

STOCHASTIC DYNAMIC DEMAND INVENTORY MODELS WITH EXPLICIT
TRANSPORTATION COSTS AND DECISIONS

A Dissertation

by

LIQING ZHANG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2011

Major Subject: Industrial Engineering

STOCHASTIC DYNAMIC DEMAND INVENTORY MODELS WITH EXPLICIT
TRANSPORTATION COSTS AND DECISIONS

A Dissertation

by

LIQING ZHANG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Sıla Çetinkaya
Committee Members,	Martin A. Wortman
	Eylem Tekin
	E. Powell Robinson
Head of Department,	Brett Peters

August 2011

Major Subject: Industrial Engineering

ABSTRACT

Stochastic Dynamic Demand Inventory Models with Explicit Transportation Costs
and Decisions. (August 2011)

Liqing Zhang, B.S.; M.S., Tsinghua University, P.R. China

Chair of Advisory Committee: Dr. Sila Çetinkaya

Recent supply chain literature and practice recognize that significant cost savings can be achieved by coordinating inventory and transportation decisions. Although the existing literature on analytical models for these decisions is very broad, there are still some challenging issues. In particular, the uncertainty of demand in a dynamic system and the structure of various practical transportation cost functions remain unexplored in detail. Taking these motivations into account, this dissertation focuses on the analytical investigation of the impact of transportation-related costs and practices on inventory decisions, as well as the integrated inventory and transportation decisions, under stochastic dynamic demand.

Considering complicated, yet realistic, transportation-related costs and practices, we develop and solve three classes of models: (1) Pure inbound inventory model impacted by transportation cost; (2) Pure outbound transportation models concerning shipment consolidation strategy; (3) Integrated inbound inventory and outbound transportation models. In broad terms, we investigate the modeling framework of vendor-customer systems for integrated inventory and transportation decisions, and we identify the optimal inbound and outbound policies for stochastic dynamic supply chain systems.

This dissertation contributes to the previous literature by exploring the impact of realistic transportation costs and practices on stochastic dynamic supply chain

systems while identifying the structural properties of the corresponding optimal inventory and/or transportation policies. Placing an emphasis on the cases of stochastic demand and dynamic planning, this research has roots in applied probability, optimal control, and stochastic dynamic programming.

To my parents and my husband

ACKNOWLEDGMENTS

I would never have been able to finish my dissertation without the guidance of my committee members, help from friends, and support from my family.

I owe my deepest appreciation to my advisor, Prof. Sıla Çetinkaya, for her excellent supervision, constant encouragement and persistent help in my research program. It is a great fortune for me to have the opportunity to work with Dr. Çetinkaya and I really enjoy the wonderful time in her group. I am deeply grateful to my committee members, Prof. Martin A. Wortman, Prof. Eylem Tekin and Prof. E. Powell Robinson, for their guidance and support throughout the course of this research. In addition, I would like to extend my gratitude to Prof. Halit Üster for his interest and suggestions on my research.

I would like to express my gratitude to my friends and colleagues, Hui Lin, Panitan Kewcharoenwong, Su Zhao, Abhilasha Katariya, Xinghua Wang, Yi Zhang, Bo Wei, Nikhil Kondati, Sung Ook Hwang, Gopalakrishnan Easwaran, Burcu Keskin, Ezgi Eren, Homarjun Agrahari, and Joaquin Torres. They generously share their knowledge and experience, and constantly provide valuable ideas and comments.

Finally, I would like to thank my parents, Hongkun Zhang and Hexiang Du for their understanding and support throughout the past years of my education. I would also like to thank my husband, Yunfei Chu. He is always there listening to me, cheering me up and offering unfailing support whenever I need it. Words are not enough to express how important and helpful he has been to me.

TABLE OF CONTENTS

CHAPTER		Page
I	INTRODUCTION	1
	I.1. Scope of the Dissertation	7
	I.1.1. Pure Inbound Inventory Model	8
	I.1.2. Pure Outbound Transportation Models	10
	I.1.3. Integrated Inbound Inventory and Outbound Transportation Model	13
	I.2. Organization of the Dissertation	14
II	LITERATURE REVIEW	15
	II.1. Integration of Inventory and Transportation Policies . . .	17
	II.1.1. Shipment Consolidation Policy	18
	II.1.2. Inventory Policy Considering Transportation Cost	21
	II.1.3. Integrated Inventory and Transportation Policy . .	25
	II.2. Stochastic Dynamic Inventory Control Problems	29
III	STOCHASTIC DYNAMIC INVENTORY PROBLEM UN- DER EXPLICIT INBOUND TRANSPORTATION COST AND CAPACITY	33
	III.1. Notation and Problem Formulation	35
	III.2. New Concepts and Basic Properties	39
	III.3. Model Analysis	42
	III.3.1. Optimal Ordering Policy	43
	III.3.2. Special Case	64
	III.4. Numerical Analysis	71
	III.4.1. Factorial Design Experiment	72
	III.4.2. Impact of System Parameters on Policy Values . .	73
	III.5. Summary	77
IV	EXACT MODELS AND OPTIMAL POLICIES FOR SHIP- MENT CONSOLIDATION: A STOCHASTIC DYNAMIC PROGRAMMING APPROACH	79
	IV.1. Common System Settings and Modeling Assumptions . .	81
	IV.2. Private Fleet Transportation without Cargo Capacity . .	83
	IV.2.1. Problem Formulation	83

CHAPTER	Page
IV.2.2. Exact Optimal Policy	86
IV.3. Single-Truck Transportation with Cargo Capacity and Fixed Cost	92
IV.3.1. Problem Formulation	93
IV.3.2. Exact Optimal Policy	95
IV.4. Common Carriage Transportation	100
IV.4.1. Problem Formulation	100
IV.4.2. Analysis of the Optimal Policies	102
IV.4.3. Computational Studies	120
IV.5. Multi-Truck Transportation with Cargo Capacity	123
IV.5.1. Problem Formulation	123
IV.5.2. Analysis of the Optimal Policy	125
IV.5.3. Single Period Problem	133
IV.6. Summary	136
 V THE VENDOR'S OPTIMAL STOCK REPLENISHMENT AND SHIPMENT SCHEDULING POLICY UNDER TEM- PORAL SHIPMENT CONSOLIDATION	 138
V.1. Notation and Problem Formulation	143
V.2. A "Single-Period" Problem	148
V.2.1. Properties of the Terminal Cost	149
V.2.2. Optimal Joint Decision for Period n	152
V.3. Optimal Joint Policy for Multi-Period Problems	170
V.4. Summary	180
 VI CONCLUSIONS AND FUTURE DIRECTIONS	 181
VI.1. Contributions	181
VI.2. Future Work	184
 REFERENCES	 187
 VITA	 200

LIST OF TABLES

TABLE		Page
1	The Cost of the Business Logistics System in Relation to Gross Domestic Product (in \$ Billion)	2
2	Data Set for Factorial Design Experiment	72
3	Data Set for System Parameter's Impact	74

LIST OF FIGURES

FIGURE	Page
1	The Generalized Replenishment Cost Function to Consider Cargo Cost and Capacity 34
2	Problem Setting of the Inbound Replenishment System 36
3	An illustration of value determination 46
4	An Illustration of (Q, \vec{s}, \vec{S}) Policy $\left(m = \left\lfloor \frac{s_n^1 - Q_n}{C} \right\rfloor + 1\right)$ 52
5	Optimal Policies for Special Case 71
6	Influence of Parameter K 74
7	Influence of Parameter C 75
8	Influence of Parameter Δ 76
9	Influence of Parameter h 76
10	Influence of Parameter p 77
11	System Settings of the Outbound Consolidation System 81
12	Private Fleet Transportation Cost without Cargo Capacity 84
13	Single-Truck Transportation Cost with Cargo Capacity 93
14	Common Carriage Transportation Cost 101
15	Example 1 of Optimal Policies for Common Carriage Transportation 122
16	Example 2 of Optimal Policies for Common Carriage Transportation 122
17	Multi-Truck Transportation Cost with Cargo Capacity 124
18	Problem Setting of the VMI System 145
19	Illustration of the Optimal Joint Policy 169

FIGURE	Page
20	A Realization of the Process 178

CHAPTER I

INTRODUCTION

It is believed that quality supply chain management keeps the company ahead of his competitors in the highly competitive world of today. To satisfy the customer demands in a timely and cost effective way, supply chain management takes into account every entity that has an impact on cost and plays a role in making the product conform to customer requirements. Indeed, total system-wide costs, from purchasing raw materials, producing items, holding inventory, to distributing finished goods, should be minimized. Therefore, the concentration is not on simply improving production planning, reducing inventories or minimizing transportation cost, but rather, on taking a systems approach of optimization (Simchi-Levi et al., 2007).

Improvement in supply chain can be particularly realized through coordinating its two main activities: inventory and transportation. The data from the 19th annual State of Logistics Report (Wilson, 2008), sponsored by the Council of Supply Chain Management Professionals, suggests that the cost of the U.S. business logistics system has continued to increase during the last decade, and it climbed to \$1.397 trillion in 2007, which doubles 1990's total logistics cost. Table 1 gives the growth in total logistics cost and its components in relation to Gross Domestic Product (GDP). From the historical data, it can be found that the combination of transportation costs and inventory costs consistently account for more than 96% of the total logistics costs, as well as around 9% of the U.S. GDP. As a result, substantial savings can be achieved through better system-wide optimization.

Coordination and integration of the inventory and transportation operations be-

This dissertation follows the style and format of *Operations Research*.

Table 1: The Cost of the Business Logistics System in Relation to Gross Domestic Product (in \$ Billion)

Year	Inventory Costs	Transportation Costs	Administrative Costs	Total Cost	Total Cost % of GDP
1990	283	351	25	659	11.4
1991	256	355	24	635	10.6
1992	237	375	24	636	10.0
1993	239	396	25	660	9.9
1994	265	420	27	712	10.1
1995	302	441	30	773	10.4
1996	303	467	31	801	10.2
1997	314	503	33	850	10.2
1998	321	529	34	884	10.1
1999	333	554	35	922	9.9
2000	374	594	39	1007	10.3
2001	320	609	37	966	9.5
2002	300	582	35	917	8.8
2003	304	607	36	947	8.6
2004	337	652	39	1028	8.8
2005	395	739	46	1180	9.5
2006	447	809	50	1306	9.9
2007	487	856	54	1397	10.1

comes especially essential when oil prices increase, as they have recently. Since oil provides the fuel that powers the majority of transportation vehicles, it plays an important role in supply chain efficiency, particularly in the transportation portion. The increase in oil prices introduces inefficiencies in terms of low capacity utilization of transportation vehicles and high unit freight cost. Simchi-Levi et al. (2008) of MIT outlines the impacts of this change. According to his analysis, when oil prices increase, transportation costs are impacted greatly. Due to the higher transportation costs, many companies prefer to shipping in larger lot sizes and less frequently; consequently, they need to pay for higher inventory carrying costs, or locate more

distribution centers. Data from the annual State of Logistics Report also demonstrated this influence. Table 1 shows that the transportation related portion of the U.S business logistics costs rose 6% in 2007 due to high fuel costs and lower demand. An accompanying phenomenon is that there is an increase of \$40 billion increase in the inventory carrying costs from 2006 to 2007 (an increase of 9%).

Ballou (1992) stated that if the inventory can be sufficiently large, the upstream replenishment and downstream transportation functions can be completely decoupled. However, in the trend towards Just-In-Time (JIT) manufacturing and Lean Production, more and more companies are trying to keep their inventory at a low level, and this again raises the importance of integrating inventory and transportation functions.

In line with this trend, this dissertation concentrates on supply chain models that involve two sets of management concerns: those related to inventory decision and those related to transportation decision. Usually, inventory and transportation decisions can be classified into three levels: (1) The strategic level decisions specify where and how many facilities or warehouses should be built, or how the material should be flow through the supply chain network; (2) The tactical level decisions includes replenishment and production decisions, inventory policies and transportation strategies that are updated once every moderate length of time; (3) The operational level decisions determine the day-to-day operations like scheduling, routing, or truck loading.

In the last two decades, the integration of inventory and transportation decisions has been investigated in both industry and academia (e.g., Bell et al., 1983; Federgruen and Zipkin, 1984; Çetinkaya and Lee, 2000; Toptal et al., 2003; Chen et al., 2005; Schwarz et al., 2006; Çetinkaya et al., 2006). According to the levels of integration, the current research works can be classified into three groups that respectively

focus on:

- Evaluating the impact of transportation costs on inventory decisions;
- Integrating inventory and transportation decisions;
- Incorporating the transporters into supply chain coordination.

The first group of research places a particular emphasis on the inclusion of transportation costs, without explicitly optimizing and coordinating inventory decisions (e.g., Lippman, 1971; Lee, 1986; Hwang et al., 1990; Ben-khedher and Yano, 1994; Mendoza and Ventura, 2008). The second group of research proposes integration, i.e., simultaneous optimization, of inventory and transportation decisions (e.g., Çetinkaya and Lee, 2000; Axsäter, 2001; Cheung and Lee, 2002; Kleywegt et al., 2004; Çetinkaya et al., 2006; Guan and Zhao, 2010; Kaya et al., 2010). In the third group, the transporters are modeled as crucial entities in the supply chain (e.g., Mutlu, 2006). In order to decrease the system-wide operating costs, the transporters have to work in coordination with their shippers. Research in the last group is still new. However, as the oil prices remain high, supply chain members tend to increase their use of third party transportation operators that are able to consolidate shipments across companies.

Although the literature on integration of inventory and transportation decisions is very broad, there are still some challenging issues remaining:

1. The Uncertainty of Demand in a Dynamic System

In most existing literature that considers the integration of inventory and transportation decisions, it is assumed that either the demand is deterministic, although it can be non-stationary, or the demand follows a stationary random distribution. Theory exists mainly for the deterministic problems. For the

stochastic problems, although the stationary policies discussed in existing papers are easy to implement and compute in practical situations, they may be suboptimal in the class of all feasible policies as they focus on steady-state behavior. In fact, uncertainty is inherent in every supply chain system, and it is caused by supply chain dynamics. In practice, most supply chain problems arise over shorter planning horizons and are non-stationary; in other words, the economic and/or distributional parameters of the system may change over time. Under such circumstances, a dynamic formulation is more appropriate.

2. The Structure of Transportation Cost

One feature that is not commonly treated in the logistic systems but is often observed in the real-life situation is the presence of various transportation alternatives. The form of the transportation cost usually depends on the type of vehicles used for transportation (Higginson, 1993). In the traditional inventory models, the transportation cost is implicitly considered to be a part of the production cost, i.e. either as a constant sum in the fixed cost, or proportional to the quantity of produced items included in the variable ordering cost. However in practice, the transportation cost usually includes both a fixed setup cost and a variable cost. The fixed setup cost consists of the administrative cost of processing an order for both members between whom the transportation takes place, and the variable cost depends on the quantity of products shipped.

The variable cost is proportional to the shipment quantity in some cases, for example, when the delivery is conducted by privately owned trucks (private carriage). With private carriage, the logistics provider uses her own fleet to dispatch the retailer orders. The main incentive for using a private fleet is to realize a more controllable and reliable transportation together with an increased vis-

ibility of the products in transit. And under some circumstances, specifically designed vehicles are required. For example, in the cold chain logistics fresh food needs to be transported in a specific climate-controlled vehicle.

Most of the time, the variable cost is not linear. For example, when transport is performed by a public, for-hire trucking company (common-carriage), a quantity discount for shipping larger lot sizes is available, and the unit freight cost decreases as the shipment quantity increases. Compared with private carriage, common carriage has the advantage of increasing efficiency in fleet utilization and maintenance as well as reducing overhead expense.

Another example is when the transportation cost mainly depends on the number of vehicles used. In other words, no matter if the transportation vehicles are fully or partly loaded, the cost is unaffected. Hence, considering real situations, great opportunities for cost savings are missing if we assume that the transportation cost is proportional to the shipment quantity or even assume it is a constant sum.

3. Exact Optimal Integrated Inventory-Transportation Policies

Since the early 1980's, integrated inventory-transportation policies have been successfully implemented in many industries. It has been demonstrated that significant savings can be realized when inventory and transportation concerns are considered jointly. To achieve the economies of scale possible in transportation, other than those in inventory, a strategy for shipment consolidation can be included. Shipment consolidation is the policy where several small loads will be dispatched as a single, combined load. From an inventory-modeling perspective, the integrated inventory-transportation problems add dispatch quantities as decision variables to the stochastic dynamic inventory models with general

ordering cost structure, which are already known to be very difficult to optimize. Relevant existing research focuses on performing cost optimization over a subset of feasible policies. These policies are easier to implement and compute for practical purposes, nevertheless, they are probably sub-optimal in the class of all feasible policies. In order to characterize exact optimal policies, Dynamic Programming techniques are required.

Recognizing these challenges and opportunities, we have the following objectives in this dissertation:

1. To build on the theoretical framework of the existing literature in the context of integrated inventory and transportation decisions.
2. To evaluate the impact of transportation costs on inbound and outbound logistics decisions.
3. To identify optimal policies for integrated inventory and transportation decisions.

Inventory decisions are tactical level decisions, whereas transportation decisions are operational level decisions. Therefore, our research in this dissertation contributes to the literature by investigating opportunities for the coordination of tactical and operational decisions.

I.1. Scope of the Dissertation

The analysis and operation of a supply chain system varies significantly depending on its characteristics. Demands at the buyers may be deterministic or stochastic. Private fleet or common carriage may be used for transportation. The private fleet transportation may be capacitated or uncapacitated. There could be one or multiple

buyers, and consequently one or multiple products distributed through the system. The planning horizon can be of one, many or an infinite number of periods. We restrict ourselves to a single product originating at a single supplier and distributed through the vendor to one or multiple buyers so as to satisfy stochastic demand over a periodic review, finite planning horizon.

More specifically, we study the following classes of problems:

1. Pure Inbound Inventory Model (PI): The vendor makes the inventory replenishment decisions on how much to order from the outside supplier.
2. Pure Outbound Transportation Models (PO): The collection depot makes the delivery schedules of order dispatches to the buyer(s).
3. Integrated Inbound Inventory and Outbound Transportation Model (IIO): The two decisions of inventory replenishment and order dispatches are coordinated.

I.1.1. Pure Inbound Inventory Model

As we discuss in detail in Chapter II, the existing literature overlooks important transportation considerations. In particular, the impact of cargo capacity and cargo cost are rarely evaluated in previous work. However, substantial savings are realizable in supply chain system when such transportation consideration is incorporated with the inventory decisions.

In Chapter III, we focus on the economies of scale possible in the vendor's inbound replenishment. We consider a single echelon inventory system composed of a single vendor that receives a single product from an outside supplier and serves a single customer with random demand. To consider the transportation costs associated with using private trucks with cargo capacity, we model the replenishment costs in

the form of

$$W(a) = KI_{[a>0]} + \left\lceil \frac{a}{C} \right\rceil \Delta, \quad (1.1)$$

where the first term (i.e., K) is a fixed cost and the second term is the total truck cost in proportion to the number of trucks used. Here C is the cargo capacity; Δ is the cargo cost; and a is the replenishment quantity. This type of cost structure is also known as multiple setup cost structure in the literature. The system is planned over a discrete and finite time horizon. The objective is to find the structure of the optimal replenishment policy so as to minimize the total expected transportation, holding and penalty costs over the finite planning horizon.

Actually, this model is a significant extension of the classic stochastic dynamic inventory model of Scarf (1960) in that it generalizes the replenishment cost function by including the multiple setup cost term which represents the inbound transportation cost. Although there are some existing studies considering the multiple setup costs in inventory systems, one common characteristic of the previous studies is that they either focus mainly on quantity policies for deterministic demand, or focus on single period problems. All works with stochastic dynamic settings fail to characterize the complete structure of the exact optimal policy.

Based on the concept of *non- K -decreasing* of Porteus (1971), we first introduce two new concepts *non- Δ -decreasing* and *non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing* in Chapter III. Then the optimal policy for any given period can be identified provided that the major part of the recursive optimality equation satisfies certain conditions. We name the optimal policy as (Q, \vec{s}, \vec{S}) policy. Using the single period result, we provide sufficient conditions under which the new policy is optimal.

I.1.2. Pure Outbound Transportation Models

Usually, transportation costs depend on the volume and size of specific shipments. Similar to inventory and production operations, economies of scale also exist in transportation. An ideal strategy would be to stock sufficient items at the collection depot so that small orders requested by a customer or geographic area can be consolidated before a delivery is made. The corresponding savings in transportation may more than offset the increased cost of holding the inventory.

Three types of consolidation policies, i.e., time-, quantity- and time-and-quantity-based consolidation policies, have been identified in the literature and widely adopted in industry. In Chapter II, we present an overview of the related research with explicit shipment consolidation considerations. Although these shipment consolidation are easy to understand and use, they are defined by the practitioners and researchers according to their experience, and might not be optimal from the perspective of cost optimization. In Chapter IV, we examine the exact structural properties of the optimal shipment consolidation policies under four different transportation scenarios.

Scenario 1: The collection depot serves a group of retailers located in close proximity. The retailers are willing to wait to receive their orders at an additional expense for the vendor to include retailer waiting and inventory holding costs. The depot consolidates the orders in order to benefit from the scale economies of transportation. It is assumed that the outbound transportation is performed by private fleet with unlimited capacity. Thus, the transportation cost is expressed as

$$\tilde{C}_P(t, d) = K_D \cdot I_{[t>0, d>0]} + K_S d + ct, \quad (1.2)$$

where the first term represents the fixed cost for a vehicle dispatch, the second term represents the fixed cost for an order delivery and the last term represents the marginal

cost. d is the number of random orders waiting to be dispatched, and t is the total weight of the consolidated load. Consider this system for finite multiple periods, we showed that a state-dependent threshold policy is analytically optimal for this model.

Scenario 2: In the previous scenario, the transportation capacity is assumed to be infinite. However, this is not the case in most industrial practices. To address the specific consideration of cargo capacity, we replace the transportation cost by

$$\tilde{C}_S(t) = K \cdot I_{[t>0]}, \quad 0 \leq t \leq C. \quad (1.3)$$

Here, we assume the depot has only one truck with capacity C , hence, the maximum dispatch quantity is C . The cost for dispatching a shipment is fixed at K regardless of whether the truck is fully or partially loaded. Since all types of costs concerned are irrelevant with the number of consolidated orders d , d is trivial in this model, and hence, can be ignored. We develop the model as a stochastic dynamic program, analyze it for multiple periods, and characterize the optimal consolidation policy as a threshold policy.

It is worth noting that the optimal dispatch quantities in scenarios 1 and 2 are either zero or the maximal possible dispatch quantity. That is to say, when there is no cargo capacity constraint, the optimal policy possesses the “clearing property”. When the cargo capacity constraint is imposed and a dispatch should be made, the optimal dispatch quantity is equal to the consolidated load, if the load does not exceed the truck capacity; otherwise, dispatch a fully loaded truck is optimal.

Scenario 3: As mentioned above, scenario 1 and 2 both assume the employment of private trucks for transportation. However, many companies in reality use common carriers to make such shipments. Common carrier freight rates also exhibit economies

of scale in transportation. A typical common carrier transportation cost is of the form

$$\tilde{C}_C(t) = \begin{cases} c_N t, & t \leq \text{WBT}, \\ c_V \text{MWT}, & \text{WBT} < t \leq \text{MWT}, \\ c_V t, & t > \text{MWT}, \end{cases} \quad (1.4)$$

where $c_N > c_V$ denote non-volume and volume freight rates. MWT is the stated minimum weight to obtain the quantity discount and WBT is the weight at which the bumping clause comes into play. An exact characterization of the optimal consolidation policy for the common carrier case is challenging due to the complexity of the transportation cost. Therefore, assuming a “clearing property”, we examine the optimality of three practical consolidation policy and provide sufficient conditions under which they are optimal for a multiple-period dynamic distribution system.

Scenario 4: We revisit the consolidation systems discussed previously, and investigate the optimal policy for the case where the depot owns sufficient trucks and each truck is capacitated. Similar to the pure inbound inventory model, the transportation cost is expressed in the form of multiple setup costs, that is

$$\tilde{C}_M(t) = K_D \cdot I_{[t>0]} + ct + \Delta \left\lceil \frac{t}{C} \right\rceil, \quad (1.5)$$

where K_D is the fixed cost for a vehicle dispatch from the depot to the retailers, c is the transportation cost per unit weight, C and Δ are the cargo capacity and cargo cost, respectively. This cost structure particularly represents the situation where the collection depot relies on private truck fleets to deliver orders in virtue of the advantages of guaranteed capacity, flexible scheduling and enhanced customer service. Again, we examine the structure of the optimal consolidation policy via a stochastic dynamic programming approach.

I.1.3. Integrated Inbound Inventory and Outbound Transportation Model

The previous models study either the pure inbound replenishment decision or the pure outbound shipment scheduling. Although the transportation and inventory costs are explicitly incorporated, the decisions are optimized separately. Consider a vendor-managed inventory (VMI) system where the vendor has flexibility over the timing and quantity of resupply at a group of retailers with stochastic demand and located in a given geographical region. Under a VMI contract of interest, employing temporal shipment consolidation strategy allows the vendor to hold smaller orders (realized stochastic demands) from the retailers and to release them in a combined shipment to realize transportation scale economies.

In the literature, the typical VMI system that requires making joint stock replenishment and shipment scheduling decisions is assumed stationary in the long run. Researchers usually focus on finding the optimal parameter values for a predefined joint policy (Çetinkaya and Lee, 2000; Axsäter, 2001). The exact optimal policy remains unknown. To fill the gap, we consider a joint stock replenishment and shipment scheduling problem in Chapter V. We formulate the problem via a stochastic dynamic programming approach and examine the exact optimal joint policies specifying, simultaneously, the vendor's inbound replenishment and outbound dispatch quantities in successive periods so that transportation economies of scale due to shipment consolidation are realized without excessive inventory holding and/or order delay. We characterize the structure of the optimal policy as a zoned, state-dependent threshold policy which is a new class of policies in multi-echelon stochastic inventory control theory.

I.2. Organization of the Dissertation

The rest of this dissertation is organized as follows. In Chapter II, we present an overview of the literature on integration of inventory and transportation consideration in relation to the models discussed in this dissertation. In Chapter III, we study an important generalization of the classical stochastic dynamic inventory problem where privately owned trucks with limited cargo capacity are used to transport the replenishment quantity. We develop a new replenishment policy and provide the conditions under which the new policy is optimal. In Chapter IV, we address the issues regarding outbound shipment consolidation policies. We consider different types of transportation costs aiming at four transportation modes and examine the structures of optimal policies for each mode. In Chapter V, a joint stock replenishment and shipment scheduling problem under a vendor-managed inventory contract is investigated. Assuming the vendor has the authority to consolidate orders requested by the retailers, we characterize the structure of the optimal joint policy. Finally, concluding remarks, potential impact of this research, and possible future research are summarized in Chapter VI.

CHAPTER II

LITERATURE REVIEW

The efficiency of transportation systems has become particularly important with increased competition in the market. In today's highly competitive environment, companies are utilizing every possibility for decreasing their cost and making their systems more efficient.

Since the early 1980's, researchers and practitioners have demonstrated that substantial savings are realizable through coordinating transportation and inventory operations in supply chain systems. There are a variety of studies in the literature about the integration of inventory and transportation; however, among these studies a significant amount of work focuses on large scale optimization problems that include Facility Location-Allocation problems, Network Design problems, and Location-Routing problems, etc (Bell et al., 1983; Golden et al., 1984). The main goal of this group of research is to develop effective heuristic algorithms to solve the large scale Mixed Integer Programming (MIP) problems. Usually, this literature only considers deterministic demand and linear transportation cost structures (the transportation cost is proportional to the shipment quantity). The transportation policies is also defined in advance. Thus, it does not render general managerial insights into operational decisions under conditions of uncertainty or related system design issues. Since this dissertation focuses on analytical models that examine the coordination of inventory and outbound shipment decisions, theoretical studies of inventory and transportation policies in single- or multi- echelon supply chain systems, in this chapter, we present a critical review on analytical models that concentrate on production/inventory decision, transportation decision, and their coordination. This literature provides insightful tools for operational decision-making and distribution

system design. Extensive literature surveys on the inventory and routing models for the freight distribution problem were given by Baita et al. (1998) and Erengüç et al. (1999).

Consider a vendor, serving one (a group of) retailer(s), i.e., customers: the vendor's inventory is depleted by the orders coming from the customer(s) and the vendor needs to decide (1) when and how much to replenish her inventory, and (2) when and how much to deliver the orders to his downstream customer(s). Analyzing such a vendor-customer system, this dissertation is related to two streams of literature. From modeling perspective, the first stream concentrates on the incorporation of transportation costs, either implicitly or explicitly, into the inventory systems. In this review, we discuss and compare the research works for the following, but not limited to, model characteristics.

Demand: In the supply chain systems of interest, the retailers/customers are considered as the ultimate destinations with either deterministic (i.e., fixed and known) or stochastic (i.e., random variables with known probability distributions) orders. For multi-retailer system, the demand distribution can be either i.i.d. across the retailers/customers or retailer/customer specific. The multi-product case is similarly treated.

Decision(s): The concerned models can be optimizing the pure inventory policy, pure transportation policy, or attempting to jointly optimizing these two policies.

Review schedule: Period-review models divide time into one or more discrete time periods. Correspondingly, information is provided and decisions are made and implemented periodically. Continuous-review models represent information, decision-making and implementation in continuous time.

Horizon: The planning horizon in the supply chain system can be either finite or infinite. Some other models involve only a single planning period.

Number of items: Although most of the literature considers the distribution of only a single item, some models incorporate multiple items.

From the perspective of methodology, the second stream concentrates on a set of inventory control problems solved by stochastic dynamic programming approaches. Although transportation issues are not considered in these problems, their technical solution procedures provide insight and support to the current research. Furthermore, the inventory costs considered in this stream of literature, in some sense, can be translated into transportation cost with no problems. Therefore, it is worthy to review them in this chapter.

The remainder of this chapter is organized as follows. In Sublevel II.1, we present a review on the literature that considers the integration of inventory and transportation policies. This literature is discussed in three groups: (1) models employing an shipment consolidation strategy; (2) models investigating inventory policies with the consideration of transportation costs; (3) models simultaneously optimize the inventory and transportation policies. In Sublevel II.2, the stream of stochastic dynamic inventory models is reviewed.

II.1. Integration of Inventory and Transportation Policies

The integration of inventory and transportation operations has been attracting attentions recently. According to the levels of integration, the current research works can be classified into three groups that respectively focus on:

1. Shipment consolidation policy;
2. Inventory policy considering transportation cost;
3. Integrated inventory and transportation policy.

The literature review in this stream is presented as below.

II.1.1. Shipment Consolidation Policy

Transportation related issues, such as the type of the carrier and the associated transportation costs, have been extensively discussed in the literature. The broad range of this literature makes it virtually impossible to present a complete review in this dissertation. Focusing on the impact of shipment scheduling decisions, our research is most related to the distribution systems employing a *shipment consolidation* strategy.

In a traditional distribution system, merchandise is dispatched immediately to the customer when the order is received. Rapid delivery service is provided in this way; however, possible savings due to the economies of scale in transportation are missing. Recognizing this, researchers and practitioners started to investigate the shipment scheduling problems with consolidation strategies in the 1980s (Masters, 1980; Jackson, 1985). Shipment consolidation can be implemented at a Third Party Logistics (3PL) provider, a consolidating warehouse or a delivery terminal on a supply chain. Under a shipment consolidation strategy, multiple orders/shipments arrived at different times, from different origins, or for different customers can be combined into single larger dispatch loads. Subsequently, total logistics costs are reduced.

There exists a significant amount of shipment consolidation literature. The majority of the early research focuses on discussing the timing of load dispatches and proposes some practical policies (e.g., Newbourn and Barrett, 1972; Pollock, 1978). The most popular policies of consolidation programs include time-based, quantity-based, and hybrid, i.e., time-and-quantity-based policies. A time-based policy releases a shipment on regular intervals, and a quantity-based policy releases a shipment whenever an economical dispatch quantity is available. Under a hybrid policy, a shipment is released either upon a predetermined shipping date or upon the accumulation of

a dispatch quantity, whichever occurs first (Higginson and Bookbinder, 1994, 1995; Çetinkaya, 2005). It is worth noting that although various shipment consolidation policies have been proposed and adopted in industry, these policies are all designed in advance. In other words, the consolidation policy that is theoretically optimal remains unknown.

