
| Optimal solution | 517.14 | 100% | 5,443 | 0% | - |
| Efficient solution | - | - | 2,699 | 100% | - |
| Solution II-C | 514.34 | 90.05% | 3,216 | 81.16% | 7 |
| Solution II-D | 488.55 | 85.54% | 3,217 | 81.12% | 6 |
| Solution II-E | 534.54 | 93.59% | 3,217 | 81.12% | 8 |

**Table 9:** Solutions with base-stock inventory policy for Problem III

| | Expected total reward | Quality | Run time (sec.) | Efficiency | Base-Stock level |
|---|---|---|---|---|---|
| Optimal solution | 477.96 | 100% | 20.96 | 0% | - |
| Efficient solution | - | - | 10.75 | 100% | - |
| Solution III-A | 433.24 | 90.05% | 12.81 | 79.48% | 6 |
| Solution III-B | 418.31 | 87.52% | 12.78 | 79.78% | 5 |
| Solution III-C | 439.45 | 91.94% | 12.77 | 79.88% | 7 |

## *4.3 Myopic Solutions of the Infinite Horizon SVMI Problem*

In this section, we study how well myopic solutions perform in the infinite horizon SVMI problem under the average reward criterion. Next we describe how we compute the myopic reward and the optimal average reward. Relevant numerical results then follow.

### 4.3.1 Myopic Reward vs. Optimal Average Reward

We assume that the infinite horizon SVMI problem has periodic reward and transition structures. To study how well myopic solutions perform in the infinite horizon SVMI problem, we compare the optimal average reward of the infinite horizon SVMI problem with the optimal expected total reward for the finite horizon SVMI problem. The latter quantity represents the average reward for the infinite horizon SVMI problem when it is solved myopically (i.e., one finite horizon at a time). Therefore, standard backward induction algorithm can be used to compute the myopic reward. To obtain the optimal average reward, we use an equivalent of Algorithm 3.3.2. The parameters of these numerical examples are chosen such that the algorithm converges. In both cases, we assign different salvage value functions and observe how they affect the relative performance of the myopic policies.

### 4.3.2 Numerical Results

We present computational results for the myopic solutions of an instance of the infinite horizon SVMI problem with two retailers (Problem IV) in Table 10. Here $e$ represents the unit salvage value at the retailers, $h$ is the unit holding cost per unit time, $c$ is the unit procurement cost, and $\tau$ is the time until the salvage value of remaining inventory is realized from the end of the horizon. In this case, we assume that $h\tau \leq 0.5c$.

For the myopic policy, by definition, the salvage value is included in the average reward. On the other hand, in calculating the optimal average reward, the salvage value is realized only once at the end of the infinite horizon. This difference gives rise to the bias towards myopic policy and this bias becomes greater as the salvage value increases. Table 10 illustrates this behavior. The cases $e = h\tau$ and $e = c$ represent low and high salvage values, respectively. When the salvage value is what we would expect, in particular $e = 0.5c$, the quality level of myopic policy for the sample problem is close to $93\%$. Based on our computational experience, there is another interesting observation which is not shown here. In particular, when myopic policy is employed and the salvage value is in normal range, there tend to be less inventory left at the retailers at the end of each finite horizon.

**Table 10:** Infinite horizon myopic solutions for Problem IV

| | Myopic average reward | Optimal average reward | Quality |
|---|---|---|---|
| $e = h\tau$ | 1,135.00 | 1,234.79 | 91.92% |
| $e = 0.5c$ | 1,147.05 | 1,237.82 | 92.67% |
| $e = c$ | 1,181.86 | 1,246.64 | 94.80% |

The computational efficiency of myopic solutions depends on how many iterations are required by the algorithm that determines the optimal policy. Only one iteration is needed to find the best myopic policy. However, the procedure needs to be done every period. On the other hand, computing the optimal policy takes several iterations but, theoretically, this is done only once. In general, a multi-period problem is solved myopically because data for future periods are difficult to obtain. Reducing the computational requirement in solving the problem is rarely a reason for using the solution approach. For these reasons, we shall not discuss the computational efficiency of myopic solutions.

## 4.4 Conclusions and Future Research

In this chapter, we have presented three heuristic solution procedures for the SVMI problem. First, the heuristic based on the structural results for inventory control, gives us great solution quality and reasonable computational efficiency in solving our sample problems. Meanwhile, the second heuristic, which has base-stock inventory policy, provides greater efficiency but less quality than the first one. Finally, we studied the performance of myopic solutions in the infinite horizon SVMI problem and illustrate the potential bias towards these solutions as a result of the salvage value of remaining inventory at the retailers.

Though the reductions in computational requirement by the suboptimal solution procedures for the SVMI problems are significant, more may be needed. One promising idea involves more aggressive partition of the state space according to strong structural results. Based on our numerical examples, the suboptimal solutions based on the structural results for inventory control still maintain great quality. This implies that further use of the structures may still be beneficial. We plan to investigate this idea further.

# CHAPTER V

# VARIATIONS OF THE SVMI PROBLEM

## *5.1 Introduction*

Many supply chains have as their components a depot, multiple retailers, and a vehicle, which transports units of a product from the depot to the retailers. A distribution problem with more than one vehicles can be transformed into multiple one-vehicle problems by, for example, solving an instance of the assignment problem in combinatorial optimization. In this chapter, we consider a distribution system of this nature and study how state information quality and vehicle routing strategy affect the operating performance of the distribution system. To do so, we first formulate the SVMI problem and its four variations as finite horizon non-homogeneous Markov decision processes. Then we compare their optimal expected total rewards analytically. For the five problems, the quality of state information ranges from one with delay to one that is current and almost always available. The vehicle route varies from one with a fixed order of retailers to one that can be determined at intersections between inventory locations. Demands for the product at the retailers are independent and random with known distributions. Furthermore, these demands are time-dependent. We use the finite horizon to represent a working day. Thus, the time dependency of demand represents the varying rate of order arrivals throughout the day.

In the SVMI problem, we assume that current inventory levels of the retailers and the vehicle are available at each decision epoch. The decision maker then decides how many units of the product to drop off at the current retailer, or pick up at the depot, and which inventory location (the depot or one of the retailers) the vehicle will visit next. The vehicle can travel to any one of the inventory locations or stay where it is. Inventory costs under consideration include the holding cost, penalty cost for lost order, and procurement cost.

There is also a transportation cost for the vehicle to travel from one inventory location to another. Finally, revenue is accrued for each filled order.

We investigate four variations of the SVMI problem. The distinguishing features among these variations are the available state information, particularly the inventory levels of the retailers, and how the vehicle route is selected. In particular, the first variation is the case in which there is a delay in obtaining state information and the vehicle visits the retailers in a fixed order. The second variation is similar to the first one but without the delay in state information. Meanwhile, in the third variation, the order of the retailers may be varied but only before the vehicle departs the depot at the start of each round of service. In these three variations, the vehicle has the option of making a stop at the depot for replenishment before travelling to the next retailer in the order. Finally, we study a variation of the SVMI problem featuring an intersection between each pair of inventory locations. At each intersection, the decision maker receives current state information and determines which of the two inventory locations accessible from that intersection to visit next.

In Chapter 2, we establish monotone relations between the optimal vehicle routing and inventory actions and inventory levels of the retailers in the SVMI problem. In this chapter, for the first two variations, we show how the optimal replenishment decision varies with inventory level of the vehicle. Then, analytically, we compare the optimal expected total rewards for the SVMI problem and its variations. As expected, improved state information and/or higher flexibility in the vehicle routing procedure increase the optimal expected total reward. Numerical results confirm our findings. Subsequently, we introduce the following notion of complementarity, as defined in Topkis (1998): two products are considered complementary if having more of one product increases the marginal value of having more of the other product. Based on our numerical results, we discuss the hypothesis that suggests a complementary relationship between the quality of state information and the flexibility in vehicle routing procedure towards improving the operating performance of the distribution system.

Given the current state of information technology, it is reasonable to assume that the vendor has access to current state information before inventory and transportation decisions are made. There has not been much study on the effects of state information quality and vehicle routing strategy on the operating performance of the distribution system. We expect this topic to be increasingly relevant as information technology keeps improving and companies thrive for even higher levels of efficiency.

This chapter is organized as follows. In the next four sections, variations of the SVMI problem are formulated and relevant theoretical results presented. Section 5.6 has the numerical results. Our discussion on the numerical results then follows in Section 5.7.

## 5.2 Variation I: The SVMI Problem with Fixed Vehicle Route and Delayed State Information

For this variation, we assume that there is a delay of one period in the state information and the vehicle visits the retailers in a fixed order. Specifically, the vehicle routing procedure is simplified as follows. At the beginning of the trip, the vehicle departs the depot for the first retailer. From each of the non-final retailers, the vehicle can either proceed to the next retailer directly or make a stop at the depot for replenishment before doing so. The vehicle returns to the depot once the final retailer is visited. Another round of service then begins. Next we formulate the problem and present some theoretical results. Unless specified otherwise, the parameters that appear in this section are defined as in the problem formulation of the SVMI problem, which is included in Chapter 2.

### 5.2.1 Problem Formulation

The state at a decision epoch at time $t$ is $s_t = (\tilde{x}, \tilde{l}, x_v, l, z)$, where $s_t \in S = X \times K \times X_v \times K \times Z$. We define $\tilde{x}$ as the row vector of inventory levels of the retailers after the inventory action was taken at the previous decision epoch. This reflects the one-period delay in the state information available to the decision maker. We let $\tilde{l}$ denote the vehicle's location at the previous decision epoch. As in the SVMI problem, $x_v$ and $l$ are the current inventory level and current location of the vehicle, respectively. We define $z$ as the number of retailers that have been visited on the current trip. It follows that $z \in Z = \{0, 1, 2, ..., N\}$.

The set of inventory actions is as follows: for $l = 0$,

$$A_t(\tilde{x}, \tilde{l}, x_v, l, z) = \{a : -(q_v - x_v) \le a \le 0\},$$

and for $l > 0$,

$$A_t(\tilde{x}, \tilde{l}, x_v, l, z) = \{a : 0 \le a \le min\{q_l - \tilde{x}_l, x_v\}\}.$$

The set of vehicle routing actions is such that, at the depot, for $z = 0$, $K_t(\tilde{x}, \tilde{l}, x_v, l, z) = \{1\}$, and for $z > 0$, $K_t(\tilde{x}, \tilde{l}, x_v, l, z) = \{z + 1\}$. At the retailers, for $1 \le l < N$, $K_t(\tilde{x}, \tilde{l}, x_v, l, z) = \{0, z + 1\}$, and for $l = N$, $K_t(\tilde{x}, \tilde{l}, x_v, l, z) = \{0\}$.

We define $p_t(x|\tilde{x}, \tilde{l}, l)$ as the probability that the vector of current inventory levels of the retailers at time $t$ is $x = (x_1, x_2, ..., x_N)$, given that the vector of inventory levels after inventory action was taken at the previous decision epoch (time $t - d_{\tilde{l}l}$) was $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_N)$. By independence, it follows that

$$p_t(x|\tilde{x}, \tilde{l}, l) = \prod_{1 \le i \le N} p_t^i(x_i|\tilde{x}_i, \tilde{l}, l),$$

where $p_t^i(x_i|\tilde{x}_i, \tilde{l}, l)$, for $i = 1, 2, ..., N$, is the transition probability for retailer $i$. This probability can be determined from the distribution of demand at retailer $i$ during the time between the previous and current decision epochs.

Conditioning on the current inventory levels of the retailers and the vehicle, the reward structure for this problem is equivalent to that of the SVMI problem. In particular, assume that the vector of current inventory levels of the retailers is $x = (x_1, x_2, ..., x_N)$. Then, for $l = 0$,

$$r_t((x, x_v, l), a, k) = \sum_{1 \le i \le N} r_t^i(x_i, l, k) - c_{lk},$$

and for $l > 0$,

$$r_t((x, x_v, l), a, k) = \tilde{r}_t^l(x_l, l, a, k) + \sum_{i \in K \setminus \{0, l\}} r_t^i(x_i, l, k) - c_{lk}.$$

By convention, we denote any time $t > T$ by $T+1$. The time-invariant terminal reward is

$$\bar{r}_{T+1}(\tilde{x}, x_v, l, z) = -c_{l0}.$$

The parameters in the reward structure are defined as in the SVMI problem.

We define a decision rule $\delta_t$ as $\delta_t : X \times K \times X_v \times K \times Z \rightarrow \tilde{A}_t \times \tilde{K}_t$, where $\tilde{A}_t = \bigcup_{\tilde{x},\tilde{l},x_v,l,z} A_t(\tilde{x}, \tilde{l}, x_v, l, z)$ and $\tilde{K}_t = \bigcup_{\tilde{x},\tilde{l},x_v,l,z} K_t(\tilde{x}, \tilde{l}, x_v, l, z)$. A policy $\pi$ is defined as $\pi = (\delta_1, \delta_2, ..., \delta_T)$. Here $\pi \in \Pi^A$, where $\Pi^A$ is the set of all deterministic Markov policies for this problem. The objective is to find a policy that maximizes the criterion:

$$v_A^\pi(\tilde{s}_1) = E_{\tilde{s}_1}^\pi \{ \sum_{1 \le j \le J} r_{t_j}(g_{t_j}, a_{t_j}, k_{t_j}) + \bar{r}_{T+1}(s_{T+1}) \},$$

where $\tilde{s}_1$ is the state at time $t = 1$ (or $t_1$) and $g_t$, for $t = 1, 2, ..., T+1$, is the vector $(x, l, x_v, l, 0)$. To compute the value of the criterion, we need the conditional probability of $x$ given $\tilde{x}$, that is $p_t(x|\tilde{x}, \tilde{l}, l)$.

### 5.2.2 Theoretical Results

Let $x'$ and $z'$ be the updated value of $x$ as a result of the inventory action being evaluated and the updated value of $z$ as a result of the vehicle routing action being evaluated, respectively. The optimality equations, including the boundary condition, for this problem are as follows:

$$\tilde{u}_t^A((\tilde{x}, \tilde{l}, x_v, l, z), k) = max_{a \in A_t(\tilde{x},\tilde{l},x_v,l,z)} \{ \sum_x p_t(x|\tilde{x}, \tilde{l}, l) f_t^A((x, \tilde{l}, x_v, l, z), a, k) \},$$

where

$$f_t^A((x, \tilde{l}, x_v, l, z), a, k) = r_t((x, x_v, l), a, k) + u_{t+d_{lk}}^A(x', l, x_v - a, k, z'),$$

$$u_t^A(\tilde{x}, \tilde{l}, x_v, l, z) = max_{k \in K_t(\tilde{x},\tilde{l},x_v,l,z)} \{ \tilde{u}_t^A((\tilde{x}, \tilde{l}, x_v, l, z), k) \},$$

and

$$u_{T+1}^A(\tilde{x}, \tilde{l}, x_v, l, z) = -c_{l0}.$$

Next we show that the optimal value function is non-decreasing in the vehicle's inventory level. The subsequent corollary then shows that, at the depot, the vehicle is always replenished to its full capacity. In the theorem that follows, we establish a desirable structure for the vehicle routing problem in this variation.

**Proposition 5.2.1.** *$u_t^A(\tilde{x}, \tilde{l}, x_v, l, z)$ is non-decreasing in $x_v$, for $t = 1, 2, ..., T+1$.*

*Proof.* Consider the following optimality equation:

$$\tilde{u}_t^A((\tilde{x}, \tilde{l}, x_v, l, z), k) = max_{a \in A_t(\tilde{x}, \tilde{l}, x_v, l, z)} \{ \sum_x p_t(x|\tilde{x}, \tilde{l}, l) f_t^A((x, \tilde{l}, x_v, l, z), a, k) \},$$

where

$$f_t^A((x, \tilde{l}, x_v, l, z), a, k) = r_t((x, x_v, l, z), a, k) + u_{t+d_{lk}}^A(x', l, x_v - a, k, z').$$

Since $u_{T+1}^A(\tilde{x}, \tilde{l}, x_v, l, z) = -c_{l0}$, it is non-decreasing in $x_v$, for all $\tilde{x}, \tilde{l}, x_v, l$ and $z$. Assume that $u_n^A(x', l, x_v - a, k, z')$ is non-decreasing in $x_v$, for $n = T, T-1, T-2, ..., t+1$. We now show that $u_t^A(\tilde{x}, \tilde{l}, x_v, l, z)$ is non-decreasing in $x_v$. By definition, for all $a \in A_t(\tilde{x}, \tilde{l}, x_v, l, z)$, $r_t((x, x_v, l, z), a, k)$ is independent and, therefore, non-decreasing in $x_v$. Also, $p_t(x|\tilde{x}, \tilde{l}, l)$ is independent of $x_v$. Furthermore, the set $A_t(\tilde{x}, \tilde{l}, x_v, l, z)$ is such that, for $x_v'' \geq x_v'$, there exists $a'' \in A_t(\tilde{x}, \tilde{l}, x_v'', l, z)$ such that $x_v'' - a'' \geq x_v' - a'$, for all $a' \in A_t(\tilde{x}, \tilde{l}, x_v', l, z)$. It follows that $\tilde{u}_t^A((\tilde{x}, \tilde{l}, x_v, l, z), k)$ is non-decreasing in $x_v$, for all $k$. This completes the induction. The desired result follows. □

The next corollary establishes an intuitive result that follows from Proposition 5.2.1.

**Corollary 5.2.2.** *At the depot, it is always optimal to replenish the vehicle to its full capacity.*

*Proof.* Consider the following optimality equation:

$$\tilde{u}_t^A((\tilde{x}, \tilde{l}, x_v, l, z), k) = max_{a \in A_t(\tilde{x}, \tilde{l}, x_v, l, z)} \{ \sum_x p_t(x|\tilde{x}, \tilde{l}, l) f_t^A((x, \tilde{l}, x_v, l, z), a, k) \},$$

where

$$f_t^A((x, \tilde{l}, x_v, l, z), a, k) = r_t((x, x_v, l), a, k) + u_{t+d_{lk}}^A(x', l, x_v - a, k, z').$$

By Proposition 5.2.1, it follows that $u^A_{t+d_{lk}}(x', l, x_v - a, k, z')$ is non-decreasing in $x_v - a$. Furthermore, at the depot, $r_t((x, x_v, l), a, k)$ is independent of $a$. As a result, $\sum_x p_t(x|\tilde{x}, \tilde{l}, l) f^A_t((x, \tilde{l}, x_v, l, z), a, k)$ is non-decreasing in $x_v - a$. The assertion of the proposition follows. □

The result of Corollary 5.2.2 can be shown to hold in the SVMI problem. Theorem 5.2.3 establishes the relationship between the optimal destination of the vehicle, when it is at one of the non-final retailers, and its inventory level.