Another focus of the early literature is to examine the performance of systems using different consolidation programs via simulation (Masters, 1980; Jackson, 1981; Cooper, 1984; Jackson, 1985; Closs and Cook, 1987; Bagchi and Davis, 1988; Pooley and Stenger, 1992; Higginson and Bookbinder, 1994). For example, Jackson (1981) compares a time-based policy to a hybrid policy and indicates that the time-based policy is more convenient to implement. Higginson and Bookbinder (1994) investigate the performance of different consolidation policies for common carriage transportation by adjusting the policy parameters in simulation studies. They assume that the shipments arrive at the collection depot at random time and with random sizes, and identify possible situations where one policy works better than the others. During the early years, neither economic justification of the practical policies nor approaches for computing the optimal consolidation policy parameters are provided.

Recently, research on shipment consolidation focuses on computing the optimal policy parameters using analytical skills. Much of the research provides optimization approaches for finding the parameters of pure consolidation practices (see, for example, Gupta and Bagchi, 1987; Minkoff, 1993; Higginson and Bookbinder, 1995; Bookbinder and Higginson, 2002; Çetinkaya and Bookbinder, 2003; Mutlu et al., 2010). For continuous-review systems, Gupta and Bagchi (1987) examine the inbound consolidation policy under a just-in-time procurement system and provide a tool to calculate the minimum cost-effective consolidation quantity by using the stochastic-clearing-system theory. Based on Gupta and Bagchi (1987)'s model, Bookbinder and

Higginson (2002) study the time-and-quantity-based policy. Instead of employing the stochastic-clearing-system theory, Çetinkaya and Bookbinder (2003) use renewal theory to analyze the time-based and the quantity-based policies for both common carriage and private fleet transportation. In their model, the orders are assumed to arrive at the depot according to a Poisson process. For private carriage, they were able to provide exact optimal solutions for the two policies. They also provide approximate solutions for common carriage, and discuss the special case with unit order sizes and provide results. Following Çetinkaya and Bookbinder (2003)'s work, Mutlu et al. (2010) study the hybrid, i.e., the time-and-quantity-based policy for the case of private carriage.

Research regarding continuous-review systems is devoted to the steady-state behavior of consolidation systems. However, most consolidation problems in real-life arise over shorter planning horizons and are non-stationary; in other words, the stationary policies discussed in the existing literature are practical, although they are probably suboptimal. To identify the exact optimal consolidation policy, Markov Decision Process (MDP) methods are adopted. Minkoff (1993) uses a MDP method to formulate the consolidation problem and proposes a heuristic for computing the dispatch policy values. Higginson and Bookbinder (1995) use MDP method to model both common and private carriage situations and identify the optimal consolidation policies via numerical study.

There are some recent papers extending the shipment consolidation schedule to a two-echelon supply chain system. In the two-echelon models, the inbound replenishment and outbound shipment consolidation decisions are simultaneously optimized. Literature in this group is discussed in Sublevel II.1.3.

II.1.2. Inventory Policy Considering Transportation Cost

In this group of works, the transportation policy is assumed to be given. Although the inventory and transportation decisions are not optimized simultaneously, the impact of the transportation operations is modeled by explicitly including a cost term representing the realistic transportation situations, for example, transportation with quantity discount, or transportation with vehicle capacity constraints. The objective of this group is to find the optimal production/inventory decisions that are directly affected by the concerned transportation cost.

II.1.2.1. Models with Deterministic Demand

Corresponding to different transportation patterns, various structures of transportation cost have been investigated in supply chain systems with deterministic demand, and they are mostly studied in single-echelon lot-sizing models. Ever since the classical dynamic lot-sizing model was introduced by Wagner and Whitin (1958), many researchers have developed extended models (Zangwill, 1966, 1969; Florian and Klein, 1971; Love, 1973; Swoveland, 1975; Chen et al., 1994; Lee et al., 2001; Hwang and Jaruphongsa, 2006) with various considerations, including concave costs, piecewise concave costs, or capacitated production/transportation.

Since our research concentrates on analyzing inventory policies under explicit general private-fleet transportation costs, we proceed with a detailed discussion of the literature that directly considers the cost of multiple setups.

In fact, the majority of existing work on multiple setup cost structure is for single-echelon lot-sizing models with deterministic demand. This literature can be classified into two streams. One stream of work examines the structures of optimal policies for periodic-review systems (e.g., Lippman, 1969a; Lee, 1989; Ben-khedher and Yano,

1994; Alp et al., 2003; Li et al., 2004; Lee, 2004; Anily and Tzur, 2005; Jaruphongsa et al., 2005; Anily and Tzur, 2006; Jaruphongsa and Lee, 2008; Hwang, 2009, 2010). Among these studies, Lippman (1969a) provides the most basic model that integrates the multiple setup costs as the inbound transportation cost. Lippman proves that there exists an optimal solution such that in each period, either beginning inventory is zero or the order quantity is a multiple of the full truckload. Lee (1989) generalizes Lippman's model by incorporating a replenishment setup cost for each order. Based on Lee's model, various extensions include the study of multiple setup cost function in the context of the applications to multi-product replenishment systems (e.g., Benkhedher and Yano, 1994; Anily and Tzur, 2005, 2006), batch production processes with stochastic lead times (e.g., Alp et al., 2003), the generalization of freight cost with a truckload discount (e.g., Li et al., 2004), the selection of transportation modes from multiple choices (e.g., Jaruphongsa et al., 2005), the consideration of constraints on replenishment quantity and replenishment time (e.g., Lee, 2004; Jaruphongsa and Lee, 2008; Hwang, 2009).

The other stream examines the optimal policies for models with continuous time scale, constant demand rate and finite/infinite time horizon, i.e., in Economic Order Quantity (EOQ) type models (e.g., Lippman, 1971; Aucamp, 1982, 1984; Lee, 1986; Hwang et al., 1990; Mendoza and Ventura, 2008). Lippman (1971) gives mathematical formulations for both infinite and finite planning horizons with ordering cost in form of multiple setup cost, and characterized the form of an optimal ordering schedule for both cases. Aucamp (1982) formulates the problem introduced by Lippman (1971) in a traditional EOQ way and provides an algorithm for solving it. Aucamp (1984) extends his model to consider discounted cash flows and demonstrates the equivalence of the modified model with his earlier one. Lee (1986) incorporated freight discounts into the model of Aucamp (1982), i.e., the cost per load decreases as the number of

truckloads increases. Lee provides a revised solution algorithm to solve the generalized model. Hwang et al. (1990) further extends Lee's model by considering an all-unit quantity discount on the purchasing cost. Recently, the case of incremental quantity discount is also studied by Mendoza and Ventura (2008).

II.1.2.2. Models with Stochastic Demand

It is worth noting that all of the studies discussed in Sublevel II.1.2.1 within the context of economic lot-sizing and transportation considerations assume deterministic demand, i.e., ignore the stochastic nature of demand. To the best of our knowledge, the existing literature that focuses on modeling transportation cost with cargo capacity in stochastic demand inventory systems only includes the work by Lippman (1969b), Iwaniec (1979), Toptal (2009) and Çalışkan Demirağ et al. (2011). Examining a single-echelon stochastic dynamic inventory problem with multiple setup ordering costs, Lippman (1969b) identifies a partial characteristic of an optimal policy, while Iwaniec (1979) provides a sufficient condition under which the full load ordering policy is optimal for a finite horizon problem. Recently, Toptal (2009) reconsiders the inventory system of Hwang et al. (1990) by formulating it in a more general form and discusses its application to the single period, single-echelon stochastic demand problem, i.e., the news-vendor problem.

Recently, Çalışkan Demirağ et al. (2011) revisit the classical stochastic dynamic inventory problem while assuming the replenishment cost is $W(a) = K_1 I_{[0 < a \leq C]} + K_2 I_{[a > C]}$. Here, parameters K_1 and K_2 satisfy $0 \leq K_1 \leq K_2$ and are called quantity-dependent fixed setup costs. They attempt to analyze the general case where $0 \leq K_1 \leq K_2$ as well as a special case where $K_1 \leq K_2 \leq 2K_1$. They rely on two concepts: namely, a new concept called C - (K_1, K_2) -convexity and an existing concept known as *strong K -convexity* developed by Gallego and Scheller-Wolf (2000). The authors

conclude that five critical points ($s \leq s'' \leq s' \leq s_1 \leq S$) divide the whole state space of x , inventory level, ($-\infty < x < \infty$), into six regions. They prove that the optimal replenishment policies for both the general and special cases of their problem possess identical characteristics except for the region $[s_1, S)$. In particular, on $[s_1, S)$ the exact value of the optimal replenishment quantity is unclear in the general case while taking value of zero in the special case. That is, although the authors characterize a simple property of the optimal policy over $[s'', s')$ for the general case, the characterization of the exact optimal policy remains incomplete overall.

Çalışkan Demirağ et al. (2011) also attempt to examine the case where $W(a) = \sum_{i=1}^n K_i I_{[C_i < a \leq C_{i+1}]}$, where $K_i \leq K_j$ and $C_i \leq C_j$ for any $i < j$, $C_1 = 0$, and $C_{n+1} = \infty$. However, they simply conclude by stating that “the optimal policy can indeed be highly complex” without any specific results. The authors proceed with assuming $C_{i+1} - C_i = C$, $K_{i+1} - K_i = K$, and $n \rightarrow \infty$. Under these assumptions, their problem is equivalent to our problem. Using an existing concept, known as (C, K) -convexity and developed by Shaoxiang (2004), Çalışkan Demirağ et al. (2011) are only able to offer a simple preliminary result that is clearly insufficient to characterize the structure of the optimal policy leaving an important gap in the literature. In Chapter III, we fill this gap by providing a complete characterization of a new class of policies which we call the (Q, \vec{s}, \vec{S}) policy. We also develop sufficient conditions for the optimality of this policy. In order to justify the difficulty associated with the optimal policy, Çalışkan Demirağ et al. (2011) solve a numerical example by complete enumeration. They compare the result with the full truckload ordering policy developed by Iwaniec (1979). Their numerical results are such that the optimal policy obtained through a complete enumeration is of the form of our (Q, \vec{s}, \vec{S}) policy whose optimality is proved here. Clearly, the policy outperforms Iwaniec’s (1979) policy.

II.1.3. Integrated Inventory and Transportation Policy

In addition to the single-echelon models discussed in Sublevel II.1.2, a recent line of research analyzes the impact of transportation cost in the context of two-echelon inventory systems. More specifically, this group of work concentrates on identifying the integrated inbound inventory and outbound transportation decisions for a vendor. Although warehouse location and vehicle routing decisions are important subcategories of inventory and transportation decisions, in this dissertation, we are only interested in the decisions regarding the timing and quantity of inbound order replenishment and outbound shipment scheduling.

II.1.3.1. Models with Deterministic Demand

The integrated inventory and transportation model has its root in the multi-echelon inventory control problem with deterministic demand. For the purpose of completion, we review the literature on multi-echelon inventory control problems where the inventory replenishment decisions of successive echelons are decided simultaneously. The reason that we are also interested in this branch of research is that the models with discount on the unit ordering/production cost can also be considered in the transportation context. Hence, they give insights into the problems with transportation policies.

The pioneers of the deterministic demand multi-echelon models are Schwarz (1973) and Goyal (1976). In their models, both the supplier and the buyer incur fixed costs for replenishing inventory and a per unit per time inventory holding cost. Infinite production rate is assumed. Later, Goyal (1988) considers an integrated inventory model with finite production rate and suggests equal sized shipments to the buyer. Based on Goyal's results, Lu (1995) designs heuristic algorithms to find the

optimal equal sized shipments. However, the equal sized shipment policy is suboptimal among all the possible policies. Goyal (1995) suggests an unequal sized shipment policy in which successive shipments of a lot increase by a factor equal to the ratio of the production rate to the demand rate. Hill (1997) relaxes Goyal's (1995) result by allowing the increasing factor between 1 and the ratio of the production rate to the demand rate. Goyal and Nebebe (2000) suggests a shipment policy with the first shipment of small size followed by several equal sized shipments of larger size. Hill (1999) derives that the globally optimal shipping policy for the single-vendor single-buyer inventory problem is a combination of equal and unequal sized shipment policy. Hoque and Goyal (2000) extends Hill's (1999) results to a case where the transport equipment between the two echelons is capacitated.

In the classical models above, the transportation costs are assumed to be constant. Recently, Zhao et al. (2004) addresses the problem of deciding the optimal ordering quantity and frequency for a system where transportation cost is assumed to be the sum of a fixed setup cost and a variable cost. They build a modified EOQ model and provided an algorithm for solving the model. Ertogral et al. (2007) incorporate transportation cost explicitly into the vendor-buyer lot-sizing problem. They employed the equal-size shipment policy and developed optimal solution procedures for solving the integrated models. All-unit-discount transportation cost structures with and without over declaration have been considered.

Multiple setup cost function that represents the transportation cost from the warehouse to the retailer is also investigated. For continuous deterministic demand, Çetinkaya and Lee (2002) characterize the properties of the optimal integrated inventory replenishment and freight consolidation policies for an inventory system consisting of a single warehouse and a single retailer and over an infinite time horizon. They provided exact solutions for the optimal shipment consolidation policy parameters

and showed that the optimal consolidation load is not constant. Notice that in their model, the impact of multiple setup costs is only evaluated for the outbound transportation, and later, Toptal et al. (2003) and Toptal and Çetinkaya (2008) consider this cost structure for both the inbound and outbound transportation from different perspectives: centralized optimization and channel coordination, respectively. Considering two modes of transportation available to the warehouse, Rieksts and Ventura (2010) proposed a heuristic algorithm for a single-warehouse multi-retailer system.

For periodic deterministic demand, Lee et al. (2003) incorporate a multiple setup cost term into the outbound transportation cost and provide a network approach to solve for the optimal integrated replenishment/shipment scheduling policy over a finite time horizon. Jaruphongsa et al. (2007) study a model similar to that of Lee et al. (2003), but they consider two modes of outbound transportation available to the warehouse, i.e., one option with a fixed setup cost and unit dispatch cost and the other with a multiple setup cost structure. In a recent paper, Jin and Muriel (2009) extend the model by including multiple retailers and incorporating multiple setups into the inbound replenishment cost. They develop exact algorithms for computing the optimal policies for both decentralized and centralized cases.

II.1.3.2. Models with Stochastic Demand

Deterministic models help us to gain insights into the dynamics of the problem. However, stochastic models provide better representations of real life applications. A growing body of literature examines different aspects of VMI systems since the late 1990s. Several authors, such as Campbell et al. (1998); Çetinkaya and Lee (2000); Axsäter (2001); Cheung and Lee (2002); Kleywegt et al. (2004); Çetinkaya et al. (2006); Schwarz et al. (2006); Toptal and Çetinkaya (2006); Gurbuz et al. (2007); Zhao et al. (2007); Çetinkaya et al. (2008); Mutlu and Çetinkaya (2010); Guan and Zhao

(2010); Kaya et al. (2010); Savaseneril and Erkip (2010) focus on analyzing inventory, shipment, and routing policies under VMI or similar multi-echelon settings. Within this line of work, Çetinkaya and Lee (2000) are the first to examine the problem of interest here while assuming a time-based shipment consolidation policy that clears the entire consolidated load on a periodic time schedule. In their model, the supplier has the power to control the inventory management of her downstream customers. The VMI contract between the supplier and the customer enables the supplier to hold small shipments requested by different customers and to dispatch a combined, larger load. The orders are assumed to follow a Poisson process, and each shipment is of unit size. Çetinkaya and Lee (2000) present analytical results for a renewal theoretic model and provide an easy-to-implement approximate solution method for this problem. Axsäter (2001) presents a procedure that optimally solves the problem.

Three recent papers, Chen et al. (2005), Çetinkaya et al. (2006) and Çetinkaya et al. (2008) have revisited the model of Çetinkaya and Lee (2000). Chen et al. (2005) investigated the integrated inventory replenishment and shipment consolidation problem by comparing the two consolidation policies: quantity-based and time-based. They showed that the quantity-based consolidation can outperform the time-based counterpart while the reverse never occurs. Çetinkaya et al. (2006) examined the case where the vendor implements quantity-based and hybrid policies for consolidating Poisson demands. Çetinkaya et al. (2008) generalized the demand to a more realistic and complicated demand process of practical interest. They study the case where a quantity-based shipment consolidation policy is in place under which a clearing decision is triggered based on a critical dispatch quantity rather than a time schedule. In addition, Toptal and Çetinkaya (2006) explore the issues about channel coordination under explicit transportation consideration and solve a single period, two-echelon inventory system where the cost structure of multiple setups is

considered for both echelons. More recently, Mutlu and Çetinkaya (2010), and Kaya et al. (2010) also revisit the problem of Çetinkaya and Lee (2000) focusing on the computation of practical policies that still rely on quantity- or time-based shipment consolidation. Such policies are not necessarily optimal within the class of all feasible policies for the simple reason that they rely on stochastic clearing assumptions. Mutlu and Çetinkaya (2010) investigate the optimal joint policy with consideration of common carriage for outbound shipment. Their results demonstrate that common carriers can also benefit from shipment consolidation in integrated inventory systems. Through numerical examples, Kaya et al. (2010) demonstrate that the exact optimal policy is complex and non-monotonic; but, they are unable to characterize the exact optimal policy.

II.2. Stochastic Dynamic Inventory Control Problems

The methodologically oriented literature on stochastic dynamic inventory systems supports the analysis in this dissertation from a technical perspective. In the interest of brevity, we review only the papers that concern the characterization of the optimal policy in a single-product, single-echelon, periodic review inventory control setting.

For a single-period case, the news-vendor model is one of the most popular models, and the solution to the news-vendor model balances the expected inventory holding costs and the shortage cost for unsatisfied demands. When extended to multiple-period cases, this model gives rise to the so-called base-stock or order-up-to policy. The base-stock policy specifies a single critical parameter that determines the optimal amount of inventory to carry in any period. The optimality of the base-stock policy is first proved for the case of the finite planning horizon by Clark and Scarf (1960). Following Clark and Scarf's work, Federgruen and Zipkin (1984a,b) showed that this

policy is still optimal in the case of the infinite planning horizon and they also provided computational methods for finding the critical parameters of the base-stock policies. Chen and Zheng (1994) and Chen (2000) generalized the results for Clark and Scarf's model by giving fixed batch size at each stage. More generalizations of Clark and Scarf's model can be found in Chen and Song (2001), Gallego and Ozer (2003, 2005) and van Houtum et al. (2007).

It is worth noting that in most of the above papers, the optimal policy is either a base-stock policy or a modified base-stock policy. This is because all these models assume linear replenishment/production costs with no fixed setup costs; in other words, their models do not exhibit economies of scale. Fixed setup costs arise in the vendor-buyer system as the sum of costs involved in setups plus the cost of processing orders. The literature that considers the fixed setup costs in a periodic review inventory model dates back to the early years of 1960s. Scarf (1960) studied a periodic review, finite horizon inventory problem under the condition that the ordering cost includes a fixed setup cost K and the one period expected holding/shortage cost is assumed to be convex. Scarf introduced the concept of K -convexity and characterized the structure of the optimal policy as the notable (s, S) policy. Under an (s, S) policy, whenever the inventory level (inventory on hand plus on order minus backorders) is below s , an order is placed to bring the inventory level up to S . The optimality of (s, S) policy for the infinite horizon problem was proved in Iglehart (1963). Veinott (1966) presented another proof for the (s, S) optimality result under different assumptions. He relaxed the convexity constraint on the one period expected holding/shortage cost to include quasi-convex functions. Building on Scarf's model, many other researchers have demonstrated the optimality of (s, S) policy in their specific settings (e.g., Schal (1976), Sethi and Cheng (1997), Gallego and Scheller-Wolf (2000)). Also, Gallego and Toktay (2004) considered a special case of the capacitated problem where all orders

are constrained to be full-capacity orders and they showed that the optimal policy is a threshold policy under the specific settings. From the computational perspectives, Veinott and Wagner (1965) and Zheng and Federgruen (1991) provided effective methods and algorithms for finding the optimal (s, S) inventory policy.

All the studies mentioned in the previous paragraph make the assumption that the ordering cost consists of a fixed setup cost and a linear variable cost. However, this is still too restrictive for some real-life inventory problems. As for generalization of cost structure, Porteus (1971, 1972) considered the case of a concave increasing ordering cost with a fixed setup cost and proves the optimality of a generalized (s, S) policy for some specific demand distributions. Lippman (1969a) considered a deterministic demand, periodic review, finite horizon inventory problem with multiple setup ordering cost. In Lippman (1969b), the stochastic demand case with multiple setup ordering cost was studied and a partial characterization of an optimal policy was obtained. With the same model as Lippman (1969b), Iwaniec (1979) also studied the stochastic dynamic inventory problem with multiple setup ordering cost structure, and provided a sufficient condition for the optimality of a full truck load ordering policy. Other studies for a periodic review, stochastic inventory models can be found in Parlar and Rempala (1992); Çetinkaya and Parlar (2004); Janakiraman and Roundy (2004); Chen et al. (2006), and van Houtum et al. (2007).

Essentially, the characterization of the structural optimal policies in stochastic dynamic program relies on the properties of some term in the recursive optimality equation. Related to our results in Chapter III, two noteworthy papers include Porteus' seminal work (see Porteus (1971, 1972)) regarding the concept of *non-K-decreasing* and optimality of generalized (s, S) policies under concave increasing replenishment costs. Traditionally, various *generalized convexity concepts* have been useful to solve stochastic dynamic inventory problems, e.g., *K-convexity* developed by

Scarf (1960), CK -convexity developed by Gallego and Scheller-Wolf (2000), (C, K) -convexity developed by Shaoxiang (2004), and C - (K_1, K_2) -convexity developed by Çalışkan Demirağ et al. (2011). Although directly useful for other purposes, these existing concepts do not suffice to characterize the structure of the optimal policy under multiple setup costs. For this reason, we introduce two new concepts called, $non-(\Delta, C)$ -decreasing and $non-(\Delta, C)_{\mathbb{N}}^K$ -decreasing, both of which build on the concept of $non-K$ -decreasing developed by Porteus (1971). These two concepts are also related to the concept of (C, K) -convexity developed by Shaoxiang (2004) who extends Scarf's (1960) model to consider a finite order capacity. The results provided by Shaoxiang (2004) require the (C, K) -convexity of the cost-to-go function over its entire domain which is somewhat restrictive. In our approach, on the other hand, instead of requiring our newly introduced concepts to apply on the whole domain of cost-to-go function, we introduce a family of functions, called \mathcal{G} in Definition 3, such that for each member of the family the concept of $non-(\Delta, C)$ -decreasing is applicable on a subset of its domain. It is easy to verify that the functions in \mathcal{G} preserve the major characteristics that pertain to the concept of (C, K) -convexity, i.e., conditions A3(a), A3(b) and the second part of condition A3(c) in our Definition 3, respectively, correspond to parts (b), (a), and (c) in Shaoxiang's (2004) Lemma 1. However, the functions in \mathcal{G} possess an extra property as specified in the first part of condition A3(c). The inclusion of this part enables us to completely characterize the optimal replenishment policy under multiple setup costs in Chapter III.

CHAPTER III

STOCHASTIC DYNAMIC INVENTORY PROBLEM UNDER EXPLICIT
INBOUND TRANSPORTATION COST AND CAPACITY

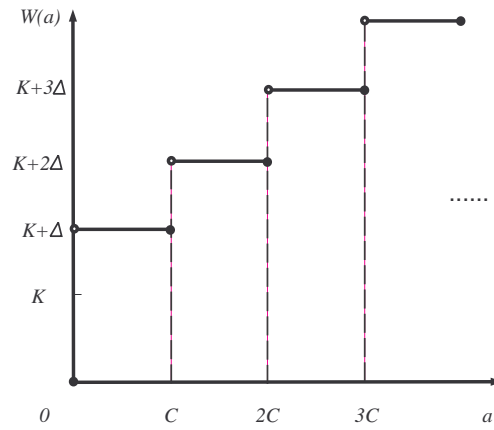
In this chapter, we generalize the classical stochastic dynamic inventory problem solved by Scarf (1960) to consider the impact of inbound transportation cost and capacity, explicitly. In Scarf's problem, the replenishment cost is presented as the summation of a fixed setup cost and a linear variable cost. Clearly, this cost structure ignores the impact of transportation cost and capacity related to delivery of replenishment orders; thereby, also ignoring possible transportation scale economies achievable via optimization. With this observation, we modify Scarf's model *by generalizing the replenishment cost function* as a means to include more information about realistic inbound transportation issues.

Specifically, we focus on the case where a private fleet of capacitated trucks are being used for inbound transportation of replenishment orders. Hence, we model the inbound transportation cost as a staircase function to represent the situation where the trucks have finite cargo capacity, denoted by C , and the transportation cost is based on the number of trucks used. Under this condition, the cargo cost, denoted by Δ (the cost for using one truck) is the same regardless of whether a truck is fully or partially loaded. This type of cost structure is also known as multiple setup cost structure in the literature (Lee, 1986). An illustration of the *generalized replenishment cost function* of interest, denoted by $W(\cdot)$, is provided in Figure 1. Letting $I_{[a>0]}$ denote the indicator function that has value 1 if $a > 0$ and 0 otherwise, we have

$$W(a) = KI_{[a>0]} + \left\lceil \frac{a}{C} \right\rceil \Delta. \quad (3.1)$$

Hence, $W(\cdot)$ includes both the traditional fixed setup cost, K , and the multiple setup cost structure representing the case of private fleet transportation considered in this research.

Figure 1: The Generalized Replenishment Cost Function to Consider Cargo Cost and Capacity



It is worth noting that this particular generalization of Scarf's model is also investigated by Lippman (1969b) and Iwaniec (1979). However, Lippman (1969b) fails to identify the structural properties of the optimal multi-period policy whereas Iwaniec (1979) provides a sufficient condition for the optimality of full cargo (full truckload) replenishment policy that is clearly suboptimal for our problem. More recently, Çalışkan Demirağ et al. (2011) revisit the full truckload policy in Iwaniec (1979) but they are also unable to provide a complete characterization of exact optimal policies under multiple setup costs. Our results extend those developed by both Lippman (1969b) and Iwaniec (1979) as well as by Çalışkan Demirağ et al. (2011) while providing a significantly enhanced characterization of optimal policies under multiple setup cost functions and stochastic demand.

This chapter is organized as follows. In Sublevel III.1, we develop a stochastic dynamic programming formulation of the problem. Several new concepts are introduced in Sublevel III.2 before presenting our structural results in Sublevel III.3. To examine the impact of system parameters, some computational study is included in Sublevel III.4. We summarize this chapter with practical insights we obtain from the optimal ordering policy as well as suggestions for possible future research, in Sublevel III.5.

III.1. Notation and Problem Formulation

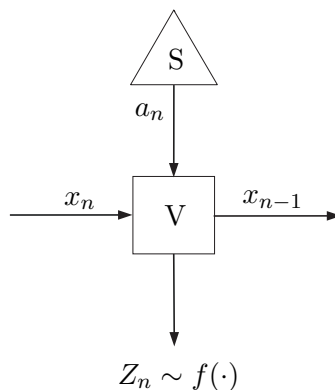
As explained above, our model shares the same system settings as the classical stochastic dynamic inventory problem solved by Scarf (1960). The only difference exists in the structure of the replenishment cost. For completeness, the problem is described as follows: A vendor (e.g., wholesaler, distributor, retailer, etc.) faces independent and identically distributed stochastic demands during a planning horizon of N periods (N is finite) and replenishes inventory from an ample external supplier, i.e., the manufacturer. At the beginning of period n ($n \leq N$), the vendor's initial inventory level x_n is observed. At this time, a replenishment quantity a_n can be placed. We assume the replenishment delivery lead times are negligible. There is a fixed setup cost, K , associated with each replenishment order. In addition, the shipments of replenishment orders from the manufacturer to the vendor are performed by the vendor's own truck fleet. The trucks have identical cargo capacity C and cargo cost Δ . There are no constraints on the replenishment quantity in each period, but when this quantity exceeds a full truckload, additional trucks are required for the transportation. Accordingly, the replenishment cost for ordering a_n units can be

represented as

$$W(a_n) = KI_{[a_n > 0]} + \left\lceil \frac{a_n}{C} \right\rceil \Delta. \quad (3.2)$$

After a_n is chosen, the demand in period n arrives at the vendor and depletes the inventory. The demand is a nonnegative random variable Z_n with the density function $f(\cdot)$. All unsatisfied demands are backordered and all excessive inventories are carried to the next period. A holding or shortage cost is charged based on the net inventory at the end of the period, i.e., $x_n + a_n - Z_n$. Future costs are discounted at a one-period discount rate β ($0 < \beta \leq 1$), and all parameters are assumed to be stationary. $L(x_n + a_n)$ denotes the expected holding and shortage cost in period n excluding the replenishment cost. The objective is to find the optimal replenishment policy for a_n so as to minimize the total expected replenishment, holding and penalty costs over the finite planning horizon.

Figure 2: Problem Setting of the Inbound Replenishment System



We define $y_n = x_n + a_n$. Thus, y_n is the number of products available upon the arrival of the order, i.e., y_n is the order-up-to level. Since there is a one-to-one correspondence between y_n and a_n , our problem can be stated as finding the optimal values of y_n so the total expected cost is minimized. The problem can be formulated

as a dynamic programming problem using backward recursion where the periods are indexed in a backward order, i.e., they occur over time in the order $N, N - 1, \dots, 0$, and period 0 is the end of the planning horizon. Figure 2 depicts the setting of the dynamic system. For notational simplicity, the subscript n is omitted on x and y in the remainder of this chapter. Before proceeding to the formulation development, let us summarize the notation introduced so far below and define some new notation that will be used throughout the rest of the chapter.

System Parameters:

N	length of the planning horizon
n	period index ($n = 0, 1, \dots, N$)
β	one-period discount factor ($0 < \beta \leq 1$)
Z_n	nonnegative demand in period n (we assume $\{Z_n\}_{0 \leq n \leq N}$ forms an i.i.d. sequence. A generic element is denoted as Z with density and distribution functions $f(\cdot)$ and $F(\cdot)$)
K	fixed setup cost
C	cargo capacity
Δ	cargo cost
$W(a)$	replenishment cost for ordering a units
h	inventory holding cost per unit per period
p	backorder penalty cost per unit per period
$g_T(x)$	terminal cost for x units of ending inventory at the end of the planning horizon
$L(y)$	one-period expected holding and shortage cost when the order-up-to level is y

States:

x_n inventory level at the beginning of period n , before a replenishment order is placed

Decisions:

a_n the amount ordered and received instantaneously in period n

y_n the order-up-to level in period n , before the demand is realized
($y_n = x_n + a_n$)

$Y_n(x)$ the optimal order-up-to level of period n with its beginning inventory x

Optimality Equation:

$V_n(x)$ the optimal expected total cost from period n to the end, when period n has x units of initial inventory

If the inventory level immediately after a replenishment arrives is y , then the one-period expected holding and shortage cost is given by

$$L(y) = \begin{cases} h \int_0^y (y - z) f(z) dz + p \int_y^\infty (z - y) f(z) dz, & y \geq 0, \\ p \int_0^\infty (z - y) f(z) dz, & y < 0. \end{cases} \quad (3.3)$$

Then the optimality equation can be written as

$$V_n(x) = \begin{cases} \min_{y \geq x} \{W(y - x) + L(y) + \beta \int_0^\infty V_{n-1}(y - z) f(z) dz\}, & n = 1, \dots, N, \\ g_T(x), & n = 0. \end{cases} \quad (3.4)$$

Subsequently, the objective is to find the optimal order-up-to level $Y_n(x)$ that minimizes the expected total cost for each period n and for any beginning inventory level x . For our purpose, we define

$$G_n(y) = L(y) + \beta \int_0^\infty V_{n-1}(y - z) f(z) dz. \quad (3.5)$$

Then the optimality equation can be rewritten as

$$V_n(x) = \begin{cases} \min_{y \geq x} \{W(y-x) + G_n(y)\}, & n = 1, \dots, N, \\ g_T(x), & n = 0. \end{cases} \quad (3.6)$$

In period n , if the beginning inventory level is x , it is optimal to place a replenishment order if and only if there exists some y greater than x with $G_n(x) > W(y-x) + G_n(y)$. If a replenishment order should be placed, it is optimal to order up to the y such that $W(y-x) + G_n(y)$ is minimized.