**Theorem 5.2.3.** *For the state $s_t = (\tilde{x}, \tilde{l}, x_v, l, z)$, in which $l \in K \backslash \{0, N\}$, assume that it is optimal for the vehicle to proceed directly to the next retailer, instead of making a stop at the depot for replenishment first. Then for the state $s'_t = (\tilde{x}, \tilde{l}, x'_v, l, z)$, in which $x'_v \geq x_v$, it is also optimal for the vehicle to proceed directly to the next retailer.*

*Proof.* For $1 \leq l < N$, $K_t(\tilde{x}, \tilde{l}, x_v, l, z) = \{0, z+1\}$. Therefore,

$$u^A_t(\tilde{x}, \tilde{l}, x_v, l, z) = max\{\tilde{u}^A_t((\tilde{x}, \tilde{l}, x_v, l, z), 0), \tilde{u}^A_t((\tilde{x}, \tilde{l}, x_v, l, z), z+1)\}.$$

Corollary 5.2.2 states that, at the depot, the vehicle is always replenished to its full capacity. Thus, we may write

$$\tilde{u}^A_t((\tilde{x}, \tilde{l}, x_v, l, z), 0) = max_{a \in A_t(\tilde{x}, \tilde{l}, x_v, l, z)}\{\sum_x p_t(x|\tilde{x}, \tilde{l}, l) w^0_t((x, \tilde{l}, x_v, l, z), a, 0)\},$$

where

$$w^0_t((x, \tilde{l}, x_v, l, z), a, 0) = r_t((x, x_v, l), a, 0) + \sum_y p_t(y|x, l, 0)\tilde{w}^0_t(y, l, q_v, 0, z),$$

in which

$$w^0_t(y, l, q_v, 0, z) = r_{t+d_{l0}}((y, q_v, 0), q_v - x_v + a, z+1) + u^A_{t+d_{l0}}(y, l, q_v, z+1, z+1).$$

Meanwhile,

$$\tilde{u}^A_t((\tilde{x}, \tilde{l}, x_v, l, z), z+1) = max_{a \in A_t(\tilde{x}, \tilde{l}, x_v, l, z)}\{\sum_x p(x|\tilde{x}, \tilde{l}, l) w^1_t((x, x_v, l, z), a, z+1)\},$$

where

$$w_t^1((x, x_v, l, z), a, z+1) = r_t((x, x_v, l), a, z+1) + u^A_{t+d_{l(z+1)}}(x', l, x_v - a, z+1, z+1),$$

in which $x'$ is the updated value of $x$ as a result of the inventory action being evaluated. It can be shown that, for all $a \in A_t(\tilde{x}, \tilde{l}, x_v, l, z)$, $\tilde{u}_t^A((\tilde{x}, \tilde{l}, x_v, l, z), z+1)$ is non-decreasing in $x_v$ and $\tilde{u}_t^A((\tilde{x}, \tilde{l}, x_v, l, z), 0)$ is independent of $x_v$. The assertion of the theorem follows. $\square$

## *5.3 Variation II: The SVMI Problem with Fixed Vehicle Route*

This variation is similar to the previous one except that, in this case, there is no delay in state information. We first formulate the problem and then present theoretical results. Similar results to those for Variation I are included. Additionally, we compare the optimal expected total rewards for the two variations. Unless specified otherwise, the parameters that appear in this section are defined as in the problem formulation of the SVMI problem.

### 5.3.1 Problem Formulation

The state at a decision epoch at time $t$ is $s_t = (x, x_v, l, z)$, where $s_t \in S = X \times X_v \times K \times Z$. As in the SVMI problem, $x$ is the vector of current inventory levels of the retailers. Furthermore, $x_v$ and $l$ are the current inventory level and current location of the vehicle, respectively. As in Variation I, $z$ is the number of retailers that have been visited on the current trip, where $z \in Z = \{0, 1, 2, ..., N\}$.

The inventory and vehicle routing action sets, $A_t(x, x_v, l, z)$ and $K_t(x, x_v, l, z)$, respectively, are as follows. For $l = 0$,

$$A_t(x, x_v, l, z) = \{a : -(q_v - x_v) \leq a \leq 0\},$$

and for $l > 0$,

$$A_t(x, x_v, l, z) = \{a : 0 \leq a \leq min\{q_l - x_l, x_v\}\}.$$

At the depot, for $z = 0$, $K_t(x, x_v, l, z) = \{1\}$ and, for $z > 0$, $K_t(x, x_v, l, z) = \{z+1\}$. At the last retailer, $K_t(x, x_v, l, z) = \{0\}$. Finally, for $l \in K \backslash \{0, N\}$, $K_t(x, x_v, l, z) =$

$\{0, z+1\}$.

The reward and transition structures are independent of $z$. So they are the same as those for the SVMI problem. In particular, for $l = 0$,

$$r_t((x, x_v, l), a, k) = \sum_{1 \le i \le N} r_t^i(x_i, l, k) - c_{lk},$$

and, for $l > 0$,

$$r_t((x, x_v, l), a, k) = \tilde{r}_t^l(x_l, l, a, k) + \sum_{i \in K \setminus \{0,l\}} r_t^i(x_i, l, k) - c_{lk}.$$

Let $T+1$ denote any time $t > T$. The terminal reward is

$$\bar{r}_{T+1}(x, x_v, l, z) = -c_{l0}.$$

The transition probability is such that, for $l = 0$,

$$p_t(y|(x, x_v, l), a, k) = \prod_{1 \le i \le N} p_t^i(y_i|x_i, l, k),$$

and for $l > 0$,

$$p_t(y|(x, x_v, l), a, k) = p_t^l(y_l|x_l + a, l, k) \prod_{i \in K \setminus \{0,l\}} p_t^i(y_i|x_i, l, k).$$

Let $\delta_t$ be a decision rule, where $\delta_t : X \times X_v \times K \times Z \rightarrow \tilde{A}_t \times \tilde{K}_t$, $\tilde{A}_t = \bigcup_{x,x_v,l,z} A_t(x, x_v, l, z)$, and $\tilde{K}_t = \bigcup_{x,x_v,l,z} K_t(x, x_v, l, z)$. A policy $\pi$ is defined as $\pi = (\delta_1, \delta_2, ..., \delta_T)$, where $\pi \in \Pi^B$. We define $\Pi^B$ as the set of all deterministic Markov policies for this problem. The objective is to find a policy that maximizes the expected total reward:

$$v_B^\pi(s_1) = E_{s_1}^\pi\{\sum_{1 \le j \le J} r_{t_j}(s_{t_j}, a_{t_j}, k_{t_j}) + \bar{r}_{T+1}(s_{T+1})\},$$

where $s_1$ is the state at time $t = 1$ (or $t_1$).

### 5.3.2 Theoretical Results

Let $z'$ be the updated value of $z$ as a result of the vehicle routing action being evaluated. The optimality equations for this problem are as follows:

$$\tilde{u}_t^B((x, x_v, l, z), k) = max_{a \in A_t(x, x_v, l, z)}\{r_t((x, x_v, l), a, k) + \sum_y p_t(y|(x, x_v, l), a, k) u_{t+d_{lk}}^B(y, x_v - a, k, z')\},$$

$$u_t^B(x, x_v, l, z) = max_{k \in K_t(x, x_v, l, z)}\{\tilde{u}_t^B((x, x_v, l, z), k)\},$$

and

$$u_{T+1}^B(x, x_v, l, z) = \bar{r}_{T+1}(x, x_v, l, z) = -c_{l0}.$$

Next we state the theoretical results similar to those for Variation I. Their proofs are also similar and, thus, omitted.

**Proposition 5.3.1.** *$u_t^B(x, x_v, l, z)$ is non-decreasing in $x_v$, for $t = 1, 2, ..., T$.*

**Corollary 5.3.2.** *At the depot, it is always optimal to replenish the vehicle to its full capacity.*

**Theorem 5.3.3.** *For the state $s_t = (x, x_v, l, z)$, in which $l \in K \backslash \{0, N\}$, assume that the optimal destination of the vehicle is the next retailer, instead of the depot. Then the optimal destination of the vehicle for the state $s_t' = (x, x_v', l, z)$, in which $x_v' \geq x_v$, is also the next retailer.*

We define $\tilde{x}$ as the row vector of inventory levels of the retailers after inventory action was taken at the previous decision epoch. Also, $\tilde{l}$ denotes the vehicle's previous location. We define $w_t^A(\tilde{x}, \tilde{l}, x_v, l, z)$ as the optimal expected reward from time $t$ to the end of the horizon for Variation I. It follows that

$$w_t^A(\tilde{x}, \tilde{l}, x_v, l, z) = u_t^A(\tilde{x}, \tilde{l}, x_v, l, z).$$

Let $w_t^B(\tilde{x}, \tilde{l}, x_v, l, z)$ be the optimal expected reward from time $t$ to the end of the horizon for Variation II, given $\tilde{x}$ and $\tilde{l}$. Since current inventory levels of the retailers are available in Variation II, we have that

$$w_t^B(\tilde{x}, \tilde{l}, x_v, l, z) = \sum_x p_t(x|\tilde{x}, \tilde{l}, l) u_t^B(x, x_v, l, z).$$

The following proposition compares $w_t^A(\tilde{x}, \tilde{l}, x_v, l, z)$ and $w_t^B(\tilde{x}, \tilde{l}, x_v, l, z)$.

**Proposition 5.3.4.** *For* $t = 1, 2, ..., T, T+1$,

$$w_t^A(\tilde{x}, \tilde{l}, x_v, l, z) \leq w_t^B(\tilde{x}, \tilde{l}, x_v, l, z),$$

*for all* $\tilde{x}$, $\tilde{l}$, $x_v$, $l$, *and* $z$.

*Proof.* We shall prove this proposition by induction. At time $T+1$, both quantities are the same, so the inequality holds. Assume that the inequality holds for times $T, T-1, ..., t+1$, we will show that it also holds at time $t$. Let $x'$ and $z'$ be the updated value of $x$ as a result of the inventory action being evaluated and the updated value of $z$ as a result of the vehicle routing action being evaluated, respectively. By their definitions, we may write $w_t^A(\tilde{x}, \tilde{l}, x_v, l, z)$ and $w_t^B(\tilde{x}, x_v, l, z)$ as follows:

$$w_t^A(\tilde{x}, \tilde{l}, x_v, l, z) = max_{k \in K_t(\tilde{x}, \tilde{l}, x_v, l, z)}\{max_{a \in A_t(\tilde{x}, \tilde{l}, x_v, l, z)}\{\sum_x p_t(x|\tilde{x}, \tilde{l}, l) g_t^A((x, x_v, l, z), a, k)\}\}, \tag{5.3.1}$$

where

$$g_t^A((x, x_v, l, z), a, k) = r_t((x, x_v, l), a, k) + w_{t+d_{lk}}^A(x', l, x_v - a, k, z')$$

and

$$w_t^B(\tilde{x}, \tilde{l}, x_v, l, z) = \sum_x p_t(x|\tilde{x}, \tilde{l}, l)[max_{k \in K_t(x, x_v, l, z)}\{max_{a \in A_t(x, x_v, l, z)}\{g_t^B((x, x_v, l, z), a, k)\}\}], \tag{5.3.2}$$

where

$$g_t^B((x, x_v, l, z), a, k) = r_t((x, x_v, l), a, k) + w_{t+d_{lk}}^B(x', x_v - a, k, z').$$

For the induction hypothesis, it follows that

$$g_t^A((x, x_v, l, z), a, k) \leq g_t^B((x, x_v, l, z), a, k),$$

for all

$$k \in \{K_t(\tilde{x}, \tilde{l}, x_v, l, z) \cap K_t(x, x_v, l, z)\},$$

and all

$$a \in \{A_t(\tilde{x}, \tilde{l}, x_v, l, z) \cap A_t(x, x_v, l, z)\}.$$

We note that

$$K_t(\tilde{x}, \tilde{l}, x_v, l, z) = K_t(x, x_v, l, z),$$

and

$$A_t(\tilde{x}, \tilde{l}, x_v, l, z) \subseteq A_t(x, x_v, l, z).$$

By inspecting equation 5.3.1 and equation 5.3.2, we have that

$$w_t^A(\tilde{x}, \tilde{l}, x_v, l, z) \leq w_t^B(\tilde{x}, \tilde{l}, x_v, l, z).$$

This completes the induction. □

By definition, $v_B^{\pi^*}(x, x_v, l, z)$ is the optimal expected total reward for Variation II, given that the current state of the system at the beginning of the horizon (t=1) is $(x, x_v, l, z)$. Let the same quantity for Variation I be $v_A^{\pi^*}((\tilde{x}, \tilde{l}, x_v, l, z)|x)$. In these two quantities, $x$ is the vector of inventory levels of the retailers at the beginning of the horizon and $\tilde{x}$ is the vector of inventory levels of the retailers before the beginning of the horizon (t=0). Also, we define $\bar{w}_1^B((\tilde{x}, \tilde{l}, x_v, l, z)|x)$ as the previously defined quantity $w_1^B(\tilde{x}, \tilde{l}, x_v, l, z)$, given that the vector of inventory levels of the retailers at time $t = 1$ is $x$. Similarly, let $\bar{w}_1^A((\tilde{x}, \tilde{l}, x_v, l, z)|x)$ denote the quantity $w_1^A(\tilde{x}, \tilde{l}, x_v, l, z)$, given that the vector of inventory levels of the retailers at $t = 1$ is $x$. Theorem 5.3.5 compares the optimal expected total rewards for Variation I and Variation II.

**Theorem 5.3.5.** *Given the same state at the beginning of the horizon, the optimal expected total reward for Variation I is not greater than that for Variation II.*

*Proof.* It can be shown that

$$v_B^{\pi^*}(x, x_v, l, z) = \bar{w}_1^B((\tilde{x}, \tilde{l}, x_v, l, z)|x),$$

and that

$$v_A^{\pi^*}((\tilde{x}, \tilde{l}, x_v, l, z)|x) = \bar{w}_1^A((\tilde{x}, \tilde{l}, x_v, l, z)|x).$$

From Proposition 5.3.4, it follows that

$$\bar{w}_1^A((\tilde{x}, \tilde{l}, x_v, l, z)|x) \leq \bar{w}_1^B((\tilde{x}, \tilde{l}, x_v, l, z)|x).$$

Consequently,

$$v_A^{\pi^*}((\tilde{x}, \tilde{l}, x_v, l, z)|x) \leq v_B^{\pi^*}(x, x_v, l, z).$$

This is exactly the assertion of the theorem. □

## 5.4 Variation III: The SVMI Problem with Pre-Determined Vehicle Route

This variation is similar to Variation II except that, in this case, the order of the retailers is not always fixed. In particular, the order of the retailers is chosen and then fixed before the vehicle departs the depot at the beginning of each round of service. The problem formulation is presented next. Then, we compare the optimal expected total reward of this variation with those of Variation II and the SVMI problem. Unless specified otherwise, the parameters that appear in this section are defined as in the problem formulation of the SVMI problem.

### 5.4.1 Problem Formulation

The state at a decision epoch at time $t$ is $s_t = (x, x_v, l, z)$, where $s_t \in S = X \times X_v \times K \times Z$. As in the SVMI problem, $x$ is the vector of current inventory levels of the retailers. Additionally, $x_v$ and $l$ are the current inventory level and current location of the vehicle, respectively. We define $z$ as the number of retailers that have been visited during the current round of service, where $z \in Z = \{0, 1, 2, ..., N\}$.

The set of inventory actions, $A_t(x, x_v, l, z)$, is such that, for $l = 0$,

$$A_t(x, x_v, l, z) = \{a : -(q_v - x_v) \leq a \leq 0\},$$

and for $l > 0$,

$$A_t(x, x_v, l, z) = \{a : 0 \le a \le min\{q_l - x_l, x_v\}\}.$$

To specify the set of vehicle routing actions, we first define $\hat{k}$ as a permissible sequence of retailers. Let $\hat{k}(i)$ denote the $i^{th}$ retailer in sequence $\hat{k}$. The set of $\hat{k}$ is

$$\hat{K} = \{\hat{k} = (\hat{k}(1), \hat{k}(2), ..., \hat{k}(N)) | \hat{k}(i) \in \{1, 2, ..., N\}, \forall i; \hat{k}(i) \neq \hat{k}(j), \forall i \neq j\}.$$

Practically, the set $\hat{K}$ could be much smaller than what the definition implies because certain sequences may be ruled out for various reasons. Let $\hat{K}_t(x, x_v, l, z)$ be the set of all permissible sequences for state $(x, x_v, l, z)$. The sequence of retailers is determined before the vehicle departs the depot for each round of service. Therefore, for $l = z = 0$, $\hat{K}_t(x, x_v, l, z) = \hat{K}$. Otherwise, $\hat{K}_t(x, x_v, l, z) = \emptyset$. The set of vehicle routing actions, which clearly depends on $\hat{k}$, is as follows. For $l = z = 0$, $K_t((x, x_v, l, z), \hat{k}) = \{\hat{k}(1)\}$. For $l = 0$ and $z > 0$, $K_t((x, x_v, l, z), \hat{k}) = \{\hat{k}(z + 1)\}$. Meanwhile, $K_t((x, x_v, l, z), \hat{k}) = \{0\}$, for $l = \hat{k}(N)$. Finally, $K_t((x, x_v, l, z), \hat{k}) = \{0, \hat{k}(z + 1)\}$, for $l \in K \backslash \{0, \hat{k}(N)\}$.

The reward and transition structures are independent of $z$ and $\hat{k}$. So they are similar to those for the SVMI problem. In particular, for $l = 0$,

$$r_t((x, x_v, l), a, k) = \sum_{1 \le i \le N} r_t^i(x_i, l, k) - c_{lk},$$

and for $l > 0$,

$$r_t((x, x_v, l), a, k) = \tilde{r}_t^l(x_l, l, a, k) + \sum_{i \in K \backslash \{0, l\}} r_t^i(x_i, l, k) - c_{lk}.$$

We denote any time $t > T$ by $T + 1$. The terminal reward is

$$\bar{r}_{T+1}(x, x_v, l, z) = -c_{l0}.$$

The transition probability is as follows: for $l = 0$,

$$p_t(y | (x, x_v, l), a, k) = \prod_{1 \le i \le N} p_t^i(y_i | x_i, l, k),$$

and for $l > 0$,

$$p_t(y|(x, x_v, l), a, k) = p_t^l(y_l|x_l + a, l, k) \prod_{i \in K \backslash \{0,l\}} p_t^i(y_i|x_i, l, k).$$

Let $\delta_t$ be a decision rule, where $\delta_t : X \times X_v \times K \times Z \to \tilde{A}_t \times \tilde{K}_t$, $\tilde{A}_t = \bigcup_{x,x_v,l,z} A_t(x, x_v, l, z)$, and $\tilde{K}_t = \bigcup_{x,x_v,l,z,\hat{k}} K_t((x, x_v, l, z), \hat{k})$. A policy $\pi$ is defined as $\pi = (\delta_1, \delta_2, ..., \delta_T)$. Let $\pi \in \Pi^C$, where $\Pi^C$ is the set of all deterministic policies for this problem. The objective is to find a policy that maximizes the expected total reward:

$$v_C^\pi(s_1) = E_{s_1}^\pi\{\sum_{1 \le j \le J} r_t(s_{t_j}, a_{t_j}, k_{t_j}) + \bar{r}_{T+1}(s_{T+1})\},$$

where $s_1$ is the state at time $t = 1$.