This optimization problem is challenging due to the discontinuity of the staircase function W . By (3.5) and (3.6), W actually impacts the structure of the cost-to-go function V and subsequently influences the function G . In order to examine the characteristics of the function G , we define new concepts and provide basic results in Sublevel III.2. In Sublevel III.3, we analyze the function G first then identify the optimal replenishment policy based on G 's characteristics.

III.2. New Concepts and Basic Properties

Before proceeding with the development of new concepts, we examine the properties of the one-period expected holding and shortage cost $L(y)$.

Proposition 1 *$L(y)$ is a convex function with a unique minimizer denoted by P , and P satisfies*

$$F(P) = \frac{p}{h+p}, \quad (3.7)$$

where F is the distribution function of the random demand.

Proof of Proposition 1: By definition, the second derivative of $L(y)$ exists and can be derived as $L''(y) = (h+p)f(y)$ if $y \geq 0$ and 0 otherwise. Since f is a density function, $f(y)$ is nonnegative for all y , hence, $L''(y) \geq 0$. In other words, $L(y)$ is convex. Let

the first derivative of $L(y)$ be zero, then the equation (3.7) is obtained. ■

To examine the characteristics of function G_n , we introduce two new concepts which are analogs of the concept of *non- K -decreasing* of Porteus (1971).

Definition (Porteus, 1971) *A function φ is non- Δ -decreasing on a domain X if $\varphi(x) \leq \varphi(y) + \Delta$ for $x, y \in X$ and $x \leq y$.*

In other words, if a function is non- Δ -decreasing, then for any point x on a domain X , no matter how much it increases, the decrease in the function value does not exceed Δ .

We call the first of our analogs non- (Δ, C) -decreasing.

Definition 1 *Given positive constants Δ and C , a function φ on a domain X is called non- (Δ, C) -decreasing at a fixed point $x_0 \in X$, if for $y \in [x_0, x_0 + C] \cap X$, $\varphi(x_0) - \varphi(y) \leq \Delta$. And the function is called non- (Δ, C) -decreasing on a set B if it is non- (Δ, C) -decreasing at any point $x \in B$.*

The intuitive interpretation of non- (Δ, C) -decreasing is that for any point x_0 on the domain X , if increased by at most C , the decrease in the function value of φ does not exceed Δ . Non- (Δ, C) -decreasing can be thought of as a relaxation of non- Δ -decreasing, only requiring it to hold at points no more than C units greater than x_0 . Note that the standard non- Δ -decreasing corresponds to non- (Δ, ∞) -decreasing in Definition 1.

We now extend this definition to what we refer as non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing.

Definition 2 *Given a nonnegative constant K , and positive constants Δ and C , a function φ on a domain X is called non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing at a fixed point $x_0 \in X$, if for $\forall m \in \mathbb{N}$ and $y \in [x_0, x_0 + mC] \cap X$, $\varphi(x_0) - \varphi(y) \leq K + m\Delta$. And the function is called non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing on a set B if it is non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing at any point $x \in B$.*

For any point x_0 on the domain X of a non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing function φ , if one increases x_0 by at most mC (m is a positive integer), the decrease in the function value φ does not exceed $K + m\Delta$. The standard non- Δ -decreasing also corresponds to non- $(\Delta, \infty)_{\mathbb{N}}^0$ -decreasing in Definition 2.

Non- (Δ, C) -decreasing and non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing functions have several useful properties.

Property 1 *If a function φ is non- (Δ, C) -decreasing on a domain X , then it is also non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing on X for any $K \geq 0$.*

Proof of Property 1: Suppose φ is non- (Δ, C) -decreasing on a domain X . Choose any $m \in \mathbb{N}$, $\forall x, y \in X$ and $y \in [x, x + mC]$, then

$$\begin{aligned} \varphi(x) - \varphi(y) &= \varphi(x) + [-\varphi(x + C) + \varphi(x + C)] + [-\varphi(x + 2C) + \varphi(x + 2C)] + \dots \\ &\quad + \left[-\varphi\left(x + \left\lfloor \frac{y-x}{C} \right\rfloor C\right) + \varphi\left(x + \left\lfloor \frac{y-x}{C} \right\rfloor C\right) \right] - \varphi(y) \\ &= [\varphi(x) - \varphi(x + C)] + [\varphi(x + C) - \varphi(x + 2C)] + \dots \\ &\quad + \left[\varphi\left(x + \left(\left\lfloor \frac{y-x}{C} \right\rfloor - 1\right)C\right) - \varphi\left(x + \left\lfloor \frac{y-x}{C} \right\rfloor C\right) \right] \\ &\quad + \left[\varphi\left(x + \left\lfloor \frac{y-x}{C} \right\rfloor C\right) - \varphi(y) \right]. \end{aligned}$$

Since φ is non- (Δ, C) -decreasing on X , each term within a pair of square brackets above is less than or equal to Δ , then

$$\varphi(x) - \varphi(y) \leq \left\lfloor \frac{y-x}{C} \right\rfloor \Delta \leq m\Delta \leq K + m\Delta,$$

hence, φ is non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing for any $K \geq 0$. ■

Property 2 *Non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing is equivalent to non- (Δ, C) -decreasing, when $K = 0$.*

Proof of Property 2: Since we have already proved Property 1, we only need to prove that a non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing function is also non- (Δ, C) -decreasing when $K = 0$. Suppose function φ is non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing with $K = 0$ on the domain X . Choose $m = 1$, then for $\forall x \in X$ and $y \in [x, x + C]$, we have $\varphi(x) - \varphi(y) \leq \Delta$. Thus, φ is non- (Δ, C) -decreasing. This completes the proof. ■

Property 3 *If f is a probability density function of a non-negative random variable, and φ is non- (Δ, C) -decreasing (or non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing) on \mathbb{R} , then the convolution $\varphi * f$ is also non- (Δ, C) -decreasing (or non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing) on \mathbb{R} .*

Proof of Property 3: First, we'll prove the property for function φ that is non- (Δ, C) -decreasing. For any $x, y \in \mathbb{R}$, and $y \in [x, x + C]$, it always holds that

$$\begin{aligned} (\varphi * f)(x) - (\varphi * f)(y) &= \int_0^{\infty} \varphi(x - z)f(z)dz - \int_0^{\infty} \varphi(y - z)f(z)dz \\ &= \int_0^{\infty} [\varphi(x - z) - \varphi(y - z)] f(z)dz \\ &\leq \int_0^{\infty} \Delta f(z)dz = \Delta. \end{aligned}$$

Therefore, $\varphi * f$ is also non- (Δ, C) -decreasing on \mathbb{R} . The case of non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing is proved in a similar manner. ■

III.3. Model Analysis

In Sublevel III.3.1, we first discuss the structure properties of the optimal ordering policy, then we provide sufficient conditions under which the proposed policy is optimal for a finite horizon problem. In Sublevel III.3.2, we analyze a special case where the optimal policy can be characterized in a simple form.

III.3.1. Optimal Ordering Policy

The recursive part of the cost-to-go function (3.6) is composed of the replenishment cost W and the function G_n . Since the structure of W is known, the crucial point to solve this problem is to analyze the characteristics of G_n . In preparation for Theorem 1, we introduce a special family of functions of interest.

Definition 3 *For given positive parameters Δ and C , define a family \mathcal{G} of function G that satisfies the following conditions:*

(A1) $G(x)$ is continuous.

(A2) $G(x) \rightarrow \infty$ as $|x| \rightarrow \infty$.

(A3) There exists $r \in \mathbb{R}$, such that

(A3.a) $G(x)$ is non- (Δ, C) -decreasing on $[r, \infty)$;

(A3.b) $G(x)$ is decreasing on $(-\infty, r]$;

(A3.c) $G(x - C) - G(x)$ is non-increasing on $(-\infty, r]$, and for any $x \in (-\infty, r]$,
 $G(x - C) - G(x) > \Delta$.

Conditions (A1) and (A2) guarantee the existence of an optimal order-up-to level corresponding to each beginning inventory. Condition (A3) will be used to show the specific features of the proposed replenishment policy.

Theorem 1 *If $G_n(\cdot) \in \mathcal{G}$, then the optimal replenishment policy in period n can be determined by the values of three sets of parameters: $(Q_n, \vec{s}_n, \vec{S}_n)$, where $\vec{s}_n = [s_n^1, s_n^2, \dots, s_n^M]$ and $\vec{S}_n = [S_n^1, S_n^2, \dots, S_n^M]$ are M -dimensional vectors. The parameters satisfy the following condition: $Q_n \leq s_n^1 < S_n^1 < s_n^2 < S_n^2 < \dots < s_n^M < S_n^M \leq s_n^1 + C$.*

The optimal order-up-to level $Y_n(x)$ can be represented as:

$$Y_n(x) = \begin{cases} x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C, & \text{if } x < Q_n \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [s_n^i, S_n^i], i = 1, \dots, M, \\ S_n^i, & \text{if } x < Q_n \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [S_n^i, s_n^{i+1}], i = 1, \dots, M - 1, \\ S_n^M, & \text{if } x < Q_n \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [S_n^M, s_n^1 + C), \\ x, & \text{if } x \geq Q_n. \end{cases} \quad (3.8)$$

Furthermore, if $K = 0$, $Q_n = s_n^1$.

Proof of Theorem 1: The proof consists of two steps. In the first step we design a procedure to find three sets of parameters for a given function that belongs to \mathcal{G} and name them $(Q_n, \vec{s}_n, \vec{S}_n)$. In the second step we prove that the optimal order-up-to level $Y_n(x)$ can be represented in these parameters.

Step 1: Determine the values of $(Q_n, \vec{s}_n, \vec{S}_n)$ parameters.

For a given function $G_n \in \mathcal{G}$, define s_n^1 as follows:

$$s_n^1 = \min \{s \in \mathbb{R} : G_n(x) \leq G_n(y) + \Delta, \text{ for } s \leq x \leq y \leq x + C\}. \quad (3.9)$$

s_n^1 can be thought of as the smallest real number such that on $[s_n^1, \infty)$ G_n is non- (Δ, C) -decreasing. Since it is assumed G_n is non- (Δ, C) -decreasing on $[r, \infty)$ for some real number r , and s_n^1 is the smallest such value, $s_n^1 \leq r$. In addition, this definition also implies

$$G_n(s_n^1) = \min_{y \in (s_n^1, s_n^1 + C]} G_n(y) + \Delta. \quad (3.10)$$

Furthermore, since $G_n(x)$ is decreasing on $(-\infty, s_n^1]$, $\lim_{x \uparrow s_n^1} G_n'(x) \leq 0$. Actually, we can show that $\lim_{x \downarrow s_n^1} G_n'(x) \leq 0$. The proof is as follows: By (3.10), there exists $y_0 \in (s_n^1, s_n^1 + C]$, such that $G_n(s_n^1) = G_n(y_0) + \Delta$. Suppose by contradiction that $\lim_{x \downarrow s_n^1} G_n'(x) > 0$, we can increase s_n^1 by a sufficiently small value and get an s_n^{1+} ,

such that $s_n^1 < s_n^{1+} < y_0$ and $G_n(s_n^1) < G_n(s_n^{1+})$. Therefore, $G_n(s_n^{1+}) > G_n(y_0) + \Delta$ which contradicts with the fact that G_n is non- (Δ, C) -decreasing at s_n^{1+} .

Also define Q_n as

$$Q_n = \min \{q \in \mathbb{R} : G_n(x) \leq G_n(y) + (K + m\Delta), \text{ for } q \leq x \leq y \leq x + mC, \forall m \in \mathbb{N}\}. \quad (3.11)$$

Similarly, Q_n is the smallest value such that on $[Q_n, \infty)$ G_n is non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing. Since G_n is non- (Δ, C) -decreasing on $[s_n^1, \infty)$, by Property 1 it is also non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing in this interval. It follows that $Q_n \leq s_n^1$, and the equality holds if and only if $K = 0$, because when $K = 0$, Property 2 implies that non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing is equivalent to non- (Δ, C) -decreasing.

To find all remaining parameters, we first define a set \mathcal{L} of points such that for any $l \in \mathcal{L}$,

- $l \in (s_n^1, s_n^1 + C]$, where s_n^1 is found in (3.9) and C is the cargo capacity;
- there exists an $\varepsilon > 0$, such that for any $x \in [l - \varepsilon, l + \varepsilon] \cap (s_n^1, s_n^1 + C]$, $G_n(x) \geq G_n(l)$, and for any $x \in [l - \varepsilon, l) \cap (s_n^1, s_n^1 + C]$, $G_n(x) > G_n(l)$.

One way to think of the points in \mathcal{L} is as follows: They are the local minimizers of function G_n over the interval of $(s_n^1, s_n^1 + C]$. Also, within a sufficiently small neighborhood of each point, function G_n is strictly decreasing to its left. Since function G_n is continuous and bounded on the compact set $[s_n^1, s_n^1 + C]$, the number of points in \mathcal{L} is finite. Also, since $\lim_{x \downarrow s_n^1} G_n'(x) \leq 0$, $\mathcal{L} \neq \emptyset$.

We apply the following method for determining the values of S_n^1, \dots, S_n^M .

[1]: $i = 1$. Let $S_n^1 = \min \mathcal{L}$ and go to [2].

[2]: Let $\mathcal{L}_{i+1} = \{l \in \mathcal{L} : l > S_n^i \text{ and } G_n(l) < G_n(S_n^i)\}$. If \mathcal{L}_{i+1} is empty, let $M = i$ and STOP. Otherwise, go to [3].

[3]: $S_n^{i+1} = \min \mathcal{L}_{i+1}$, $i = i + 1$ and go to [2].

Obviously,

$$S_n^M = \arg \min_{y \in (s_n^1, s_n^1 + C]} G_n(y). \quad (3.12)$$

When we have the sequence of $\{S_n^i\}_{i=1, \dots, M}$, let

$$\mathcal{T}_i = \{x : x \in (S_n^{i-1}, S_n^i), G_n(x) = G_n(S_n^{i-1})\}$$

for $i = 2, \dots, M$. Defining $s_n^i = \max \mathcal{T}_i$ for $i = 2, \dots, M$, we have all parameter values determined. According to the choice of these values, it is guaranteed that $G_n(x)$ is non-increasing on $(s_n^i, S_n^i]$ for $i = 1, \dots, M$, and $Q_n \leq s_n^1 < S_n^1 < s_n^2 < S_n^2 < \dots < s_n^M < S_n^M \leq s_n^1 + C$.

Figure 3: An illustration of value determination

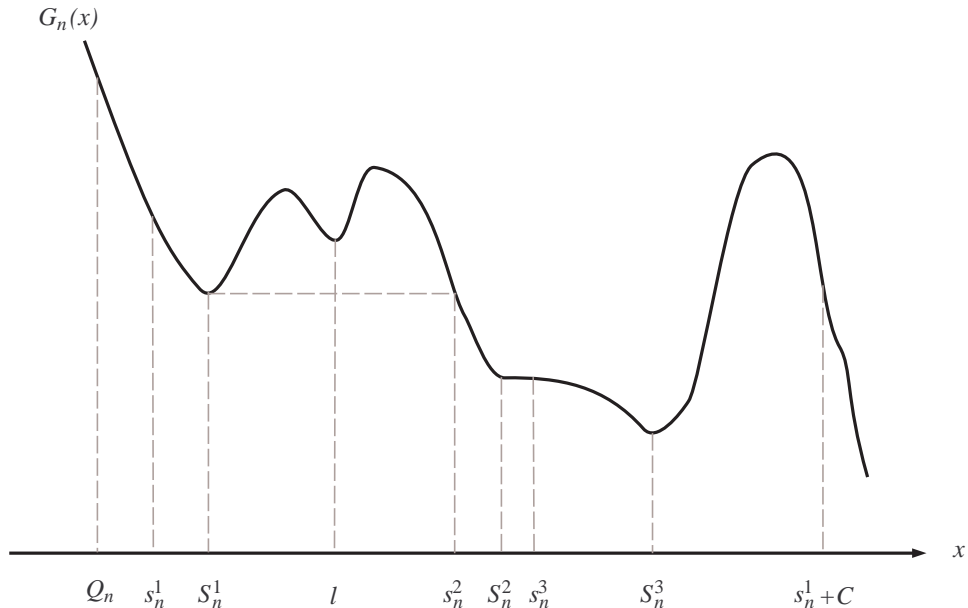


Figure 3 gives an illustration of the value determination procedure. Given a function G_n , we first use equations (3.9) and (3.11) to determine the values of s_n^1 and

Q_n . Second, we look at the region $(s_n^1, s_n^1 + C]$, and find set $\mathcal{L} = \{l_1, l_2, \dots, l_J\}$, where J is a finite positive integer ($\mathcal{L} = \{S_n^1, l, S_n^2, S_n^3\}$ in this example). Then applying the methods above, we can choose all the other values shown on the figure. Note that, in this example l is only a point of \mathcal{L} but not chosen as one of the policy parameters.

Step 2: Identify the optimal order-up-to level.

The vendor needs to make a decision of how much to order at the beginning of each period based on its beginning inventory level. Now suppose that the beginning inventory is x (in the following proof, x is assumed to be fixed), let's examine the structure of the optimal replenishment policy by looking at the optimal order-up-to level $Y_n(x)$.

First, let $u_n(y|x)$ represent the cost of ordering up to y in period n when the beginning inventory of period n is x and optimal decisions are made onward, i.e.,

$$u_n(y|x) = W(y-x) + L(y) + \beta \int_0^\infty V_{n-1}(y-z)f(z)dz = W(y-x) + G_n(y). \quad (3.13)$$

Then, we can rewrite the optimality equation (3.6) as

$$V_n(x) = \begin{cases} \min_{y \geq x} \{u_n(y|x)\}, & n = 1, \dots, N, \\ g_T(x), & n = 0. \end{cases} \quad (3.14)$$

We need to discuss on the value of the given number x .

Case 1: $x \geq Q_n$.

In this region G_n is non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing. If one chooses to order up to $y > x$, let $m = \lceil \frac{y-x}{C} \rceil$, then $y \in (x, x + mC]$ and $G_n(x) \leq G_n(y) + K + \lceil \frac{y-x}{C} \rceil \Delta$. It follows directly that $u_n(x|x) \leq u_n(y|x)$ for any $y > x$. This inequality implies that when the initial inventory level $x \geq Q_n$, it is never optimal to place an order, i.e., the optimal order up to level is $Y_n(x) = x$.

Case 2: $x < Q_n$.

Case 2.1: Consider the order-up-to level y that satisfies $y \in \left(x, x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C\right]$.

- **Case 2.1.1:** If $x = x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C$, this is an empty set.
- **Case 2.1.2:** If $x \neq x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C$, $G_n(y)$ is decreasing in this region, because $x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C \leq s_n^1$ for all $x < s_n^1$ and $s_n^1 \leq r$, where r is a real number to the left of which function G_n is decreasing. Note that the replenishment cost W is a staircase function, thus, it is piecewise constant and left continuous. It can now be seen that for any given beginning inventory level $x < Q_n$ and order-up-to level $y \in \left(x, x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C\right]$, $u_n(y|x) = W(y-x) + G_n(y)$ is piecewise decreasing and left continuous. Hence, if we want to find the minimizer of $u_n(y|x)$ over this region, we only need to consider the breakpoints, i.e., the points in the set $\left\{y : y = x + mC, m = 1, \dots, \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor\right\}$.

Now, by Condition (A3.c) $G_n(y-C) - G_n(y) > \Delta$ for $y \in \left(x, x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C\right]$, we have $G_n(x+C) + K + \Delta > G_n(x+2C) + K + 2\Delta > \dots > G_n\left(x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C\right) + K + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor \Delta$. Equivalently,

$$u_n(x+C|x) > u_n(x+2C|x) > \dots > u_n\left(x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C|x\right).$$

Thus, $u_n\left(x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C|x\right) \leq u_n(y|x)$ for any $y \in \left(x, x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C\right]$, and $x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C$ is a candidate for the optimal order-up-to level.

Case 2.2: Consider $y \in \left(x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C, s_n^1\right]$.

- **Case 2.2.1:** If $s_n^1 = x + mC$ for some m , we can find that $m = \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor = \left\lceil \frac{s_n^1 - x}{C} \right\rceil$. Thus, $x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C = s_n^1$, and this is an empty set.
- **Case 2.2.2:** If $s_n^1 \neq x + mC$ for any m , we have $x + \left\lfloor \frac{s_n^1 - x}{C} \right\rfloor C < s_n^1$ and for any

$$y \in \left(x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C, s_n^1 \right]$$

$$\begin{aligned} u_n(y|x) &= G_n(y) + K + \left\lceil \frac{s_n^1 - x}{C} \right\rceil \Delta \\ &\geq G_n(s_n^1) + K + \left\lceil \frac{s_n^1 - x}{C} \right\rceil \Delta = u_n(s_n^1|x). \end{aligned}$$

Thus, $u_n(s_n^1|x) \leq u_n(y|x)$ for any $y \in \left(x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C, s_n^1 \right]$, and s_n^1 is a candidate for the optimal order-up-to level.

Case 2.3: Consider $y \in [s_n^1, \infty)$.

By Condition (A3.a), for $y \geq s_n^1 + C$,

$$G_n(y - C) + W(y - C - x) \leq G_n(y) + \Delta + W(y - C - x) = G_n(y) + W(y - x).$$

This implies that if the order-up-to level y is at least $s_n^1 + C$, we can achieve a lower total expected cost by decreasing y by C . In other words, if the order-up-to level has to be greater than or equal to s_n^1 , we only need to search the region $[s_n^1, s_n^1 + C)$ for the one with an minimal cost. Furthermore, since $s_n^1 \leq x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C < s_n^1 + C$ for any $x < s_n^1$, and G_n is non- (Δ, C) -decreasing on $[s_n^1, \infty)$, for $y \in \left(x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C, s_n^1 + C \right)$, $G_n\left(x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C\right) \leq G_n(y) + \Delta$. Hence, we can further shrink the search region to the interval $\left[s_n^1, x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \right]$.

The analysis for **Case 2** implies that when $x \leq Q_n$, the optimal order-up-to level has to belong to the set $\{x\} \cup \left\{ x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \right\} \cup \left[s_n^1, x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \right]$. However, if $y = x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C$, according to Condition (A3) and the definition of s_n^1 , we can find $y_2 \in \left[x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C, x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \right]$, such that $u_n(y_2|x) \leq u_n\left(x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C|x\right)$. Similarly, if $y = x$, we can also find $y_3 \in \left[x, x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \right]$, such that $u_n(y_3|x) \leq G_n(x) = u_n(x|x)$. Thus, the optimal order-up-to level should be within the region of $\left[s_n^1, x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \right]$.

For any $y \in \left[s_n^1, x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \right]$,

$$\begin{aligned} G_n(y) + W(y - x) &= G_n(y) + K + \left\lceil \frac{y - x}{C} \right\rceil \Delta \\ &\geq \min_{z \in \left[s_n^1, x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \right]} G_n(z) + K + \left\lceil \frac{s_n^1 - x}{C} \right\rceil \Delta. \end{aligned}$$

Consequently, for $x < Q_n$, the optimal order-up-to level is

$$\begin{aligned} Y_n(x) &= \arg \min_{z \in \left[s_n^1, x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \right]} G_n(z) \\ &= \begin{cases} x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C, & \text{if } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [s_n^1, S_n^1), \\ S_n^1, & \text{if } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [S_n^1, s_n^2), \\ x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C, & \text{if } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [s_n^2, S_n^2), \\ S_n^2, & \text{if } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [S_n^2, s_n^3), \\ \dots & \\ S_n^M, & \text{if } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [S_n^M, s_n^1 + C). \end{cases} \end{aligned}$$

In conclusion, the optimal order-up-to level is

$$Y_n(x) = \begin{cases} x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C, & \text{if } x < Q_n \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [s_n^i, S_n^i), i = 1, \dots, M, \\ S_n^i, & \text{if } x < Q_n \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [S_n^i, s_n^{i+1}), i = 1, \dots, M - 1, \\ S_n^M, & \text{if } x < Q_n \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [S_n^M, s_n^1 + C), \\ x, & \text{if } x \geq Q_n, \end{cases}$$

and this completes the proof of Theorem 1. ■

Following Theorem 1 directly, we have the optimal replenishment quantity $a_n^*(x)$ and the optimal expected total cost $V_n(x)$ as

$$a_n^*(x) = \begin{cases} \left\lceil \frac{s_n^1 - x}{C} \right\rceil C, & \text{if } x < Q_n \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [s_n^i, S_n^i], i = 1, \dots, M, \\ S_n^i - x, & \text{if } x < Q_n \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [S_n^i, s_n^{i+1}), i = 1, \dots, M - 1, \\ S_n^M - x, & \text{if } x < Q_n \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [S_n^M, s_n^1 + C), \\ 0, & \text{if } x \geq Q_n, \end{cases} \quad (3.15)$$

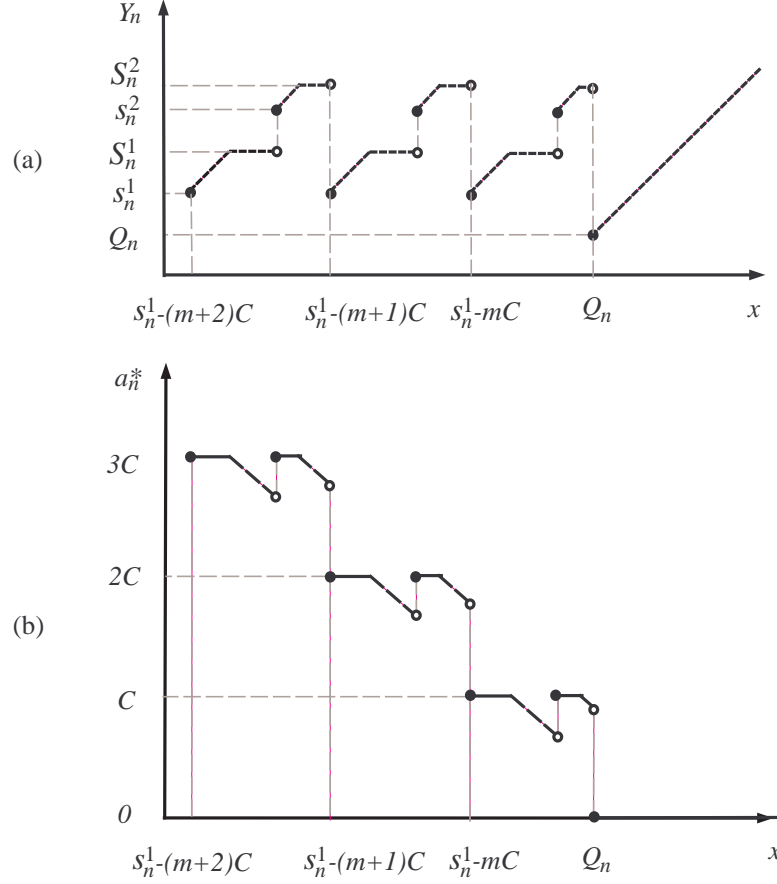
and

$$V_n(x) = \begin{cases} K + \left\lceil \frac{s_n^1 - x}{C} \right\rceil \Delta + G_n \left(x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \right), & \text{if } x < Q_n \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [s_n^i, S_n^i], i = 1, \dots, M, \\ K + \left\lceil \frac{s_n^1 - x}{C} \right\rceil \Delta + G_n (S_n^i), & \text{if } x < Q_n \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [S_n^i, s_n^{i+1}), i = 1, \dots, M - 1, \\ K + \left\lceil \frac{s_n^1 - x}{C} \right\rceil \Delta + G_n (S_n^M), & \text{if } x < Q_n \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [S_n^M, s_n^1 + C), \\ G_n(x), & \text{if } x \geq Q_n. \end{cases} \quad (3.16)$$

Theorem 1 states that if $G_n \in \mathcal{G}$, then there exists a threshold value Q_n such that it is optimal to idle if the beginning inventory x is at least Q_n . Otherwise, the optimal replenishment quantity can be determined by x and the parameters of s_n^1, \dots, s_n^M and S_n^1, \dots, S_n^M . We call this replenishment policy (Q, \vec{s}, \vec{S}) policy. The algorithm for solving the values of the policy parameters can be found in the proof of Theorem 1.

Figure 4 gives an illustration of the (Q, \vec{s}, \vec{S}) policy. In this figure, x denotes the beginning inventory level before a replenishment order is placed. The sawtooth-like

Figure 4: An Illustration of (Q, \vec{s}, \vec{S}) Policy $\left(m = \left\lfloor \frac{s_n^1 - Q_n}{C} \right\rfloor + 1\right)$



function in (a) represents the optimal order-up-to level and the decreasing staircase-like function in (b) represents the optimal replenishment quantity. From this figure, we can see that depending on the beginning inventory level x , sometimes the optimal replenishment quantity is an integral multiple of the cargo capacity, i.e., Full TruckLoad (FTL) quantity, and sometimes it includes one partially loaded truck, i.e., Less than TruckLoad (LTL) quantity. On an interval of beginning inventory level x , if the optimal replenishment decision includes an LTL quantity, then the corresponding order-up-to level is constant.

The (Q, \vec{s}, \vec{S}) policy has the following interpretations: (1) Q_n is the minimal beginning inventory level that does not require replenishment in period n ; (2) Elements of \vec{s}_n are the threshold values of the optimal order-up-to level whose corresponding optimal replenishment quantity changes from LTL to FTL. s_n^1 is the minimal order-up-to level if a replenishment is necessary in period n ; (3) Elements of \vec{S}_n are the threshold values of the optimal order-up-to level whose corresponding optimal replenishment quantity changes from FTL to LTL. S_n^M is the maximal order-up-to level if a replenishment is necessary in period n ; (4) M is the maximal number of LTL order-up-to levels.

The sufficient condition in Theorem 1 requires function G_n to belong to the family of functions \mathcal{G} . However, in general, successive G_n functions are not guaranteed to satisfy this condition. It is necessary to check the function G_n for each period to decide whether the (Q, \vec{s}, \vec{S}) policy is optimal for that period. In the next theorem, we give sufficient conditions to ensure that G_n belongs to \mathcal{G} for any period. To this end, we introduce the following conditions:

(B1) $K = 0$.

(B2) $g_T(x) = 0$.

(B3) The one-period expected holding and shortage cost L defined by (3.3) satisfies

$$L(P - C) - L(P) > \Delta, \text{ where } P \text{ is the unique minimizer determined by (3.7).}$$

Condition (B1) implies that the administrative cost of processing the replenishment order is trivial compared with the transportation cost, hence, it can be omitted. Condition (B2) means that if there is excess inventory or outstanding backorders at the end of the planning horizon, no costs are incurred. Condition (B3) indicates that the one-period expected holding and shortage cost for replenishing the inventory level up to P is less than that for replenishing one less full truckload, where P is the order-up-to level that minimizes the one-period expected holding and shortage

cost. In other words, if the vendor orders up-to P instead of $P - C$, although the incremental cost Δ needs to be incurred for the usage of an extra truck, the expected saving in current period's holding and shortage cost is more than that.

The following three lemmas are used to prove Theorem 2. In Lemmas 2 and Lemma 3, a point, say PT , is the “first increasing point” of function φ when PT is the greatest number such that φ is non-increasing on $(-\infty, PT]$.

Lemma 1 *Given a continuous and convex function $\varphi(x)$, and a positive constant C , $\varphi(x - C) - \varphi(x)$ is non-increasing in x .*

Proof of Lemma 1: We need to prove that $\varphi(x - C) - \varphi(x)$ is non-increasing in x if $\varphi(x)$ is continuous and convex.