### 5.4.2 Theoretical Results

Let $z'$ be the updated value of $z$ as a result of the vehicle routing action evaluated. We now present the optimality equations for this variation. For $l \in K \backslash \{0, \hat{k}(N)\}$,

$$\tilde{u}_t^C((x, x_v, l, z), \hat{k}, k) = max_{a \in A_t(x,x_v,l,z)}\{\bar{w}_t((x, x_v, l, z), \hat{k}, k, a)\},$$

where

$$\begin{aligned}\bar{w}_t((x, x_v, l, z), \hat{k}, k, a) =& r_t((x, x_v, l), a, k) \\ &+ \sum_y p_t(y|(x, x_v, l), a, k)\bar{u}_{t+d_{lk}}^C((y, x_v - a, k, z'), \hat{k}),\end{aligned}$$

and

$$\bar{u}_t^C((x, x_v, l, z), \hat{k}) = max_{k \in K_t((x,x_v,l,z),\hat{k})}\{\tilde{u}_t^C((x, x_v, l, z), \hat{k}, k)\}.$$

For $l = \hat{k}(N)$,

$$\tilde{u}_t^C((x, x_v, l, z), \hat{k}, k) = max_{a \in A_t(x,x_v,l,z)}\{w_t((x, x_v, l, z), \hat{k}, k, a)\},$$

where

$$w_t((x, x_v, l, z), \hat{k}, k, a) = r_t((x, x_v, l), a, k) + \sum_y p_t(y|(x, x_v, l), a, k)u_{t+d_{lk}}^C(y, x_v - a, k, z'),$$

and

$$\bar{u}_t^C((x, x_v, l, z), \hat{k}) = max_{k \in K_t((x, x_v, l, z), \hat{k})}\{\tilde{u}_t^C((x, x_v, l, z), \hat{k}, k)\}.$$

Finally, for $l = 0$,

$$\tilde{u}_t^C((x, x_v, l, z), \hat{k}, k) = max_{a \in A_t(x, x_v, l, z)}\{\bar{w}_t((x, x_v, l, z), \hat{k}, k, a)\},$$

$$\bar{u}_t^C((x, x_v, l, z), \hat{k}) = max_{k \in K_t((x, x_v, l, z), \hat{k})}\{\tilde{u}_t^C((x, x_v, l, z), \hat{k}, k)\},$$

and

$$u_t^C(x, x_v, l, z) = max_{\hat{k} \in \hat{K}_t(x, x_v, l, z)}\{\bar{u}_t^C((x, x_v, l, z), \hat{k})\}.$$

At the end of the horizon,

$$u_{T+1}^C(x, x_v, l, z) = \bar{u}_{T+1}^C((x, x_v, l, z), \hat{k}) = \bar{r}_{T+1}(x, x_v, l, z) = -c_{l0}.$$

Next we compare the optimal expected total reward for this variation with that for Variation II and the SVMI problem.

**Theorem 5.4.1.** *Given the same state at the beginning of the horizon, the optimal expected total reward for Variation II is not greater than that for Variation III.*

*Proof.* For each $\pi \in \Pi^B$, there exists an equivalent policy $\pi' \in \Pi^C$. That is, $\Pi^B \subseteq \Pi^C$. One way to show this is to let $\hat{K}_t(x, x_v, l, z) = \{(1, 2, ..., N)\}$, for $l = z = 0$ in Variation III. That is, the vehicle visits the same (fixed) order of retailers as in Variation II. The assertion of the proposition follows. □

**Theorem 5.4.2.** *Given the same state at the beginning of the horizon, the optimal expected total reward for Variation III in not greater than that for the SVMI problem.*

*Proof.* It can be shown that the the sets of inventory actions for both problems are the same. However, the set of vehicle routing actions for Variation III is a subset of that for the SVMI problem. As a result, $\Pi^C \subseteq \Pi$. The result follows. □

## 5.5 Variation IV: The SVMI Problem with Route-Variable Intersections

In this setting, we assume that there is an intersection between each pair of inventory locations. When the vehicle is at each intersection, current inventory levels of the retailers are available. The decision maker then decides where the vehicle will travel to next. This must be one of the two inventory locations accessible from the intersection. No inventory is taken at an intersection. Current inventory levels are also available when the vehicle is at each of the inventory locations. Here inventory action is taken and then the vehicle can travel to any one of the intersections. The problem formulation is presented next. Subsequently, we compare the optimal expected total reward of this variation with that of the SVMI problem. Unless specified otherwise, the parameters that appear in this section are defined as in the problem formulation of the SVMI problem.

### 5.5.1 Problem Formulation

The state at a decision epoch at time $t$ is $s_t = (x, x_v, l)$, where $s_t \in S = X \times X_v \times L$. As in the SVMI problem, $x$ is the vector of current inventory levels of the retailers. Moreover, $x_v$ and $l$ are the current inventory level and current location of the vehicle, respectively. We define the set of locations as $L = K \cup I$, where $K$ is the set of inventory locations and $I$ is the set of intersections. We assume that there is an intersection between each pair of inventory locations. In particular, $I = \{i_{jk} | j, k \in K\}$.

Based on the vehicle's location, there are two types of decision epochs. To help us distinguish them, we define $t_j$ as the time of the $j^{th}$ decision epoch at which the vehicle is at an inventory location. Let $t'_j$, where $t_j < t'_j < t_{j+1}$, be the time of the $j^{th}$ decision epoch at which the vehicle is at an intersection. Let us recall that $d_{lk}$, for $l, k \in K$, is the travel time from inventory location $l$ to inventory location $k$. Thus, if the vehicle is at inventory location $l$ at time $t_j$, then, at time $t'_j$, it will be at an intersection. If, at time $t'_j$, inventory location $k$ is chosen as the vehicle destination, then the next decision epoch will occur at time $t_{j+1} = t + d_{lk}$.

Beyond this point, a decision epoch that occurs at time $t$ implies that the vehicle is at one of the inventory locations. Furthermore, from time $t$, the next decision epoch will occur at time $t'$ and the vehicle will be at an intersection. Let $J$ be the random integer such that $t_J \leq T$ and $t_{J+1} > T$. Consequently, the decision epochs for this problem are $t_1, t'_1, t_2, t'_2, ..., t_J, t'_J$.

We now specify the action sets. Since no inventory action is taken at an intersection, for $l \in I$, we have that $A_{t'}(x, x_v, l) = \emptyset$. For $l \in K$, if $l = 0$,

$$A_t(x, x_v, l) = \{a : -(q_v - x_v) \leq a \leq 0\},$$

and if $l > 0$,

$$A_t(x, x_v, l) = \{a : 0 \leq a \leq min\{q_l - x_l, x_v\}\}.$$

For $l \in K$, the set of vehicle routing actions is $K_t^D(x, x_v, l) = I$. We assume that the vehicle returns to the depot at time $t_{J+1}$. When the vehicle is at an intersection, say $l = i_{jk} \in I$, then the set vehicle routing actions is $K_{t'}^D(x, x_v, l) = \{j, k\}$. That is, from an intersection, the vehicle can travel to only the two inventory locations associated with that intersection.

Let us recall that $c_{lk}$, for $l, k \in K$, denotes the transportation cost for the vehicle to travel from inventory location $l$ to inventory location $k$. We make the following assumption regarding $c_{lk}$ and $d_{lk}$, for $l, k \in K$.

**Assumption 5.5.1.** *For all $j, k, l \in K$, suppose that the vehicle travels from inventory location $l$ to intersection $i_{jk}$. Then $c_{lj} = c_{lk}$ and $d_{lj} = d_{lk}$.*

Suppose the vehicle is at inventory location $l$ at time $t_j$. By the above assumption, the vehicle routing action taken at the subsequent intersection (at time $t'_j$) does not affect the reward accrued during the current interval (between time $t_j$ and time $t_{j+1}$). So we may assume that this reward is accrued before the vehicle arrives at the intersection. As a result, the reward structure for this problem is similar to that for the SVMI problem. In particular, from inventory location $l$, if intersection $i_{jk}$ is visited next, then the reward accrued during the interval is $r_t((x, x_v, l), a, k) = r_t((x, x_v, l), a, j)$. As in the SVMI problem, $r_t((x, x_v, l), a, k)$ is the reward accrued between time $t$ and time $t + d_{lk}$, where $a$ is the inventory action taken at time $t$, and $k$ is the next inventory location to visit. For $l = 0$,

$$r_t((x, x_v, l), a, k) = \sum_{1 \leq i \leq N} r_t^i(x_i, l, k) - c_{lk}.$$

For $l > 0$,

$$r_t((x, x_v, l), a, k) = \tilde{r}_t^l(x_l, l, a, k) + \sum_{i \in K \setminus \{0,l\}} r_t^i(x_i, l, k) - c_{lk}.$$

By convention, we denote any time $t > T$ as $T + 1$. The terminal reward is

$$\bar{r}_{T+1}(x, x_v, l, z) = -c_{l0}.$$

For $l \in K$, we define $p_t^1(y|(x, x_v, l), a, k)$ as the probability that the vector of inventory levels of the retailers at time $t'$ is $y$, given that the state at time $t$ is $(x, x_v, l)$ and inventory action $a$ and vehicle routing action $k$ (an intersection) are taken. Meanwhile, $p_{t'}^2(y|(x, x_v, l), k)$, in which $l \in I$, is defined as the probability that, at time $t + d_{lk}$, the vector of inventory levels of the retailers is $y$, given that the state at time $t'$ is $(x, x_v, l)$ and vehicle routing action $k$ (an inventory location) is chosen. For $l = 0$,

$$p_t^1(y|(x, x_v, l), a, k) = \prod_{1 \leq i \leq N} p_t^i(y_i|x_i, l, k),$$

where $p_t^i(y_i|x_i, l, k)$ is the transition probability at retailer $i$ from time $t$ to time $t'$. For $l > 0$,

$$p_t^1(y|(x, x_v, l), a, k) = p_t^l(y_l|x_l + a, l, k) \prod_{i \in K \setminus \{0,l\}} p_t^i(y_i|x_i, l, k),$$

where $p_t^l(y_l|x_l + a, l, k)$ is the transition probability at the current retailer from time $t$ to time $t'$. Meanwhile,

$$p_{t'}^2(y|(x, x_v, l), k) = \prod_{1 \leq i \leq N} p_{t'}^i(y_i|x_i, l, k),$$

where $p^i_{t'}(y_i|x_i, k)$ is the transition probability for retailer $i$ from time $t'$ to time $t + d_{lk}$.

Let $\delta_t : X \times X_v \times K \to \tilde{A}_t \times \tilde{K}_t$ be a decision rule for the decision epoch at which the vehicle is at an inventory location. In this case, $\tilde{A}_t = \bigcup_{x,x_v,l} A_t(x, x_v, l)$ and $\tilde{K}_t = \bigcup_{x,x_v,l} K^D_t(x, x_v, l)$. A decision rule for the decision epoch at which the vehicle is at an intersection is defined as $\delta_{t'} : X \times X_v \times I \to \tilde{K}_{t'}$, where $\tilde{K}_{t'} = \bigcup_{x,x_v,l} K^D_{t'}(x, x_v, l)$. A policy $\pi$ is defined as

$$\pi = (\delta_1, \delta_{1'}, \delta_2, \delta_{2'}, ..., \delta_T, \delta_{T'}).$$

Let $\pi \in \Pi^D$, where $\Pi^D$ is the set of all deterministic Markov policies for this problem. The objective is to find a policy that maximize the following criterion:

$$v^\pi_D(s_1) = E^\pi_{s_1}\{\sum_{1 \le j \le J} r_{t_j}(s_{t_j}, a_{t_j}, k_{t_j}) + \bar{r}_{T+1}(s_{T+1})\},$$

where $s_1$ is the state at time $t = 1$ (or $t_1$).

### 5.5.2 Theoretical Results

We now present the optimality equations for this problem. For $l \in K$,

$$\tilde{u}^D_t((x, x_v, l), i_{jk}) = max_{a \in A_t(x,x_v,l)}\{r_t((x, x_v, l), a, j) + \sum_y p^1_t(y|(x, x_v, l), a, i_{jk}) u^D_{t'}(y, x_v - a, i_{jk})\},$$

and

$$u^D_t(x, x_v, l) = max_{i_{jk} \in K^D_t(x,x_v,l)}\{\tilde{u}^D_t((x, x_v, l), i_{jk})\}.$$

For $l \in I$,

$$\tilde{u}^D_{t'}((x, x_v, l), k) = \sum_y p^2_{t'}(y|(x, x_v, l), k) u^D_{t+d_{lk}}(y, x_v, k),$$

and

$$u^D_{t'}(x, x_v, l) = max_{k \in K^D_{t'}(x,x_v,l)}\{\tilde{u}^D_{t'}((x, x_v, l), k)\}.$$

At the end of the horizon,

$$u^D_{T+1}(x, x_v, l) = \bar{r}_{T+1}(x, x_v, l) = -c_{l0}.$$

The following theorem compares the optimal expected reward for Variation IV with that for the SVMI problem.

**Theorem 5.5.2.** *Given the same state at the beginning of the horizon, the optimal expected total reward for Variation IV in not less than that for the SVMI problem.*

*Proof.* Let us refer to $u_t^*(x, x_v, l)$ in the optimality equations for the SVMI problem. We shall prove that, for $t = 1, 2, ..., T+1$,

$$u_t^*(x, x_v, l) \leq u_t^D(x, x_v, l).$$

This implies the assertion of the theorem. We shall prove the inequality by induction. By the boundary conditions for both problems, the inequality holds for time $T+1$. Assume that it holds for times $T, T-1, ..., t+1$. We shall prove that the inequality is true for time $t$. We may write $u_t^*(x, x_v, l)$ as

$$u_t^*(x, x_v, l) = max_{k \in K_t(x, x_v, l)}\{max_{a \in A_t(x, x_v, l)}\{w_t((x, x_v, l), a, k)\}\},$$

where

$$w_t((x, x_v, l), a, k) = r_t((x, x_v, l), a, k) + \sum_y p_t(y|(x, x_v, l), a, k) u^*_{t+d_{lk}}(y, x_v - a, k).$$

Meanwhile,

$$u_t^D(x, x_v, l) = max_{i_{jk} \in K_t^D(x, x_v, l)}\{max_{a \in A_t(x, x_v, l)}\{w_t^D((x, x_v, l), a, i_{jk})\},$$

where

$$w_t^D((x, x_v, l), a, i_{jk}) = r_t((x, x_v, l), a, k) + \sum_{x'} p_t^1(x'|(x, x_v, l), a, i_{jk}) g_t^D(x', x_v - a, i_{jk}),$$

in which

$$g_t^D(x', x_v - a, i_{jk}) = max_{k \in K^D_{t'}(x', x_v - a, i_{jk})}\{\sum_y p_{t'}^2(y|(x', x_v - a, k), i_{jk}) u^D_{t+d_{lk}}(y, x_v - a, k)\}.$$

Because no inventory action is taken at an intersection,

$$p_t(y|(x,x_v,l),a,k) = \sum_{x'} p_t^1(x'|(x,x_v,l),a,k) p_{t'}^2(y|(x',x_v-a,k'),k).$$

The induction hypothesis and the above results imply that

$$w_t((x,x_v,l),a,k) \le max\{w_t((x,x_v,l),a,j), w_t((x,x_v,l),a,k)\} \le w_t^D((x,x_v,l),a,i_{jk}),$$

for all $i_{jk} \in K_t^D(x,x_v,l)$.

It follows that

$$u_t^*(x,x_v,l) \le u_t^D(x,x_v,l).$$

This completes the induction. □

## 5.6 Numerical Results

In this section, we present numerical results that compare the optimal expected total rewards for the SVMI problem and its variations. In all five problems, we consider a distribution system with three retailers. Two demand distributions are applied. They are the uniform and discrete normal distributions. We present the optimal expected total rewards for the five problems in Table 11 and Figure 1.

**Table 11:** Optimal expected total rewards for the SVMI problem and its variations

| | Optimal expected total reward for uniform demand | Optimal expected total reward for discrete normal demand |
|---|---|---|
| Variation I | 3,323.25 | 3,055.15 |
| Variation II | 4,149.50 | 3,868.70 |
| Variation III | 4,156.27 | 3,869.14 |
| SVMI Problem | 4,605.36 | 4,276.06 |
| Variation IV | 4,716.38 | 4,337.65 |

From the numerical results, the optimal expected total rewards for the five problems rank in the order that we expected. In particular, in ascending order, there are the optimal expected total rewards for Variation I, Variation II, Variation III, the SVMI problem, and

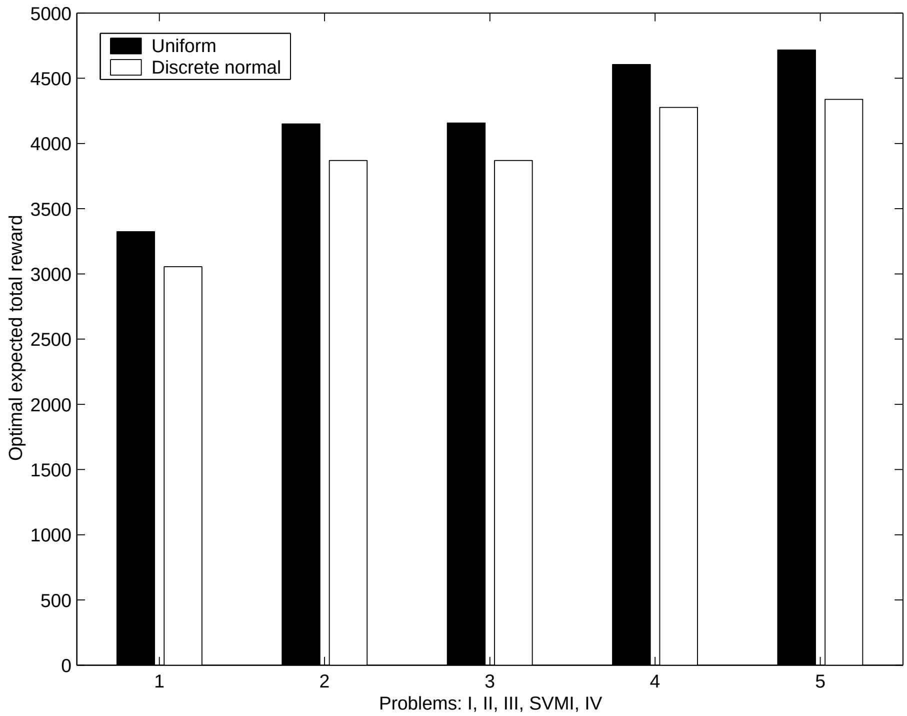


**Figure 1:** Optimal expected total rewards for the SVMI problem and its variations

Variation IV. This supports our analytical results which imply that the optimal expected total reward increases as the quality of state information and/or the flexibility in vehicle routing procedure increase.

There is another interesting observation of the numerical results: the difference between the rewards for Variation I and Variation II and the difference between the rewards for Variation III and the SVMI problem are relatively large. We shall discuss this observation further in the next section. For each of the five problems, the optimal expected total reward for the uniform demand is greater than that for the discrete normal demand. This supports the generally-true hypothesis for inventory problems that the optimal expected total reward is non-increasing in the demand variability.*

*Exceptions to this hypothesis are rare. For interested readers, Ridder et al. (1998) present sufficient conditions for the cost to decrease as the demand variability increases in the Newsvendor's problem.

## 5.7 Discussion

Based on the above numerical results, we shall attempt to establish a qualitative relationship between state information quality and vehicle routing strategy in the distribution system that we have described. Let us propose the following hypothesis: in general, the greater use of improved state information, particularly by adding more flexibility to the vehicle routing procedure, results in increasingly better operating performance of the distribution system. In this case, we use the optimal expected total reward as the measure of operating performance of the distribution system.

The numerical results are used as evidence that supports, but does not proves, the above hypothesis. From Figure 1, the difference between the optimal expected total rewards for Variation II and Variation III is small relative to the difference between those for Variation III and the SVMI problem. In all four problems, current inventory levels of the retailers are available at each decision epoch. Meanwhile, the vehicle routing procedure in the SVMI problem is much more flexible than that in Variation III, which is slightly more flexible than that in Variation II. This shows that our numerical examples supports the above hypothesis.

Even though the vehicle routing procedure in Variation IV is somewhat more flexible than that of the SVMI problem, the difference between the optimal expected total rewards for the two problems are relatively small. This represents a limit in the use of better state information to improve the operating performance of the distribution system. In practice, whether or not Variation IV will be implemented would depend on the projected operating costs and benefits.

Finally, it can be argued that eliminating the delay in state information improves the operating performance of the distribution system significantly. This is because of the marked difference between the optimal expected total reward for Variation I and that for Variation II. In this case, a delay of one period in the state information has strong negative impact on the value of the optimality criterion.

## 5.8 Conclusions and Future Research

In this chapter, we have shown how state information quality and vehicle routing strategy affect the operating performance of a distribution system. In particular, improved state information and increased flexibility in the vehicle routing procedure both result in higher optimal expected total reward for the distribution problem. Furthermore, based on our numerical results, it can be argued that there is a complementary relationship between the two factors. However, there is a limit to the use of better state information to increase the optimal expected total reward via increased flexibility in the vehicle routing procedure. Finally, our numerical examples implies that delay in state information has strong negative impact on the operating performance of the distribution system.

More numerical examples are needed to support our hypothesis. It is interesting to find out if we can prove this result analytically. Another interesting question is how significant our theoretical results will be in actual distribution systems. In practice, we expect to experience issues beyond those considered in our study that affect management's decisions on the appropriate state information system and vehicle routing strategy for their distribution system.

# CHAPTER VI

# THE COCA-COLA DISTRIBUTION PROBLEM: A CASE STUDY

Although the stochastic vendor managed inventory (SVMI) problem is rather general and applicable in practice, we believe it is important that an actual distribution problem is included in our study. We are particularly interested in showing that structural results similar to those for the SVMI problem can be established in an actual distribution problem. This led us to consider local companies that could provide us with such example. Subsequently, we had telephone conversations with the logistics team at the Coca-Cola Enterprises, Inc. about their distribution problems. One of the problems facing the company shares some characteristics with the SVMI problem and its variations. The problem is studied in this chapter. In particular, we consider the problem of producing soft-drink products at the cannery and then delivering them to the distribution centers, which face stochastic demands. We formulate the problem as a finite horizon non-homogeneous Markov decision process. The problem objective is to find a policy that maximizes the expected total reward, which includes revenue, transportation cost, and inventory costs. We simplify the vehicle routing procedure by considering only one distribution center. For the inventory control, we show that the optimal delivery actions vary monotonically with inventory levels of the products at the distribution center.

This chapter is organized as follows. First, we provide more details of the problem in Section 6.1. Section 6.2 is the problem formulation. In the subsequent sections, we present theoretical results, which include the optimality equations and structural results for delivery control.

## 6.1 Problem Description

There is a cannery, where soda cans are produced and packaged into six-packs and twelve-packs. We shall restrict our attention to only the two types of products. From the cannery, a vehicle transports the products to the distribution center. In practice, a cannery serves multiple distribution centers. However, for reasons to be specified later, we will include only one distribution center in our analysis. Demand for each product (from retail outlets) at the distribution center is history-independent, stochastic, and time-dependent.

The actual distribution problem facing the company also involves the delivery of products to the retail outlets. However, because of the following reasons, we shall not consider this part of the problem. First of all, each retailer manages its own inventory and independently incurs the associated costs. Also, as we were told, the vehicle that delivers soft-drink products to the retailers usually follows a fixed route and pre-determined replenishment times. Finally, the vehicle operator knows in advance the order quantities for the products from each retailer. As a result, we shall focus our attention on the two-level supply chain, consisting of the cannery and the distribution center.

Based on our conversations with the company's logistics team, deliveries of the six-packs and twelve-packs from the cannery to the distribution centers are made on a periodic basis. In particular, each distribution center receives a shipment every few days and the delivery vehicle follows a fixed route. Because of this predictability, we can simplifly the problem by including only one distribution center in the problem formulation. This is a reasonable assumption because, from the decision maker's perspective, the distribution centers are treated almost identically. The assumption helps us focus our attention on the inventory decisions. In particular, the resulting problem involves managing the inventories of soda cans at the cannery, soft-drink products at the cannery, and soft-drink products at the distribution center.

We now describe the series of events in each time period. At the cannery, the decision maker decides whether or not to begin production of soda cans in the current period. Production occurs in a fixed lot size and takes less than one period to complete. We refer to the required amount of time as the production delay. From the available soda cans, decisions are made on how many six-packs and twelve-packs to be packaged and how many units of each product to deliver to the distribution center. Packaging and delivery take less than one period to complete. We refer to the lead times as packaging and delivery delays, correspondingly.

## 6.2 *Problem Formulation*

Because of the production, packaging, and delivery delays, the state for the Coca-Cola distribution problem includes the numbers of cans being produced, units of each product being packaged, and units of each product being delivered to the distribution center. We define the state at time $t$ as $s_t = (\hat{i}, \hat{q})$, where $\hat{i} = (i_c^c, i_6^c, i_{12}^c, i_6^d, i_{12}^d)$ and $\hat{q} = (q_c^c, q_6^p, q_{12}^p, q_6^d, q_{12}^d)$. Here $i_c^c$ is the inventory level of soda cans at the cannery, $i_6^c$ ($i_{12}^c$) is the inventory level of six-packs (twelve-packs) at the cannery, and $i_6^d$ ($i_{12}^d$) is the inventory level of the six-packs (twelve-packs) at the distribution center. Meanwhile, $q_c^c$ is the number of soda cans being produced, $q_6^p$ ($q_{12}^p$) is the number of six-packs (twelve-packs) being packaged at the cannery, and $q_6^d$ ($q_{12}^d$) is the number of six-packs (twelve-packs) being delivered to the distribution center.

We define $M_x^y$, for $x \in \{c, 6, 12\}$ and $y \in \{c, d\}$, as the capacity for product $x$ at location $y$. Specifically, as in the definitions of $\hat{i}$ and $\hat{q}$, product $c$ means the soda cans, product $6$ means the six-packs, and product $12$ means the twelve-packs. Meanwhile, location $c$ is the cannery and location $d$ is the distribution center. Let $(\hat{i}, \hat{q}) \in I \times Q$, where $I = I_c^c \times I_6^c \times I_{12}^c \times I_6^d \times I_{12}^d$. Here $I_c^c = \{0, 1, 2, ..., M_c^c\}$, $I_6^c = \{0, 1, 2, ..., M_6^c\}$, $I_{12}^c = \{0, 1, 2, ..., M_{12}^c\}$, $I_6^d = \{0, 1, 2, ..., M_6^d\}$, and $I_{12}^d = \{0, 1, 2, ..., M_{12}^d\}$. Meanwhile, $Q = \{0, b\} \times Q_6^p \times Q_{12}^p \times Q_6^d \times Q_{12}^d$, where $b$ is the fixed production lot size, $Q_6^p = I_6^c$, $Q_{12}^p = I_{12}^c$, $Q_6^d = I_6^d$, and $Q_{12}^d = I_{12}^d$.

Inventory decisions are made at the cannery. They are the production (canning), packaging, and delivery actions. Let $a_1$ be a production action. Precisely, $a_1 = 1$ ($a_1 = 0$) means production starts (does not start) in the current period. We assume that production occurs in a fixed lot size of $b$ soda cans. Let us recall that $M_c^c$ denotes the maximum capacity for the soda cans at the cannery. The set of $a_1$, denoted by $A_t^1(\hat{i}, \hat{q})$, is such that $A_t^1(\hat{i}, \hat{q}) = \{0, 1\}$, if $q_c^c = 0$ and $i_c^c \leq M_c^c - b$. Otherwise, $A_t^1(\hat{i}, \hat{q}) = \{0\}$. We define $a_2$ ($b_2$) as the number of six-packs (twelve-packs) to be packaged at the cannery. As aforementioned, $M_6^c$ ($M_{12}^c$) represents the storage capacity for the six-packs (twelve-packs) at this location. The set of packaging actions are as follows:

$$a_2 \in A_t^2(\hat{i}, \hat{q}) = \{0, 1, 2, ..., min\{\lfloor i_c^c/6 \rfloor, M_6^c - i_6^c - q_6^c\}\},$$

and

$$b_2 \in B_t^2(\hat{i}, \hat{q}, a_2) = \{0, 1, 2, ..., min\{\lfloor (i_c^c - a_2)/12 \rfloor, M_{12}^c - i_{12}^c - q_{12}^c\}\}.$$

Note that $B_t^2$ is dependent on $a_2$. Let $a_3$ ($b_3$) be the number of six-packs (twelve-packs) to be delivered from the cannery to the distribution center. We have defined $M_6^d$ ($M_{12}^d$) as the storage capacity for the six-packs (twelve-packs) at the distribution center. The sets of delivery actions are

$$a_3 \in A_t^3(\hat{i}, \hat{q}) = \{0, 1, 2, ..., min\{i_6^c, M_6^d - i_6^d - q_6^d\}\},$$

and

$$b_3 \in B_t^3(\hat{i}, \hat{q}) = \{0, 1, 2, ..., min\{i_{12}^c, M_{12}^d - i_{12}^d - q_{12}^d\}\}.$$

Costs (per unit) under consideration include the production cost (c), packaging costs ($g_6, g_{12}$), delivery costs ($d_6, d_{12}$), holding costs ($h_c^c, h_6^c, h_{12}^c, h_6^d, h_{12}^d$), and shortage costs ($s_6, s_{12}$). Revenues per unit for the six-packs and twelve-packs are denoted by $w_6$ and $w_{12}$, respectively. Given that the state at time $t$ is $(\hat{i}, \hat{q})$ and the set of actions $\hat{a} = (a_1, a_2, b_2, a_3, b_3)$

is taken, the reward accrued between time $t$ and time $t+1$ is

$$\begin{aligned} r_t((\hat{i},\hat{q}),\hat{a}) = &- ca_1 - g_6 a_2 - g_{12} b_2 - d_6 a_3 - d_{12} b_3 - h_c(i_c^c + q_c^c - 6a_2 - 12b_2) \\ &- h_6^c(i_6^c + q_6^p - a_3) - h_{12}^c(i_{12}^c + q_{12}^p - b_3) - h_6^d(i_6^d + q_6^d) - h_{12}^d(i_{12}^d + q_{12}^d) \\ &+ w_6 E[min\{(i_6^d + q_6^d), D_t^6\}] + w_{12} E[min\{(i_{12}^d + q_{12}^d), D_t^{12}\}] \\ &- s_6 E[max\{D_t^6 - (i_6^d + q_6^d), 0\}] - s_{12} E[max\{D_t^{12} - (i_{12}^d + q_{12}^d), 0\}], \end{aligned}$$

where $\hat{a} = (a_1, a_2, b_2, a_3, b_3)$ and a random integer variable $D_t^6$ ($D_t^{12}$) is the number of six-packs (twelve-packs) demanded at the distribution center between time $t$ and time $t+1$. At the end of the horizon, the terminal reward is

$$\bar{r}_{T+1}(\hat{i},\hat{q}) = e_c^c i_c^c + e_6^c i_6^c + e_{12}^c i_{12}^c + e_6^d i_6^d + e_{12}^d i_{12}^d,$$

where $e_x^y$, for $x \in \{c, 6, 12\}$ and $y \in \{c, d\}$, is the per-unit salvage value for product $x$ at location $y$.

Let $\hat{j} = (j_c^c, j_6^c, j_{12}^c, j_6^d, j_{12}^d)$. We define $p_t(\hat{j}|(\hat{i},\hat{q}),\hat{a})$ as the probability that the vector of inventory levels at time $t+1$ is $\hat{j}$, given that the state at time $t$ is $(\hat{i},\hat{q})$ and the set of actions $\hat{a} = (a_1, a_2, b_2, a_3, b_3)$ is taken. The transitions of $\hat{q}$ and some elements in $\hat{j}$ are deterministic. In particular, at time $t+1$, $q_c^c = a_1 b$, $q_6^c = a_2$, $q_{12}^c = b_2$, $q_6^d = a_3$, and $q_{12}^d = b_3$. Furthermore, $j_c^c = i_c^c + q_c^c - 6a_2 - 12b_{12}$, $j_6^c = i_6^c + q_6^p - a_3$, and $j_{12}^c = i_{12}^c + q_{12}^p - b_3$. Because orders for the two products arrive at the distribution center, we have that

$$p_t(\hat{j}|(\hat{i},\hat{q}),\hat{a}) = \tilde{p}_t(j_6^d | i_6^d + q_6^d)\tilde{p}_t(j_{12}^d | i_{12}^d + q_{12}^d),$$

where $\tilde{p}_t(j_6^d | i_6^d + q_6^d)$ and $\tilde{p}_t(j_{12}^d | i_{12}^d + q_{12}^d)$ are the transition probabilities at the distribution center of the six-packs and twelve-packs, respectively.

Let $\hat{d}_t : I \times Q \to A$ be a decision rule, where

$$A = \bigcup_{t,\hat{i},\hat{q}} A_t^1(\hat{i},\hat{q}) \times A_t^2(\hat{i},\hat{q}) \times B_t^2(\hat{i},\hat{q}) \times A_t^3(\hat{i},\hat{q}) \times B_t^3(\hat{i},\hat{q}).$$

In particular, $\hat{d}_t(\hat{i},\hat{q}) = (dta_1, dta_2, dtb_2, dta_3, dtb_3)$, where $dta_1$ specifies $a_1$ (the production action), $dta_2$ specifies $a_2$ (the packaging action for the six-packs), and so on. A policy

$\pi$ is defined as $\pi = (\hat{d}_1, \hat{d}_2, ..., \hat{d}_T)$. An optimal policy maximizes the following criterion: given that the state at time $t = 1$ is $(\hat{i}_1, \hat{q}_1)$,

$$v_T^\pi(\hat{i}_1, \hat{q}_1) = \sum_{1 \le t \le T} r_t((\hat{i}_t, \hat{q}_t), \hat{a}_t) + \bar{r}_{T+1}(\hat{i}_{T+1}, \hat{q}_{T+1}).$$

It follows from the history-indenpendent property of the demands that the model is Markovian. Furthermore, the state space and the action sets are finite. As a result, we can restrict our attention to deterministic Markov policies.

## 6.3 Preliminary Results

Let $\hat{q}'$ be the updated value of $\hat{q}$ for the next period. The optimality equations for this problem are as follows:

$$u_t^3((\hat{i}, \hat{q}), a_1, a_2, b_2) = max_{(a_3,b_3) \in A_t^3(\hat{i},\hat{q}) \times B_t^3(\hat{i},\hat{q})} \{r_t((\hat{i}, \hat{q}), \hat{a}) + \sum_{\hat{j}} p_t(\hat{j}|(\hat{i}, \hat{q}), \hat{a}) u_{t+1}^1(\hat{j}, \hat{q}')\}$$

$$u_t^2((\hat{i}, \hat{q}), a_1) = max_{(a_2,b_2) \in A_t^2(\hat{i},\hat{q}) \times B_t^2(\hat{i},\hat{q},a_2)} \{u_t^3((\hat{i}, \hat{q}), a_1, a_2, b_2)\}$$

$$u_t^1(\hat{i}, \hat{q}) = max_{a_1 \in A_t^1(\hat{i},\hat{q})} \{u_t^2((\hat{i}, \hat{q}), a_1)\}$$

$$u_{T+1}^1(\hat{i}, \hat{q}) = \bar{r}_{T+1}(\hat{i}, \hat{q}) = e_c^c i_c^c + e_6^c i_6^c + e_{12}^c i_{12}^c + e_6^d i_6^d + e_{12}^d i_{12}^d$$

We shall denote an optimal pair of delivery actions, which are determined in the first optimality equation, as $a_3^*((\hat{i}, \hat{q}), a_1, a_2, b_2)$ and $b_3^*((\hat{i}, \hat{q}), a_1, a_2, b_2)$. Similarly, $a_2^*((\hat{i}, \hat{q}), a_1)$ and $b_2^*((\hat{i}, \hat{q}), a_1)$ represent optimal packaging actions for the six-packs and twelve-packs in the second optimality equation. Finally, $a_1^*(\hat{i}, \hat{q})$ denotes an optimal production action as determined in the third optimality equation. Let

$$w_t^3((\hat{i}, \hat{q}), \hat{a}) = r_t((\hat{i}, \hat{q}), \hat{a}) + \sum_{\hat{j}} p_t(\hat{j}|(\hat{i}, \hat{q}), \hat{a}) u_{t+1}^1(\hat{j}, \hat{q}').$$

It follows that

$$u_t^3((\hat{i}, \hat{q}), a_1, a_2, b_2) = max_{(a_3,b_3) \in A_t^3(\hat{i},\hat{q}) \times B_t^3(\hat{i},\hat{q})} \{w_t^3((\hat{i}, \hat{q}), \hat{a})\}.$$

Because the demands for the two products at the distribution center are independent, $a_3^*((\hat{i},\hat{q}),a_1,a_2,b_2)$ and $b_3^*((\hat{i},\hat{q}),a_1,a_2,b_2)$ are independent of each other. Therefore, they can be determined independently. On the other hand, $a_2^*((\hat{i},\hat{q}),a_1)$ and $b_2^*((\hat{i},\hat{q}),a_1)$ are dependent because a packaging action for the six-packs affects the set of packaging actions for the twelve-packs. Finally, optimal packaging and delivery actions affect the production decision. We shall not attempt to establish structural result for production control as there are only two possible production actions.

## *6.4 Structural Results for Delivery Control*

In this section, we show how the optimal delivery actions $a_3^*$ and $b_3^*$ vary with $i_6^d$ and $i_{12}^d$, respectively. The structural results and their sufficient conditions are presented in Subsection 6.4.1. Then, in Subsection 6.4.2, we show how the parameters of the distribution problem satisfy these sufficient conditions. Finally, we summarize the structural results for delivery control in Subsection 6.4.3.

### 6.4.1 Sufficient Conditions for the Structural Results

Proposition 6.4.1 specifies how optimal delivery actions vary with inventory levels of the products at the distribution center, given that a condition on $w_t^3((\hat{i},\hat{q}),\hat{a})$ is satisfied.