Suppose there are two real numbers x_1 and x_2 , and $x_1 < x_2$, then for a given constant C ,

$$\min \{x_1, x_2, x_1 - C, x_2 - C\} = x_1 - C \quad \text{and} \quad \max \{x_1, x_2, x_1 - C, x_2 - C\} = x_2.$$

If the four values of x_1 , x_2 , $x_1 - C$ and $x_2 - C$ are compared with each other, the situation can only be

$$x_1 - C < x_1 \leq x_2 - C < x_2 \quad \text{or} \quad x_1 - C < x_2 - C \leq x_1 < x_2.$$

If $x_1 - C < x_1 \leq x_2 - C < x_2$ is the case, let's define two real numbers ξ and η as follows:

$$\xi = \frac{\varphi(x_1) - \varphi(x_1 - C)}{C} \quad \text{and} \quad \eta = \frac{\varphi(x_2) - \varphi(x_2 - C)}{C}.$$

Letting $(D_l\varphi)(x)$ and $(D_r\varphi)(x)$ be the left- and right- derivatives of φ at x defined as

$$(D_l\varphi)(x) := \lim_{\delta \downarrow 0} \frac{\varphi(x + \delta) - \varphi(x)}{\delta} \quad \text{and} \quad (D_r\varphi)(x) := \lim_{\delta \downarrow 0} \frac{\varphi(x + \delta) - \varphi(x)}{\delta},$$

we have

$$(D_r\varphi)(x_1 - C) \leq \xi \leq (D_l\varphi)(x_1) \quad \text{and} \quad (D_r\varphi)(x_2 - C) \leq \eta \leq (D_l\varphi)(x_2).$$

Since $\varphi(x)$ is convex and $x_1 \leq x_2 - C$ as assumed,

$$(D_l\varphi)(x_1) \leq (D_r\varphi)(x_1) \leq (D_r\varphi)(x_2 - C).$$

Thus $\xi \leq \eta$, which implies

$$\varphi(x_1 - C) - \varphi(x_1) \geq \varphi(x_2 - C) - \varphi(x_2).$$

The case of $x_1 - C < x_2 - C \leq x_1 < x_2$ is showed in a similar manner. Thus, $\varphi(x - C) - \varphi(x)$ is non-increasing in x , and Lemma 1 is proved. \blacksquare

Lemma 2 *If Conditions (B1)–(B3) are satisfied, and for period n , $G_n \in \mathcal{G}$ with the critical number r_n in Condition (A3) taking the value of P (the unique minimizer of function L), then the cost-to-go function $V_n(x)$ satisfies:*

(C1) $V_n(x)$ is continuous.

(C2) $V_n(x) \rightarrow \infty$ as $|x| \rightarrow \infty$.

(C3) $V_n(x)$ is non- (Δ, C) -decreasing on \mathbb{R} .

(C4) V_n 's first increasing point $P_n \geq P$, and $V_n(x)$ is non-increasing on $(-\infty, P_n]$.

(C5) $V_n(x - C) - V_n(x) = \Delta$ on $(-\infty, P]$.

Proof of Lemma 2: Since $G_n \in \mathcal{G}$, the optimal order-up-to level $Y_n(x)$ for period n can be determined according to Theorem 1. Furthermore, with the Condition (B1) that $K = 0$, we have $Q_n = s_n^1$ and the optimal order-up-to level (3.8) can be rewritten

as:

$$Y_n(x) = \begin{cases} x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C, & \text{if } x < s_n^1 \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [s_n^i, S_n^i), i = 1, \dots, M, \\ S_n^i, & \text{if } x < s_n^1 \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [S_n^i, s_n^{i+1}), i = 1, \dots, M - 1, \\ S_n^M, & \text{if } x < s_n^1 \text{ and } x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C \in [S_n^M, s_n^1 + C), \\ x, & \text{if } x \geq s_n^1. \end{cases} \quad (3.17)$$

The corresponding cost-to-go function is

$$V_n(x) = W(Y_n(x) - x) + G_n(Y_n(x)).$$

Now, let's examine the Conditions (C1)-(C5) one by one.

Condition (C1):

The proof is developed by showing that $V_n(x)$ is continuous on $[s_n^1, \infty)$ and $(-\infty, s_n^1)$ respectively, and also continuous at the breakpoint s_n^1 .

(1). $V_n(x)$ is continuous on $[s_n^1, \infty)$.

According to (3.17), for $x \geq s_n^1$ we have $Y_n(x) = x$, thus $V_n(x) = G_n(x)$. Obviously, $V_n(x)$ is continuous, since G_n belongs to \mathcal{G} and is continuous.

(2). $V_n(x)$ is continuous on $(-\infty, s_n^1)$.

Breaking this segment into pieces with length C , we have

$$(-\infty, s_n^1) = \bigcup_{j \in \mathbb{N}} [s_n^1 - jC, s_n^1 - (j-1)C). \quad (3.18)$$

Consider x on any one of these pieces, say, $x \in [s_n^1 - mC, s_n^1 - (m-1)C)$ where m is an arbitrary natural number, then $m-1 < \frac{s_n^1 - x}{C} \leq m$, hence, $\left\lceil \frac{s_n^1 - x}{C} \right\rceil = m$. Using the policy values of \vec{s}_n, \vec{S}_n , we can further divide $[s_n^1 - mC, s_n^1 - (m-1)C)$ into

$$\left(\bigcup_{i=1}^M [s_n^i - mC, S_n^i - mC) \right) \cup \left(\bigcup_{i=1}^{M-1} [S_n^i - mC, s_n^{i+1} - mC) \right)$$

$$\bigcup [S_n^M - mC, s_n^1 - (m-1)C]. \quad (3.19)$$

Case 1: If $x \in [s_n^i - mC, S_n^i - mC)$, $i = 1, \dots, M$, according to (3.17) we obtain $Y_n(x) = x + \left\lceil \frac{s_n^1 - x}{C} \right\rceil C = x + mC$. Then

$$V_n(x) = m\Delta + G_n(x + mC), \quad (3.20)$$

and it is continuous on this interval.

Case 2: If $x \in [S_n^i - mC, s_n^{i+1} - mC)$, $i = 1, \dots, M-1$, according to (3.17) we have $Y_n(x) = S_n^i$. Then

$$V_n(x) = m\Delta + G_n(S_n^i), \quad (3.21)$$

and $V_n(x)$ is constant hence continuous on this interval.

Case 3: If $x \in [S_n^M - mC, s_n^1 - (m-1)C)$, similarly, $Y_n(x) = S_n^M$. Then

$$V_n(x) = m\Delta + G_n(S_n^M), \quad (3.22)$$

and $V_n(x)$ is also constant and continuous on this interval.

Since we have already shown that $V_n(x)$ is continuous on each subset of (3.19), in order to have $V_n(x)$ continuous on $(-\infty, s_n^1)$, we still need to examine its continuity at the breakpoints $\{s_n^i - mC, S_n^i - mC\}_{i=1, \dots, M}$ and $m \in \mathbb{N}$. The function φ is continuous at a given point x_0 if and only if the limit of $\varphi(x)$ as x approaches x_0 exists and is equal to $\varphi(x_0)$.

Part 1: Examining breakpoints $\{s_n^1 - mC\}_{m \in \mathbb{N}}$, by (3.22), we have the left-handed limit of function $V_n(x)$ at $s_n^1 - mC$ as $\lim_{x \uparrow (s_n^1 - mC)} V_n(x) = (m+1)\Delta + G_n(S_n^M)$, and by (3.20), $V_n(x)$ is right continuous at $s_n^1 - mC$ and $V_n(s_n^1 - mC) = \lim_{x \downarrow (s_n^1 - mC)} V_n(x) = m\Delta + G_n(s_n^1)$. Substituting (3.12) in equation (3.10), we have

$$G_n(s_n^1) = G_n(S_n^M) + \Delta, \quad (3.23)$$

hence, $\lim_{x \uparrow (s_n^1 - mC)} V_n(x) = \lim_{x \downarrow (s_n^1 - mC)} V_n(x) = V_n(s_n^1 - mC)$.

Part 2: For breakpoints $\{s_n^i - mC\}_{i=2, \dots, M, m \in \mathbb{N}}$, (3.21) and (3.20) give us the following results: $\lim_{x \uparrow (s_n^i - mC)} V_n(x) = m\Delta + G_n(S_n^{i-1})$. and

$$V_n(s_n^i - mC) = \lim_{x \downarrow (s_n^i - mC)} V_n(x) = m\Delta + G_n(s_n^i).$$

By the choice of s_n^i , we have

$$G_n(s_n^i) = G_n(S_n^{i-1}), \quad (3.24)$$

hence, $\lim_{x \uparrow (s_n^i - mC)} V_n(x) = \lim_{x \downarrow (s_n^i - mC)} V_n(x) = V_n(s_n^i - mC)$.

Part 3: For breakpoints $\{S_n^i - mC\}_{i=1, \dots, M, m \in \mathbb{N}}$, similarly to the previous two cases, using (3.20), (3.21) and (3.22), we have $\lim_{x \uparrow (S_n^i - mC)} V_n(x) = m\Delta + G_n(S_n^i)$, and $V_n(S_n^i - mC) = \lim_{x \downarrow (S_n^i - mC)} V_n(x) = m\Delta + G_n(S_n^i)$. Thus, $\lim_{x \uparrow (S_n^i - mC)} V_n(x) = \lim_{x \downarrow (S_n^i - mC)} V_n(x) = V_n(S_n^i - mC)$.

(3). $V_n(x)$ is continuous at s_n^1 .

Still, we find the left-handed limit of $V_n(x)$ at s_n^1 as $\lim_{x \uparrow s_n^1} V_n(x) = G_n(S_n^M) + \Delta$. Since $V_n(x)$ is right continuous at s_n^1 , $V_n(s_n^1) = \lim_{x \downarrow s_n^1} V_n(x) = G_n(s_n^1)$. By (3.23), we have $\lim_{x \uparrow s_n^1} V_n(x) = \lim_{x \downarrow s_n^1} V_n(x) = V_n(s_n^1)$. Thus V_n is continuous at s_n^1 , and hence continuous on the whole real line.

Condition (C2):

We need to prove $V_n(x) \rightarrow \infty$ as $|x| \rightarrow \infty$. Since $V_n(x) = G_n(x)$ when $x \geq s_n^1$, and $G_n(x) \in \mathcal{G}$, it is obvious that $V_n(x) \rightarrow \infty$ as $x \rightarrow \infty$. And when x goes to $-\infty$, (3.16) implies

$$\begin{aligned} \lim_{x \rightarrow -\infty} V_n(x) &= \lim_{x \rightarrow -\infty} \left\lceil \frac{s_n^1 - x}{C} \right\rceil \Delta + G_n(\bar{x}) \\ &= \infty + G_n(\bar{x}) = \infty, \end{aligned}$$

where $G_n(\bar{x})$ is finite because $\bar{x} \in [s_n^1, s_n^1 + C]$.

Condition (C3):

We need to prove $V_n(x)$ is non- (Δ, C) -decreasing on \mathbb{R} . Similarly to the proof of Condition (C1), we divide the real number field into analyzable sub-pieces.

(1). $V_n(x)$ is non- (Δ, C) -decreasing on $[s_n^1, \infty)$.

It is known that $G_n(x)$ is non- (Δ, C) -decreasing on $[s_n^1, \infty)$. Since $V_n(x) = G_n(x)$ on $[s_n^1, \infty)$, $V_n(x)$ is also non- (Δ, C) -decreasing on $[s_n^1, \infty)$.

(2). $V_n(x)$ is non- (Δ, C) -decreasing on $(-\infty, s_n^1)$.

Similarly to the arguments in proof of Condition (C1), we also break this segment into pieces with length C , i.e.,

$$(-\infty, s_n^1) = \bigcup_{j \in \mathbb{N}} [s_n^1 - jC, s_n^1 - (j-1)C]. \quad (3.25)$$

And each piece $[s_n^1 - mC, s_n^1 - (m-1)C]$ is again rewritten as (3.19) for $m = 1, \dots, M$. Note that when $x + mC \in (s_n^i, S_n^i]$, $G_n(x + mC)$ is non-increasing as x increases, hence, by (3.20), (3.21) and (3.22) we see that $V_n(x)$ is non-increasing on $(-\infty, s_n^1)$.

Case 1: If $x \in (-\infty, s_n^1 - C)$, for $y \in [x, x+C]$ we have $V_n(x) - V_n(y) \leq V_n(x) - V_n(x+C) = \Delta$, which gives the result that $V_n(x)$ is non- (Δ, C) -decreasing on $(-\infty, s_n^1 - C)$.

Case 2: If $x \in [s_n^i - C, S_n^i - C)$, $i = 1, \dots, M$, $V_n(x) = G_n(x + C) + \Delta$. Since $V_n(x)$ is continuous and non-increasing on $(-\infty, s_n^1]$,

$$\begin{aligned} \min_{y \in [x, x+C]} V_n(y) &= \min_{y \in [s_n^1, x+C]} V_n(y) \\ &= \min_{y \in [s_n^1, x+C]} G_n(y) = G_n(x + C). \end{aligned}$$

Thus, $V_n(x) = \min_{y \in [x, x+C]} V_n(y) + \Delta \leq V_n(y) + \Delta$ for $y \in [x, x + C]$, consequently, $V_n(x)$ is non- (Δ, C) -decreasing on this region.

Case 3: If $x \in [S_n^i - C, s_n^{i+1} - C]$ for $i = 1, \dots, M-1$, or $x \in [S_n^M - C, s_n^1]$ for $i = M$, then $V_n(x) = G_n(S_n^i) + \Delta$ for $i = 1, \dots, M$, respectively. Since $V_n(x)$ is continuous and non-increasing on $(-\infty, s_n^1]$,

$$\begin{aligned} \min_{y \in [x, x+C]} V_n(y) &= \min_{y \in [s_n^1, x+C]} V_n(y) \\ &= \min_{y \in [s_n^1, x+C]} G_n(y) = G_n(S_n^i). \end{aligned}$$

Thus, $V_n(x) = \min_{y \in [x, x+C]} V_n(y) + \Delta \leq V_n(y) + \Delta$ for $y \in [x, x+C]$, consequently, $V_n(x)$ is non- (Δ, C) -decreasing on this region. This completes the proof of Condition (C3).

Condition (C4):

We need to show the first increasing point P_n of V_n satisfies $P_n \geq P$. Since $G_n \in \mathcal{G}$, and on $[P, \infty)$, $G_n(x)$ is non- (Δ, C) -decreasing, we know $s_n^1 \leq P$, because s_n^1 is the smallest value that ensures $G_n(x)$ is non- (Δ, C) -decreasing on $[s_n^1, \infty)$. While from (3.17), $V_n(x) = G_n(x)$ when $x \geq s_n^1$, and $V_n(x)$ is non-increasing on $(-\infty, s_n^1)$, that means the first increasing point P_n of $V_n(x)$ is exactly the same as that of $G_n(x)$, which is greater than or equal to P , hence, on $(-\infty, P_n)$, $V_n(x)$ is non-increasing.

Condition (C5):

Since $G_n(x - C) - G_n(x) > \Delta$ on $(-\infty, P]$, $P < s_n^1 + C$. Also, the determination procedure of S_n^1 guarantees that $G_n(x)$ is decreasing on $(-\infty, S_n^1]$. Considering that $S_n^1 \leq s_n^1 + C$, we discuss two cases: (1) if $S_n^1 < s_n^1 + C$, then for any $x > S_n^1$, $G_n(x)$ is not strictly decreasing on $(-\infty, x]$, consequently, $P \leq S_n^1$; (2) if $S_n^1 = s_n^1 + C$, then $P < S_n^1$. Therefore, $s_n^1 \leq P \leq S_n^1$.

Using (3.20), (3.21), (3.22) and $V_n(x) = G_n(x)$ for $x \in [s_n^1, P]$, for any $x \leq P$, we have $V_n(x - C) - V_n(x) = \Delta$ by simple algebra. This completes the proof of Lemma 2. ■

Lemma 3 *If $f(\cdot)$ is the pdf of a nonnegative random variable, and for period n , the cost-to-go function $V_n(x)$ satisfies Conditions (C1)–(C5), then the convolution $(V_n * f)(x)$ satisfies:*

(D1) $(V_n * f)(x)$ is continuous.

(D2) $(V_n * f)(x) \rightarrow \infty$ as $|x| \rightarrow \infty$.

(D3) $V_n * f$ is non- (Δ, C) -decreasing on \mathbb{R} .

(D4) $(V_n * f)$'s first increasing point $P_n^f \geq P$.

(D5) $(V_n * f)(x - C) - (V_n * f)(x) = \Delta$ on $(-\infty, P]$.

Proof of Lemma 3: We prove this by examining Conditions (D1)–(D5) one by one.

Conditions (D1), (D2), and (D3):

Suppose $f(\cdot)$ is the pdf of a nonnegative random variable, since V_n is continuous and $V_n(x) \rightarrow \infty$ as $|x| \rightarrow \infty$, Conditions (D1) and (D2) are immediate. Condition (D3) is also immediate by Property 3.

Condition (D4):

Since V_n 's first increasing point $P_n \geq P$, then $V_n(x) \geq V_n(y)$ for $\forall x < y \leq P_n$. The difference that we are interested in is

$$(V_n * f)(x) - (V_n * f)(y) = \int_0^\infty [V_n(x - t) - V_n(y - t)] f(t) dt.$$

Note that, since the quantity of the demand is nonnegative, i.e., $t \geq 0$, thus, $x - t < y - t \leq P_n$. We still have $V_n(x - t) \geq V_n(y - t)$, which implies $(V_n * f)(x) - (V_n * f)(y) \geq 0$. Hence, $V_n * f$ is also non-increasing on $(-\infty, P_n]$, and this implies that its first increasing point $P_n^f \geq P_n \geq P$.

Condition (D5):

For $x \in (-\infty, P]$, by Condition (C5),

$$(V_n * f)(x - C) - (V_n * f)(x) = \int_0^\infty [V_n(x - C - t) - V_n(x - t)] f(t) dt = \Delta.$$

Thus, Lemma 3 is proved. ■

Based on Lemmas 1 through 3, we have Theorem 2 below. Specifically, we show that G_n preserves Conditions (A1)–(A3) for any period under assumptions (B1)–(B3). Consequently, the (Q, \vec{s}, \vec{S}) policy is optimal for a problem over a finite planning horizon.

Theorem 2 *If Conditions (B1)–(B3) are satisfied, then a $(Q_n, \vec{s}_n, \vec{S}_n)$ policy is optimal in any period for a finite horizon problem. Especially, $Q_n = s_n^1$.*

Proof of Theorem 2: It is sufficient to show that for all $n \leq N$, $G_n \in \mathcal{G}$. Note that $Q_n = s_n^1$ directly follows the result of Theorem 1. To obtain $G_n \in \mathcal{G}$, we need to show that Conditions (A1)–(A3) of Definition 3 hold for any n . In particular, we purposely choose P as the real number r in Condition (A3) for all periods, where P is the unique minimizer of L . The proof is by induction.

- When $n = 1$, since the terminal cost $g_T(x) = 0$, $G_1(x) = L(x)$. By the definition (3.3) of $L(\cdot)$, (A1) and (A2) hold. Since P is the unique minimizer of L , by Condition (B3) and Lemma 1, (A3.a)–(A3.c) hold with $r_1 = P$. Hence, $G_1(x) \in \mathcal{G}$.
- Suppose $G_n(x) \in \mathcal{G}$ with $r_n = P$ for $n = k$, we seek to prove that $G_{k+1}(x) \in \mathcal{G}$ with $r_{k+1} = P$. Since $G_k \in \mathcal{G}$ with $r_k = P$, by Conditions (B1)–(B3) and Lemma 2, $V_k(x)$ satisfies Conditions (C1)–(C5), and in turn, $V_k * f$ satisfies Conditions (D1)–(D5) by Lemma 3. By definition,

$$G_{k+1} = L(x) + \beta(V_k * f)(x). \quad (3.26)$$

Since $L(x)$ and $(V_k * f)(x)$ are both continuous and go to ∞ as $|x| \rightarrow \infty$, $G_{k+1}(x)$ is also continuous and goes to ∞ as $|x| \rightarrow \infty$. Thus (A1) and (A2) hold. We show (A3.a)–(A3.c) hold with $r_{k+1} = P$ one by one.

Condition (A3.a):

Since $V_k * f$ is non- (Δ, C) -decreasing on \mathbb{R} , it is also non- (Δ, C) -decreasing on $[P, \infty)$, i.e., for any $x \in [P, \infty)$ and $y \in [x, x + C]$,

$$(V_k * f)(x) \leq (V_k * f)(y) + \Delta.$$

Since $0 < \beta \leq 1$,

$$\beta(V_k * f)(x) \leq \beta(V_k * f)(y) + \beta\Delta \leq \beta(V_k * f)(y) + \Delta. \quad (3.27)$$

On the other hand, since $L(x)$ is increasing on $[P, \infty)$, for $y \in [x, x + C]$,

$$L(x) \leq L(y). \quad (3.28)$$

Inequalities (3.27) and (3.28) together imply that for $x \in [P, \infty)$ and $y \in [x, x + C]$,

$$L(x) + \beta(V_k * f)(x) \leq L(y) + \beta(V_k * f)(y) + \Delta.$$

Substituting (3.26) into the above inequality, we conclude that G_{k+1} is non- (Δ, C) -decreasing on $[P, \infty)$.

Condition (A3.b):

We know that P is the unique minimizer of L , and the first increasing point P_n^f of $V_k * f$ is greater than P . Therefore, both $L(x)$ and $(V_k * f)(x)$ are non-increasing on $(-\infty, P]$, and $L(x)$ is strictly decreasing on this interval. Hence, $G_{k+1}(x) = L(x) + \beta(V_k * f)(x)$ is decreasing on $(-\infty, P]$.

Condition (A3.c):

Since $V_k * f$ is non-increasing on $(-\infty, P]$ and $0 < \beta \leq 1$, for $x \in (-\infty, P]$,

$$\beta[(V_k * f)(x - C) - (V_k * f)(x)] \geq 0. \quad (3.29)$$

Also, by Condition (B3) and Lemma 1, for $x \in (-\infty, P]$,

$$L(x - C) - L(x) > \Delta. \quad (3.30)$$

Thus, inequalities (3.29) and (3.30) together imply that for $x \in (-\infty, P]$,

$$[L(x - C) + \beta(V_k * f)(x - C)] - [L(x) + \beta(V_k * f)(x)] > \Delta$$

By (3.26), the above inequality implies $G_{k+1}(x - C) - G_{k+1}(x) > \Delta$ on $[P, \infty)$.

Finally, since the left hand sides of equations (3.29) and (3.30) are both non-increasing on $x \in (-\infty, P]$, then $G_{k+1}(x - C) - G_{k+1}(x)$ is also non-increasing on this interval. Therefore, $G_{k+1} \in \mathcal{G}$ with $r_{k+1} = P$, and this completes the proof. ■

Actually, Conditions (B1)–(B3) that guarantee the optimality of the proposed policy are encountered commonly both in practice and the existing literature. As noted earlier, when the administrative cost for processing a replenishment is negligible compared to the cost for using each truck, we can easily assume $K = 0$; and when the terminal cost can be omitted (e.g., Veinott’s terminal Conditions hold) then we can assume $g_T = 0$. These assumptions are especially reasonable in the current economy considering that oil prices remain high and, hence, transportation costs dominate the administrative and terminal cost terms. Generally, Condition (B3) is perhaps the most restrictive in the sense that it may or may not hold depending on the model parameters; but, it is easy to verify nonetheless.

III.3.2. Special Case

It is worth noting that Lippman (1969b) provided an optimal replenishment policy in his Theorem 11, for a single-period problem. Actually, the policy Lippman devised

is a special case of the (Q, \vec{s}, \vec{S}) policy that we proposed in this research. Specifically, the dimension M of \vec{s} and \vec{S} equals to one in Lippman's policy. A restatement of Lippman's theorem in our notation is expressed thus:

Theorem (Lippman, 1969b): *When the replenishment cost is in the form of multiple setup cost where the cargo capacity and cargo cost are C and Δ respectively, if G_n is convex for any n , and the fixed setup cost $K = 0$, then the optimal order-up-to level is given by*

$$Y_n(x) = \begin{cases} \min(S_n, x + C \lceil \frac{s_n - x}{C} \rceil), & \text{if } x < s_n, \\ x, & \text{if } x \geq s_n. \end{cases} \quad (3.31)$$

Where, S_n is the minimizer of G_n , and s_n is the largest number less than or equal to S_n such that

$$G_n(s_n) = \Delta + G_n(\min(s_n + C, S_n)). \quad (3.32)$$

We simply call the policy with this structure (Q, s, S) policy, where $Q = s$.

In the following theorem, we provide a sufficient condition that guarantees the optimality of the (Q, s, S) policy for a finite horizon problem. This necessitates a lemma whose proof is provided.

Lemma 4 *If for period n , G_n is convex, and the fixed setup cost $K = 0$, then $V_n(x) - V_n(x - C)$ is non-decreasing. Furthermore, for any positive integer m , $V_n(x) - V_n(x - mC)$ is non-decreasing.*

Proof of Lemma 4: Since G_n is convex, and $K = 0$, by Theorem (Lippman, 1969b), the optimal replenishment policy can be determined by (3.31). And the cost-to-go

function can be expressed as:

$$V_n(x) = \begin{cases} G_n \left(\min \left(S_n, x + \left\lceil \frac{s_n - x}{C} \right\rceil C \right) \right) + \left\lceil \frac{s_n - x}{C} \right\rceil \Delta, & \text{if } x < s_n, \\ G_n(x), & \text{if } x \geq s_n. \end{cases} \quad (3.33)$$

It is worth noting that the S_n in this lemma denotes the minimizer of the convex function $G_n(x)$, thus it has a different meaning than the S_n^i in Theorem 1.

Case 1: When $x < s_n$, $V_n(x) - V_n(x - C)$ is constant.

$$\begin{aligned} & V_n(x) - V_n(x - C) \\ &= G_n \left(\min \left(S_n, x + \left\lceil \frac{s_n - x}{C} \right\rceil C \right) \right) + \left\lceil \frac{s_n - x}{C} \right\rceil \Delta \\ &\quad - G_n \left(\min \left(S_n, x - C + \left\lceil \frac{s_n - x + C}{C} \right\rceil C \right) \right) - \left\lceil \frac{s_n - x + C}{C} \right\rceil \Delta \\ &= G_n \left(\min \left(S_n, x + \left\lceil \frac{s_n - x}{C} \right\rceil C \right) \right) + \left\lceil \frac{s_n - x}{C} \right\rceil \Delta \\ &\quad - G_n \left(\min \left(S_n, x + \left\lceil \frac{s_n - x}{C} \right\rceil C \right) \right) - \left(\left\lceil \frac{s_n - x}{C} \right\rceil + 1 \right) \Delta \\ &= -\Delta. \end{aligned}$$

Case 2: When $x \geq s_n$, $V_n(x) - V_n(x - C)$ is non-decreasing.

Case 2.1: $s_n + C \leq S_n$.

- **Case 2.1.1:** For $x \in [s_n, s_n + C)$, $0 < s_n - x + C \leq C$ and $\left\lceil \frac{s_n - x + C}{C} \right\rceil = 1$. Then,

$$\begin{aligned} & V_n(x) - V_n(x - C) \\ &= G_n(x) - G_n \left(\min \left(S_n, x - C + \left\lceil \frac{s_n - x + C}{C} \right\rceil C \right) \right) - \left\lceil \frac{s_n - x + C}{C} \right\rceil \Delta \\ &= G_n(x) - G_n(\min(S_n, x)) - \Delta = G_n(x) - G_n(x) - \Delta = -\Delta. \end{aligned}$$

- **Case 2.1.2:** For $x \in [s_n + C, \infty)$, $x > s_n$ and $x - C \geq s_n$, hence, $V_n(x) - V_n(x - C) = G_n(x) - G_n(x - C)$. Since G_n is convex, by Lemma 1 $G_n(x) - G_n(x - C)$ is

non-decreasing, hence $V_n(x) - V_n(x - C)$ is also non-decreasing. In addition, by definition (3.32) of s_n , when $s_n + C \leq S_n$, we have $G_n(s_n + C) - G_n(s_n) = -\Delta$. Thus, $V_n(x) - V_n(x - C) \geq V_n(s_n + C) - V_n(s_n) = -\Delta$ which completes the proof that $V_n(x) - V_n(x - C)$ is non-decreasing for $x \geq s_n$ in Case 1.

Case 2.2: $S_n < s_n + C$.

- **Case 2.2.1:** For $x \in [s_n, S_n]$, it can be found that $V_n(x) - V_n(x - C) = -\Delta$.
- **Case 2.2.2:** For $x \in [S_n, s_n + C)$,

$$\begin{aligned} & V_n(x) - V_n(x - C) \\ &= G_n(x) - G_n\left(\min\left(S_n, x - C + \left\lceil \frac{s_n - x + C}{C} \right\rceil C\right)\right) - \left\lceil \frac{s_n - x + C}{C} \right\rceil \Delta \\ &= G_n(x) - G_n(\min(S_n, x)) - \Delta = G_n(x) - G_n(S_n) - \Delta. \end{aligned}$$

Since G_n is convex, and S_n is the global minimum, $G_n(x)$ is non-decreasing on $[S_n, \infty)$, and hence, $V_n(x) - V_n(x - C)$ is non-decreasing on $[S_n, s_n + C)$ and

$$-\Delta \leq V_n(x) - V_n(x - C) < G_n(s_n + C) - G_n(S_n) - \Delta.$$

- **Case 2.2.3:** For $x \in [s_n + C, \infty)$, $x \geq s_n$ and $x - C \geq s_n$, hence $V_n(x) - V_n(x - C) = G_n(x) - G_n(x - C)$. Also by Lemma 1 and the assumption that $G_n(x)$ is convex, $V_n(x) - V_n(x - C)$ is also non-decreasing on $[s_n + C, \infty)$. By (3.32), for $x \in [s_n + C, \infty)$

$$\begin{aligned} V_n(x) - V_n(x - C) &\geq \min_{x \in [s_n + C, \infty)} V_n(x) - V_n(x - C) \\ &= G_n(s_n + C) - G_n(s_n) \\ &= G_n(s_n + C) - G_n(S_n) - \Delta. \end{aligned}$$

Thus, no matter $S_n \geq s_n + C$ or $S_n > s_n + C$, as far as $G_n(x)$ is convex, $V_n(x) - V_n(x - C)$ is non-decreasing. Furthermore,

$$\begin{aligned} V_n(x) - V_n(x - mC) &= [V_n(x) - V_n(x - C)] + [V_n(x - C) - V_n(x - 2C)] + \dots \\ &\quad + [V_n(x - (m - 1)C) - V_n(x - mC)]. \end{aligned}$$

It is obvious that $V_n(x) - V_n(x - mC)$ is the sum of m non-decreasing functions, thus, it is also non-decreasing for any positive integer m . This completes the proof of Lemma 4 is complete. \blacksquare

Theorem 3 *If the demand follows a Uniform(0, mC) distribution where $m \in \mathbb{N}$, the terminal cost $g_T(x)$ is convex, and $K = 0$, for each period n , the optimal replenishment quantity can be determined by equation (3.31), that is restated here:*

$$Y_n(x) = \begin{cases} \min(S_n, x + C \lceil \frac{s_n - x}{C} \rceil), & \text{if } x < s_n, \\ x, & \text{if } x \geq s_n. \end{cases}$$

Proof of Theorem 3: Since we already restated the result from Lippman (1969b), Theorem (Lippman, 1969b), it is sufficient for us to prove that G_n is convex for all n under the assumptions and do it by induction. Note that for a Uniform(0, mC) distribution, the probability density function $f(z)$ can be written as $f(z) = \frac{1}{mC}I_{[0, mC]}(z)$.

- When $n = 1$,

$$G_1(y) = L(y) + \beta \int_0^{mC} g_T(y - z) \frac{1}{mC} dz.$$

Since $L(y)$ and $g_T(y)$ are both convex, $G_1(y)$ is convex.