**Proposition 6.4.1.** *Assume that* $w_t^3((\hat{i},\hat{q}),\hat{a})$ *is sub-additive in* $(i_6^d,a_3)$*. Then there exists* $a_3^*((\hat{i},\hat{q}),a_1,a_2,b_2)$ *which is non-increasing in* $i_6^d$*.*

*Proof.* By assumption, for each pair of $i_6^d=i'$ and $i_6^d=i''$ such that $i''\geq i'$ and for each pair of

$$a_3',a_3''\in\{A_t^3((i_c^c,i_6^c,i_{12}^c,i',i_{12}^d),\hat{q})\cap A_t^3((i_c^c,i_6^c,i_{12}^c,i'',i_{12}^d),\hat{q})\}$$

such that $a_3''\geq a_3'$, we have that

$$w_t^3(((i_c^c,i_6^c,i_{12}^c,i',i_{12}^d),\hat{q}),a_1,a_2,b_2,a_3'',b_3)-w_t^3(((i_c^c,i_6^c,i_{12}^c,i',i_{12}^d),\hat{q}),a_1,a_2,b_2,a_3',b_3)$$
$$\geq w_t^3(((i_c^c,i_6^c,i_{12}^c,i'',i_{12}^d),\hat{q}),a_1,a_2,b_2,a_3'',b_3)-w_t^3(((i_c^c,i_6^c,i_{12}^c,i'',i_{12}^d),\hat{q}),a_1,a_2,b_2,a_3',b_3).$$

By Lemma 2.3.7,

$$min\{argmax_{a_3 \in A_t^3((i_c^c, i_6^c, i_{12}^c, i', i_{12}^d), \hat{q})} w_t^3(((i_c^c, i_6^c, i_{12}^c, i', i_{12}^d), \hat{q}), a_1, a_2, b_2, a_3, b_3)\}$$
$$\geq min\{argmax_{a_3 \in A_t^3((i_c^c, i_6^c, i_{12}^c, i'', i_{12}^d), \hat{q})} w_t^3(((i_c^c, i_6^c, i_{12}^c, i'', i_{12}^d), \hat{q}), a_1, a_2, b_2, a_3, b_3)\}.$$

We note that

$$A_t^3((i_c^c, i_6^c, i_{12}^c, i'', i_{12}^d), \hat{q}) \subseteq A_t^3((i_c^c, i_6^c, i_{12}^c, i', i_{12}^d), \hat{q}).$$

Furthermore, as aforementioned, $a_3^*((\hat{i}, \hat{q}), a_1, a_2, b_2)$ and $b_3^*((\hat{i}, \hat{q}), a_1, a_2, b_2)$ can be determined independently. It follows that the arguments in this proof hold regardless of the value of $b_3$. The assertion of the proposition follows. □

Let $\hat{j} = (j_c^c, j_6^c, j_{12}^c, j_6^d, j_{12}^d)$. Theorem 6.4.2 states the sufficient conditions for the above structural result on the reward and transition structures of the distribution problem. This result is comparable to Theorem 4.7.4 in Puterman's book.

**Theorem 6.4.2.** *Assume that the following conditions hold:*

1. $r_t((\hat{i}, \hat{q}), \hat{a})$ *is non-decreasing in* $i_6^d$*, for all* $a_3$*, for* $t = 1, 2, ..., T$*,*
2. $\sum_{k \leq j_6^d \leq M_6^d} p_t(\hat{j}|(\hat{i}, \hat{q}), \hat{a})$ *is non-decreasing in* $i_6^d$*, for all* $k$*, for all* $a_3$*, for* $t = 1, 2, ..., T$*,*
3. $\bar{r}_{T+1}(\hat{i}, \hat{q})$ *is non-decreasing in* $i_6^d$*,*
4. $r_t((\hat{i}, \hat{q}), \hat{a})$ *is sub-additive in* $(i_6^d, a_3)$*, for* $t = 1, 2, ..., T$*, and*
5. $\sum_{k \leq j_6^d \leq M_6^d} p_t(\hat{j}|(\hat{i}, \hat{q}), \hat{a})$ *is sub-additive in* $(i_6^d, a_3)$*, for all* $k$*, for* $t = 1, 2, ..., T$*.*

*Then there exists* $a_3^*((\hat{i}, \hat{q}), a_1, a_2, b_2)$ *which is non-increasing in* $i_6^d$*.*

*Proof.* First, by induction on $t$, we show that $u_t^1(\hat{j}, \hat{q}')$ is non-decreasing in $i_6^d$. By (iii), $u_{T+1}^1(\hat{i}, \hat{q})$ is non-decreasing in $i_6^d$. Assume that this is true for times $T, T-1, T-2, ..., t+1$. This with condition (ii) and Lemma 2.3.10 in Chapter 2 imply that $\sum_{\hat{j}} p_t(\hat{j}|(\hat{i}, \hat{q}), \hat{a}) u_{t+1}^1(\hat{j}, \hat{q}')$ is non-decreasing in $i_6^d$. Condition (i) then implies that $u_t^1(\hat{j}, \hat{q}')$ is non-decreasing in $i_6^d$. By this result, condition (v), and Lemma 2.3.10 in Chapter 2, we have that the second term in the definition of $w_t^3((\hat{i},\hat{q}),\hat{a})$ is sub-additive in $(i_6^d, a_3)$. This and condition (iv) then implies that $w_t^3((\hat{i},\hat{q}),\hat{a})$ is sub-additive in $(i_6^d, a_3)$. The result of the theorem follows from Proposition 6.4.1. $\square$

The following results establish monotone relations between the optimal delivery action for the twelve-packs, $b_3^*$, and inventory level of the twelve-packs at the distribution center, $i_{12}^d$. Proposition 6.4.3 and Theorem 6.4.4 are equivalent to Proposition 6.4.1 and Theorem 6.4.2, respectively. Their proofs are analogous and, thus, omitted.

**Proposition 6.4.3.** *Assume that $w_t^3((\hat{i},\hat{q}),\hat{a})$ is sub-additive in $(i_{12}^d, b_3)$. Then there exists $b_3^*((\hat{i},\hat{q}), a_1, a_2, b_2)$ which is non-increasing in $i_{12}^d$.*

**Theorem 6.4.4.** *Assume that the following conditions hold:*

1. *$r_t((\hat{i},\hat{q}),\hat{a})$ is non-decreasing in $i_{12}^d$, for all $b_3$, for $t = 1, 2, ..., T$,*
2. *$\sum_{k \le j_{12}^d \le M_{12}^d} p_t(\hat{j}|(\hat{i},\hat{q}),\hat{a})$ is non-decreasing in $i_{12}^d$, for all $k$, for all $b_3$, for $t = 1, 2, ..., T$,*
3. *$\bar{r}_{T+1}(\hat{i},\hat{q})$ is non-decreasing in $i_{12}^d$,*
4. *$r_t((\hat{i},\hat{q}),\hat{a})$ is sub-additive in $(i_{12}^d, b_3)$, for $t = 1, 2, ..., T$, and*
5. *$\sum_{k \le j_{12}^d \le M_{12}^d} p_t(\hat{j}|(\hat{i},\hat{q}),\hat{a})$ is sub-additive in $(i_{12}^d, b_3)$, for all $k$, for $t = 1, 2, ..., T$.*

*Then there exists $b_3^*((\hat{i},\hat{q}), a_1, a_2, b_2)$ which is non-increasing in $i_{12}^d$.*

### 6.4.2 Sufficient Conditions on the Problem Parameters

Here we establish the sufficient conditions on the problem parameters that imply the conditions of Theorem 6.4.2. Similar conditions for Theorem 6.4.4 can also be established and shall be omitted here. It is relatively straightforward to show that $\bar{r}_{T+1}(\hat{i},\hat{q})$ is non-decreasing in $i_6^d$. The following corollary shows that, with a condition on the demand distribution, $r_t((\hat{i},\hat{q}),\hat{a})$ is non-decreasing in $i_6^d$. Let $\bar{F}_t^6(i_6^d) = 1 - F_t^6(i_6^d)$, where $F_t^6(i_6^d)$

is the cumulative demand distribution for the six-packs at the distribution center between time $t$ and time $t+1$.

**Corollary 6.4.5.** *Assume that* $\bar{F}_t^6(M_6^d) \geq h_6^d/(w_6+s_6)$*. Then,* $r_t((\hat{i},\hat{q}),\hat{a})$ *is non-decreasing in* $i_6^d$*.*

*Proof.* Consider the expression for $r_t((\hat{i},\hat{q}),\hat{a})$. Let $f_t(i_6^d) = r_t((\hat{i},\hat{q}),\hat{a})$, with other parameters besides $i_6^d$ fixed. It can be shown that

$$f_t(i_6^d+1) - f_t(i_6^d) = (w_6+s_6)\bar{F}_t^6(i_6^d+q_6^d) - h_6^d.$$

When the condition of the corollary holds, the left hand side of the above equation is non-negative. The assertion of the proposition follows. □

The next corollary proves that condition (ii) of Theorem 6.4.2 holds.

**Corollary 6.4.6.** $\sum_{k \leq j_6^d \leq M_6^d} p_t(\hat{j}|(\hat{i},\hat{q}),\hat{a})$ *is non-decreasing in* $i_6^d$*, for all* $k$*.*

*Proof.* It can be shown that

$$\sum_{k \leq j_6^d \leq M_6^d} p_t(\hat{j}|(\hat{i},\hat{q}),\hat{a}) = \sum_{k \leq j_6^d \leq (i_6^d+q_6^d)} p_t(\hat{j}|(\hat{i},\hat{q}),\hat{a}) = F_t^6(i_6^d+q_6^d-k).$$

The assertion of the corollary follows. □

Corollary 6.4.7 and 6.4.8 prove the sub-additive property of the reward and transition structures, respectively.

**Corollary 6.4.7.** $r_t((\hat{i},\hat{q}),\hat{a})$ *is sub-additive in* $(i_6^d, a_3)$*, for* $t = 1, 2, ..., T$*.*

*Proof.* From the proof of Corollary 6.4.5, we have that

$$f_t(i_6^d+1) - f_t(i_6^d) = (w_6+s_6)\bar{F}_t^6(i_6^d+q_6^d) - h_6^d.$$

Because the above quantity is independent of $a_3$, the result follows. □

**Corollary 6.4.8.** $\sum_{k \leq j_6^d \leq M_6^d} p_t(\hat{j}|(\hat{i},\hat{q}),\hat{a})$ *is sub-additive in* $(i_6^d, a_3)$*, for all* $k$*, for* $t = 1, 2, ..., T$*.*

*Proof.* From the proof of Corollary 6.4.6, we have that

$$\sum_{k \le j_6^d \le M_6^d} p_t(\hat{j}|(\hat{i},\hat{q}),\hat{a}) = \sum_{k \le j_6^d \le (i_6^d + q_6^d)} p_t(\hat{j}|(\hat{i},\hat{q}),\hat{a}) = F_t^6(i_6^d + q_6^d - k).$$

And since $F_t^6(i_6^d + 1 + q_6^d - k) - F_t^6(i_6^d + q_6^d - k)$ is independent of $a_3$. The result follows. □

### 6.4.3 Summary of the Structural Results

In summary, the structural results for delivery control state (informally) that less units of each product are to be delivered to the distribution center as the inventory level of the product at the distribution center increases. This result applies to both the six-packs and twelve-packs. Based on the above results, theorem and Theorem summarize the structural results for delivery control for the six-packs and twelve-packs, respectively.

**Theorem 6.4.9.** *Assume that* $\bar{F}_t^6(M_6^d) \ge h_6^d/(w_6 + s_6)$*, for* $t = 1, 2, ..., T$*. Then there exists* $a_3^*((\hat{i},\hat{q}), a_1, a_2, b_2)$ *which is non-increasing in* $i_6^d$*.*

**Theorem 6.4.10.** *Assume that* $\bar{F}_t^{12}(M_{12}^d) \ge h_{12}^d/(w_{12} + s_{12})$*, for* $t = 1, 2, ..., T$*. Then there exists* $b_3^*((\hat{i},\hat{q}), a_1, a_2, b_2)$ *which is non-increasing in* $i_{12}^d$*.*

## *6.5 Conclusions and Future Research*

In this chapter, we have shown that structural results similar to those for the SVMI problem can be established in an actual distribution problem, despite the different settings. Furthermore, the resulting sufficient conditions on the problem parameters for the two problems are similar. This allows us to conclude that the approach we use in proving the structural results for the SVMI problem, especially those for the inventory control, is robust in the sense that it can be applied to different distribution problems.

Other research questions were proposed during our conversations with the logistics team at the Coca-Cola Enterprises, Inc. For instance, we proposed a study on the the effects of information sharing on the operating performance of the distribution system. In particular, under full information sharing environment, we planned to study the improvement in the operating performance of a new distribution system in which the packaging of soft-drink products is done at the distribution center (instead of the cannery). Another research question involves how the company can best increase the flexibility of its vehicle routing procedure. Because of the currently gloomy state of the economy, the company's priorities have changed and these research topics are not being attended to. In our view, these questions will be an important part of any future research on the distribution strategy of the company.

# CHAPTER VII

# A NUMERICAL STUDY OF THE PERFORMANCE MEASURES IN THE SVMI PROBLEM

In this chapter, we illustrate, by numerous examples, how some problem parameters affect the customer service level and the optimal expected total reward in the SVMI problem. These problem parameters include the revenue and costs per unit (reward parameters), demand variance, and distance from the depot. Specifically, we first show how the ratio of the sum of the revenue and penalty cost to the holding cost affects customer service level. Then we examine the effect of demand variance on the optimal expected total reward. Additionally, we study how the distance from the depot affects both the customer service level and the optimal expected total reward.

Besides profits, customer service level is very important for many supply chain vendors. In this chapter, we consider both the optimal expected total reward and customer service level as the performance measures in our numerical study. In a distribution system with stochastic demand, the system's capability to fill customer orders depends largely on the capacities of the retailers and the vehicle and on the distances between the depot and the retailers. Meanwhile, reward parameters can determine an optimal customer service level for the distribution system. Let us recall that, in the SVMI problem, the decision maker maximizes the expected total reward. In all of our numerical examples, the optimality criterion is still the optimal expected total reward. Customer service level is just another performance measure. We solved all of our numerical examples by standard backward induction algorithms.

The optimal expected total reward for a distribution system depends on the demand distribution, e.g., its variance, and the distances from the depot to the retailers. The latter determines how often the retailers can be replenished, which directly affects the optimal expected total reward. Our numerical examples will show this relationship. Additionally, we shall present the numerical results that illustrate both positive and negative correlations between demand variance and the optimal expected total reward. The positive correlation occurs when the profit margin of filling an order is very high.

This chapter is organized as follows. The four sections of numerical analysis are presented in the order described in the first paragraph above. In each of these sections, we present and discuss our numerical results. In Section 7.5, we add some remarks and summarize the numerical study.

## 7.1 The Effect of Reward Parameters on Customer Service Level

For various businesses in the retail industry, customer service level is very important, especially for highly substitutable products. In this section, we illustrate the relationship between customer service level and the reward parameters in the SVMI problem. These parameters include the holding cost, revenue, and penalty cost.

Based on the reward functions first described in Chapter 2, the vendor receives a revenue and avoids a penalty cost for filling an order that arrives at a retailer. This can be considered as the benefit of having one more unit of inventory. Meanwhile, having an additional unit of inventory incurs a holding cost. For these reasons, an appropriate parameter of interest is the ratio of the sum of revenue and penalty cost to the holding cost. We shall refer to this ratio as the margin ratio and denote it by $m_r$. That is,

$$m_r = (b^1 + b^2)/h,$$

where $b^1$ is the revenue per filled order, $b^2$ is the penalty cost per lost order, and $h$ is the holding cost per unit of inventory per period.

We define customer service level ($CSL$) as the ratio of the expected total number of customers served ($ENCS$) to the sum of $ENCS$ and the expected total number of customers lost ($ENCL$). That is,

$$CSL = ENCS/(ENCS + ENCL),$$

where $ENCS$ and $ENCL$ are computed in a similar manner to the expected total reward. That is, the computations of these parameters in the backward induction algorithm are based on the inventory levels, the inventory action, the realized demand values, and the probability of those demand values. In particular, let the demand at a (non-current) retailer $i$ between the current and next decision epochs be represented by a random variable $D_i$. Then, at retailer $i$, the number of customers served ($NCS_i$) and the number of customers lost ($NCL_i$) between the current and next decision epochs are as follows:

$$NCS_i = min\{D_i, x_i\},$$

and

$$NCL_i = max\{D_i - x_i, 0\},$$

where $x_i$ is the inventory level of retailer $i$. At the current retailer $l$,

$$NCS_l = min\{D_l, x_l + a\},$$

and

$$NCL_l = max\{D_l - (x_l + a), 0\},$$

where $a$ is the number of units of inventory added to the current retailer.

To obtain the expected total number of customer served ($ENCS$) and the expected total number of customers lost ($ENCL$), we consider the demand distributions at each retailer (to get the expected values of the numbers of customers served and lost for that retailer). Then we consider all retailers together and keep track of these parameters for the entire length of the problem horizon.

For the rest of this section, we show, by examples, how the margin ratio affects the customer service level, as the decision maker continues to maximize the expected total reward. The parameters that we use include four sets of demand distributions (four combinations of discretized uniform and normal distributions), four sets of distances between inventory locations (four combinations of one and two unit distances), three numbers of retailers ($N \in \{1, 2, 3\}$), and five margin ratios ($m_r \in \{0.15, 1.5, 15, 37.5, 75\}$). In order to study the effect of reward parameters on customer service level, we do not restrict ourselves to common values of margin ratio. Instead, we use a wide range of values of margin ratio. This idea is comparable to that for the numerical examples presented in Federgruen and Zipkin (1984b). We solved all possible combinations of these parameters and measured the customer service level.

Since the SVMI problem has finite horizon (15 time periods in our examples), the starting state affects the customer service level. Therefore, we present our numerical examples for both the empty starting state (all retailers are empty, Table 12) and full starting state (all retailers are full, Table 13). This will allow us to make conclusive statements about our numerical results. In each of the tables presented in this chapter, we define the mean as the average value of the performance measure of interest for all 240 cases solved in that table.

**Table 12:** The effect of margin ratio on customer service level (empty starting state)

| Margin ratio | Cases | Customer service level | Percent of mean |
|---|---|---|---|
| 0.15 | 48 | 0 | 0 |
| 1.5 | 48 | 0.15 | 33 |
| 15 | 48 | 0.70 | 156 |
| 37.5 | 48 | 0.70 | 156 |
| 75 | 48 | 0.70 | 156 |
| All cases | 240 | 0.45 | 100 |

In Table 12, we note that customer service level increases with the margin ratio. The customer service level of 0 for the margin ratio of 0.15 implies that no inventory is added to the retailer(s). For the margin ratios of 15, 37.5 and 75, the customer service level at

0.70 is limited by the distribution system's capability to replenish the retailer(s) and by the empty starting state. For the second margin ratio of 1.5, the customer service level of 0.15 implies that units of inventory are added to the retailer(s) in those 48 cases.