- Now, suppose $G_n(y)$ is convex for $n = k$, we need to show that $G_{k+1}(y)$ is also convex. By definition

$$G_{k+1}(y) = L(y) + \beta \int_0^{mC} V_k(y - z) \frac{1}{mC} dz.$$

Obviously, G_{k+1} is differentiable. Taking the first derivative of $G_{k+1}(y)$, we have

$$\begin{aligned}
 G'_{k+1}(y) &= \frac{d}{dy} \left(L(y) + \beta \int_0^{mC} V_k(y-z) \frac{1}{mC} dz \right) \\
 &= L'(y) + \frac{d}{dy} \left(\beta \int_0^{mC} V_k(y-z) \frac{1}{mC} dz \right) \\
 &= L'(y) + \frac{\beta}{mC} \int_0^{mC} \frac{d}{dy} V_k(y-z) dz \\
 &= L'(y) + \frac{\beta}{mC} [V_k(y) - V_k(y - mC)].
 \end{aligned}$$

According to Lemma 4, $V_k(y) - V_k(y - mC)$ is non-decreasing. In addition, $L'(y)$ is also non-decreasing because the first derivative of a convex function is non-decreasing. Thus, $G'_{k+1}(y)$ is non-decreasing, and hence, $G_{k+1}(y)$ is convex. Since $G_n(y)$ is convex for all n , the optimal replenishment policy for any period n follows (3.31). ■

According to Theorem 3, there are three possible values of $Y_n(x)$. For a given beginning inventory level x , the optimal order-up-to level $Y_n(x)$ can be x , S_n , or $x + C \lceil \frac{s_n - x}{C} \rceil$. As far as x is less than a threshold value s_n , it is optimal for the vendor to wait until at least the next period to replenish inventory. If x exceeds s_n , then a replenishment is necessary. The order-up-to level depends on the relation between s_n and S_n .

Case 1: $S_n > s_n + C$

Since the operator $\lceil \cdot \rceil$ rounds up a number, it is always true that

$$\frac{s_n - x}{C} \leq \left\lceil \frac{s_n - x}{C} \right\rceil < \frac{s_n - x}{C} + 1.$$

Equivalently, $s_n \leq x + C \lceil \frac{s_n - x}{C} \rceil < s_n + C$. Considering that $S_n > s_n + C$ in this case, we have

$$x + C \left\lceil \frac{s_n - x}{C} \right\rceil < S_n.$$

Consequently,

$$Y_n(x) = \begin{cases} x + C \lceil \frac{s_n - x}{C} \rceil, & \text{if } x < s_n, \\ x, & \text{if } x \geq s_n. \end{cases}$$

The result of Case 1 can be interpreted so that when $S_n > s_n + C$, if the vendor needs to replenish inventory, it is always optimal to order an FTL quantity to bring the inventory level just above s_n .

Case 2: $S_n \leq s_n + C$

In this case, parameters s_n and S_n satisfy $s_n - kC < S_n - kC < s_n - (k-1)C < S_n - (k-1)C < \dots < s_n < S_n$. For $x \in [s_n - kC, s_n - (k-1)C)$, we have

$$k-1 < \frac{s_n - x}{C} \leq k \text{ and } \lceil \frac{s_n - x}{C} \rceil = k.$$

Case 2.1: $s_n - kC \leq x \leq S_n - kC$

Since $x + kC \leq S_n$, $Y_n(x) = x + kC = x + \lceil \frac{s_n - x}{C} \rceil C$.

Case 2.2: $S_n - kC < x < s_n - (k-1)C$

Since $x + kC > S_n$, $Y_n(x) = S_n$.

Thus,

$$Y_n(x) = \begin{cases} x + C \lceil \frac{s_n - x}{C} \rceil, & \text{if } x < s_n, \text{ and } s_n \leq x + C \lceil \frac{s_n - x}{C} \rceil \leq S_n \\ S_n, & \text{if } x < s_n, \text{ and } S_n < x + C \lceil \frac{s_n - x}{C} \rceil < s_n + C \\ x, & \text{if } x \geq s_n. \end{cases}$$

The result of Case 2 implies that when $S_n \leq s_n + C$, if a replenishment is necessary, sometimes the vendor should order an FTL quantity, and sometimes the optimal replenishment quantity includes an LTL quantity. The optimal order-up-to level and the optimal order quantity can be depicted as shown in Figure 5. In addition, the size of the LTL quantity can be expressed as $(S_n - x) - \lfloor \frac{S_n - x}{C} \rfloor C$. It

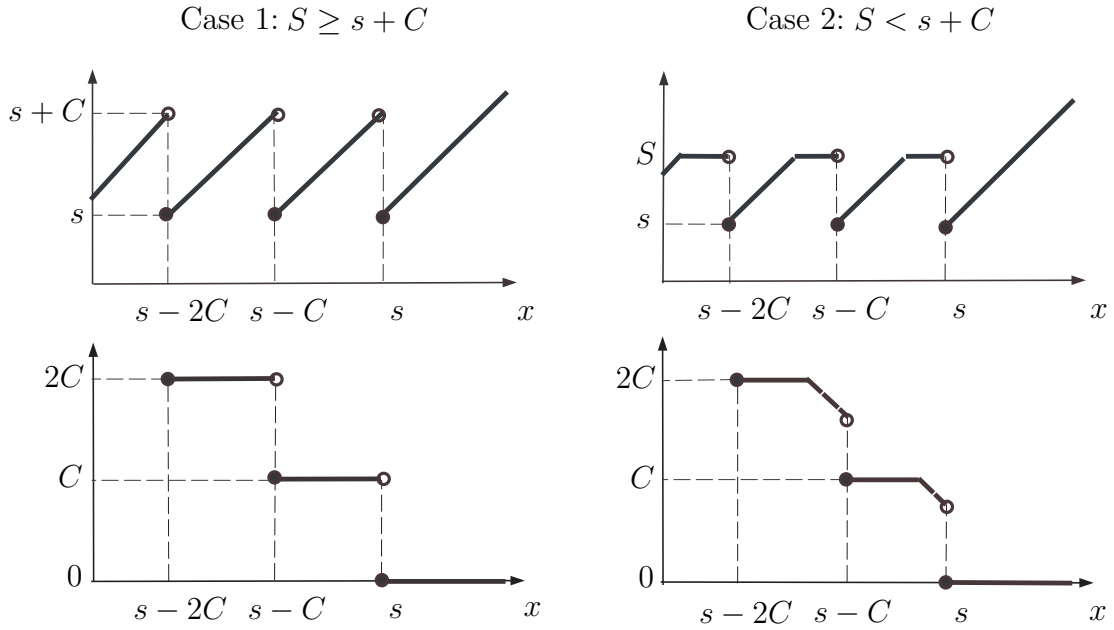
can be verified that it satisfies

$$(S_n - x) - \left\lfloor \frac{S_n - x}{C} \right\rfloor C = (S_n - x) - (k - 1)C$$

$$> S_n - s_n.$$

This means if a truck is partially loaded, its load must exceed $S_n - s_n$.

Figure 5: Optimal Policies for Special Case



III.4. Numerical Analysis

In this sublevel, we first conduct a factorial design experiment to test the optimality of (Q, \vec{s}, \vec{S}) policy numerically, then, we proceed with numerical tests to examine the influence of the system parameters on the values of the policy.

III.4.1. Factorial Design Experiment

Considering a problem of 5 periods, we conducted a factorial design experiment for four types of demand distributions, including Uniform $(0, b)$ distribution, Exponential (α) distribution, Poisson (λ) distribution and Gamma (k, θ) distribution. The data we use is shown in Table 2, where μ represents the mean of the random demand. Obviously, the parameters of the Uniform, Exponential and Poisson distributions can be uniquely determined by μ , due to the fact that each of them has only one parameter. Since the Gamma distribution has two parameters, i.e., k the shape parameter and θ the scale parameter, we let $\theta = 1$ be fixed and choose value of k to make $k\theta = \mu$.

Table 2: Data Set for Factorial Design Experiment

Parameters	Values			
K	0	75	150	300
C	100	200	400	
Δ	45	90	180	
h	1	2	4	
p	8	16	32	
μ	10	20	40	

We generated $4 \times 3^5 = 972$ instances for each type of demand distribution. We use complete enumeration to find the optimal values of the decision variables for each period and each instance. In all the cases we studied, the optimal policy is in form of (Q, \vec{s}, \vec{S}) policy, even if the fixed setup cost K takes positive value. In Sublevel 4, we have provided a sufficient condition for the optimality of (Q, \vec{s}, \vec{S}) policy, however, we

cannot obtain a sufficient condition for the most general setting here currently, and the parameter values of the optimal policy cannot be presented in closed form yet.

There's an interesting observation that, both \vec{s}_n and \vec{S}_n are always 1-dimensional if the demand is Exponentially distributed or Uniformly distributed. Since the optimal order policies for these two types of distributions show a simpler form, we tried to analyze the results of the factorial design experiments in order to find out the upper bound and lower bound for the (Q_n, s_n, S_n) parameters, for these two types of demand distributions. And we found that the values of Q_n, s_n, S_n are not monotonically decreasing or increasing in n . However, when the terminal cost is 0, we have all the Q_n, s_n, S_n values greater than or equal to their counterparts in the last period, i.e., period 0.

III.4.2. Impact of System Parameters on Policy Values

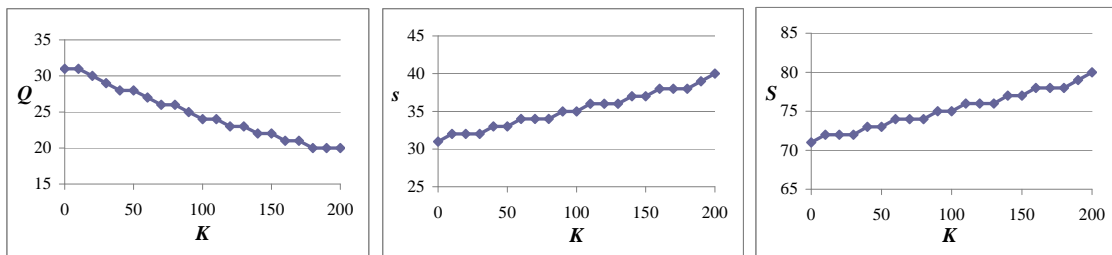
We also conducted some numerical tests to analyze the influence of the system parameters on the (Q_n, s_n, S_n) policy values for demand distribution of Uniform(0, b), where $b = 80$. We generated a basic problem and a set of compare problems. The basic problem uses the parameters in the second column of the Table 3, and all the compare problems use the same parameter values as the basic problem except one that uses the value range in the third column.

It is worth mentioning that, in most cases the optimal policy converges very fast, and after some iterations the optimal policy for the starting period, i.e., period 5, seems to be a reasonably good approximation for a stationary optimal policy for a system with much more planning periods. Therefore, in this sublevel, we examine the impact of the economic parameters on the values of the optimal policy for period 5. We use the terms 'increasing' and 'decreasing' in the weak sense that they refer to 'non-decreasing' and 'non-increasing', respectively, in the following discussion.

Table 3: Data Set for System Parameter's Impact

Parameters	Basic Problem	Analyzed problem
Fixed setup cost: K	50	0-200
Truck capacity: C	400	1-400
Unit truck cost: Δ	180	50-200
Inventory holding cost: h	4	0-50
Back-order penalty cost: p	16	0-50

- Fixed setup cost: K

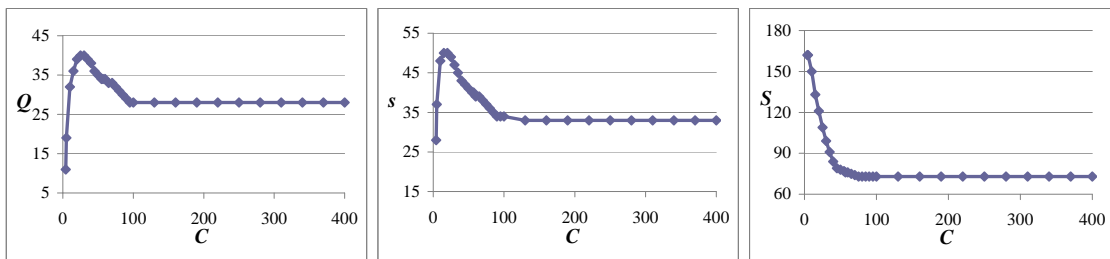
Figure 6: Influence of Parameter K 

In Figure 6, we provide the optimal values of (Q, s, S) policy for different values of fixed setup cost K . We observe that when the fixed setup cost K increases, the value of Q decreases, and s, S increase, all in a roughly linear relation. As known, if the fixed setup cost K increases, ordering more products each time can compensate the increments in the average fixed setup cost per item. Since all other parameters remain unchanged, especially the distribution of the demand is the same, when the quantity of each order increases, the order frequency needs to decrease. Such adjustment can be realized by decreasing the threshold value

Q for non-ordering, and at the same time increasing the minimal order-up-to level s and the maximal order-up-to level S . We also note that a higher value of K prefers full truck load to less than truckload.

- Truck capacity: C

Figure 7: Influence of Parameter C



In Figure 7, the impact of the truck capacity C on the optimal policy is given. When C increases, both Q and s increase first and then decrease. The maximal order-up-to level S decreases as a response to a larger truck capacity. When C is very small compared with the expected demand, an ordering policy with only full truckload is more likely to be optimal. While when C is sufficiently large, all the values of (Q, s, S) converges to their limits respectively. Under such circumstances, the problem is very similar to a classic stochastic inventory problem where the ordering cost is linear with a fixed setup cost $K + \Delta$, and the famous (s, S) policy is known to be optimal for such a classic problem. Hence, for our problem with large truck capacity, a less than full truckload order quantity is optimal for most of the time and the order-up-to level S does not depend on the value of C .

- Unit truck cost: Δ

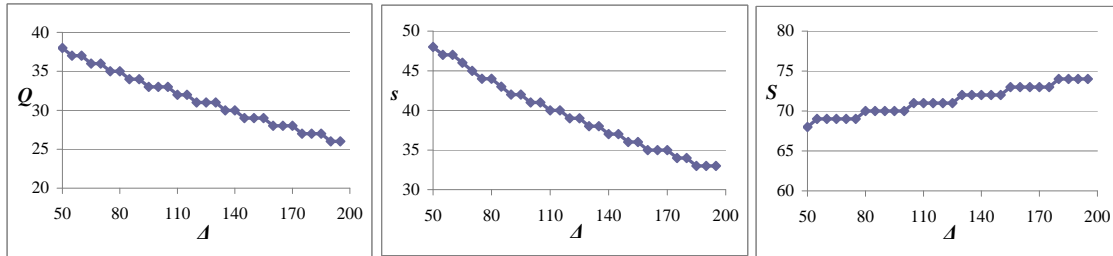
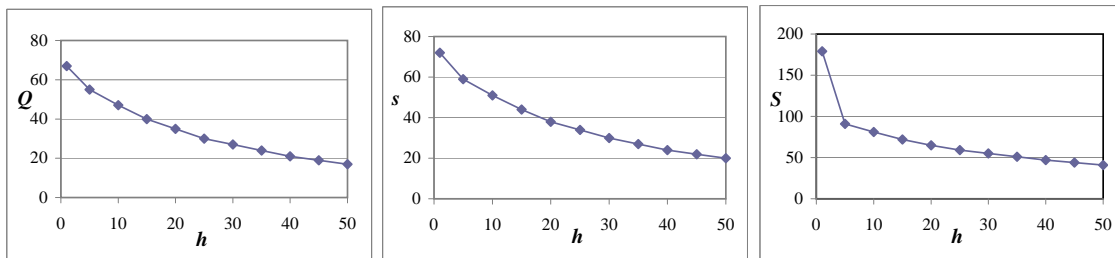
Figure 8: Influence of Parameter Δ 

Figure 8 depicts how the optimal policy parameters (Q, s, S) change for different values of unit truck cost Δ . We note that as Δ gets larger, the reorder point Q and the minimal order-up-to level s decrease, while the maximal order-up-to level S increases. Since we assume the truck capacity $C = 400$ which is sufficiently large, it is guaranteed the optimal ordering policy is a combination of full truckload and less than truckload order decisions. Figure 8 suggests that the differences $S - s$ and $S - Q$ both increase, which implies a higher value of unit truck cost prefers full truckload, similarly to the fixed setup cost K .

- Inventory holding cost: h

Figure 9: Influence of Parameter h 

From Figure 9, we observe that all values of (Q, s, S) decrease as the inventory

holding cost h increases. That is because when the inventory holding cost increases, it is better to decrease the order-up-to level in order to avoid the excessive carrying cost. When h is sufficiently large, the (Q, s, S) policy is similar to an order-up-to 0 policy, which means we always order the quantity that is backordered in the previous period.

- Backorder penalty cost: p

Figure 10: Influence of Parameter p

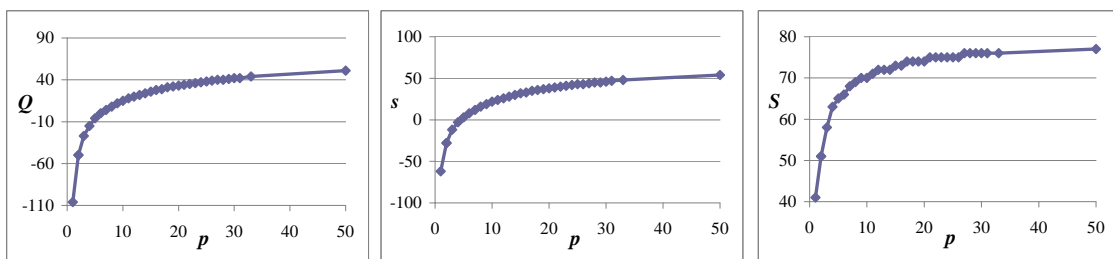


Figure 10 shows the influence in the policy values by increasing the backorder penalty cost p . Opposite to the case of increasing the inventory holding cost h , Q, s, S all increase as p increases. When the penalty cost for delayed order increases, it is better to increase the order-up-to level to avoid high penalty cost. And if p is sufficiently large, the values of Q, s, S will be very close to each other, then the (Q, s, S) policy is similar to a base stock policy.

III.5. Summary

In this chapter, we study a single-echelon, single-product stochastic dynamic inventory problem, with the replenishment cost consisting of a fixed ordering setup cost and an inbound transportation cost. The inbound transportation cost only depends

on the number of capacitated trucks that are used for shipping products, hence it is a staircase function. We propose a new ordering policy: (Q, \vec{s}, \vec{S}) policy. We provide a sufficient condition and prove the optimality of the proposed policy under this condition. During the proof, we introduce a new concept non- (Δ, C) -decreasing, which is analogous to the existing concept of non- K -decreasing. For a more special case when the demand is uniform distributed on $(0, mC)$ where m is an integer and C is the truck capacity, we prove the optimality of the (Q, \vec{s}, \vec{S}) policy that shows in a simple form. We also conduct tests to numerically verify that under general conditions, the optimal inventory policies still have a structure of the (Q, \vec{s}, \vec{S}) policy. Sensitivity analysis is performed to evaluate the impact of system parameters on the values of the policy parameters.

CHAPTER IV

EXACT MODELS AND OPTIMAL POLICIES FOR SHIPMENT
CONSOLIDATION: A STOCHASTIC DYNAMIC PROGRAMMING APPROACH

In Chapter III, we model the inbound inventory control of a vendor who uses vehicles with limited capacity to transport replenishment orders from the outside supplier to its own warehouse. In this chapter, we investigate the outbound distribution schedules under different transportation situations. As mentioned in Chapter II, economies of scale in transportation are particularly realizable under a shipment consolidation strategy. Although various shipment consolidation policies, e.g., time-, quantity- and time-and-quantity- based consolidation policies, have been proposed and adopted in industry, these policies are all designed in advance. And the existing research results on shipment consolidation either rely on the predetermined practical policies or come from numerical studies. Therefore, the consolidation policy that is theoretically optimal remains unknown. The objectives of the current research are to develop the consolidation models by using the stochastic dynamic programming approach and to characterize the structural properties of the optimal consolidation policies. Different from the stationary and practical consolidation policies, both the time and quantity parameters are implicit in our model; in other words, a consolidated load may be dispatched with a larger amount than the maximum holding quantity earlier than the predetermined shipping date.

Specifically, we consider a periodic-review distribution system. In the system, a collection depot is responsible for delivering orders to a number of retailers located in a geographical region over a discrete and finite time horizon. During each period, random orders/shipments from the retailers arrive at the collection depot and gather

into a large load. At the end of that period, the depot reviews the size of the consolidated load and decides how much to dispatch to the retailers. There are two types of costs associated with this system, the transportation cost for delivering the retailer orders and the customer waiting cost for delaying their fulfillments.

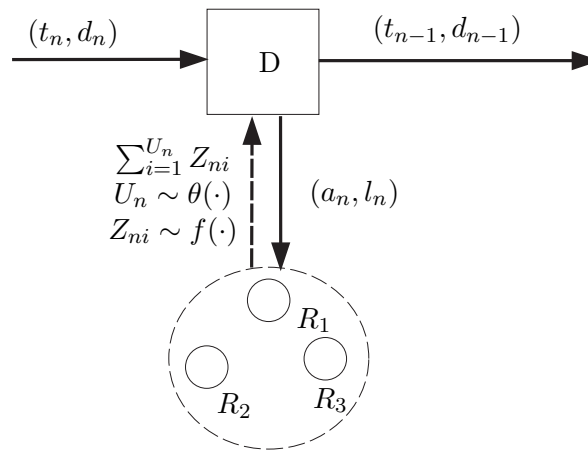
We concentrate on the theoretical analysis for four models with different transportation modes: the private fleet transportation with unlimited cargo capacity, the single-truck transportation with cargo capacity and fixed cost, the common carriage transportation, and the multiple-truck transportation with cargo capacity. The transportation cost structures in the first and the third models are the same as used in models of Çetinkaya and Bookbinder (2003). To model the cargo capacity and cargo cost in the last model, the transportation cost is presented in the term of multiple setup costs. We show the structures of the exact optimal consolidation policy for the first two models, provide sufficient conditions of some practical policies for the third model with the “clearing property” assumption, and characterize the rough structure of the optimal policy for the last model.

The plan of this chapter is as follows. Common system settings of the four models are provided in Sublevel IV.1. In Sublevel IV.2, we formulate the optimal control problem for private fleet transportation with unlimited cargo capacity and provide the structural results of the optimal consolidation policy. The model of single capacitated truck with fixed cost is studied in Sublevel IV.3. In Sublevel IV.4, we proceed with analyzing the scenario of common-carriage. We analytically investigate the optimality of some practical policies, and provide sufficient conditions under which these policies are optimal. Computational studies are presented to illustrate the complexity of the exact optimal policy. In Sublevel IV.5, we investigate the model of multiple-truck transportation with limited cargo capacity. Conclusions and directions for future research are given in Sublevel IV.6.

IV.1. Common System Settings and Modeling Assumptions

As is mentioned, the four models studied in this chapter share the same system settings. The only difference exists in the structure of the transportation cost. For this reason, we provide the common setting and modeling assumptions in this sublevel.

Figure 11: System Settings of the Outbound Consolidation System



In this chapter we study a periodic-review consolidation system where a depot collects and delivers random orders of a single-product to multiple retailers over a finite horizon of N periods. An illustration is given in Figure 11, and a summary of basic notation is provided below.

System Parameters:

- N length of the planning horizon
- n period index ($n = 0, 1, \dots, N$)

U_n	random number of orders received in period n (we assume $\{U_n\}_{0 \leq n \leq N}$ forms an i.i.d. sequence with a generic element denoted by U . The density and distribution functions of U are $\theta(\cdot)$ and $\Theta(\cdot)$)
Z_{ni}	random weight of the i th order in period n (we denote Z_i as a generic element with density and distribution functions $f(\cdot)$ and $F(\cdot)$)
$\tilde{C}(t, d)$	transportation cost for delivering d orders with total weight t
w	waiting cost of consolidated load per unit weight per period

States:

t_n	the total weight of consolidated orders in period n
d_n	the total number of orders consolidated in period n

Decisions:

a_n	weight of consolidated load dispatched in period n
l_n	number of consolidated orders dispatched in period n

Optimality Equation:

$V_n(t, d)$	the optimal expected total cost from period n to the end, when period n has d orders with total weight t before a dispatch decision is made
-------------	---

In Figure 11, D and R_i represent the collection depot and the retailers, respectively. During period n ($n \leq N$), a random number U_n of orders to be shipped to retailers arrive at the depot. The number U_n has a discrete distribution with a density function $\theta(\cdot)$ and a distribution function $\Theta(\cdot)$. U_n s are independent and identically distributed (*i.i.d.*). Furthermore, each order weighs a random quantity. We denote the weight of the i th order in period n as Z_{ni} , which follows a distribution function $F(\cdot)$ and a probability density function $f(\cdot)$. Z_{ni} s are also *i.i.d.* and independent of U_n . That is, the weight of each order is independent of the number of orders received.

By the end of period n , the depot reviews the number of consolidated orders d_n and the total weight t_n . Based on d_n and t_n , the depot needs to determine the dispatch number l_n and dispatch weight a_n of consolidated orders. A transportation cost $\tilde{C}(\cdot)$ is incurred for a positive dispatch quantity. All remaining consolidated load incurs a linear waiting cost w per unit weight per period. The waiting cost represents an opportunity loss in delayed receipt of revenue as well as a goodwill penalty. It is assumed that by the end of the planning horizon, the depot should satisfy as many consolidated orders as possible.

The objective of the research is to identify the structure of the optimal consolidation policies that minimize the total expected transportation and waiting costs over a finite planning horizon. In the following four sublevels, we discuss the models with different modes of transportation, i.e., private fleet transportation without cargo capacity constraint, single-truck transportation with cargo capacity and fixed cost, common carriage transportation, and multiple-truck transportation with cargo capacity.

IV.2. Private Fleet Transportation without Cargo Capacity

In this sublevel, we consider the model where the orders are delivered to the retailers by the depot's private truck. We assume the truck has unlimited capacity, so that all the orders can be loaded on a single truck.

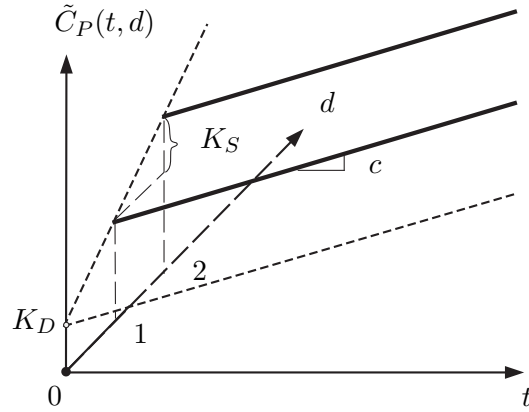
IV.2.1. Problem Formulation

Borrowed from Çetinkaya and Bookbinder (2003), the transportation cost under private carriage includes a fixed cost K_D for a vehicle dispatch from the depot to the retailers, a fixed cost K_S for an order delivery, and a variable transportation cost c

per unit weight. Thus, the transportation cost for shipping d units of orders with total weight t can be expressed as equation (4.1) and illustrated as Figure 12.

$$\tilde{C}_P(t, d) = K_D \cdot I_{[t>0, d>0]} + K_S d + ct. \quad (4.1)$$

Figure 12: Private Fleet Transportation Cost without Cargo Capacity



With its system setting described in Sublevel IV.1, this model can be formulated as a dynamic program by using a backward recursion method. Backward recursion means that period N is done first, and period 0 is the end of the planning horizon. For notational simplicity, the subscript of n is omitted on t_n and d_n in the remaining part of this chapter. Notice that when there is no order consolidated at the depot ($d = 0$), the total weight t must be 0 and vice versa. Therefore, the values of t and d must satisfy conditions: $t = d = 0$, or $t \in (0, \infty)$ and $d \in \mathbb{N}$. Denoting $\Omega_{(t,d)}$ as the decision space for a given state (t, d) , we have $\Omega_{(t,d)} = \{(0, 0)\} \cup \{(0, t] \times \{1, \dots, d\}\}$.

Then the optimality equation can be written as

$$V_n(t, d) = \begin{cases} \min_{(a,l) \in \Omega_{(t,d)}} \left\{ \begin{array}{l} \tilde{C}_P(a, l) + w(t - a) \\ + E \left(V_{n-1} \left(t - a + \sum_{i=1}^U Z_i, d - l + U \right) \right) \end{array} \right\}, & n = 1, \dots, N, \\ \tilde{C}_P(t, d), & n = 0, \end{cases} \quad (4.2)$$

where

$$\begin{aligned} E \left(V_{n-1} \left(t - a + \sum_{i=1}^U Z_i, d - l + U \right) \right) \\ = \sum_{u=0}^{\infty} \left[\int_0^{\infty} V_{n-1}(t - a + z, d - l + u) f_u(z) dz \right] \theta(u). \end{aligned}$$

$f_u(x)$ is the density function of the random variable $\sum_{i=1}^u Z_i$ which represents the total weight of u orders. The distribution function of $\sum_{i=1}^u Z_i$, $F_u(x)$, is obtained by taking u -fold convolution of $F(x)$ with itself.

Let

$$v_n(a, l|t, d) := \tilde{C}_P(a, l) + w(t - a) + E \left(V_{n-1} \left(t - a + \sum_{i=1}^U Z_i, d - l + U \right) \right)$$

represent the expected cost for dispatching l orders with total weight a in period n , when there are d orders with total weight t consolidated at the depot. Notice that when $a > 0$ and $l > 0$, $v_n(a, l|t, d)$ represents the cost for dispatching a positive quantity. When $a = l = 0$, $v_n(a, l|t, d)$ represents the cost for consolidating the orders. Based on this observation, we define the decision space $\Omega_{(t,d)}$ as the union of two disjoint sub-spaces. Specifically, $\Omega_{(t,d)} = \Omega_{(t,d)}^1 \cup \Omega_{(t,d)}^2$, where $\Omega_{(t,d)}^1 = \{(0, 0)\}$ denotes the decision space for consolidating the orders, and $\Omega_{(t,d)}^2 = (0, t] \times \{1, \dots, d\}$

denotes the decision space for dispatching a positive quantity. Accordingly, we define

$$u_n^1(t, d) = \min_{(a,l) \in \Omega_{(t,d)}^1} v_n(a, l|t, d) = v_n(0, 0|t, d), \text{ for } t \geq 0, d \in \mathbb{Z}^+$$

and

$$u_n^2(t, d) = \min_{(a,l) \in \Omega_{(t,d)}^2} v_n(a, l|t, d), \text{ for } t > 0, d \in \mathbb{N}.$$

It follows that $u_n^1(t, d)$ represents the minimum expected accumulated cost for consolidating orders in period n and planning optimally onward, and $u_n^2(t, d)$ represents the minimum expected accumulated cost for dispatching the consolidated load. In fact, the optimal decision of period n can be obtained by optimizing these two cases individually and choosing the one which gives a lower cost, i.e., for $n = 1, \dots, N$,

$$V_n(t, d) = \begin{cases} u_n^1(t, d), & \text{if } t = 0, d = 0, \\ \min \{u_n^1(t, d), u_n^2(t, d)\}, & \text{if } t > 0, d \in \mathbb{N}. \end{cases} \quad (4.3)$$

In period n ($n = 1, \dots, N$), if $V_n(t, d) = u_n^1(t, d)$, then it is optimal to consolidate the load, i.e., the optimal dispatch quantity $a_n^* = 0$ and $l_n^* = 0$. Otherwise, the collection depot should deliver a shipment to the retailers.

IV.2.2. Exact Optimal Policy

Before the consolidation decision is made in period n , if no order is consolidated ($t = d = 0$), we obviously choose to consolidate the orders. For cases where $t \in (0, \infty)$ and $d \in \mathbb{N}$, a careful examination is required. To analyze the exact optimal consolidation policy, we identify a set of conditions such that if the two-dimensional value function $V_{n-1}(t, d)$ satisfies these conditions, then the structure of the optimal dispatch policy in period n can be characterized.

Proposition 2 *If $V_{n-1}(t, d)$, the value function of period $n - 1$ satisfies the following*

condition:

(A1) $V_{n-1}(t, d)$ is jointly concave in t and d .