**Table 13:** The effect of margin ratio on customer service level (full starting state)

| Margin ratio | Cases | Customer service level | Percent of mean |
|---|---|---|---|
| 0.15 | 48 | 0.83 | 91 |
| 1.5 | 48 | 0.83 | 91 |
| 15 | 48 | 0.98 | 106 |
| 37.5 | 48 | 0.98 | 106 |
| 75 | 48 | 0.98 | 106 |
| All cases | 240 | 0.92 | 100 |

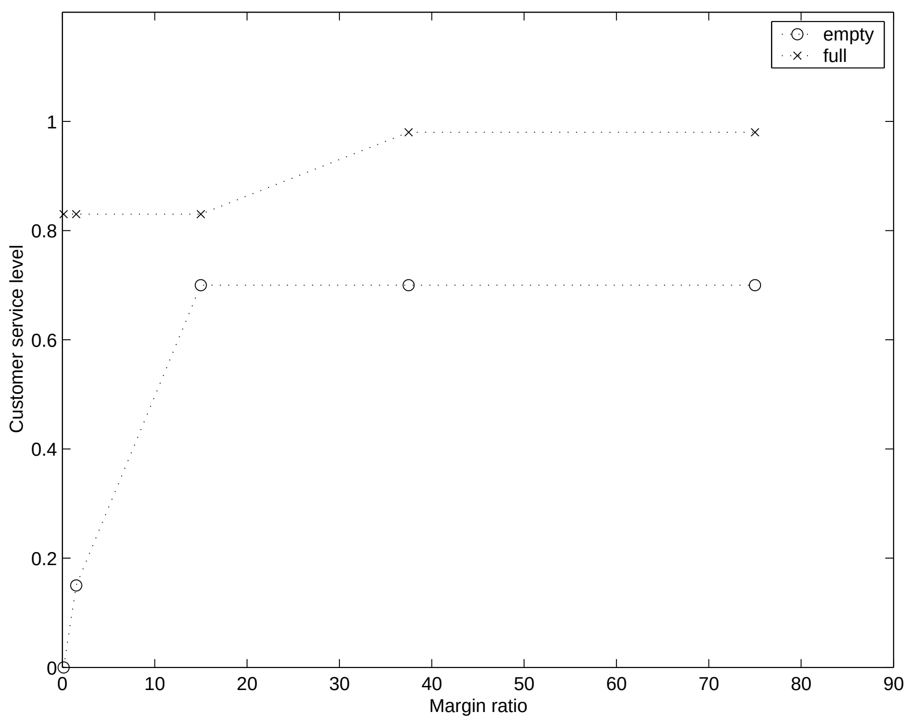


**Figure 2:** The effect of margin ratio on customer service level

The results in Table 13 are somewhat similar to those in Table 12. Here the margin ratio also has a positive correlation with customer service level. For the margin ratios of 15, 37.5, and 75, the full starting state helps improve the customer service level to 0.98.

Additionally, the full starting state also maintains the customer service level for the first two margin ratios at 0.83, even though no inventory may not have been added to the retailer(s). Figure 2 illustrates the results in Table 12 and Table 13.

## 7.2 The Effect of Demand Variance on the Optimal Expected Total Reward

In this section, we study how the variance of the demand at the retailers affects the optimal expected total reward. Three numbers of retailers are considered ($N \in \{1, 2, 3\}$). In each of the cases below, we assume that the demand distributions at the retailers are the same discretized normal distribution. Five different demand distributions are used (same mean but different variances). The mean is two units and the variances are as stated in Table 14. We state the the mean and the variance for a one-period interval, where an interval is the time between two successive decision epochs. If an interval is longer than one period, the appropriate convolution of the demand distribution is computed and applied. To expand our sample set, we consider four different sets of distances between the inventory locations and four different sets of reward parameters. As in the previous section, all possible combinations of these sets of parameters are solved. We present the numerical results in Table 14 and Figure 3.

**Table 14:** The effect of demand variance on the optimal expected total reward

| Demand variance | Cases | Optimal expected total reward | Percent of mean |
|---|---|---|---|
| 0 | 48 | 988.5 | 62 |
| 2 | 48 | 1476.7 | 92 |
| 4 | 48 | 1820.9 | 113 |
| 6 | 48 | 1884.5 | 118 |
| 8 | 48 | 1844.6 | 115 |
| All cases | 240 | 1603.04 | 100 |

For each of the above cases, we computed the optimal expected total reward by averaging the optimal expected total rewards for all possible starting states. At first, the results

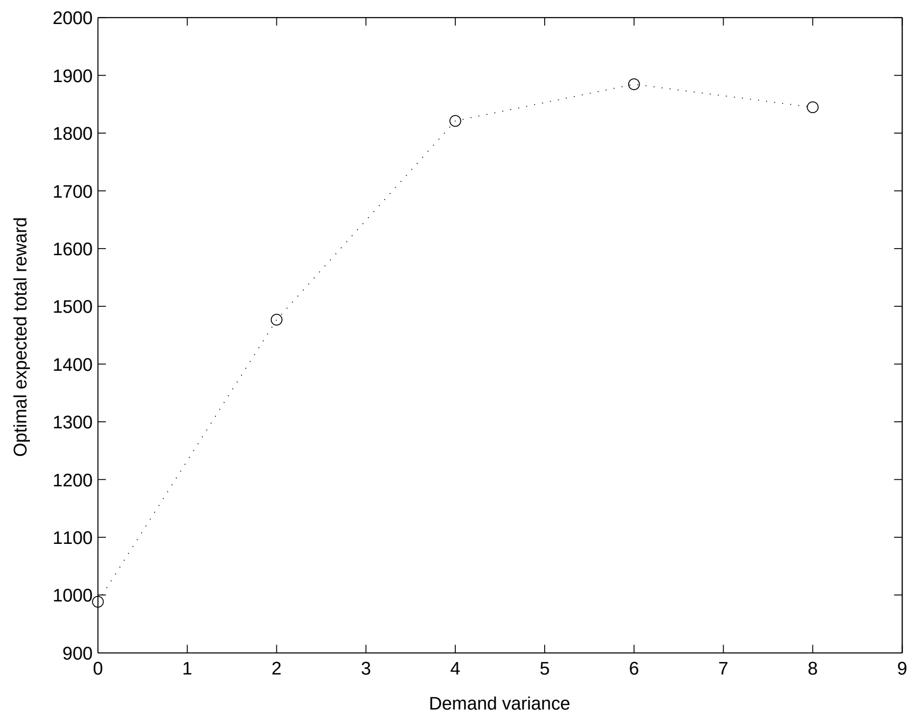


**Figure 3:** The effect of demand variance on the optimal expected total reward

in Table 14 seem surprising. In particular, for the first four variances, the optimal expected total reward increases with demand variance. In general, higher demand variance reduces the effectiveness of the decision maker in making inventory decisions. In our examples, we set the reward parameters so that the profit margin for filling an order is very high. As demand variance increases, there is a higher probability that more orders are filled and, consequently, this increases the optimal expected total reward. However, when the demand variance increases from 6 to 8, we begins to observe the negative correlation between demand variance and the optimal expected total reward that we initially expected. Some readers might wonder why our numerical results do not follow the main theoretical result in White and Harrington (1980), which states that larger demand variance reduces the optimal expected total reward. The reason is that the premise of the theoretical result does not hold in the SVMI problem. Specifically, we do not have a concave expected cost-to-go function. As a result, our numerical results are different from what the paper implies. This is what makes Table 14 interesting.

## 7.3 The Effect of Distance from the Depot on Customer Service Level

How far a retailer is from the depot affects the customer service level at the retailer. In order to study this relationship, we consider five different distances from the depot. In our sample sets, three numbers of retailers are used ($N \in \{1, 2, 3\}$). In each of the two-retailer and three-retailer cases, we assume that the retailers are at the same distance from the depot. Furthermore, our parameters include four different sets of reward parameters and four different sets of demand distributions (four combinations of discretized uniform and normal distributions). All combinations of these sets of parameters are solved. We present the numerical results in Table 15 and Figure 4.

**Table 15:** The effect of distance from the depot on customer service level

| Distance | Cases | Customer service level | Percent of mean |
|---|---|---|---|
| 1 | 48 | 0.78 | 129 |
| 2 | 48 | 0.75 | 124 |
| 3 | 48 | 0.60 | 99 |
| 4 | 48 | 0.48 | 79 |
| 5 | 48 | 0.42 | 69 |
| Total | 240 | 0.61 | 100 |

In Table 15, we observe a negative correlation between distance from the depot and customer service level. A direct result of longer distance from the depot to the retailer(s) is the longer time it takes the vehicle to travel from the depot to the retailer(s) for replenishment. Consequently, more orders are lost. This directly results in lower customer service level. We would like to point out that, in order to increase the computational efficiency, we reduced the capacities of the retailers (from 10 units to 7 units) and the vehicle (from 30 units to 21 units) in our sample problems from their original values. This resulted in the highest customer service level in Table 15 being only 0.78.

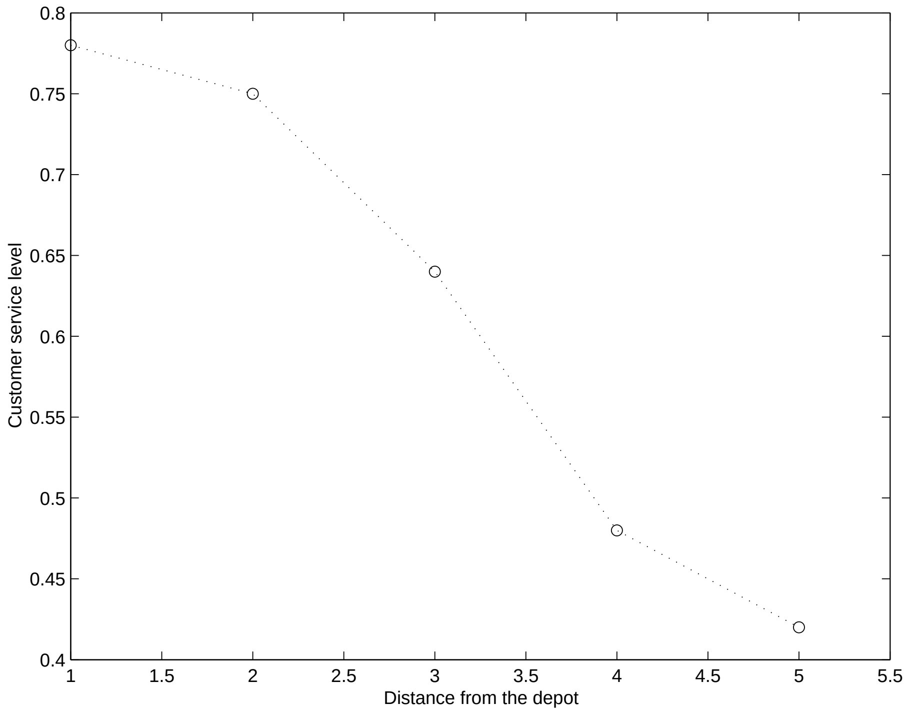


**Figure 4:** The effect of distance from the depot on customer service level

## *7.4 The Effect of Distance from the Depot on the Optimal Expected Total Reward*

In this section, we present the numerical examples that illustrate how distance from the depot affects the optimal expected total reward. We use the same sets of parameters as in the previous section. In particular, the sets of parameters that we use include three numbers of retailers ($N \in \{1, 2, 3\}$), four sets of demand distributions, four sets of reward parameters, and five distances from the depot. The results are presented in Table 16 and Figure 5.

As in Section 7.2, for each of the above cases, we computed the optimal expected total reward by averaging the optimal expected total rewards for all possible starting states. The results in Table 16 are intuitive. Long distance from the depot limits the replenishment capability of the vehicle. This results in less orders being filled and, ultimately, lower optimal expected total reward. What may be surprising is the somewhat small (percentage)

**Table 16:** The effect of distance from the depot on the optimal expected total reward

| Distance | Cases | Optimal expected total reward | Percent of mean |
|---|---|---|---|
| 1 | 48 | 49,903.81 | 109 |
| 2 | 48 | 46,576.83 | 102 |
| 3 | 48 | 45,492.65 | 99 |
| 4 | 48 | 44,033.42 | 97 |
| 5 | 48 | 42,064.25 | 93 |
| All cases | 240 | 45,614.19 | 100 |

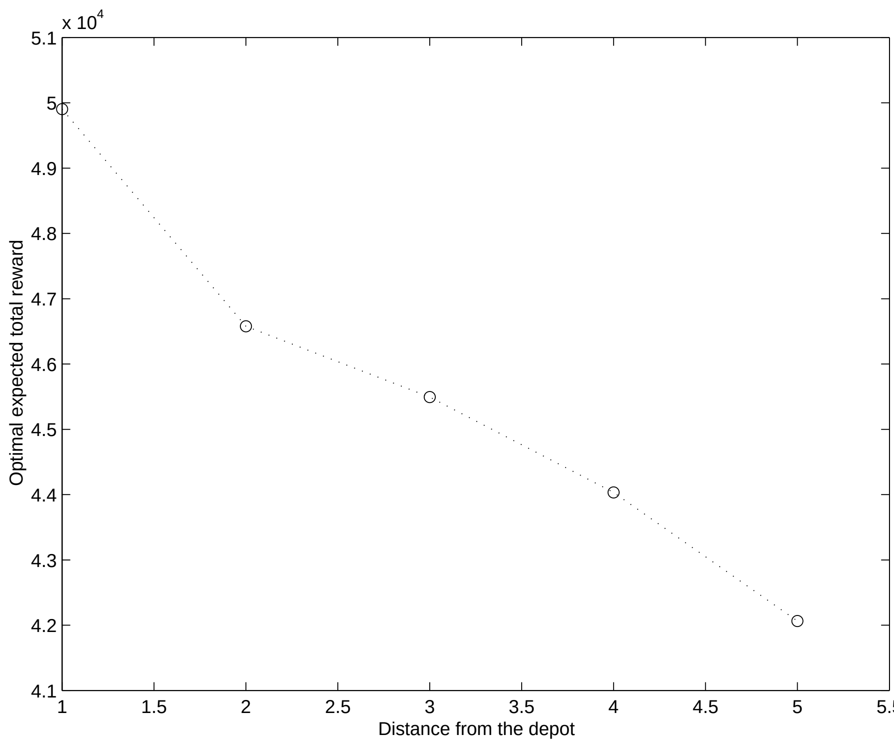


**Figure 5:** The effect of distance from the depot on the optimal expected total reward

reductions in the optimal expected total reward as the distance from the depot increases. In Table 16, as the distance from the depot doubles (from 2 to 4 units), the optimal expected total reward decreases by only 5 percentage points (from 102 to 97 percent of the mean value). This could be because the optimal inventory levels at the retailers are relatively low. Therefore, the longer distance from the depot does not severely limit the vehicle from maintaining those inventory levels at the retailers.

## 7.5 Conclusions

In general, the numerical results presented in this chapter are intuitive. Specifically, higher profit margin leads to more orders being filled and, consequently, higher customer service level. We also observed that high profit margin can lead to the positive correlation between demand variance and the optimal expected total reward. Additionally, as one would expect, longer distance from the depot reduces both customer service level and the optimal expected total reward.

We limited our numerical examples to the SVMI problems with one, two, and three retailers because of computational reasons. Four-retailer SVMI problems were solved in Chapter 2 and Chapter 4. It took almost 18 hours of CPU time to solve that problem by standard backward induction algorithm. Given the number of problems we wanted to solve in the numerical study, it is simply not practical to include larger problems. For even larger problems, there is also a coding difficulty. In particular, we used arrays to represent the set of states. As the problem gets larger, the arrays need to be bigger and this causes a memory problem in the computer.

The theoretical results that we present in earlier chapters apply to SVMI problems with any (integral) number of retailers. Some might wonder how the qualitative results of the numerical study will hold in the SVMI problems with more than three retailers. Our argument goes as follows. In any SVMI problem, all retailers are similar in the sense that they are all cost centers. The number of retailers does not change how inventory and vehicle routing decisions are made. How the reward and cost are computed also remains the same. There is no reason to expect the qualitative implications of our numerical study to change significantly as the problem gets larger.

Based on our computational experience, customer service level is not quite as interesting as we first thought. In particular, it is a matter of yes or no but not how much. That is, if servicing a customer is profitable, the decision maker will do anything to service as many customers as possible. This is the same thing as maximizing the expected total reward.

# CHAPTER VIII

# CONCLUSIONS AND FUTURE RESEARCH

We have presented our study of the stochastic vendor managed inventory (SVMI) problem, its variations, a related case study, and a numerical study. Specifically, in Chapter 2, we formulated the finite horizon SVMI problem and showed that the optimal vehicle routing and inventory actions vary monotonically with inventory levels of the retailers. These results and the algorithms were extended to the infinite horizon SVMI problem in Chapter 3. Suboptimal solutions of the SVMI problem are the topic of Chapter 4. These heuristic solution procedures include one based on the structural results for inventory control, another with base-stock inventory policy, and myopic solutions of the infinite horizon SVMI problem. In Chapter 5, we investigated four variations of the SVMI problem. Our analytical results imply that the optimal expected total reward increases as the available state information improves and/or the vehicle routing procedure becomes more flexible. In Chapter 6, we presented a case study involving the distribution problem at the Coca-Cola Enterprises, Inc. The previous chapter contains a numerical study of the performance measures in the SVMI problem. In this chapter, we first summarize important results and state our concluding remarks. Subsequently, some future research questions are proposed.

## *8.1 Conclusions*

The SVMI problem consists of a depot, multiple retailers, and a vehicle. The vehicle is used to distribute units of a product from the depot to the retailers, which face stochastic demand. Our primary objective was to study how to manage the integrated vehicle routing and inventory control problem such that the expected total reward is maximized. To do so, we formulated the SVMI problem as a finite horizon non-homogeneous Markov decision process. Then we established the structural results that would help the decision makers improve the operating performance of the distribution system. Also important is the computational advantage that these monotone relations provide. Most of the structures for the inventory control in the SVMI problem invoke relatively weak sufficient conditions on the problem parameters. Based on our computational experience, these structures help reduce the run time in solving the SVMI problem by almost $50\%$.

For the vehicle routing, we showed that the optimal destination of the vehicle (from the depot) varies monotonically with inventory levels of the retailers. In particular, the main structural result for vehicle routing states (informally) that a particular retailer continues to be preferred to other retailers as the vehicle's destination if the inventory level of that retailer decreases and/or inventory levels of the other retailers increase. This result is intuitive and its sufficient conditions require the assumption on travel times and a moderate condition on the demand distributions. In our numerical examples, applying the monotone structure in the backward induction algorithms helps reduce the run time by about one third.

Optimal inventory actions also have monotone relations with inventory levels of the retailers. In particular, as inventory levels of the non-current retailers increase, more units are to be deposited at the current retailer. This result holds with relatively weak sufficient conditions on the demand distributions. Assuming a stronger condition on the demand at the current retailer, more units are to be added to the current retailer as inventory level of the current retailer decreases. When applied simultaneously, these two results help reduce the run time by about $40\%$.

Assuming periodicity in the reward and transition structures, we formulated the infinite horizon SVMI problem as an infinite horizon periodic Markov decision process. We showed that this stochastic process has an equivalent infinite horizon stationary Markov decision process. Previously established results for the SVMI problem, plus the algorithms, were extended to the infinite horizon SVMI problem (with the expected total discounted reward criterion). We also discussed the multiperiod SVMI problem, to which most of the theoretical results for the SVMI problem directly apply.

To further study the computational requirements in solving the SVMI problem, we developed three heuristic solution procedures. First, based on the structural results for inventory control, we further assumed that the optimal inventory actions were piecewise linear in inventory levels of the retailers. The resulting algorithm maintains great solution quality and provides good computational efficiency in our numerical examples. The second heuristic solution has base-stock inventory policy. In this case, the target inventory levels are determined in a similar manner to that for the Newsboy's problem. This results in a more efficient algorithm than the previous heuristic but with less solution quality. Finally, we studied myopic solutions of the infinite horizon SVMI problem and illustrated the bias towards myopic policies as a result of the salvage value of remaining inventory at the retailers.

There are certain characteristics of the SVMI problem that may not be true in practice. First of all, current state information may not be available. For instance, it could take a significant amount of time to gather data. The result is a delay in the state information available to the decision maker. Moreover, in practice, the vehicle routes in many distribution systems are not allowed to vary as much as that in the SVMI problem. On the other hand, it is also possible that, with currently available inventory information, the vehicle can change its route after it has departed an inventory location but not yet arrived at its original destination. To capture these scenarios, we studied four variations of the SVMI problem and compared their performances.

We formulated the four variations of the SVMI problem as finite horizon non-homogeneous Markov decision processes. Then we compared the optimal expected total rewards for the four variations and the SVMI problem analytically. The results are intuitive and they imply that if the quality of state information, especially its timeliness and availability, improves and/or the vehicle routing procedure becomes more flexible, then the optimal expected total reward increases. Based on our computational experience, we proposed the hypothesis on a complementary relationship between the quality of state information and the flexibility in vehicle routing procedure towards improving the operating performance of the distribution system.

The distribution problem at the Coca-Cola Enterprises, Inc. gave us an opportunity to study an actual problem that has some similarities with the SVMI problem and its variations. In particular, this is a problem of producing soft-drink products at the cannery and distributing them to the distribution center. By a similar approach to that used in the SVMI problem, we showed that the optimal delivery actions have monotone relations with inventory levels of the products at the distribution center.

In Chapter 7, we presented our numerical study of the performance measures in the SVMI problem. In particular, by solving a large number of instances of the SVMI problem, we observed the relationship between certain problem parameters and two performance measures, which are customer service level and the optimal expected total reward. Our numerical results are generally intuitive. Precisely, we observed a positive correlation between the margin ratio and customer service level. Additionally, our numerical examples illustrate the inverse effects that distance from the depot has on customer service level and the optimal expected total reward. Meanwhile, demand variance has both positive and negative correlations with the optimal expected total reward.

## *8.2 Future Research*

For the SVMI problem, the structural results that we have established can provide the insight and intuition for decision makers and the computational benefits for problem solvers. Even though applying these results reduces the run time by about one half, more savings may be needed. We believe this can be achieved by some heuristic methods based on the structural results that we obtained. Further study is needed to determine how best to use these results to develop new solution techniques.

We can extend the SVMI problem to allow the vehicle to pick up the product at the retailers. This increases the replenishment flexibility. In many instances, tt might be less expensive to replenish inventory at a retailer with units of the product from a nearby retailer than by requiring the vehicle to return to the depot. We believe that this is an interesting extension of the SVMI problem and it may be suitable for a distribution system with multiple clusters of retailers.

Several real-world problems can be modelled as a SVMI problem or one of its variations. Other researchers will find our results helpful when they study actual distribution systems or supply chains that have similar characteristics. Future research problems include the applications that we have mentioned such as industrial gas distribution, money distribution to banks and automated teller machines, material or part handling on plant floor, and so on.

Performance comparisons between competing operating strategies are an aspect of our study that has great potential. When a company decides on its inventory information system and vehicle routing strategy, theoretical and numerical results on different operating strategies similar to ours can be very useful. Based on our numerical results, a delay in the available state information and certain restrictions on the vehicle route have strong negative impact on the expected total reward. This observation prompts further study, especially in quantifying the effects for different distribution systems.

# APPENDIX A

# THE SVMI PROBLEM WITH BACKLOGGING

In the SVMI problem that we analyzed in Chapter 2, orders that arrive at an empty store are lost. In this appendix, we shall consider a variation of the SVMI problem in which unfilled orders at each retailer are backlogged and then satisfied as units of the product become available at the retailer. Our primary objective is to identify what the sufficient conditions on the problem parameters are for the structural results for vehicle routing and inventory control. We begin by stating the following assumption on the number of orders that can be backlogged at each retailer.

**Assumption A.0.1.** *For each $i \in \{1, 2, ..., N\}$, the maximum number of orders that can be backlogged at retailer $i$ is $B_i$, where $0 < B_i < \infty$.*

This is a reasonable assumption because it is unrealistic for a retailer with finite capacity, and which is replenished by a vehicle with finite capacity, to satisfy an extremely large number of orders. Because backlogging is allowed, the inventory level at each retailer can be negative. However, from the above assumption, it follows that the state space for this variation of the SVMI problem remains finite. As a result, most of the analyses in Chapter 2 are applicable. Those that may change will be discussed later in the appendix.

Section A.1 is the problem formulation. In Section A.2 and Section A.3, we present the structural results for vehicle routing and inventory control, respectively. In the latter two sections, sufficient conditions on the problem parameters are included. In all three sections, we shall focus on parts of the analyses in Chapter 2 that change as a result of backlogging. Unless specified otherwise, the parameters that appear here are as defined as in Chapter 2.

## A.1 Problem Formulation

The state at time $t$ is $s_t = (x, x_v, l)$, where $x = (x_1, x_2, ..., x_N)$. In this case, for $i = 1, 2, ..., N$, $x_i \in X_i$, where $X_i = \{-B_i, -B_i + 1, ..., -1, 0, 1, ..., q_i\}$. By allowing unfilled orders to be backlogged, the state space becomes significantly bigger. The set of inventory actions and the set of vehicle routing actions are as previously defined.

Let us recall the following definition of the current reward function. For $l = 0$,

$$r_t((x, x_v, l), a, k) = \sum_{1 \le i \le N} r_t^i(x_i, l, k) - c_{lk},$$

and for $l > 0$,

$$r_t((x, x_v, l), a, k) = \tilde{r}_t^l(x_l, l, a, k) + \sum_{i \in K \setminus \{0, l\}} r_t^i(x_i, l, k) - c_{lk}.$$

If $x_i < 0$, then

$$r_t^i(x_i, l, k) = -b_i^2 E[D_{t,i}^{l,k}].$$

Otherwise, the original definition is true. That is, for $x_i \ge 0$,

$$r_t^i(x_i, l, k) = -h_i d_{lk} x_i + b_i^1 E[min\{D_{t,i}^{l,k}, x_i\}] - b_i^2 E[max\{0, D_{t,i}^{l,k} - x_i\}].$$

Similarly, for $\tilde{x}_l > 0$, where $\tilde{x}_l = x_l + a$,

$$\tilde{r}_t^l(x_l, l, a, k) = -h_l d_{lk} \tilde{x}_l + b_l^1 E[min\{D_{t,l}^{l,k}, \tilde{x}_l\}] - b_l^2 E[max\{0, D_{t,l}^{l,k} - \tilde{x}_l\}] - b_l^3 a.$$

If $\tilde{x} < 0$, then

$$\tilde{r}_t^l(x_l, l, a, k) = -b_l^2 E[D_{t,l}^{l,k}] - b_l^3 a.$$

The terminal reward is modified as follows:

$$\bar{r}_{T+1}(x, x_v, l) = \sum_{1 \le j \le N} (e_j - h_j \tau) max\{x_j, 0\} - c_{l0}.$$

The definition of $p_t(y|(x, x_v, l), a, k)$ remains the same. However, the state transitions change such that, given that actions $a$ and $k$ are taken at time $t$,

$$x_l(t + d_{lk}) = max\{x_l(t) + a - D_{t,l}^{l,k}, -B_l\},$$

and for $i \in K\backslash\{0, l\}$,

$$x_i(t + d_{lk}) = max\{x_i(t) - D^{l,k}_{t,i}, -B_i\}.$$

The decision rule, policy, and optimality criterion are as previously defined. We note that, since the model is Markovian and the state space and action sets are finite, we can restrict our attention to deterministic Markov policies.

## *A.2 Structural Results for Vehicle Routing*

Backlogging of unfilled orders does not affect the optimality equations for the SVMI problem. Furthermore, the theoretical results in Subsection 2.4.1 remain valid for the SVMI problem with backlogging. Finally, the sufficient conditions on the problem parameters, as presented in Subsection 2.4.2 are not significantly affected by backlogging. We shall discuss these results one by one next. Subsequently, we shall summarize the structural results for vehicle routing in the SVMI problem with backlogging.

The proof of Corollary 2.4.8 is still true for the backlogging model. Therefore, this result remains valid. Similarly, Corollary 2.4.10 still holds for the case of backlogging. Furthermore, for the SVMI problem with backlogging, it can be shown that $\bar{r}_{T+1}((x_i, x_i^c), x_v, l)$ is non-decreasing in $x_i$, for all $i \in \{1, 2, ..., N\}$.

From Proposition 2.4.9, for $x_i \geq 0$, where $i \in \{1, 2, ..., N\}$ and $i \neq l$, assume that $(b_i^1 + b_i^2)\bar{F}^{l,k}_{t,i}(q_i) \geq h_i d_{lk}$. Then, $r_t(((x_i, x_i^c), x_v, l), a, k)$ is non-decreasing in $x_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$. For $x_i \leq 0$,

$$r_t^i(x_i, l, k) = -b_i^2 E[D^{l,k}_{t,i}],$$

which is independent in $x_i$. Thus, no additional condition is required for $r_t(((x_i, x_i^c), x_v, l), a, k)$ to be non-decreasing in $x_i$. In conclusion, we may state that if $(b_i^1 + b_i^2)\bar{F}^{l,k}_{t,i}(q_i) \geq h_i d_{lk}$, then, $r_t(((x_i, x_i^c), x_v, l), a, k)$ is non-decreasing in $x_i$, for all $a \in A_t((x_i, x_i^c), x_v, l)$. Similarly, because, for $x_l < 0$,

$$\tilde{r}_t^l(x_l, l, a, k) = -b_l^2 E[D^{l,k}_{t,l}] - b_l^3 a,$$

the above arguments apply for the case $i = l$.

In conclusion, the sufficient conditions on the problem parameters remain valid in the backlogging model. We summarize the structural results for the vehicle routing in the SVMI problem with backlogging in Theorem A.2.1.

**Theorem A.2.1.** *For all $l, k \in K$, let $d_{lk} = 1$. Furthermore, for all $n \in \{1, 2, ..., N\}$, assume that*

$$(b_n^1 + b_n^2)\bar{F}_{t,n}^{l,k}(q_n) \geq h_n d_{lk},$$

*for all $l, k \in K$, and for $t = 1, 2, ..., T$. At the depot, if, for an $i \in \{1, 2, ..., N\}$, an optimal (vehicle routing) action for the state $\tilde{s}_t = ((\tilde{x}_i, \tilde{x}_i^c), x_v, l)$ is to go to retailer $i$, then an optimal action for state $s_t = ((x_i, x_i^c), x_v, l)$, in which $x_i \leq \tilde{x}_i$ and $x_j \geq \tilde{x}_j$, for all $j \in \{1, 2, ..., N\}$, $j \neq i$, is to go to retailer $i$ or the depot.*

## *A.3 Structural Results for Inventory Control*

It is relatively straight-forward to verify that the results in Subsection 2.5.1 are applicable to the SVMI problem with backlogging. This is because none of the proofs requires the inventory levels of the retailers to be non-negative. We shall consider the results in Subsection 2.5.2 as follows. Corollary 2.5.9 remains valid in this case. That is, for all $l \in \{1, 2, ..., N\}$, $r_t(((x_i, x_i^c), x_v, l), a, k)$ is super-additive in $(x_i, a)$, for all $i \in K \backslash \{0, l\}$, for all $k \in K$, and for $t = 1, 2, ..., T$. This result requires no condition on the problem parameters. The proof of Corollary 2.5.10 is true for the backlogging model. Similarly, it can be shown that Corollary 2.5.11 and Proposition 2.5.12 are still valid. Consequently, we may conclude that the sufficient conditions on the problem parameters for the structural results for inventory control are not affected by the backlogging unfilled orders.

Theorem A.3.1 and Theorem A.3.2 summarize the structural results for the SVMI problem with backlogging.

**Theorem A.3.1.** *For all* $n \in \{1, 2, ..., N\}$*, assume that*

$$(b_n^1 + b_n^2)\bar{F}_{t,n}^{l,k}(q_n) \geq h_n d_{lk},$$

*for all* $l, k \in K$*, and for* $t = 1, 2, ..., T$*. Then there exits* $a^*(k)$ *for the state* $s_t = ((x_i, x_i^c), x_v, l)$ *which is non-decreasing in* $x_i^c$*.*

**Theorem A.3.2.** *For all* $n \in \{1, 2, ..., N\}$*, assume that*

$$(b_n^1 + b_n^2)\bar{F}_{t,n}^{l,k}(q_n) \geq h_n d_{lk},$$

*for all* $l, k \in K$*, and for* $t = 1, 2, ..., T$*. Furthermore, assume that, for all* $k \in K$*, the demand at retailer* $l$ *between time* $t$ *and* $t + d_{lk}$*, that is* $D_{t,l}^{l,k}$*, has non-increasing probability mass function. Then there exists* $a^*(k)$ *for the state* $s_t = ((x_l, x_l^c), x_v, l)$ *which is non-decreasing in* $x_i^c$ *and non-increasing in* $x_l$*.*

# APPENDIX B

# APPLICABLE DEMAND PROCESSES

In this appendix, we discuss the properties of the demand processes that are sufficient for the SVMI model to be Markovian. Additionally, we refer to additional properties of the demand processes that are sufficient conditions for the structural results for vehicle routing and inventory control. The following definitions are generally well-known. We remark that Definition B.0.4 is a similar to that for a continuous time stochastic process presented in Ross (1996).

**Definition B.0.3.** *A discrete time stochastic process* $\{X(t), t = 1, 2, ...\}$ *is said to be* *Markovian* *if*

$$Pr\{X(t+1) = j | X(t) = i, X(t-1) = i_1, X(t-2) = i_2, ..., X(1) = i_{t-1}\} = P_{ij}.$$

**Definition B.0.4.** *A discrete time stochastic process* $\{Y(t), t = 1, 2, ...\}$ *is said to have* *independent increments* *if, for all* $t_0 < t_1 < ... < t_n$*, the random variables*

$$Y(t_1) - Y(t_0), Y(t_2) - Y(t_1), ..., Y(t_n) - Y(t_{n-1})$$

*are independent.*

For the SVMI problem, we assume that demand processes at the retailers are independent of each other. At each retailer, the demand process is assumed to be history-independent. For this assumption to be satisfied, it is sufficient that the stochastic process $\{Y_i(t), t = 1, 2, ...\}$, where $Y_i(t)$ is the number of orders that arrive at retailer $i$ from time $t = 1$ to time $t$, has independent increments. If the demand processes are history-independent, then it follows that the state transitions of the SVMI problem are history-independent. That is, the SVMI model is Markovian.

In Chapter 2, we establish the structural results for the SVMI problem by specifying additional properties of the demand processes that are sufficient for those structural results. In Subsection 2.4.3 (Subsection 2.5.3), we summarize the structural results for vehicle routing (the structural results for inventory control). Sufficient conditions on the demand distributions are also included in the two subsections.

# APPENDIX C

# ADDITIONAL NUMERICAL EXAMPLES

To consider the importance of Assumption 2.4.1 in the main structural result for vehicle routing, we now present additional numerical examples. Let us recall that Assumption 2.4.1 implies equal travel times between all pairs of retailers. We consider a simple instance of the SVMI problem with two retailers, each of which has five-unit capacity. The demand distributions at the two retailers are the same (time-invariant and uniform between 0 and 6). Tables 17 and 18 present the optimal vehicle's destinations for different inventory levels of the two retailers. In this case, the vehicle's current location is the depot and its inventory level is 20 units (its capacity). The decision epoch of interest occurs 10 time units before the end of the problem horizon.

For the example in Table 18, we assume that the travel time from the depot to the second retailer is two time units. Other travel times for this example and all travel times for the example in Table 17 are assumed to be one time unit. Consequently, Assumption 2.4.1 holds in the example in Table 17 but not in the example in Table 18. Both examples have the same the reward structure. We define the revenue and cost parameters such that the second retailer is significantly more profitable than the first one.

**Table 17:** The vehicle's optimal destinations from the depot for the two-retailer SVMI problem with Assumption 2.4.1

| | $x_2 = 0$ | $x_2 = 1$ | $x_2 = 2$ | $x_2 = 3$ | $x_2 = 4$ | $x_2 = 5$ |
|---|---|---|---|---|---|---|
| $x_1 = 5$ | 2 | 2 | 2 | 2 | 2 | 2 |
| $x_1 = 4$ | 2 | 2 | 2 | 2 | 2 | 2 |
| $x_1 = 3$ | 2 | 2 | 2 | 2 | 2 | 1 |
| $x_1 = 2$ | 2 | 2 | 2 | 2 | 2 | 1 |
| $x_1 = 1$ | 2 | 2 | 2 | 2 | 1 | 1 |
| $x_1 = 0$ | 2 | 2 | 2 | 2 | 1 | 1 |

**Table 18:** The vehicle's optimal destinations from the depot for the two-retailer SVMI problem without Assumption 2.4.1

| | $x_2 = 0$ | $x_2 = 1$ | $x_2 = 2$ | $x_2 = 3$ | $x_2 = 4$ | $x_2 = 5$ |
|---|---|---|---|---|---|---|
| $x_1 = 5$ | 2 | 2 | 2 | 2 | 2 | 2 |
| $x_1 = 4$ | 2 | 2 | 2 | 2 | 2 | 2 |
| $x_1 = 3$ | 2 | 2 | 2 | 2 | 2 | 2 |
| $x_1 = 2$ | 2 | 2 | 2 | 2 | 2 | 2 |
| $x_1 = 1$ | 1 | 1 | 1 | 1 | 1 | 2 |
| $x_1 = 0$ | 1 | 1 | 1 | 1 | 1 | 1 |

In Table 17, it is easy to verify that the structural result for vehicle routing holds. In particular, each retailer continues to be an optimal destination for the vehicle as its inventory level decreases and/or inventory level of the other retailer increases. Meanwhile, there is a contradiction to the structural result in Table 18. In particular, the optimal destination of the vehicle changes from retailer 1 to retailer 2, as inventory level of retailer 2 increases from 4 to 5 units and inventory level of retailer 1 stays at 1 unit. This contradiction suggests the importance of Assumption 2.4.1 in the structural result for vehicle routing.

# APPENDIX D

# PARAMETERS FOR THE NUMERICAL EXAMPLES

We have used numerous numerical examples in this work to support and extend our theoretical results. In some cases, the numerical results are significant by themselves. For reference, the parameters that were used in our numerical examples are summarized here. The parameters for the numerical results in Chapter 2 (and Appendix C), Chapter 4, Chapter 5, and Chapter 7 are presented in Section D.1, Section D.2, Section D.3, and Section D.4, respectively. Unless stated otherwise, the parameters in this appendix are defined as in the chapters that they first appeared.