(A2) For any fixed $t > 0$, $V_{n-1}(t, d)$ is non-decreasing in $d \in \mathbb{N}$ and $\lim_{d \rightarrow \infty} V_{n-1}(t, d) - V_{n-1}(t, d-1) = K_S$. For any fixed $d > 0$, $V_{n-1}(t, d)$ is non-decreasing in $t > 0$ and $\lim_{t \rightarrow \infty} \frac{\partial V_{n-1}(t, d)}{\partial t} = c$.

(A3) For any fixed $t > 0$, $V_{n-1}(t, d) - K_S d$ is non-decreasing in d . For any fixed $d > 0$, $V_{n-1}(t, d) - ct$ is non-decreasing in t .

then $u_n^2(t, d) = v_n(t, d|t, d)$. Furthermore, the threshold policy $\{(a, l)_n^*(t, d)\}$ given by

$$(a, l)_n^*(t, d) = \begin{cases} (0, 0), & \text{if } d < S_t^n, \\ (t, d), & \text{if } d \geq S_t^n, \end{cases} \quad (4.4)$$

where S_t^n is a function of t , is optimal for period n .

Proof of Proposition 2 According to Condition (A1), since $V_{n-1}(t, d)$ is jointly concave in t and d , its partial derivative (with respect to t) and difference (with respect to d) are both non-increasing. Due to Condition (A2), for any given state (t, d) , we have

$$V_{n-1}(t, d) - V_{n-1}(t, d-1) \geq \lim_{i \rightarrow \infty} V_{n-1}(t, i) - V_{n-1}(t, i-1) = K_S, \quad (4.5)$$

and

$$\frac{\partial V_{n-1}(t, d)}{\partial t} \geq \lim_{t \rightarrow \infty} \frac{\partial V_{n-1}(t, d)}{\partial t} = c. \quad (4.6)$$

If l orders with total weight a ($(a, l) \in \Omega_{(t, d)}^2$) are dispatched, then

$$\begin{aligned}
v_n(a, l|t, d) - v_n(a, l-1|t, d) &= K_S + E \left(V_{n-1} \left(t - a + \sum_{i=1}^U Z_i, d - l + U \right) \right. \\
&\quad \left. - V_{n-1} \left(t - a + \sum_{i=1}^U Z_i, d - (l-1) + U \right) \right) \\
&= K_S + \sum_{u=0}^{\infty} \left[\int_0^{\infty} (V_{n-1}(t - a + z, d - l + u) \right. \\
&\quad \left. - V_{n-1}(t - a + z, d - (l-1) + u)) f_u(z) dz \right] \theta(u) \\
&\leq K_S - \sum_{u=0}^{\infty} \left[\int_0^{\infty} K_S f_u(z) dz \right] \theta(u) = 0,
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial v_n(a, l|t, d)}{\partial a} &= c - w + \frac{\partial E \left(V_{n-1} \left(t - a + \sum_{i=1}^U Z_i, d - l + U \right) \right)}{\partial a} \\
&= c - w + \sum_{u=0}^{\infty} \left[\int_0^{\infty} \frac{\partial V_{n-1}(t - a + z, d - l + u)}{\partial a} f_u(z) dz \right] \theta(u) \\
&\leq c - w - \sum_{u=0}^{\infty} \left[\int_0^{\infty} c f_u(z) dz \right] \theta(u) = -w.
\end{aligned}$$

The above inequalities hold because of (4.5) and (4.6), respectively. Therefore, $v_n(a, l|t, d)$ is non-increasing in l for a fixed a and also non-increasing in a for a fixed l on the bounded domain $\Omega_{(t,d)}^2$. This implies that $v_n(a, l|t, d)$ takes its minimum at $a = t$ and $l = d$, i.e., $u_n^2(t, d) = \min_{(a,l) \in \Omega_{(t,d)}^2} v_n(a, l|t, d) = v_n(t, d|t, d)$.

Based on the previous analysis, we can write $u_n^1(t, d)$ and $u_n^2(t, d)$ explicitly as

$$\begin{aligned}
u_n^1(t, d) &= wt + E \left(V_{n-1} \left(t + \sum_{i=1}^U Z_i, d + U \right) \right) \\
u_n^2(t, d) &= \tilde{C}_P(t, d) + E \left(V_{n-1} \left(\sum_{i=1}^U Z_i, U \right) \right).
\end{aligned}$$

For a fixed $t > 0$,

$$\begin{aligned} u_n^1(t, d) - K_S d &= wt + E \left(V_{n-1} \left(t + \sum_{i=1}^U Z_i, d + U \right) \right) - K_S d \\ &= wt + E \left(V_{n-1} \left(t + \sum_{i=1}^U Z_i, d + U \right) - K_S d \right) \end{aligned}$$

is non-decreasing in d according to Condition (A3). However, it is obvious that $u_n^2(t, d) - K_S d = K_D + ct + E \left(V_{n-1} \left(\sum_{i=1}^U Z_i, U \right) \right)$ is constant in d for a fixed $t > 0$. Since $V_n(t, d) - K_S d = \min \{u_n^1(t, d) - K_S d, u_n^2(t, d) - K_S d\}$, there must exist a critical point $S_t^n \geq 0$ such that

$$S_t^n = \min \{d \in \mathbb{Z}^+ : u_n^1(t, d) - K_S d \geq u_n^2(t, d) - K_S d\}.$$

For all $d < S_t^n$, $u_n^1(t, d) < u_n^2(t, d)$ and consolidating the orders in period n is optimal, i.e., $(a, l)_n^*(t, d) = (0, 0)$. For all $d \geq S_t^n$, $u_n^1(t, d) \geq u_n^2(t, d)$, which means it is optimal to dispatch the consolidated load, i.e., $(a, l)_n^*(t, d) = (t, d)$. In summary, the optimal consolidation policy is given by equation (4.4). Then Proposition 2 is proved. ■

Proposition 2 tells us that if the value function of period $n-1$ satisfies Conditions (A1)–(A3), and a positive size of shipment has to be dispatched, then the optimal dispatch decision in period n is to dispatch the entire consolidated load. As a result, a threshold policy is optimal for period n . In detail, for any state (t, d) , we can find a state-dependent threshold value S_t^n , so that if $d < S_t^n$, it is optimal to consolidate the orders; otherwise, dispatching the entire load is optimal. Equivalently, we can also find a threshold value S_d^n so that if $t \geq S_d^n$, it is optimal to dispatch all the orders; otherwise, consolidating is more preferable. In fact, such a threshold policy is optimal under a multiple-period setting, regardless of the parameter settings.

Theorem 4 *For the finite horizon problem, the threshold policy $\{(a, l)_n^*(t, d)\}$ given by (4.4) is optimal for period n .*

Proof of Theorem 4 It is sufficient to prove $V_n(t, d)$ satisfies Conditions (A1)–(A3), for $n = 0, \dots, N$. We prove this by induction.

- When $n = 0$, $V_0(t, d) = \tilde{C}_P(t, d) = K_D \cdot I_{[t>0, d>0]} + K_S d + ct$.

Condition (A1):

To show that $V_0(t, d)$ is jointly concave in t and d , we need to show that for any two different points (t_1, d_1) and (t_2, d_2) in $\Omega = \{(0, 0)\} \cup [(0, \infty) \times \mathbb{N}]$, and any $\lambda \in [0, 1]$,

$$V_0((1 - \lambda)(t_1, d_1) + \lambda(t_2, d_2)) \geq (1 - \lambda)V_0(t_1, d_1) + \lambda V_0(t_2, d_2).$$

Case 1: When neither of (t_1, d_1) and (t_2, d_2) equals to $(0, 0)$,

$$\begin{aligned} V_0((1 - \lambda)(t_1, d_1) + \lambda(t_2, d_2)) &= V_0((1 - \lambda)t_1 + \lambda t_2, (1 - \lambda)d_1 + \lambda d_2) \\ &= K_D + K_S((1 - \lambda)d_1 + \lambda d_2) + c((1 - \lambda)t_1 + \lambda t_2) \\ &= (1 - \lambda)V_0(t_1, d_1) + \lambda V_0(t_2, d_2). \end{aligned}$$

Case 2: When one of (t_1, d_1) and (t_2, d_2) , say (t_2, d_2) , is equal to $(0, 0)$,

$$\begin{aligned} V_0((1 - \lambda)(t_1, d_1) + \lambda(0, 0)) &= V_0((1 - \lambda)t_1, (1 - \lambda)d_1) \\ &= K_D + K_S((1 - \lambda)d_1) + c((1 - \lambda)t_1) > (1 - \lambda)V_0(t_1, d_1) + \lambda V_0(0, 0). \end{aligned}$$

Hence, $V_0(t, d)$ is jointly concave in t and d .

Condition (A2):

For a fixed $t > 0$, $V_0(t, d) = K_D + ct + K_S d$ is obviously increasing in d , and $\lim_{d \rightarrow \infty} (V_0(t, d) - V_0(t, d - 1)) = K_S$. And for a fixed $d > 0$, $V_0(t, d) = K_D + K_S d + ct$ is obviously increasing in t , and $\lim_{t \rightarrow \infty} \frac{\partial V_0(t, d)}{\partial t} = c$.

Condition (A3):

For any fixed consolidated weight $t > 0$, $V_0(t, d) - K_S d = K_D + ct$ is constant, and

hence, non-decreasing in d . Similarly, for any fixed number $d > 0$, $V_0(t, d) - ct = K_D + K_S d$ is constant, and hence, non-decreasing in t .

- Suppose the statements hold for $n = k - 1$. We need to show for $n = k$. Since $V_{k-1}(t, d)$ satisfies Conditions (A1)–(A3), by Proposition 2,

$$u_k^1(t, d) = v_k(0, 0|t, d) = wt + E \left(V_{k-1} \left(t + \sum_{i=1}^U Z_i, d + U \right) \right),$$

$$u_k^2(t, d) = v_k(t, d|t, d) = \tilde{C}_P(t, d) + E \left(V_{k-1} \left(\sum_{i=1}^U Z_i, U \right) \right).$$

Condition (A1):

It is easy to see that both $u_k^1(t, d)$ and $u_k^2(t, d)$ are jointly concave in t and d ; therefore, their minimum $V_k(t, d)$ is also jointly concave in t and d .

Condition (A2):

Obviously, both $u_k^1(t, d)$ and $u_k^2(t, d)$ are non-decreasing in t for a fixed $d > 0$, as well as non-decreasing in d for a fixed $t > 0$. As a result, $V_k(t, d) = \min \{u_k^1(t, d), u_k^2(t, d)\}$ is also non-decreasing in t for a fixed $d > 0$ and non-decreasing in d for a fixed $t > 0$. We still need to find the limits of $V_k(t, d)$'s partial derivative and difference. Since we assume $V_{k-1}(t, d)$ satisfies Conditions (A1)–(A3), by Proposition 2, when t is fixed and d is sufficiently large, it is always optimal to dispatch the consolidated load, i.e., $\lim_{d \rightarrow \infty} V_k(t, d) = \lim_{d \rightarrow \infty} u_k^2(t, d)$. Thus,

$$\begin{aligned} \lim_{d \rightarrow \infty} (V_k(t, d) - V_k(t, d - 1)) &= \lim_{d \rightarrow \infty} (u_k^2(t, d) - u_k^2(t, d - 1)) \\ &= \lim_{d \rightarrow \infty} (v_k(t, d|t, d) - v_k(t, d - 1|t, d - 1)) = K_S. \end{aligned}$$

Similarly, when d is fixed and t goes to ∞ , it is always optimal to dispatch the entire load, i.e., $\lim_{t \rightarrow \infty} V_k(t, d) = \lim_{t \rightarrow \infty} u_k^1(t, d)$, and $\lim_{t \rightarrow \infty} \frac{\partial V_k(t, d)}{\partial t} = c$.

Condition (A3):

For a fixed $t > 0$, $V_k(t, d) - K_S d = \min \{u_k^1(t, d) - K_S d, u_k^2(t, d) - K_S d\}$. It can be computed that $u_k^1(t, d) - K_S d = wt + E \left(V_{k-1} \left(t + \sum_{i=1}^U Z_i, d + U \right) - K_S d \right)$ is non-decreasing in d while $u_k^2(t, d) - K_S d = K_D + ct + E \left(V_{k-1} \left(\sum_{i=1}^U Z_i, U \right) \right)$ is constant in d for fixed $t > 0$. Therefore, in period k , for any fixed consolidated weight $t > 0$, $V_k(t, d) - K_S d$ is non-decreasing in d . Similarly, it can be proved that for any fixed number of consolidated orders $d > 0$, $V_k(t, d) - ct$ is non-decreasing in t .

Thus, $V_k(t, d)$ also satisfies Conditions (A1)–(A3), and the proof is complete. ■

It is interesting to find that according to Theorem 4, the optimal decision of any period is either not to dispatch or to dispatch all. In other words, the optimal policy possesses the “clearing property”. Dispatching partial of the consolidated load is never economical, because in the case of private fleet transportation with unlimited cargo capacity, the major influential part of the transportation cost is the fixed cost for a vehicle dispatch. Hence, separately dispatching a consolidated load only incurs additional waiting cost and does not lead to any savings.

IV.3. Single-Truck Transportation with Cargo Capacity and Fixed Cost

In Sublevel IV.2, we consider the stochastic dynamic distribution system with out-bound transportation performed by the private fleet owned by the collection depot. It is worth noting that the transportation capacity in that model is assumed to be infinite, which might not be realistic for industrial practices. Therefore, it is crucial for successful decision-makers to explicitly take into account the capacity constraint on the truck. In this sublevel, we consider the case where the shipment is delivered on a single capacitated truck. When the truck is used, no matter whether it is fully

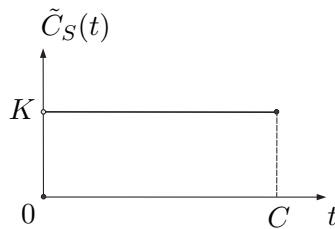
or partially loaded, a fixed cost is incurred. We examine the optimal consolidation policy.

IV.3.1. Problem Formulation

We assume the single truck has a finite capacity C with a fixed usage cost K for dispatching a shipment, regardless of the dispatch quantity and the corresponding number of orders. Since the customer waiting cost only depends on the weight of the consolidated load t , the number of consolidated orders d is trivial in this case, and hence, can be ignored in this model. Then the transportation cost for dispatching t units of consolidation load is given by equation (4.7) and illustrated as Figure 13.

$$\tilde{C}_S(t) = K \cdot I_{[t>0]}, \quad 0 \leq t \leq C. \quad (4.7)$$

Figure 13: Single-Truck Transportation Cost with Cargo Capacity



Due to the existence of the capacity constraint, the dispatch quantity can only be within 0 and C , and any quantity exceeding C is infeasible. Therefore, it is not always viable to satisfy all retailer demands by the end of the planning horizon. For this reason, we assume that at the end of the planning horizon, as many as possible consolidated orders should be dispatched, and all the remaining load still incurs waiting cost w per unit weight like any earlier period. Considering this distribution system for N periods, we are interested in the optimal dispatch quantity in each

period. Since the number of consolidated orders is ignored, multiple orders received during one period can be considered as a single combined order. Denoting Z as the weight of a random combined order and letting $f(\cdot)$ and $F(\cdot)$ be the density and distribution functions, we can write the optimality equation as follows:

$$V_n(t) = \begin{cases} \min_{0 \leq a \leq \min(t, C)} \left\{ \tilde{C}_S(a) + w(t-a) + E(V_{n-1}(t-a+Z)) \right\}, & n = 1, \dots, N \\ \tilde{C}_S(t) + w \cdot (t - \min(t, C)), & n = 0. \end{cases} \quad (4.8)$$

For period $n = 1, \dots, N$, when the consolidated load $t = 0$, there is nothing to dispatch, hence, the decision is to be idle. For $t > 0$, a consolidation decision needs to be made. Similar to the discussion in Sublevel IV.2, we can examine two cases: do not dispatch and dispatch, separately, and then choose the one with a lower cost. Denote $v_n(a|t) = \tilde{C}_S(a) + w(t-a) + E(V_{n-1}(t-a+Z))$, we define the corresponding minimum expected accumulated cost functions as

$$u_n^1(t) = v_n(0|t) = wt + E(V_{n-1}(t+Z)), \text{ for } t \geq 0, \quad (4.9)$$

$$\begin{aligned} u_n^2(t) &= \min_{0 < a \leq \min(t, C)} v_n(a|t) \\ &= \min_{0 < a \leq \min(t, C)} \{K + w(t-a) + E(V_{n-1}(t-a+Z))\}, \text{ for } t > 0. \end{aligned} \quad (4.10)$$

Then, the optimality equation of period n can be rewritten as

$$V_n(t) = \begin{cases} u_n^1(t), & \text{if } t = 0, \\ \min \{u_n^1(t), u_n^2(t)\}, & \text{if } t > 0. \end{cases}$$

IV.3.2. Exact Optimal Policy

Since the optimality equation is in a recursive form, it is essential to examine the properties of the value function in each period in order to identify the exact optimal consolidation policy.

Proposition 3 *If $V_{n-1}(t)$, the value function of period $n - 1$ satisfies the following conditions:*

(B1) $V_{n-1}(t)$ is non-decreasing in t for any n ;

(B2) $V_{n-1}(t) - V_{n-1}(t + C)$ is non-increasing in t for $t > 0$;

(B3) $\lim_{t \rightarrow \infty} V_{n-1}(t) - V_{n-1}(t + C) = -n \cdot wC$,

then the optimal decision in period n can be determined as follows: (1) if $K - (n + 1)wC \geq 0$, then to consolidate the orders for any observed size of load is optimal; (2) if $K - (n + 1)wC < 0$, then the optimal policy is in the form of a threshold policy defined by a parameter S_n , such that it is optimal to consolidate when $t < S_n$ and to dispatch $\min(t, C)$ otherwise.

Proof of Proposition 3 Since for $t > 0$,

$$V_n(t) = \min \{u_n^1(t), u_n^2(t)\} = \min \{0, u_n^2(t) - u_n^1(t)\} + u_n^1(t),$$

when $u_n^2(t) - u_n^1(t) > 0$, $V_n(t) = u_n^2(t)$; otherwise, $V_n(t) = u_n^1(t)$. Therefore, we can identify the optimal consolidation policy by examining the value of $u_n^2(t) - u_n^1(t)$.

First, according to Condition (B1), the cost $K + w(t - a) + E(V_{n-1}(t - a + Z))$ associated with dispatching quantity a for consolidated load t is non-increasing in a for a given state t , therefore $a^* = \min(t, C)$. This implies that if a shipment has to be dispatched, then the larger the dispatch quantity, the lower the cost. And $u_n^2(t)$ can be rewritten as $u_n^2(t) = v_n(\min(t, C)|t)$.

Case 1: When $0 < t \leq C$, $u_n^2(t) = K + E(V_{n-1}(Z))$, then

$$u_n^2(t) - u_n^1(t) = K + E(V_{n-1}(Z)) - tw - E(V_{n-1}(t + Z)).$$

By Condition (B1), $V_{n-1}(t+z)$ is non-decreasing in t for a known z . Since expectation is a linear operator, $u_n^2(t) - u_n^1(t)$ is non-increasing in t .

Case 2: When $t > C$, $u_n^2(t) = K + w(t - C) + E(V_{n-1}(t - C + Z))$, then

$$\begin{aligned} u_n^2(t) - u_n^1(t) &= K + w(t - C) + E(V_{n-1}(t - C + Z)) - tw - E(V_{n-1}(t + Z)) \\ &= K - wC + [E(V_{n-1}(t - C + Z)) - E(V_{n-1}(t + Z))] \end{aligned}$$

By Condition (B2) and the linearity of the expectation operator, $E(V_{n-1}(t - C + Z)) - E(V_{n-1}(t + Z))$ is non-increasing in t and so is $u_n^2(t) - u_n^1(t)$.

Summarizing Case 1 and 2, we obtain that $u_n^2(t) - u_n^1(t)$ is non-increasing in t for $t > 0$. Since both $u_n^1(t)$ and $u_n^2(t)$ are continuous in t , $u_n^2(t) - u_n^1(t)$ is continuous in t , too. In addition, notice that by Condition (B3),

$$\lim_{t \rightarrow \infty} u_n^2(t) - u_n^1(t) = K - wC + (-n \cdot wC) = K - (n + 1) \cdot wC,$$

$$\lim_{t \rightarrow 0} u_n^2(t) - u_n^1(t) = K + E(V_{n-1}(Z)) - E(V_{n-1}(Z)) = K > 0.$$

Therefore, $u_n^2(t) - u_n^1(t)$ is non-increasing on $t \in (0, \infty)$ with $\inf_{t>0}(u_n^2(t) - u_n^1(t)) = K - (n + 1) \cdot wC$ and $\sup_{t>0}(u_n^2(t) - u_n^1(t)) = K$.

Considering that when $t = 0$, $V_n(t) = u_n^1(t)$. When $K - (n + 1) \cdot wC \geq 0$, it means $u_n^2(t) \geq u_n^1(t)$ for any $t > 0$, therefore, to consolidate the load is always optimal. On the other hand, when $K - (n + 1) \cdot wC < 0$, there must exist a critical number $S_n > 0$, such that $u_n^1(t) < u_n^2(t)$ for $0 < t < S_n$ and $u_n^1(t) \geq u_n^2(t)$ for $t \geq S_n$. Therefore, a threshold policy is optimal, and when a shipment should be released, $a_n^* = \min(t, C)$.

The proof is complete. ■

Proposition 3 identifies the optimal policy for a specific period n under some assumptions of the value function of period $n - 1$. There are two possible structures for the optimal policy. One is to consolidate the orders regardless of its size, and the other is a threshold policy. Which structure is optimal in period n depends on the value of $K - (n + 1)wC$. If $K - (n + 1)wC \geq 0$, then to consolidate is optimal, otherwise, a threshold policy is optimal. It is worth mentioning that this result is built on the structural assumptions of the value function in period $n - 1$, which seems restrictive and complicated to examine. However, under careful analysis, we can prove that the value function of each period always satisfies those conditions in Proposition 3. Therefore, in each period of a finite horizon problem, the optimal consolidation decision follows the results of Proposition 3.

Theorem 5 *For a finite horizon problem, the optimal consolidation policy for period n can be described as: (1) if $K - (n + 1)wC \geq 0$, then to consolidate the orders for any observed size of load is optimal; (2) if $K - (n + 1)wC < 0$, then the optimal policy is in the form of a threshold policy defined by a parameter S_n , such that it is optimal to consolidate when $t < S_n$ and to dispatch $\min(t, C)$ otherwise.*

Proof of Theorem 5 It is sufficient to prove that the value function $V_n(t)$ satisfies Conditions (B1)–(B3) for $n = 0, \dots, N$. We prove this by induction.

- when $n = 0$, $V_0(t) = K \cdot I_{[t>0]} + w \cdot (t - \min(t, C)) = K \cdot I_{[t>0]} + w \cdot \max(0, t - C)$.

Condition (B1):

Obviously, $V_0(t)$ is non-decreasing in t .

Conditions (B2) and (B3):

For $t > 0$, $V_n(t) = K + w \cdot \max(0, t - C)$. Particularly, when $0 < t \leq C$, $V_0(t) - V_0(t + C) = K - K - wt = -wt$ which is decreasing in t . When

$t > C$, $V_0(t) - V_0(t + C) = K + w(t - C) - K - wt = -wC$ which is also non-increasing in t . Therefore, $V_0(t) - V_0(t + C)$ is non-increasing in t for $t > 0$. And, $\lim_{t \rightarrow \infty} V_0(t) - V_0(t + C) = \lim_{t \rightarrow \infty} K + w(t - C) - K - wt = -wC$.

- Suppose the value function of period $k - 1$ satisfies Conditions (B1)–(B3), we need to examine the value function of period k . According to Proposition 3, either consolidating orders or a threshold policy is optimal in period k .

Condition (B1):

Since $V_{k-1}(t)$ is non-decreasing in t by hypothesis, $u_k^1(t)$ in (4.9) is strictly increasing in t . Also, in the proof of Proposition 3 we have demonstrated that when $V_{k-1}(t)$ satisfies Conditions B1-B3, $u_k^2(t) = v_k(\min(t, C)|t) = K + w \cdot \max(0, t - C) + E(V_{k-1}(\max(0, t - C) + Z))$. Thus, $u_k^2(t)$ is also non-decreasing in t . Thus, as the minimum of $u_k^1(t)$ and $u_k^2(t)$, $V_k(t)$ is non-decreasing in t .

Conditions (B2) and (B3):

Case 1: When $K - (k + 1)wC \geq 0$, Proposition 3 implies consolidating orders is always optimal in period k , i.e., $u_k^1(t) \leq u_k^2(t)$ for any $t \geq 0$, and $V_k(t) = u_k^1(t) = wt + E(V_{k-1}(t + Z))$. Then,

$$V_k(t) - V_k(t + C) = -wC + E(V_{k-1}(t + Z)) - E(V_{k-1}(t + C + Z))$$

is non-increasing in t for $t > 0$. And

$$\begin{aligned} \lim_{t \rightarrow \infty} V_k(t) - V_k(t + C) &= \lim_{t \rightarrow \infty} -wC + E(V_{k-1}(t + Z)) - E(V_{k-1}(t + C + Z)) \\ &= -wC - k \cdot wC = -(k + 1) \cdot wC. \end{aligned}$$

Case 2: When $K - (k + 1)wC < 0$, a threshold policy is optimal in period k ,

i.e., there exists an S_k such that $V_k(t) = \begin{cases} u_k^1(t), & \text{if } t < S_k, \\ u_k^2(t), & \text{if } t \geq S_k. \end{cases}$

Case 2.1: If $V_k(t) - V_k(t + C) = u_k^1(t) - u_k^1(t + C)$, $V_k(t) - V_k(t + C)$ is non-increasing in t like we showed in Case 1.

Case 2.2: If $V_k(t) - V_k(t + C) = u_k^2(t) - u_k^2(t + C)$, we need to examine its value on two ranges separately. (1) When $0 < t \leq C$, $u_k^2(t) - u_k^2(t + C) = E(V_{k-1}(Z)) - wt - E(v_{k-1}(t + Z))$ is non-increasing in t . (2) When $t > C$, $u_k^2(t) - u_k^2(t + C) = -wC + [E(V_{k-1}(t - C + Z)) - E(V_{k-1}(t + Z))]$ is also non-increasing in t .

Case 2.3: If $V_k(t) - V_k(t + C) = u_k^1(t) - u_k^2(t + C)$, $V_k(t) - V_k(t + C) = wt + E(V_{k-1}(t + Z)) - K - wt - E(V_{k-1}(t + Z)) = -K$ is non-increasing in t .

Thus, $V_k(t) - V_k(t + C)$ is non-increasing in t . In addition, Since

$$\begin{aligned} \lim_{t \rightarrow \infty} V_k(t) - V_k(t + C) &= \lim_{t \rightarrow \infty} u_k^2(t) - u_k^2(t + C) \\ &= \lim_{t \rightarrow \infty} -wC + [E(V_{k-1}(t - C + Z)) - E(V_{k-1}(t + Z))] \\ &= -wC - k \cdot wC = -(k + 1) \cdot wC \end{aligned}$$

Then, $V_k(t)$ satisfies Conditions (B1)–(B3), and the proof is complete. ■

Theorem 5 suggests that when a set of parameters K , C and w are given, for period $i = 0, \dots, \lfloor \frac{K}{wC} \rfloor - 1$, since $K - (i + 1)wC \geq 0$, the optimal decision is to consolidate the orders regardless of the consolidated load. For period $i \geq \lfloor \frac{K}{wC} \rfloor$, the optimal policy is a threshold policy. Particularly, when $K < wC$, a threshold policy is optimal in each period of a finite horizon problem.

IV.4. Common Carriage Transportation

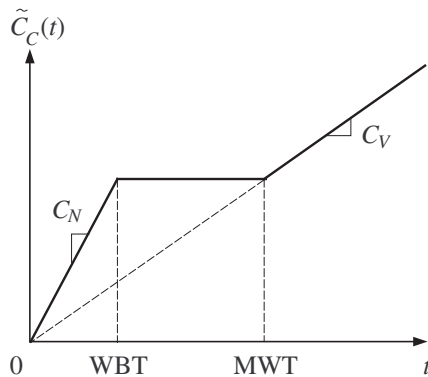
For the model of private fleet transportation, savings occur basically through spreading the fixed dispatch cost over a large dispatch quantity. In this sublevel, we consider the model where the transportation is performed by a common carrier. We examine three consolidation policies and provide sufficient conditions under which these policies are optimal.

IV.4.1. Problem Formulation

The current model modifies the one in Sublevel IV.3 by employing a common carrier to ship the orders. A common carrier can be an outside trucking company who offers a freight discount for dispatching in large quantities. Usually, the transportation cost only depends on the total weight of the shipment. As a result, the number of orders consolidated is also insignificant and can be ignored again. The cost structure we consider in this research also comes from the model by Çetinkaya and Bookbinder (2003). Denote $\tilde{C}_C(t)$ as the transportation cost for dispatching a shipment of total weight t . Then the common-carrier's tariff function in the simplest case is given by equation (4.11) and illustrated as Figure 14.

$$\tilde{C}_C(t) = \begin{cases} c_N t, & t \leq \text{WBT}, \\ c_V \text{MWT}, & \text{WBT} < t \leq \text{MWT}, \\ c_V t, & t > \text{MWT}. \end{cases} \quad (4.11)$$

In equation (4.11), C_V and C_N are the volume and non-volume freight rates, respectively. We assume $C_N > C_V + w$, because consolidation strategy will never perform better than an immediate dispatch policy when $C_N \leq C_V + w$. MWT is the stated minimum weight to obtain the quantity discount and WBT is the weight at

Figure 14: Common Carriage Transportation Cost

which the bumping clause comes into play. Denoting Z as the weight of a random combined order and letting $f(\cdot)$ and $F(\cdot)$ be the density and distribution functions, we can write the optimality equation as follows:

$$V_n(t) = \begin{cases} \min_{0 \leq a \leq t} \left\{ \tilde{C}_C(a) + w(t-a) + E(V_{n-1}(t-a+Z)) \right\}, & n = 1, \dots, N, \\ \tilde{C}_C(t), & n = 0. \end{cases} \quad (4.12)$$

For general parameter settings, the optimal policy is complicated and unknown. We investigate the exact optimal policy with some numerical examples in Sublevel IV.4.3. In Sublevel IV.4.2, we analyze the problem under the assumption that the dispatch quantity is either zero or equal to the consolidated load. This assumption allows us to examine the structure of the value function to derive some structural results.

IV.4.2. Analysis of the Optimal Policies

Assuming the “clearing property” of the consolidation policy, we can write the corresponding optimality equation as

$$V_n(t) = \begin{cases} \min \left\{ wt + E(V_{n-1}(t+Z)), \tilde{C}_C(t) + E(V_{n-1}(Z)) \right\}, & n = 1, \dots, N, \\ \tilde{C}_C(t), & n = 0. \end{cases} \quad (4.13)$$

Similar to the formulation of the private fleet cases, we define

$$u_n^1(t) = wt + E(V_{n-1}(t+Z)) \quad \text{and} \quad u_n^2(t) = \tilde{C}_C(t) + E(V_{n-1}(Z)).$$

Then, the optimality equation of period $n = 1, \dots, N$ can be rewritten as

$$V_n(t) = \begin{cases} u_n^1(t), & \text{if } t = 0, \\ \min \{u_n^1(t), u_n^2(t)\}, & \text{if } t > 0. \end{cases} \quad (4.14)$$

Although the optimal consolidation policies for this problem have been numerically examined by Higginson and Bookbinder (1995), the analytical results remain unknown due to the non-linear nature of the transportation cost. In this sublevel, we examine the optimality of some practical policies, e.g., the immediate dispatch policy and the threshold policy. We also propose a new policy and call it the (S_L, S_U) consolidation policy. Under the (S_L, S_U) policy, if the observed consolidated load t is less than S_L or greater than S_U , it is optimal to dispatch all the waiting orders; otherwise, it is optimal to consolidate until at least the arrival of the next combined order. We provide a sufficient condition under which the (S_L, S_U) policy is optimal for a finite horizon problem.