## D.1 Parameters for the Numerical Examples in Chapter 2 and Appendix C

In Table 1 and Table 2, we illustrated the structural results for vehicle routing and inventory control for the SVMI problem. The following parameters were used:

1. Number of retailer is 2.

2. Demand distribution is the following (the discrete version of an exponential distribution):
   $P[0] = 0.3, P[1] = 0.2, P[2] = P[3] = P[4] = P[5] = 0.1, P[6] = P[7] = ... = P[15] = 0.01$.

3. Capacities of the retailers and the vehicle are 10 units and 20 units, respectively.

4. Time horizon length is 15 periods.

5. The travel distances (or travel times) between all pairs of inventory locations are 1 unit.

6. The reward parameters are as follows: $e_j = 30$, $h_j = 0.5$, $b_1^1 = 90$, $b_2^1 = 200$, $b_j^2 = 10$, and $b_j^3 = 5$, for $j \in \{1, 2\}$.

In Table 3, Table 4, and Table 5, we showed how the structural results for the SVMI problem help improve the computational efficiency. These numerical examples share the following parameters:

1. Capacities of the retailers and the vehicle are 5 units and $5 * N$ units, respectively. ($N$ denotes the number of retailers in each example.)

2. Demand distribution is the discrete version of $U[0, 6]$.

3. Time horizon length is 15 periods.

4. The travel distances (or travel times) between all pairs of inventory locations are 1 unit.

5. The reward parameters are as follows: $e_j = 30$, $h_j = 0.5$, $b_j^1 = 100$, $b_j^2 = 10$, and $b_j^3 = 5$, for $j \in \{1, 2, ..., N\}$.

Other parameters for the numerical examples in each table are explicitly stated in that table.

Table 17 and Table 18, in Appendix C, present an example of the case when the structural result for vehicle routing does not hold, when the assumption on travel distances between inventory locations is not satisfied. We used the following parameters in the example:

1. Number of retailers is 2.

2. Demand distribution is the discrete version of $U[0, 6]$.

3. Capacities of the retailers and the vehicle are 5 units and 20 units, respectively.

4. Time horizon length is 15 periods.

5. The travel distances (or travel times) between the inventory locations are as stated in Appendix C.

6. The reward parameters are as follows: $e_j = 30$, $h_j = 0.5$, $b_1^1 = 90$, $b_2^1 = 200$, $b_j^2 = 10$, and $b_j^3 = 5$, for $j \in \{1, 2\}$.

## D.2 Parameters for the Numerical Examples in Chapter 4

In Chapter 4, we present some heuristics for solving the SVMI problem. For Table 6 and Table 7, which show the computational efficiency of the first heuristic, the following parameters were used in Problem I:

1. Number of retailers is 4.

2. Capacities of the retailers and the vehicle are 10 units and 24 units, respectively.

3. Demand distribution is the discrete version of $U[0, 11]$.

4. The travel distances (or travel times) between all pairs of inventory locations are 1 unit.

5. Time horizon length is 15 units.

6. The reward parameters are as follows: $e_j = 15$, $h_j = 1$, $b_j^1 = 24$, $b_j^2 = 2$, and $b_j^3 = 5$, for $j \in \{1, 2, 3\}$.

In Problem II, the parameters are as follows:

1. Number of retailers is 3.

2. Capacities of the retailers and the vehicle are 10 units and 30 units, respectively.

3. Demand distribution is the discrete version of $U[0, 11]$.

4. The travel distances (or travel times) between all pairs of inventory locations are 1 unit.

5. Time horizon length is 15 units.

6. The reward parameters are as follows: $e_j = 15$, $h_j = 1$, $b_j^1 = 24$, $b_j^2 = 2$, and $b_j^3 = 5$, for $j = \{1, 2, 3\}$.

The numerical results in Table 8 involves solving Problem II, with the above parameters by the second heuristic. The following parameters were used for Problem III in Table 9:

1. Number of retailers is 2.

2. Capacities of the retailers and the vehicle are 10 units and 20 units, respectively.

3. Demand distribution is the discrete version of $U[0, 11]$.

4. The travel distances (or travel times) between all pairs of inventory locations are 1 unit.

5. Time horizon length is 15 units.

6. The reward parameters are as follows: $e_j = 15$, $h_j = 1$, $b_j^1 = 24$, $b_j^2 = 2$, and $b_j^3 = 5$, for $j \in \{1, 2, 3\}$.

In Table 10, we solved the infinite horizon SVMI problem. The parameters for this problem are as follows:

1. Number of retailers is 2.

2. Capacities of the retailers and the vehicle are 5 units and 10 units, respectively.

3. The travel distances (or travel times) between all pairs of inventory locations are 1 unit.

4. Demand distribution is the discrete version of $U(0, 6)$.

5. The reward parameters are as follows: $c_j = b_j^2 = 10$, $h_j = 5$, $b_j^1 = 100$, and $b_j^3 = 50$, for $j \in \{1, 2\}$. (The unit salvage values at the two retailers are as stated in Table 10.)

## D.3 Parameters for the Numerical Examples in Chapter 5

In Table 11, we compare the optimal expected total rewards of different variations of the SVMI problem. The following parameters were used in these numerical results:

1. Number of retailers is 3.
2. Capacities of the retailers and the vehicle are 5 units and 20 units, respectively.
3. Demand distribution is the discrete version of $U[0, 7]$.
4. The travel distances (or travel times) between all pairs of inventory locations are 1 unit.
5. Time horizon length is 15 units.
6. The reward parameters are as follows: $e_j = 30$, $h_j = 0.5$, $b_j^1 = 100$, $b_j^2 = 10$, and $b_j^3 = 5$, for $j \in \{1, 2, 3\}$.

## D.4 Parameters for the Numerical Examples in Chapter 7

The numerical examples in this chapter share some of the parameters. In particular, the capacities of the retailers and the vehicle are 10 units and $N * 10$ units, respectively. Note that $N$ is the number of retailers in each example. We use the time horizon length of 15 periods in each of our numerical examples.

Table 12 and Table 13 present the effect of margin ratio on customer service level. The parameters used in these two tables are as follows:

1. Numbers of retailers are 1,2, and 3.

2. Sets of distances (or travel times) between inventory locations are $\{d_{ij} = 1, i, j \in 0, 1, 2, 3\}, \{d_{00} = 1, d_{01} = 2, d_{02} = 1, d_{03} = 1, d_{11} = 1, d_{12} = 1, d_{13} = 1, d_{22} = 1, d_{23} = 1, d_{33} = 1\}, \{d_{00} = 1, d_{01} = 2, d_{02} = 2, d_{03} = 1, d_{11} = 1, d_{12} = 1, d_{13} = 1, d_{22} = 1, d_{23} = 1, d_{33} = 1\}$, and $\{d_{00} = 1, d_{01} = 2, d_{02} = 2, d_{03} = 1, d_{11} = 1, d_{12} = 2, d_{13} = 1, d_{22} = 1, d_{23} = 1, d_{33} = 1\}$. Note that, for $l, k \in \{0, 1, 2, 3\}$, $d_{lk} = d_{kl}$.

3. Sets of demand distributions are (the discrete versions of), where $Q_i$ denotes the demand distribution at retailer $i$, $\{Q_1 = U[0, 1{*}d], Q_2 = U[0, 1{*}d], Q_3 = U[0, 1{*}d]\}$, $\{Q_1 = N[0.5 * d, 0.5 * d], Q_2 = N[0.5 * d, 0.5 * d], Q_3 = N[0.5 * d, 0.5 * d]\}$, $\{Q_1 = U[0, 1 * d], Q_2 = N[0.5 * d, 0.5 * d], Q_3 = U[0, 1 * d]\}$ and $\{Q_1 = N[0.5 * d, 0.5 * d], Q_2 = U[0, 1 * d], Q_3 = N[0.5 * d, 0.5 * d]\}$. Here $d$ denotes the travel time between the current and next inventory location of the vehicle.

4. The margin ratios are 0.15 ($b_j^1 = 2, b_j^2 = 1$), 1.5 ($b_j^1 = 20, b_j^2 = 10$), 15 ($b_j^1 = 200, b_j^2 = 100$), 37.5 ($b_j^1 = 500, b_j^2 = 250$), and 75 ($b_j^1 = 1000, b_j^2 = 500$). Other reward parameters are $e_j = 5$, $h_j = 20$, and $b_j^3 = 7$, for $j \in \{1, 2, 3\}$.

In Table 14, the following parameters were used to show the effect of demand variance on the optimal expected total reward:

1. Numbers of retailers are 1, 2, and 3.

2. Sets of distance (or travel times) between inventory locations are $\{d_{ij} = 1, i, j \in 0, 1, 2, 3\}, \{d_{00} = 1, d_{01} = 2, d_{02} = 1, d_{03} = 1, d_{11} = 1, d_{12} = 1, d_{13} = 1, d_{22} = 1, d_{23} = 1, d_{33} = 1\}, \{d_{00} = 1, d_{01} = 2, d_{02} = 2, d_{03} = 1, d_{11} = 1, d_{12} = 1, d_{13} = 1, d_{22} = 1, d_{23} = 1, d_{33} = 1\}$, and $\{d_{00} = 1, d_{01} = 2, d_{02} = 2, d_{03} = 1, d_{11} = 1, d_{12} = 2, d_{13} = 1, d_{22} = 1, d_{23} = 1, d_{33} = 1\}$. Note that, for $l, k \in \{0, 1, 2, 3\}$, $d_{lk} = d_{kl}$.

3. Sets of reward parameters are $\{b_j^1 = 10, b_j^2 = 5, b_j^3 = 7, e_j = 5, h_j = 1\}$, $\{b_j^1 =$

$50, b_j^2 = 5, b_j^3 = 7, e_j = 5, h_j = 1\}$, $\{b_j^1 = 100, b_j^2 = 10, b_j^3 = 7, e_j = 5, h_j = 1\}$, and $\{b_j^1 = 100, b_j^2 = 50, b_j^3 = 7, e_j = 5, h_j = 1\}$, for $j \in \{1, 2, 3\}$.

4. Demand distributions are the discrete versions of $N[1*d, 0*d]$, $N[1*d, 1*d]$, $N[1*d, 2*d]$, $N[1*d, 3*d]$, and $N[1*d, 4*d]$. Here $d$ denotes the travel time between the current and next inventory location of the vehicle.

Table 15 and Table 16 show the effect of distance from the depot on customer service level and on the optimal expected total reward, respectively. The parameters for those numerical results are as follows:

1. Numbers of retailers are 1, 2, and 3.

2. Sets of demand distributions are (the discrete versions of), where $Q_i$ denotes the demand distribution at retailer $i$, $\{Q_1 = U[0, 1*d], Q_2 = U[0, 1*d], Q_3 = U[0, 1*d]\}$, $\{Q_1 = N[0.5*d, 0.5*d], Q_2 = N[0.5*d, 0.5*d], Q_3 = N[0.5*d, 0.5*d]\}$, $\{Q_1 = U[0, 1*d], Q_2 = N[0.5*d, 0.5*d], Q_3 = U[0, 1*d]\}$ and $\{Q_1 = N[0.5*d, 0.5*d], Q_2 = U[0, 1*d], Q_3 = N[0.5*d, 0.5*d]\}$. Here $d$ denotes the travel time between the current and next inventory location of the vehicle.

3. Sets of reward parameters are $\{b_j^1 = 200, b_j^2 = 100, b_j^3 = 7, e_j = 5, h_j = 1\}$, $\{b_j^1 = 250, b_j^2 = 100, b_j^3 = 7, e_j = 5, h_j = 1\}$, $\{b_j^1 = 500, b_j^2 = 200, b_j^3 = 7, e_j = 5, h_j = 1\}$, and $\{b_j^1 = 500, b_j^2 = 250, b_j^3 = 7, e_j = 5, h_j = 1\}$, for $j \in \{1, 2, 3\}$.

4. Sets of distances (or travel times) between inventory locations are $\{d_{ij} = 1, i, j \in 0, 1, 2, 3\}$, $\{d_{00} = 1, d_{01} = 2, d_{02} = 2, d_{03} = 2, d_{11} = 1, d_{12} = 1, d_{13} = 1, d_{22} = 1, d_{23} = 1, d_{33} = 1\}$, $\{d_{00} = 1, d_{01} = 3, d_{02} = 3, d_{03} = 3, d_{11} = 1, d_{12} = 1, d_{13} = 1, d_{22} = 1, d_{23} = 1, d_{33} = 1\}$, $\{d_{00} = 1, d_{01} = 4, d_{02} = 4, d_{03} = 4, d_{11} = 1, d_{12} = 1, d_{13} = 1, d_{22} = 1, d_{23} = 1, d_{33} = 1\}$, and $\{d_{00} = 1, d_{01} = 5, d_{02} = 5, d_{03} = 5, d_{11} = 1, d_{12} = 1, d_{13} = 1, d_{22} = 1, d_{23} = 1, d_{33} = 1\}$. Note that, for $l, k \in \{0, 1, 2, 3\}$, $d_{lk} = d_{kl}$.

# REFERENCES


[1] BALINSKI, M. L., "A Competitive (Dual) Simplex Method for the Assignment Problem," *Mathematical Programming*, vol. 34, pp. 125–141, 1986.

[2] BERTAZZI, L. and SPERANZA, M. G., "Models and Algorithms for the Minimization of Inventory and Transportation Costs: A Survey," in *New Trends in Distribution Logistics* (STAHLY, P., ed.), pp. 137–157, New York, NY: Springer, 1999.

[3] BERTSIMAS, D. J., "A Vehicle Routing Problem with Stochastic Demand," *Operations Research*, vol. 40, pp. 574–585, 1992.

[4] BERTSIMAS, D. J., CHERVI, P., and PETERSON, M., "Computational Approaches to Stochastic Vehicle Routing," *Transportation Science*, vol. 29, pp. 342–352, 1995.

[5] CACHON, G. P. and FISHER, M., "Supply Chain Inventory Management and the Value of Shared Information," *Management Science*, vol. 46, pp. 1032–1048, 2000.

[6] CETINKAYA, S. and LEE, C. Y., "Stock Replenishment and Shipment Scheduling for Vendor-Managed Inventory Systems," *Management Science*, vol. 46, pp. 217–232, 2000.

[7] CHIEN, T., BALAKRISNAN, A., and WONG, R., "An Integrated Inventory Allocation and Vehicle Routing Problem," *Transportation Science*, vol. 23, pp. 67–76, 1989.

[8] CHRISTOFIDES, N., "Vehicle Routing," in *The Traveling Salesman Problem. A Guide Tour of Combinatorial Optimization* (LAWLER, E. L., LENSTRA, J. K., KAN, A. R., and SHMOYS, D. B., eds.), pp. 431–448, Chichester, U.K.: Wiley, 1985.

[9] DROR, M. and BALL, M., "Inventory/Routing: Reduction from an Annual to a Short Period Problem," *Naval Research Logistics Quarterly*, vol. 34, pp. 891–905, 1987.

[10] FEDERGRUEN, A. and ZIPKIN, P., "A Combined Vehicle Routing and Inventory Allocation Problem," *Operations Research*, vol. 32, pp. 1019–1037, 1984.

[11] FEDERGRUEN, A. and ZIPKIN, P., "An Efficient Algorithm for Computing Optimal (s,S) Policies," *Operations Research*, vol. 32, pp. 1268–1285, 1984.

[12] GAVIRNENI, S., KAPUSCINSKI, R., and TAYER, S., "Value of Information in Capacitated Supply Chains," *Management Science*, vol. 45, pp. 16–24, 1999.

[13] GENDREAU, M., LAPORTE, G., and SEGUIN, R., "An Exact Algorithm for the Vehicle Routing Problem with Stochastic Demands and Customers," *Transportation Science*, vol. 29, pp. 143–155, 1995.

[14] GENDREAU, M., LAPORTE, G., and SEGUIN, R., "A Tabu Search Heuristic for the Vehicle Routing Problem with Stochastic Demands and Customers," *Operations Research*, vol. 44, pp. 469–477, 1996.

[15] GILLIES, S., O'BRIEN, P., ROBERTS, N., SMART, J., and STOVER, L., *Coca-Cola Enterprises: Full Service Business Strategy*. Project report, School of Industrial and Systems Engineering, Georgia Instititute of Technology, Atlanta, GA, 1997.

[16] GOLDEN, B. L. and ASSAD, A. A., *Vehicle Routing: Methods and Studies*. Amsterdam: North-Holland, 1988.

[17] GOLDFARB, D., "Efficient Dual Simplex Algorithms for the Assignment Problem," *Mathematical Programming*, vol. 33, pp. 187–203, 1985.

[18] HAMMOND, J., "Quick Response in Retail/Manufacturing Channels," in *Globalization, Technology, and Competition: The Fusion of Computers and Telecommunications in the 1990's* (BRADLEY, S., HAUSMAN, J., and NOLAN, R., eds.), pp. 185–214, Boston, MA: Harvard Business School Press, 1993.

[19] KLEYWEGT, A. J., NORI, V., and SAVELSBERGH, M. W. P., "The Stochastic Vehicle Routing Problem with Direct Deliveries," *Transportation Science*, vol. 36, pp. 94–118, 2002.

[20] LAPORTE, G., "The Vehicle Routing Problem: An Overview of Exact and Approximate Algorithms," *European Journal of Operational Research*, vol. 59, pp. 345–358, 1992.

[21] LEE, H. L., SO, K. C., and TANG, C. S., "The Value of Information Sharing in a Two-Level Supply Chain," *Management Science*, vol. 46, pp. 626–643, 2000.

[22] PUTERMAN, M. L., *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY: Wiley, 1994.

[23] RIDDER, A., LAAN, E. V. D., and SALOMON, M., "How Larger Demand Variability May Lead to Lower Costs in the Newsvendor Problem," *Operations Research*, vol. 46, pp. 934–936, 1998.

[24] ROSS, S. M., *Stochastic Processes*. New York, NY: Wiley, 1996.

[25] SERFOZO, R. F., "Monotone Optimal Policies for Markov Decision Processes," *Mathematical Programming Study*, vol. 6, pp. 202–215, 1976.

[26] SRINIVASAN, K., KEKRE, S., and MUKHOPADHYAY, T., "Impact of EDI on JIT Shipments," *Management Science*, vol. 40, pp. 83–90, 1994.

[27] TOPKIS, D. M., *Supermodularity and Complementarity*. Princeton, NJ: Princeton University Press, 1998.

[28] WHITE, C. C. and WHITE, D. J., "Markov Decision Processes," *European Journal of Operational Research*, vol. 39, pp. 1–16, 1989.

[29] WHITE, C. and HARRINGTON, D. P., "Application of Jensen's Inequality to Adaptive Suboptimal Design," *Journal of Optimization Theory and Applications*, vol. 32, pp. 89–99, 1980.

[30] YANG, W. H., MATHER, K., and BALLOU, R., "Stochastic Vehicle Routing Problem with Restocking," *Transportation Science*, vol. 34, pp. 99–112, 2000.