Before the analytical results are given, an important property of the value function for the common carriage case is described in Property 4.

Property 4 $V_n(t)$ is continuous and non-decreasing for $n = 0, \dots, N$.

Proof of Property 4 We prove this by induction. Notice that, when $n = 0$, $V_0(t) = \tilde{C}_C(t)$, and it is obviously continuous and non-decreasing in t . Now, suppose $V_{k-1}(t)$ is continuous and non-decreasing in t . We need to show that $V_k(t)$ is also continuous and non-decreasing in t . Since $V_{k-1}(t)$ is continuous and non-decreasing in t , $u_k^1(t) = wt + E(V_{k-1}(t + Z))$ is also continuous and strictly increasing. Also, $u_k^2(t) = \tilde{C}_C(t) + E(V_{k-1}(Z))$ is continuous and non-decreasing. Hence, $V_k(t) = \min\{u_k^1(t), u_k^2(t)\}$ is continuous and non-decreasing in t . Therefore, the proof is complete. ■

IV.4.2.1. Optimality of the Immediate Dispatch Policy

In this sublevel, we will provide a sufficient condition under which dispatching all outstanding demands is preferable to consolidating orders in all periods of a finite horizon problem.

Proposition 4 Let Z denote the weight of a random combined order received in a single period. If

$$\tilde{C}_C(t) + E\left(\tilde{C}_C(Z)\right) \leq wt + E\left(\tilde{C}_C(t + Z)\right) \quad (4.15)$$

for $\forall t \in [0, WBT]$, dispatching the consolidated load is always the optimal decision in each period no matter how many demands are held at the end of that period. In other words, shipment consolidation strategy is not favorable under this condition.

Proof of Proposition 4 Since all outstanding demands should be dispatched by the end of the planning horizon, $V_0(t) = \tilde{C}_C(t)$ and $a_0^*(t) = t$. Therefore, we only need to examine the dispatch decisions for periods 1 through N by induction.

- When $n = 1$, we have $u_1^1(t) = wt + E\left(\tilde{C}_C(t + Z)\right)$ and $u_1^2(t) = \tilde{C}_C(t) + E\left(\tilde{C}_C(Z)\right)$. By examining the first derivatives of $u_1^1(t)$ and $u_1^2(t)$ on $t \in$

(WBT, ∞) , we see that $(u_1^1)'(t) > (u_1^2)'(t)$ on (WBT, ∞) . Specifically,

Case 1: When $t \in (WBT, MWT)$, $u_1^2(t)$ is constant, so that $(u_1^2)'(t) = 0$. On the other hand, since $\tilde{C}_C(t)$ is non-decreasing in t , $E(\tilde{C}_C(t+Z))$ is also non-decreasing in t , hence, $(u_1^1)'(t) = w + \left(E(\tilde{C}_C(t+Z))\right)' \geq w$. Thus, $(u_1^1)'(t) > (u_1^2)'(t)$ on (WBT, MWT) .

Case 2: When $t \in [MWT, \infty)$, $(u_1^1)'(t) = w + \left(E(\tilde{C}_C(t+Z))\right)' = w + C_V$ and $(u_1^2)'(t) = C_V$. Therefore, $(u_1^1)'(t) > (u_1^2)'(t)$ on $[MWT, \infty)$.

In fact, equation (4.15) implies $u_1^1(t) \geq u_1^2(t)$ on $(0, WBT]$. Provided that $u_1^1(WBT) \geq u_1^2(WBT)$ and $(u_1^1)'(t) > (u_1^2)'(t)$ on (WBT, ∞) , it can be derived that $u_1^1(t) \geq u_1^2(t)$ on (WBT, ∞) . As a result, $u_1^1(t) \geq u_1^2(t)$ for $t \in (0, \infty)$. Also, $u_1^1(0) = u_1^2(0)$, and hence, $V_1(t) = \min\{u_1^1(t), u_1^2(t)\} = u_1^2(t)$. In other words, it is always optimal to dispatch the consolidated load in period 1.

- Suppose that dispatching all the consolidated load is always optimal in period $k-1$, i.e., $V_{k-1}(t) = \tilde{C}_C(t) + E(V_{k-2}(Z))$, we need to show the optimality of an immediate dispatch policy for period k . Notice that,

$$\begin{aligned} u_k^1(t) &= wt + E(V_{k-1}(t+Z)) \\ &= wt + E\left(\tilde{C}_C(t+Z)\right) + E(V_{k-2}(Z)) = u_1^1(t) + E(V_{k-2}(Z)), \end{aligned}$$

$$\begin{aligned} u_k^2(t) &= \tilde{C}_C(t) + E(V_{k-1}(Z)) \\ &= \tilde{C}_C(t) + E\left(\tilde{C}_C(Z)\right) + E(V_{k-2}(Z)) = u_1^2(t) + E(V_{k-2}(Z)). \end{aligned}$$

Therefore, $V_k(t) = \min\{u_k^1(t), u_k^2(t)\} = u_1^2(t) + E(V_{k-2}(Z)) = u_k^2(t)$, and to dispatch the consolidated load is always optimal for period k , too.

Therefore, with the assumption (4.15), the depot should dispatch the consolidated demands in each period. ■

According to equation (4.14), the optimality equation for period 1 can be written as

$$V_1(t) = \min \left\{ wt + E \left(\tilde{C}_C(t + Z) \right), \tilde{C}_C(t) + E \left(\tilde{C}_C(Z) \right) \right\}. \quad (4.16)$$

Within the braces of expression (4.16), the first term represents the expected cost-to-go for dispatching the entire consolidated load, and the second term represents the expected cost-to-go for continuing to consolidate the orders. Proposition 4 says that if dispatching the consolidated load is optimal in period 1 for the weight less than or equal to WBT , then it is optimal to dispatch the consolidated load in each period and disregard how many demands are held at the end of that period.

For special cases of random demands, the sufficient conditions can be expressed explicitly. For example,

(1). Bernoulli (p) Demand

When the weight of random demands in all periods are *i.i.d.* and follow the Bernoulli(p) distribution, i.e., each order includes one unit with a possibility of p , and the possibility of no order arriving in the period is $(1 - p)$. For $\forall t \in [0, WBT]$, we have

$$\begin{cases} wt + E \left(\tilde{C}_C(t + Z) \right) = wt + p \cdot \tilde{C}_C(t + 1) + (1 - p) \cdot \tilde{C}_C(t). \\ \tilde{C}_C(t) + E \left(\tilde{C}_C(Z) \right) = \tilde{C}_C(t) + p \cdot \tilde{C}_C(1), \end{cases}$$

In order to satisfy the condition (4.15), we require

$$\tilde{C}_C(t) + p \cdot \tilde{C}_C(1) \leq wt + p \cdot \tilde{C}_C(t + 1) + (1 - p) \cdot \tilde{C}_C(t).$$

Thus,

$$\tilde{C}_C(t+1) - \tilde{C}_C(t) + \frac{wt}{p} \geq \tilde{C}_C(1) = C_N$$

When $t \in [0, WBT - 1]$, the left hand side is at least C_N , so the inequality holds. When $t = WBT$, the left hand side is $\frac{w \cdot WBT}{p}$, then we have the sufficient condition for optimality of the immediate dispatch policy in a finite horizon problem with Bernoulli(p) random demand as:

$$\frac{w \cdot WBT}{C_N} \geq p. \quad (4.17)$$

(2). Uniform $(0, b)$ Demand

A complete analysis of the uniform distributed demand is computationally challenging due to the relationship between the parameters b , WBT and MWT . Here, we only analyze the simplest case where $b \leq \min\{WBT, MWT - WBT\}$. For $\forall t \in [0, WBT]$,

$$\begin{cases} u_1^1(t) = wt + E\left(\tilde{C}_C(t+Z)\right) = wt + \int_0^b \tilde{C}_C(t+z) \frac{1}{b} dz. \\ u_1^2(t) = \tilde{C}_C(t) + E\left(\tilde{C}_C(Z)\right) = C_N t + \int_0^b \tilde{C}_C(z) \frac{1}{b} dz, \end{cases}$$

$$u_1^2(t) - u_1^1(t) = (C_N - w)t - \int_0^b \tilde{C}_C(t+z) \frac{1}{b} dz + \int_0^b \tilde{C}_C(z) \frac{1}{b} dz.$$

Case 1: $0 \leq t \leq WBT - b$

$$\begin{aligned} u_1^2(t) - u_1^1(t) &= (C_N - w)t - \int_0^b C_N(t+z) \frac{1}{b} dz + \int_0^b C_N z \frac{1}{b} dz \\ &= (C_N - w)t - \int_0^b C_N t \frac{1}{b} dz \\ &= (C_N - w)t - C_N t = -wt \leq 0. \end{aligned}$$

Case 2: $WBT - b < t \leq WBT$

$$\begin{aligned}
u_1^2(t) - u_1^1(t) &= (C_N - w)t + \int_0^b C_N z \frac{1}{b} dz - \int_0^{WBT-t} C_N(t+z) \frac{1}{b} dz \\
&\quad - \int_{WBT-t}^b C_N \cdot WBT \frac{1}{b} dz \\
&= (C_N - w)t + \frac{C_N}{2b} [b^2 - (WBT^2 - t^2) - 2WBT(b - WBT + t)] \\
&= \frac{C_N}{2b} \left([t - (WBT - b)]^2 - \frac{2bwt}{C_N} \right).
\end{aligned}$$

Solve the equation $[t - (WBT - b)]^2 - \frac{2bwt}{C_N} = 0$ to obtain two roots for t as

$$\begin{aligned}
t_1^* &= (WBT - b) + \frac{bw}{C_N} + \sqrt{\frac{b^2w^2}{C_N^2} + 2(WBT - b)\frac{bw}{C_N}}, \\
t_2^* &= (WBT - b) + \frac{bw}{C_N} - \sqrt{\frac{b^2w^2}{C_N^2} + 2(WBT - b)\frac{bw}{C_N}}.
\end{aligned}$$

When $u_1^2(t) - u_1^1(t) \leq 0$, it is sufficient to have $t_2^* \leq WBT - b < t \leq WBT \leq t_1^*$.

And it is easy to verify that the left inequality always holds, and the right inequality holds if and only if $b \leq \frac{2 \cdot WBT \cdot w}{C_N}$. Thus, we have a sufficient condition

for the uniform demand case as

$$b \leq \min \left\{ WBT, MWT - WBT, \frac{2 \cdot WBT \cdot h}{C_N} \right\}.$$

IV.4.2.2. Optimality of the Threshold Policy

A threshold policy is easy to understand and convenient to apply in practice. Under certain conditions, a threshold policy is optimal for the common carriage consolidation problems.

Lemma 5 *Suppose $f_1(x)$ and $f_2(x)$ are two continuous and non-decreasing functions on $[0, \infty)$. If the following conditions:*

(C1) $f_2(x) = \tilde{C}_C(x) + M$, where M is a nonnegative constant;

$$(C2) \quad f_1(0) = f_2(0);$$

$$(C3) \quad C_N = f_2'(x) \geq f_1'(x) \text{ on } [0, WBT];$$

$$(C4) \quad f_2'(x) \leq f_1'(x) \leq C_N \text{ on } (WBT, \infty);$$

$$(C5) \quad f_1(x) \geq f_2(x) \text{ on } [MWT, \infty),$$

are satisfied, then $g(x) = \min \{f_1(x), f_2(x)\}$ is continuous and non-decreasing in x .

Furthermore, there exists a number $S \in [WBT, MWT]$, such that,

$$g(x) = \begin{cases} f_1(x), & \text{if } x \in [0, S], \\ f_2(x), & \text{if } x \in (S, \infty). \end{cases} \quad (4.18)$$

Proof of Lemma 5 Since $f_1(x)$ and $f_2(x)$ are both continuous and non-decreasing, obviously, $g(x)$ is continuous and non-decreasing. Divide the real line into three segments as follows and examine the function $g(x)$ on each segment:

Segment 1: On $[0, WBT]$, by Conditions (C2) and (C3), $f_1(x) \leq f_2(x)$, and hence, $g(x) = f_1(x)$.

Segment 2: On $(WBT, MWT]$, $f_1(x)$ is non-decreasing while $f_2(x)$ is constant (by Condition (C1)). Since $f_1(MWT) \geq f_2(MWT)$ (by Condition (C5)) and $f_1(WBT) \leq f_2(WBT)$, there must exist a value $S \in [WBT, MWT]$ such that on $(WBT, S]$, $f_1(x) < f_2(x)$ and $g(x) = f_1(x)$; on $(S, MWT]$, $f_1(x) \geq f_2(x)$ and $g(x) = f_2(x)$.

Segment 3: On (MWT, ∞) , according to Condition (C5), $g(x) = f_2(x)$.

In summary, $g(x)$ can be described by $f_1(x)$ and $f_2(x)$ as equation (4.18). ■

In Lemma 5, if we replace $f_1(x)$ and $f_2(x)$ with the two value functions for dispatching the orders or consolidating in period n , then $g(x)$ actually represents the optimality equation of period n . And following equation (4.18), we can determine that in this period, a threshold policy is optimal for the system. Although Lemma 5 only gives the structure of the optimal policy for a single period problem, Proposition

5 provides a sufficient condition under which the threshold policy is optimal for a finite horizon problem.

Proposition 5 *If the following inequalities are satisfied, i.e.,*

$$E\left(\tilde{C}_C(Z)\right) - C_V \cdot \mu_Z \leq w \cdot MWT, \text{ and} \quad (4.19)$$

$$1 - F(MWT) \geq \frac{w}{C_N - C_V}, \quad (4.20)$$

where Z denotes the weight of a random combined order received in a single period, and μ_Z represents the mean, then for a finite horizon problem, a threshold policy is optimal in each period. That is, for any period n , there exists a value S_T^n such that if the weight of the consolidated load $t_n \geq S_T^n$, it is optimal to dispatch the entire load; otherwise, continuing to consolidate the demands is preferable.

Proof of Proposition 5 The main idea is to show that if the costs for two alternative decisions (consolidate or dispatch) are denoted by f_1 and f_2 , they satisfy the Conditions (C1)–(C5) in Lemma 5. The proof is conducted by induction.

- When $n = 0$, $V_0(t) = \tilde{C}_C(t)$. Without violating the definition of $V_0(t)$, let $u_0^1(t) = C_N \cdot t$ and $u_0^2(t) = \tilde{C}_C(t)$. Then $u_0^1(t)$ and $u_0^2(t)$ are both continuous and non-decreasing on $[0, \infty)$. It is easy to see that Conditions (C1)–(C5) are all satisfied.
- Suppose $u_{k-1}^1(t)$ and $u_{k-1}^2(t)$, the cost functions of two alternatives in period $k - 1$, satisfy Conditions (C1)–(C5). We need to show that the conditions are also satisfied by $u_k^1(t)$ and $u_k^2(t)$. First,

$$\begin{cases} u_k^1(t) = wt + E(V_{k-1}(t + Z)) = wt + \int_0^\infty V_{k-1}(t + z)f(z)dz, \\ u_k^2(t) = \tilde{C}_C(t) + E(V_{k-1}(Z)) = \tilde{C}_C(t) + \int_0^\infty V_{k-1}(z)f(z)dz. \end{cases}$$

By Property 4, it is immediate that $u_k^1(t)$ and $u_k^2(t)$ are both continuous and non-decreasing.

Conditions (C1) and (C2):

Let $M = \int_0^\infty V_{k-1}(z)f(z)dz$. Then M is nonnegative and $u_k^2(t) = \tilde{C}_C(t) + M$. Also, $u_k^1(0) = \int_0^\infty V_{k-1}(z)f(z)dz = u_k^2(0)$.

Condition (C3):

Since we assume that $u_{k-1}^1(t)$ and $u_{k-1}^2(t)$ satisfy Conditions (C1)–(C5), by Lemma 5 there exists a number $S_{k-1} \in [WBT, MWT]$ such that $V_{k-1}(t) = u_{k-1}^1(t)$ if $t \in [0, S_{k-1}]$, and $V_{k-1}(t) = u_{k-1}^2(t)$ otherwise. Then for $t \in [0, WBT]$, the first derivative of $u_k^1(t)$ is

$$\begin{aligned}
(u_k^1)'(t) &= w + \int_0^\infty V'_{k-1}(t+z)f(z)dz \\
&= w + \int_0^{S_{k-1}-t} V'_{k-1}(t+z)f(z)dz + \int_{S_{k-1}-t}^\infty V'_{k-1}(t+z)f(z)dz \\
&= w + \int_0^{S_{k-1}-t} (u_{k-1}^1)'(t+z)f(z)dz + \int_{S_{k-1}-t}^\infty (u_{k-1}^2)'(t+z)f(z)dz \\
&= w + \int_0^{S_{k-1}-t} (u_{k-1}^1)'(t+z)f(z)dz + \int_{MWT-t}^\infty (u_{k-1}^2)'(t+z)f(z)dz \\
&\quad \left(\text{Since } (u_{k-1}^2)'(t) = 0 \text{ on } (WBT, MWT) \right) \\
&\leq w + \int_0^{MWT-t} (u_{k-1}^1)'(t+z)f(z)dz + \int_{MWT-t}^\infty (u_{k-1}^2)'(t+z)f(z)dz \\
&\leq w + \int_0^{MWT-t} C_N f(z)dz + \int_{MWT-t}^\infty C_V f(z)dz \\
&= w + \int_0^{MWT} C_N f(z)dz + \int_{MWT}^\infty C_V f(z)dz - \int_{MWT-t}^{MWT} (C_N - C_V)f(z)dz \\
&\leq w + \int_0^{MWT} C_N f(z)dz + \int_{MWT}^\infty C_V f(z)dz \\
&= w + C_N - \int_{MWT}^\infty (C_N - C_V)f(z)dz \leq w + C_N - w \\
&\quad \text{(by equation (4.20))}
\end{aligned}$$

$$= C_N = (u_k^2)'(t).$$

Condition (C4):

We examine the relationship between $(u_k^1)'(t)$ and $(u_k^2)'(t)$ on two mutually disjoint complementary subsets of (WBT, ∞) .

Case 1: For $t \in (WBT, MWT)$, by the definition of $u_k^2(t)$, $(u_k^2)'(t) = 0$. And $(u_k^1)'(t) = w + \int_0^\infty V'_{k-1}(t+z)f(z)dz \geq w$. Thus, $(u_k^1)'(t) \geq (u_k^2)'(t)$.

Case 2: For $t \in [MWT, \infty)$, $(u_k^2)'(t) = C_V$. And

$$\begin{aligned} (u_k^1)'(t) &= w + \int_0^\infty V'_{k-1}(t+z)f(z)dz \\ &= w + \int_0^\infty (u_{k-1}^2)'(t+z)f(z)dz = w + C_V. \end{aligned}$$

Thus, $(u_k^1)'(t) \geq (u_k^2)'(t)$.

At this point we have proved the inequality $(u_k^1)'(t) \geq (u_k^2)'(t)$ on (WBT, ∞) .

To complete the examination of Condition (C4), we still need to show $(u_k^1)'(t) \leq C_N$ on (WBT, ∞) . Note that, $(u_k^1)'(t) = w + \int_0^\infty V'_{k-1}(t+z)f(z)dz$, then

Case 1: If $t \in (WBT, S_{k-1})$, we have

$$\begin{aligned} (u_k^1)'(t) &= w + \int_0^{S_{k-1}-t} V'_{k-1}(t+z)f(z)dz + \int_{S_{k-1}-t}^\infty V'_{k-1}(t+z)f(z)dz \\ &= w + \int_0^{S_{k-1}-t} (u_{k-1}^1)'(t+z)f(z)dz + \int_{S_{k-1}-t}^\infty (u_{k-1}^2)'(t+z)f(z)dz \\ &= w + \int_0^{S_{k-1}-t} (u_{k-1}^1)'(t+z)f(z)dz + \int_{MWT-t}^\infty (u_{k-1}^2)'(t+z)f(z)dz \\ &\leq w + \int_0^{MWT-t} (u_{k-1}^1)'(t+z)f(z)dz + \int_{MWT-t}^\infty (u_{k-1}^2)'(t+z)f(z)dz \\ &\leq w + \int_0^{MWT-t} C_N f(z)dz + \int_{MWT-t}^\infty C_V f(z)dz \\ &= w + \int_0^{MWT} C_N f(z)dz + \int_{MWT}^\infty C_V f(z)dz - \int_{MWT-t}^{MWT} (C_N - C_V)f(z)dz \end{aligned}$$

$$\begin{aligned}
&\leq w + \int_0^{MWT} C_N f(z) dz + \int_{MWT}^{\infty} C_V f(z) dz \\
&= w + C_N - \int_{MWT}^{\infty} (C_N - C_V) f(z) dz \leq w + C_N - w = C_N.
\end{aligned}$$

Case 2: If $t \in [S_{k-1}, MWT)$, we have

$$\begin{aligned}
(u_k^1)'(t) &= w + \int_0^{MWT-t} V'_{k-1}(t+z) f(z) dz + \int_{MWT-t}^{\infty} V'_{k-1}(t+z) f(z) dz \\
&= w + \int_0^{MWT-t} (u_{k-1}^2)'(t+z) f(z) dz + \int_{MWT-t}^{\infty} (u_{k-1}^2)'(t+z) f(z) dz \\
&= w + 0 + \int_{MWT-t}^{\infty} C_V f(z) dz \leq w + C_V < C_N.
\end{aligned}$$

Case 3: If $t \in [MWT, \infty)$, we have

$$(u_k^1)'(t) = w + \int_0^{\infty} V'_{k-1}(t+z) f(z) dz = w + \int_0^{\infty} C_V f(z) dz = w + C_V < C_N.$$

Hence, $(u_k^1)'(t) \leq C_N$ on (WBT, ∞) .

Condition (C5):

Since we have already shown that on (WBT, ∞) , $(u_k^1)'(t) \geq (u_k^2)'(t)$, in order to show $u_k^1(x) \geq u_k^2(x)$ on $[MWT, \infty)$, it is sufficient to show $u_k^1(MWT) \geq u_k^2(MWT)$.

$$\begin{aligned}
&u_k^1(MWT) - u_k^2(MWT) \\
&= w \cdot MWT + \int_0^{\infty} V_{k-1}(MWT+z) f(z) dz - C_V \cdot MWT - \int_0^{\infty} V_{k-1}(z) f(z) dz \\
&\geq w \cdot MWT + \int_0^{\infty} V_{k-1}(MWT+z) f(z) dz - C_V \cdot MWT - \int_0^{\infty} u_{k-1}^2(z) f(z) dz \\
&= w \cdot MWT + \int_0^{\infty} u_{k-1}^2(MWT+z) f(z) dz - C_V \cdot MWT - \int_0^{\infty} u_{k-1}^2(z) f(z) dz \\
&= w \cdot MWT + \int_0^{\infty} [C_V(MWT+z) + u_{k-1}^2(0)] f(z) dz - C_V \cdot MWT \\
&\quad - \int_0^{\infty} [\tilde{C}_C(z) + u_{k-1}^2(0)] f(z) dz
\end{aligned}$$

$$= w \cdot MWT - \int_0^\infty [\tilde{C}_C(z) - C_V \cdot z] f(z) dz \geq 0.$$

The last inequality holds due to the inequality (4.19). Therefore, $u_k^1(t) \geq u_k^2(t)$ on $[MWT, \infty)$.

In summary, the cost functions $u_n^1(t)$ and $u_n^2(t)$ of two alternatives in each period always satisfy the Conditions (C1)–(C5). By applying Lemma 5 to each period, the optimality of a threshold policy in a finite horizon setting is demonstrated. ■

In inequality (4.19), $E(\tilde{C}_C(Z))$ is the expected cost for shipping a combined random demand if the cost structure (4.11) is employed. $C_V \cdot \mu_Z$ is the expected cost if this demand is shipped at the volume freight rate C_V . Since C_V is the lowest feasible freight rate, the term on the left hand side of the inequality (4.19) represents the saving if a random demand is dispatched at the volume rate instead of following the cost scheme (4.11). Condition (4.19) in Proposition 5 requires such saving should not exceed the cost of holding MWT weight of retailer orders for one period.

Condition (4.20) means that the probability for the random order's weight exceeding MWT should be at least $\frac{w}{C_N - C_V}$. Obviously, we need to make sure that the right-hand side of inequality (4.20) is less than 1. Fortunately, $w < C_N - C_V$ is assumed for practical purposes. In detail, C_N is the highest cost for dispatching one unit weight of demand if it is dispatched without being held, and $w + C_V$ is the lowest cost for dispatching that unit if it has to be consolidated. If $w + C_V \geq C_N$, shipment consolidation strategy is never financially justifiable, i.e., it will never perform better than an immediate dispatch policy.

For special cases of random demands, the sufficient conditions also can be expressed explicitly. For example,

- (1). Discrete demand with density $P(Z = b) = p$ and $P(Z = 0) = 1 - p$

In order to validate the inequality (4.20), we must have $b \geq MWT$. Then

conditions (4.19) and (4.20) can be written as:

$$\tilde{C}_C(b) \cdot p - C_V b \cdot p \leq w \cdot MWT,$$

and

$$(C_N - C_V)p \geq w.$$

Therefore, we have the sufficient condition

$$b \geq MWT \text{ and } (C_N - C_V)p \geq w.$$

(2). Uniform $(0, b)$ demand

Same as the previous case, b should be at least MWT , and we can rewrite the condition (4.19) as $\frac{1}{MWT} \int_0^\infty \tilde{C}_C(z) f(z) dz - \int_0^\infty C_V z f(z) dz \leq w$. The left hand side of the inequality is

$$\begin{aligned} LHS &= \frac{1}{MWT} \left[\int_0^{WBT} (C_N - C_V) z \frac{1}{b} dz + \int_{WBT}^{MWT} C_V (MWT - z) \frac{1}{b} dz \right] \\ &= \frac{1}{MWT} \left[\frac{C_N - C_V}{2b} \cdot WBT^2 + \frac{C_V (MWT^2 - MWT \cdot WBT)}{b} \right. \\ &\quad \left. - \frac{C_V}{2b} (MWT^2 - WBT^2) \right] \\ &= \frac{1}{2b \cdot MWT} [C_V \cdot MWT^2 - C_N \cdot WBT^2] \\ &= \frac{C_V}{2b} (MWT - WBT). \end{aligned}$$

Similarly, we rewrite (4.20) as $\int_{MWT}^\infty (C_N - C_V) f(z) dz \geq w$. Then its left hand side is

$$LHS = \int_{MWT}^b (C_N - C_V) \frac{1}{b} dz = \frac{C_N - C_V}{b} (b - MWT).$$

Hence, we obtain the sufficient conditions as
$$\begin{cases} b \geq MWT, \\ \frac{C_V}{2b}(MWT - WBT) \leq h, \\ \frac{C_N - C_v}{b}(b - MWT) \geq h. \end{cases}$$

IV.4.2.3. Optimality of (S_L, S_U) Policy

For the common carriage consolidation problem, Higginson and Bookbinder (1995) identify the triangularity of the optimal policies numerically. The triangular property means “shipment is the preferred action with very large or very small accumulated weights.” In the current research, we define a policy with such feature as the (S_L, S_U) policy. Specifically, if the consolidated load t in period n satisfies $t < S_L^n$ or $t \geq S_U^n$, then it is optimal to dispatch all the held demands; otherwise, it is optimal to continue to consolidate until at least the arrival of the next order. We provide sufficient conditions under which the (S_L, S_U) policy is optimal for a finite horizon problem.

Lemma 6 *Suppose $f_1(x)$ and $f_2(x)$ are two continuous functions on $[0, \infty)$. If the following conditions:*

($\widehat{C1}$) $f_2(x) = \tilde{C}_C(x) + M$, where M is a nonnegative constant;

($\widehat{C2}$) $f_1(x)$ is strictly increasing and $f_1(0) = f_2(0)$;

($\widehat{C3}$) there exists a value $\bar{x} \in [WBT, MWT]$, s.t.

($\widehat{C3.a}$) $f_1(x)$ is concave on $[0, \bar{x}]$,

($\widehat{C3.b}$) $f_1(\bar{x}) \geq f_2(\bar{x}) = M + C_N \cdot WBT$, and

($\widehat{C3.c}$) $f_1'(x) \geq f_2'(x)$ for $x \in [\bar{x}, \infty)$;

are satisfied, then the function $g(x) = \min\{f_1(x), f_2(x)\}$ is continuous and non-decreasing in x and also concave on $[0, MWT]$. In addition, there exist two values of

S_L and S_U such that $0 \leq S_L \leq S_U \leq \bar{x}$ and

$$g(x) = \begin{cases} f_2(x), & \text{if } x \in [0, S_L), \\ f_1(x), & \text{if } x \in [S_L, S_U], \\ f_2(x), & \text{if } x \in (S_U, \infty). \end{cases} \quad (4.21)$$

Proof of Lemma 6 By Conditions $(\widehat{C1})$ and $(\widehat{C2})$, $f_1(x)$ and $f_2(x)$ are both non-decreasing and continuous, thus, their minimum $g(x)$ is obviously non-decreasing and continuous.

By Conditions $(\widehat{C1})$ and $(\widehat{C3.a})$, $f_1(x)$ and $f_2(x)$ are both concave on $[0, \bar{x}]$, then $g(x)$ is also concave on $[0, \bar{x}]$. In addition, Conditions $(\widehat{C3.b})$ and $\widehat{C3.c}$ together imply that for $x \in (\bar{x}, MWT]$, $f_1(x) \geq f_2(x)$. Since $f_2(x)$ is constant on $(\bar{x}, MWT]$, we have $g(x) = f_2(x)$ and $g(x)$ is constant on $(\bar{x}, MWT]$. Therefore, $g(x)$ is concave on $[0, MWT]$.

To show the existence of S_L and S_U , we first show that other than at the origin, $f_1(x)$ and $f_2(x)$ have at most two intersections. Using values of WBT and \bar{x} , we divide the region $(0, \infty)$ into three segments and examine each segment for possible occurrence of intersections.

Segment 1: On $(0, WBT]$, $f_2(x)$ is linear according to Condition $(\widehat{C1})$ and the definition of $\tilde{C}_C(\cdot)$. Also, by Condition $(\widehat{C3.a})$, $f_1(x)$ is concave on $[0, WBT]$, hence, $(f_1 - f_2)(x) := f_1(x) - f_2(x)$ is concave on $[0, WBT]$. Since a concave function can take value of 0 at most twice, and we already have $(f_1 - f_2)(0) = 0$, there exists at most ONE point $x^* \in (0, WBT]$ such that $(f_1 - f_2)(x^*) = 0$. Thus, $f_1(x)$ and $f_2(x)$ intersect at most once on $(0, WBT]$.

Segment 2: On $(WBT, \bar{x}]$, $f_2(x)$ is constant. According to Condition $(\widehat{C2})$, $f_1(x)$ is strictly increasing, hence, $(f_1 - f_2)(x)$ is also strictly increasing on $(WBT, \bar{x}]$. For this reason, it can take value of 0 at most once, which means on this segment $f_1(x)$

and $f_2(x)$ intersect at most once.

Segment 3: On (\bar{x}, ∞) , by Conditions $(\widehat{C3}.b)$ and $(\widehat{C3}.c)$, we have $f_1(x) \geq f_2(x)$ on (\bar{x}, ∞) . Hence, they do not intersect on this segment.

Therefore, on $(0, \infty)$, $f_1(x)$ and $f_2(x)$ intersect at most twice. There are three possible cases necessary to be examined:

Case 1: If $f_1(x)$ and $f_2(x)$ do not intersect other than at the origin, by Condition $(\widehat{C3}.b)$ it must be $f_1(x) > f_2(x)$ on $(0, \infty)$. Then $g(x) = f_2(x)$. If we let $S_L = S_U = 0$, equation (4.21) holds.

Case 2: If $f_1(x)$ and $f_2(x)$ intersect once on $(0, \infty)$, by Conditions $(\widehat{C3}.a)$ and $(\widehat{C3}.b)$, the intersection must occur at a point located in $[WBT, \bar{x}]$ such that on its left hand side $f_1(x) < f_2(x)$ and on its right hand side $f_1(x) > f_2(x)$. If we let S_U be the intersection point, and let $S_L = 0$, then equation (4.21) holds.

Case 3: If $f_1(x)$ and $f_2(x)$ intersect twice on $(0, \infty)$, then just let S_L be the first intersection point ($0 < S_L \leq WBT$) and S_U be the second one ($WBT < S_U \leq \bar{x}$). Guaranteed by Condition $(\widehat{C3}.b)$, the equation (4.21) again holds. The proof is complete. ■

Similar to Lemma 5, Lemma 6 proves the optimality of a (S_L, S_U) policy for a single period problem under certain conditions. Sufficient conditions that guarantee the optimality of the (S_L, S_U) policy in a multi-period problem is provided in Proposition 6.

Proposition 6 *If the demand has a Uniform $[0, \alpha]$ distribution, and the parameter α satisfies*

$$\alpha \leq \min \left\{ \frac{w \cdot MWT}{w + C_N/2}, MWT - WBT, WBT \right\}, \quad (4.22)$$

then for a finite horizon problem, an (S_L, S_U) policy is optimal in each period. That is, for any period n , there exists a pair of parameters S_L^n and S_U^n such that if the

weight of the consolidated load $t < S_L^n$ or $t > S_U^n$, it is optimal to dispatch the entire load; otherwise, continuing to consolidate the orders is preferable.

Proof of Proposition 6 We need to show that for any $n = 0, \dots, N$, the costs for two alternatives, i.e., $u_n^1(t)$ and $u_n^2(t)$ always satisfy the Conditions $(\widehat{C1})$ – $(\widehat{C3})$. We prove this by induction.

- When $n = 0$, the terminal cost $V_0(t) = \tilde{C}_C(t)$ can be rewritten as $V_0(t) = \min \{u_0^1(t), u_0^2(t)\}$, where $u_0^1(t) = 2C_N \cdot t$ and $u_0^2(t) = \tilde{C}_C(t)$. It is obvious that both $u_0^1(t)$ and $u_0^2(t)$ are continuous.

Conditions $(\widehat{C1})$ and $(\widehat{C2})$:

$u_0^2(t) = \tilde{C}_C(t) + M$, here $M = 0$ is nonnegative. $u_0^1(t) = 2C_N \cdot t$ is strictly increasing and $u_0^1(0) = 0 = u_0^2(0)$.

Condition $(\widehat{C3})$:

Let $\bar{t} = MWT - \alpha$. Since $\alpha \leq MWT - WBT$, $\bar{t} = MWT - \alpha \geq WBT$, so $\bar{t} \in [WBT, MWT]$ is satisfied, and

- **Condition $(\widehat{C3.a})$:** $u_0^1(t) = 2C_N \cdot t$ on $[0, \bar{t}]$. $(u_0^1)''(t) = 0$, thus $u_0^1(t)$ is concave on $[0, \bar{t}]$;
 - **Condition $(\widehat{C3.b})$:** $u_0^1(\bar{t}) = 2C_N \cdot \bar{t} > C_N \cdot \bar{t} \geq C_N \cdot WBT = u_0^2(\bar{t})$;
 - **Condition $(\widehat{C3.c})$:** For $t \in [\bar{t}, \infty)$, $(u_0^1)'(t) = 2C_N > C_N > C_V \geq (u_0^2)'(t)$.
- Suppose in period $k - 1$, $u_{k-1}^1(t)$ and $u_{k-1}^2(t)$ satisfy Conditions $(\widehat{C1})$ – $(\widehat{C3})$, we want to show $u_k^1(t)$ and $u_k^2(t)$ also satisfy Conditions $(\widehat{C1})$ – $(\widehat{C3})$. First, by their definition, $u_k^1(t)$ and $u_k^2(t)$ are both continuous.

Conditions $(\widehat{C1})$ and $(\widehat{C2})$:

$u_k^2(t) = \tilde{C}_C(t) + E(V_{k-1}(Z))$, where $M = E(V_{k-1}(Z))$ is a nonnegative constant.

By Property 4, $V_{k-1}(t)$ is non-decreasing, then $E(V_{k-1}(t+Z))$ is also non-decreasing. Since wt is strictly increasing, $u_k^1(t) = wt + E(V_{k-1}(t+Z))$ is strictly increasing and $u_k^1(0) = E(V_{k-1}(Z)) = u_k^2(0)$.

Condition $(\widehat{C3})$:

In each period, we always let $\bar{t} = MWT - \alpha$. Then,

- **Condition $(\widehat{C3.a})$:** By Lemma 6, when $u_{k-1}^1(t)$ and $u_{k-1}^2(t)$ satisfy Conditions $(\widehat{C1})$ – $(\widehat{C3})$, $V_{k-1}(t)$ is concave on $[0, MWT]$, hence, its first derivative $V'_{k-1}(t)$ is decreasing on $[0, MWT]$. Hence, $(u_k^1)'(t) = w + \int_0^\alpha V'_{k-1}(t+z) \frac{1}{\alpha} dz$ is also decreasing in t , which implies that $u_k^1(t)$ is concave on $[0, \bar{t}]$.
- **Condition $(\widehat{C3.b})$:** We can compute the function values of $u_k^1(t)$ and $u_k^2(t)$ at \bar{t} as follows. Note that $\bar{t} = MWT - \alpha$.

$$\begin{aligned} u_k^1(\bar{t}) &= w\bar{t} + \int_0^\alpha V_{k-1}(\bar{t} + z) \frac{1}{\alpha} dz \\ &= w\bar{t} + \int_0^\alpha (V_{k-1}(0) + C_N \cdot WBT) \frac{1}{\alpha} dz \\ &\quad (\text{Since } V_{k-1}(t) = u_{k-1}^2(t) \text{ on } [\bar{t}, MWT]) \\ &= w \cdot (MWT - \alpha) + V_{k-1}(0) + C_N \cdot WBT. \end{aligned}$$

Since $\alpha \leq \frac{w \cdot MWT}{w + C_N/2}$, we have $w \cdot (MWT - \alpha) \geq \frac{C_N \cdot \alpha}{2} = \int_0^\alpha C_N \cdot z \cdot \frac{1}{\alpha} dz$.

Thus,

$$\begin{aligned} u_k^1(\bar{t}) &\geq V_{k-1}(0) + C_N \cdot WBT + \int_0^\alpha C_N \cdot z \cdot \frac{1}{\alpha} dz \\ &= C_N \cdot WBT + \int_0^\alpha (V_{k-1}(0) + C_N \cdot z) \frac{1}{\alpha} dz \\ &= C_N \cdot WBT + \int_0^\alpha (u_{k-1}^2(0) + C_N \cdot z) \frac{1}{\alpha} dz \\ &= C_N \cdot WBT + \int_0^\alpha u_{k-1}^2(z) \frac{1}{\alpha} dz \geq C_N \cdot WBT + \int_0^\alpha V_{k-1}(z) \frac{1}{\alpha} dz \end{aligned}$$

$$= u_k^2(\bar{t}).$$

- **Condition** ($\widehat{C3.c}$): We examine this on two segments. On $[\bar{t}, MWT]$, we have $(u_k^2)'(t) = 0$ and $(u_k^1)'(t) \geq w > 0$, thus, $(u_k^1)'(t) \geq (u_k^2)'(t)$. On $[MWT, \infty)$, $(u_k^2)'(t) = C_V$, and $(u_k^1)'(t) = w + \int_0^\alpha V'_{k-1}(t+z) \frac{1}{\alpha} dz = w + \int_0^\alpha C_V \frac{1}{\alpha} dz = w + C_V > (u_k^2)'(t)$.

Thus, $u_k^1(t)$ and $u_k^2(t)$ also satisfy Conditions ($\widehat{C1}$)–($\widehat{C3}$), and hence, a (S_L, S_U) policy is optimal for the dispatch decision in each period. ■

For the case of uniformly distributed demand, Proposition 6 gives a sufficient condition under which the (S_L, S_U) policy is optimal for a finite horizon problem. For the (S_L, S_U) policy, $0 \leq S_L \leq S_U < \infty$, where S_L and S_U can be 0. It is worth noting that the (S_L, S_U) policy becomes a threshold policy when $S_L = 0$ and $S_U > 0$. And when $S_L = S_U = 0$, the policy is equivalent to an immediate shipment policy. In the (S_L, S_U) policy with $0 < S_L < S_U$, sometimes it is optimal to dispatch a consolidated load with weight less than S_L . The reason is that when the reviewed consolidated load is so low that a large weight that is good enough to receive shipment discount cannot be attained within a reasonable holding periods, it is preferable to ship this small load at a higher cost immediately.

IV.4.3. Computational Studies

In Sublevel IV.4.2, the optimal consolidation policies are examined under the assumption that at each decision epoch, the depot should choose between dispatching the entire consolidated load immediately, or continuing to consolidate until at least next period. The structural (S_L, S_U) policies we identify is defined by two parameters S_L and S_U . With some special values of S_L and S_U , this policy is reduced to the imme-

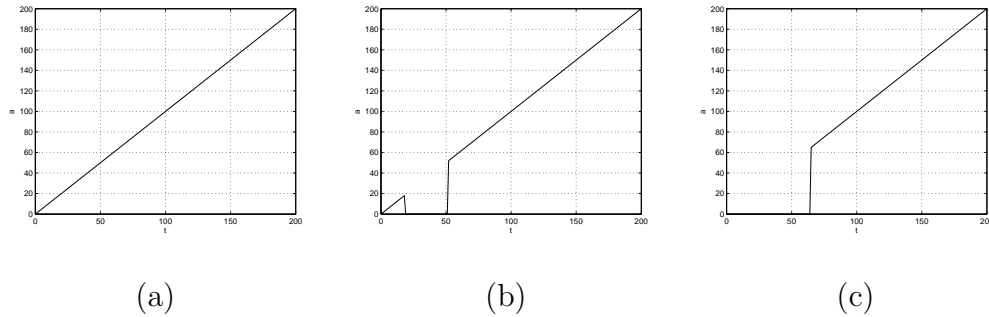
mediate shipment policy or a threshold policy. These policies are reasonable and easy to implement, however, the “clearing property” is restrictive and the limitation on the dispatch quantity can be suboptimal. Allowing partial load dispatch in this sublevel (corresponding to the formulation (4.12)), we show with numerical examples that the exact optimal policy for common carriage can be very complex.

Example 1 We demonstrate in this example that the policies discussed in Sublevel IV.4.2 can also be exactly optimal for formulation (4.12)). Considering a 4-period stochastic dynamic system, we compute the optimal policy through complete enumerations for state space on the interval $[0, 200]$ for three sets of parameters:

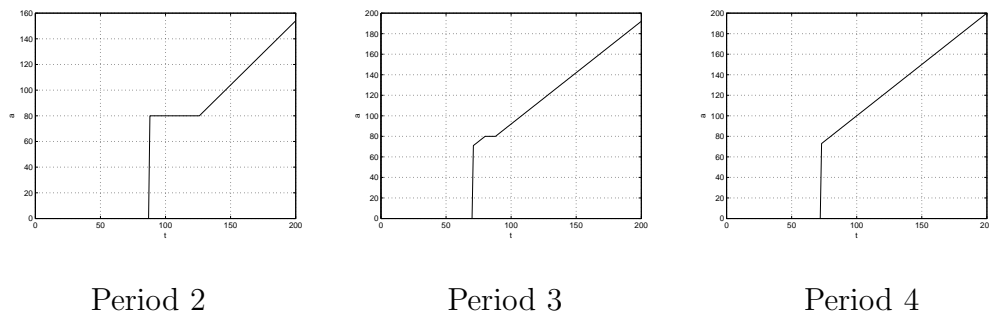
- (a). $C_N = 8, C_V = 4, MWT = 80, WBT = 40, w = 3, \text{Demand} \sim \text{Uniform}[0, 25]$;
- (b). $C_N = 8, C_V = 4, MWT = 80, WBT = 40, w = 3, \text{Demand} \sim \text{Uniform}[0, 50]$;
- (c). $C_N = 16, C_V = 8, MWT = 80, WBT = 40, w = 3, \text{Demand} \sim \text{Uniform}[0, 50]$.

For setting (a), an immediate shipment policy is optimal for each period. Setting (b) only changes setting (a)’s order distribution from $\text{Uniform}[0, 25]$ to $\text{Uniform}[0, 50]$. However, the optimal policy becomes a (S_L, S_U) Policy for each period. Modified from setting (b) by doubling the freight transportation cost C_N and C_V , setting (c) exhibits a threshold policy at optimality. The optimal dispatch quantity in period 4 for each setting is illustrated in Figure 15, where the horizontal coordinate represents the consolidated load t and the vertical coordinate represents the dispatch quantity. In these settings, the exact optimal policies do possess the “clearing property” although it is not assumed.

Example 2 For this example, we assume $C_N = 16, C_V = 8, MWT = 80, WBT = 40, w = 1, \text{order weight} \sim \text{Uniform}[0, 50]$. The only difference between setting (c) of Example 1 and this setting is the per-unit, per-period waiting cost w . Also

Figure 15: Example 1 of Optimal Policies for Common Carriage Transportation

testing this system for $N = 4$ periods, we find that the optimal policies for different periods actually have different structures. Figure 16 depicts the dispatch quantities of period 2, 3 and 4 respectively (the optimal policies for period 1 and 2 have the same structure).

Figure 16: Example 2 of Optimal Policies for Common Carriage Transportation

In Example 2, when the consolidated load is of smaller size, in each period, it is optimal to be idle. However, when the consolidated load is relatively large, the dispatch policy differs period by period. And the “clearing property” is obviously not satisfied in period 2 and 3, because there exist ranges of consolidated load on which the optimal dispatch quantity is constant. That means part of the load is left at the

depot and to be delivered together with some later orders in a later period.

From Example 1 and 2, we observe that the optimal policy structure can be very different from one instance to another. After testing many additional parameter settings, we fail to provide an explicit characterization of the optimal policy. The existence of multiple possible optimal policy structures is due to the complexity of the common carriage transportation cost.

IV.5. Multi-Truck Transportation with Cargo Capacity

In Sublevel IV.3, we consider the stochastic dynamic distribution system where the transportation is performed by a single capacitated truck. However, in other cases, the collection depot may possess a fleet of trucks that are available to deliver the shipments. Hence, in this sublevel, we consider the situation of multiple trucks. Formulating the transportation cost in the structure of multiple setup costs like we do in Chapter III, we examine the optimal consolidation policy of interest.

IV.5.1. Problem Formulation

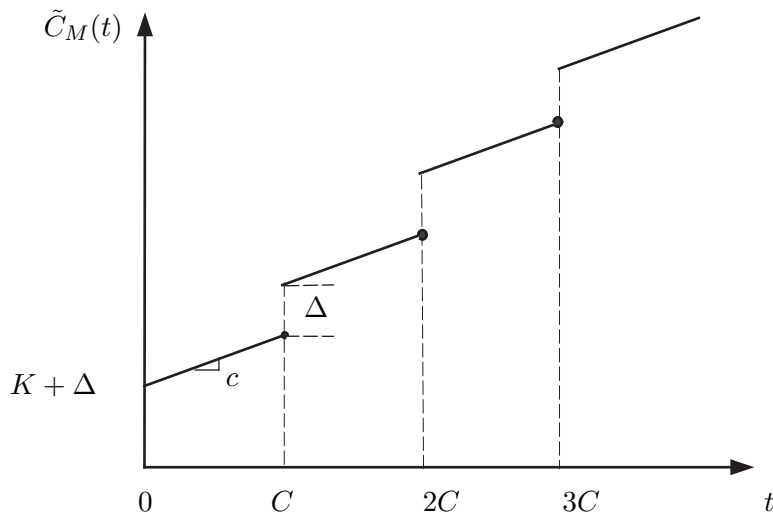
Note that the dispatch quantity of the model in Sublevel IV.3 is restricted below the cargo capacity C . In this sublevel, whenever the dispatch quantity exceeds the cargo capacity, another identical truck is available for shipping the extra load. Therefore, the transportation cost only depends on the number of the trucks used. The transportation cost for dispatching a shipment of total weight t can be presented as

$$\tilde{C}_M(t) = K_D \cdot I_{[t>0]} + ct + \Delta \left\lceil \frac{t}{C} \right\rceil, \quad (4.23)$$

where K_D is the fixed cost for a vehicle dispatch from the depot to the retailers, c is the transportation cost per unit weight, and Δ is the cost for using one truck.

We assume the trucks are identical with truck capacity C . Figure 17 illustrates the structure of this transportation cost.

Figure 17: Multi-Truck Transportation Cost with Cargo Capacity



In addition, since all demands should be satisfied by the end of the planning horizon, and the dispatch decision is made after the arrival of each demand, the unit transportation cost c will not affect the decision. Thus, we assume $c = 0$ to simplify the analysis, and hence, $\tilde{C}_M(t)$ can be represented as

$$\tilde{C}_M(t) = K_D \cdot I_{[t>0]} + \Delta \left\lceil \frac{t}{C} \right\rceil. \quad (4.24)$$

Letting t be the total weight of consolidated orders just before the dispatch decision is made, we can write the optimality equation as:

$$V_n(t) = \begin{cases} \min_{0 \leq a \leq t} \left\{ \tilde{C}_M(a) + w(t - a) + E(V_{n-1}(t - a + Z)) \right\}, & n = 1, \dots, N, \\ \tilde{C}_M(t), & n = 0. \end{cases} \quad (4.25)$$

This stochastic dynamic problem is challenging due to the fact that the term $\lceil \frac{t}{C} \rceil$ in the transportation cost introduces a piecewise component with discontinuities at the integer multiples of C . Hence, we also consider the case where the depot either dispatches all the consolidated demands or does not make a shipment. Then the optimality equation under the assumption can be written as follows.

$$V_n(t) = \begin{cases} \min \left\{ wt + E(V_{n-1}(t + Z)), \tilde{C}_M(t) + E(V_{n-1}(Z)) \right\}, & n = 1, \dots, N, \\ \tilde{C}_M(t), & n = 0. \end{cases} \quad (4.26)$$

Denote $u_n^1(t)$ and $u_n^2(t)$ as the cost-to-goes for dispatching and consolidating demands in period n , respectively, i.e.,

$$u_n^1(t) = wt + E(V_{n-1}(t + Z)), \text{ and } u_n^2(t) = \tilde{C}_M(t) + E(V_{n-1}(Z)).$$

We can rewrite the optimality equation of period $n = 1, \dots, N$ as

$$V_n(t) = \begin{cases} u_n^1(t), & \text{if } t = 0, \\ \min \{u_n^1(t), u_n^2(t)\}, & \text{if } t > 0. \end{cases} \quad (4.27)$$

IV.5.2. Analysis of the Optimal Policy

Since $E(V_{n-1}(Z))$ is constant, it is obvious that $u_n^2(t)$ is a step function and is left continuous at any break point mC where m is a nonnegative integer. Important properties of the value function $V_n(t)$ are provided below. Based on these properties, we develop the analysis on the structure of the optimal consolidation policy.

Property 5 *For any period n , $V_n(t)$ is non-decreasing in t and $V_n(t+C) - V_n(t) \geq \Delta$ for any t .*

Proof of Property 5: We prove this by induction.

- When $n = 0$, $V_0(t) = \tilde{C}_M(t) = K_D \cdot I_{[t>0]} + \lceil \frac{t}{C} \rceil \cdot \Delta$. Obviously, $V_0(t)$ is non-decreasing in t . In addition, when $t = 0$, $V_0(t + C) - V_0(t) = V_0(C) - V_0(0) = K_D + \Delta$. When $t > 0$, $V_0(t + C) - V_0(t) = \Delta$. Therefore, $V_0(t + C) - V_0(t) \geq \Delta$ for any t .
- Suppose $V_{k-1}(t)$ possesses the properties described in Property 5, we need to show for $V_k(t)$. First, recall that $u_k^1(t) = wt + E(V_{k-1}(t + Z))$ and $u_k^2(t) = K_D \cdot I_{[t>0]} + \lceil \frac{t}{C} \rceil \cdot \Delta + E(V_{k-1}(Z))$. Since $u_k^1(t)$ and $u_k^2(t)$ are both non-decreasing in t , $V_k(t) = \min(u_k^1(t), u_k^2(t))$ is also non-decreasing in t . Second, since we have $V_{k-1}(t + C) - V_{k-1}(t) \geq \Delta$ for any t ,

$$u_k^1(t + C) - u_k^1(t) = wC + E(V_{k-1}(t + C + Z) - V_{k-1}(t + Z)) \geq wC + \Delta.$$

and

$$u_k^2(t + C) - u_k^2(t) = \begin{cases} K_D + \Delta, & \text{if } t = 0, \\ \Delta, & \text{if } t > 0, \end{cases}$$

Then,

$$\begin{aligned} V_k(t + C) - V_k(t) &= \min(u_k^1(t + C), u_k^2(t + C)) - \min(u_k^1(t), u_k^2(t)) \\ &\geq \min(u_k^1(t) + \Delta, u_k^2(t) + \Delta) - \min(u_k^1(t), u_k^2(t)) \\ &= \min(u_k^1(t), u_k^2(t)) + \Delta - \min(u_k^1(t), u_k^2(t)) = \Delta. \end{aligned}$$

Thus, for any period n , $V_n(t)$ is non-decreasing in t and $V_n(t + C) - V_n(t) \geq \Delta$ for any t . The proof is complete. ■

Property 5 means that at the end of period n , when the total weight of the consolidated load is observed, the extra cost for having another full-truck load (C units) of outstanding demands exceeds the cargo cost Δ . Due to Property 5 and the

definitions of $u_n^1(t)$ and $u_n^2(t)$, for $t > 0$,

$$\begin{cases} u_n^1(t+C) - u_n^1(t) > \Delta, \\ u_n^2(t+C) - u_n^2(t) = \Delta. \end{cases} \quad (4.28)$$

Expression (4.28) means that if the depot chooses to consolidate the orders, then the extra cost-to-go for consolidating exactly one more truck load is greater than the cargo cost. If the depot has to dispatch a shipment with a positive quantity, then the extra cost-to-go for dispatching exactly one more truck load is equal to the cargo cost. Furthermore, we have the following result.

Lemma 7 *Suppose m is a positive integer,*

(1). *If $u_n^1((m-1)C) < u_n^2(mC) \leq u_n^1(mC)$, then*

$$u_n^2((m+1)C) \leq u_n^1(mC) < u_n^1((m+1)C)$$

or

$$u_n^1(mC) < u_n^2((m+1)C) \leq u_n^1((m+1)C);$$

(2). *If $u_n^1((m-1)C) \geq u_n^2(mC)$, then*

$$u_n^1(mC) \geq u_n^2((m+1)C).$$

Proof of Lemma 7:

(1). By (4.28), when m is a positive integer, $u_n^2((m+1)C) = u_n^2(mC) + \Delta$. Since $u_n^1((m-1)C) < u_n^2(mC) \leq u_n^1(mC)$, equivalently,

$$u_n^1((m-1)C) + \Delta < u_n^2((m+1)C) \leq u_n^1(mC) + \Delta.$$

Case 1: If $u_n^2((m+1)C) \leq u_n^1(mC)$, obviously,

$$u_n^2((m+1)C) \leq u_n^1(mC) < u_n^1((m+1)C).$$

Case 2: If $u_n^2((m+1)C) > u_n^1(mC)$, since $u_n^1(mC) \geq u_n^2(mC)$ and $u_n^2((m+1)C) = u_n^2(mC) + \Delta$, it follows that $u_n^2((m+1)C) \leq u_n^1(mC) + \Delta$. However, by (4.28), $u_n^1((m+1)C) > u_n^1(mC) + \Delta$, thus,

$$u_n^1(mC) < u_n^2((m+1)C) \leq u_n^1((m+1)C).$$

(2). By (4.28), $u_n^1(mC) > u_n^1((m-1)C) + \Delta \geq u_n^2(mC) + \Delta = u_n^2((m+1)C)$. ■

Lemma 7 gives the relationship between costs of alternatives with full truck loads of consolidated demands. Specifically, in period n , if dispatching m fully loaded trucks realizes a lower cost-to-go than holding them, then the cost-to-go for dispatching $m+1$ fully loaded trucks will not exceed the cost-to-go for holding the same amount. In addition, if dispatching m fully loaded trucks costs more than holding $(m-1)C$ units of demands (holding one less truck load), then the comparison between the cost-to-goes of dispatching $(m+1)C$ units and holding mC units is inconclusive. On the other hand, if dispatching m fully loaded trucks costs less than holding one less truck load, then dispatching $m+1$ fully loaded trucks also costs less than holding mC units.

For a positive consolidated load, we can divide the range into pieces with equal length of C , i.e., $(0, \infty) = \cup_{i=1}^{\infty} ((i-1)C, iC]$. Compare the cost-to-goes of two options: $u_n^1(t)$ (consolidating the orders) and $u_n^2(t)$ (dispatching the consolidated load) on interval $((m-1)C, mC]$, where m is a natural number. It is worth noting that we only need to compare three values $u_n^1(mC)$, $u_n^2((m-1)C)$ and $u_n^2(mC)$. That is because $u_n^1(t)$ is constant and $u_n^2(t)$ is strictly increasing on this interval. There are three cases:

Case 1: $u_n^1((m-1)C) < u_n^2(mC) \leq u_n^1(mC)$.

Since $u_n^2(t) = u_n^2(mC)$ on this interval, and $u_n^1(t)$ is continuous and strictly increasing, there exists a critical point S such that on $((m-1)C, S]$, $u_n^2(t) \geq u_n^1(t)$ and on $(S, mC]$, $u_n^2(t) \leq u_n^1(t)$. Hence, it is optimal to consolidate the order on $((m-1)C, S]$ and to dispatch the load on $(S, mC]$.

Case 2: $u_n^2(mC) \leq u_n^1((m-1)C) < u_n^1(mC)$.

This implies $u_n^2(t) \leq u_n^1(t)$ for any $t \in ((m-1)C, mC]$; hence, it is optimal to dispatch the load.

Case 3: $u_n^1((m-1)C) < u_n^1(mC) \leq u_n^2(mC)$.

This implies $u_n^1(t) \leq u_n^2(t)$ for any $t \in ((m-1)C, mC]$; hence, it is optimal to consolidate the demands.

Consequently, statement (1) of Lemma 7 says that if on interval $((m-1)C, mC]$, $u_n^1(t)$ and $u_n^2(t)$ satisfy the condition of Case 1, then on $(mC, (m+1)C]$, their relationship can be either Case 1 or Case 2. Meanwhile, statement (2) of Lemma 7 says that if on $((m-1)C, mC]$, $u_n^1(t)$ and $u_n^2(t)$ satisfy the condition of Case 2, then on $(mC, (m+1)C]$, their relationship can only be Case 2. In this way, the property of the optimal dispatch policy is characterized in Proposition 7.

Proposition 7 *For any period n , there exist parameters TL_n and TU_n , such that $TL_n \leq TU_n$ and*

$$\begin{cases} u_n^1(t) \leq u_n^2(t), & \text{if } 0 < t \leq TL_n, \\ u_n^1(t) \geq u_n^2(t), & \text{if } t \geq TU_n. \end{cases} \quad (4.29)$$

Proof of Proposition 7: For $n = 0$, simple let $TL_0 = TU_0 = 0$. For $n = 1, \dots, N$, we prove Proposition 7 by induction.

- When $n = 1$,

$$\begin{cases} u_1^1(t) = wt + E(K_D \cdot I_{[t+Z>0]} + \lceil \frac{t+Z}{C} \rceil \cdot \Delta) . \\ u_1^2(t) = K_D \cdot I_{[t>0]} + \lceil \frac{t}{C} \rceil \cdot \Delta + E(K_D \cdot I_{[Z>0]} + \lceil \frac{Z}{C} \rceil \cdot \Delta) , \end{cases}$$

Since $u_1^1(0) = u_1^2(0)$, we are only interested in the optimal decision for the consolidated load $t \in (0, \infty)$. Notice that when $t > 0$, the indicator function $I_{[t>0]}$ and $I_{[t+Z>0]}$ are both equal to 1. Therefore, when $t > 0$,

$$\begin{cases} u_1^1(t) = wt + K_D + E(\lceil \frac{t+Z}{C} \rceil) \cdot \Delta \\ u_1^2(t) = K_D + \lceil \frac{t}{C} \rceil \cdot \Delta + K_D \cdot P(Z > 0) + E(\lceil \frac{Z}{C} \rceil) \cdot \Delta \end{cases}$$

It is obvious that $\lceil \frac{t+Z}{C} \rceil \leq \lceil \frac{t}{C} \rceil + \lceil \frac{Z}{C} \rceil$, and hence, $E(\lceil \frac{t+Z}{C} \rceil) \leq \lceil \frac{t}{C} \rceil + E(\lceil \frac{Z}{C} \rceil)$.

Define

$$TL_1 := \frac{K_D \cdot P(Z > 0)}{w}, \quad (4.30)$$

then, if $0 < t \leq TL_1$, $u_1^1(t) \leq u_1^2(t)$.

On another hand, when the consolidated load t is not an integer multiple of the cargo capacity C , we can write $t = mC + \varepsilon$, where m is a non-negative integer and $0 < \varepsilon < C$. Accordingly,

$$\lceil \frac{t}{C} \rceil + \lceil \frac{Z}{C} \rceil = (m+1) + \lceil \frac{Z}{C} \rceil = \lceil \frac{mC+Z}{C} \rceil + 1 \leq \lceil \frac{t+Z}{C} \rceil + 1.$$

It directly follows that, $\lceil \frac{t}{C} \rceil + E(\lceil \frac{Z}{C} \rceil) \leq E(\lceil \frac{t+Z}{C} \rceil) + 1$. Therefore, if we define

$$TU_1 := \frac{K_D \cdot P(Z > 0) + \Delta}{w}, \quad (4.31)$$

when $t \geq TU_1$, $u_1^1(t) \geq u_1^2(t)$.

- Suppose equation (4.29) holds for period $k - 1$, we need to prove for period k .

By the assumption on period $k - 1$, when $t \geq TU_{k-1}$, $u_{k-1}^1(t) \geq u_{k-1}^2(t)$. Then we have

$$V_{k-1}(t) = u_{k-1}^2(t) = K_D + \left\lceil \frac{t}{C} \right\rceil \cdot \Delta + E(V_{k-2}(Z)).$$

Then let us consider period k for $t > TU_{k-1}$,

$$u_k^2(t) = K_D + \left\lceil \frac{t}{C} \right\rceil \cdot \Delta + E(V_{k-1}(Z)), \text{ and}$$

$$\begin{aligned} u_k^1(t) &= wt + E(V_{k-1}(t + Z)) = wt + \int_0^\infty V_{k-1}(t + z) f(z) dz \\ &= wt + \int_0^\infty \left[K_D + \left\lceil \frac{t+z}{C} \right\rceil \Delta + E(V_{k-2}(Z)) \right] f(z) dz \\ &= wt + K_D + E(V_{k-2}(Z)) + \int_0^\infty \left\lceil \frac{t+z}{C} \right\rceil \Delta f(z) dz \\ &= wt + K_D + E(V_{k-2}(Z)) + \int_0^{\lceil \frac{t}{C} \rceil C - t} \left\lceil \frac{t}{C} \right\rceil \Delta f(z) dz \\ &\quad + \sum_{i=1}^\infty \int_{\lceil \frac{t}{C} \rceil C - t + (i-1)C}^{\lceil \frac{t}{C} \rceil C - t + iC} \left(\left\lceil \frac{t}{C} \right\rceil + i \right) \Delta f(z) dz \\ &= wt + K_D + E(V_{k-2}(Z)) + \int_0^\infty \left\lceil \frac{t}{C} \right\rceil \Delta f(z) dz + \sum_{i=0}^\infty \int_{\lceil \frac{t}{C} \rceil C - t + iC}^\infty \Delta f(z) dz \\ &= wt + K_D + E(V_{k-2}(Z)) + \left\lceil \frac{t}{C} \right\rceil \Delta + \Delta \cdot \sum_{i=0}^\infty \int_{\lceil \frac{t}{C} \rceil C - t + iC}^\infty f(z) dz. \end{aligned}$$

To examine the value of $u_k^1(t)$, it is worth noting that

$$\int_{(i+1)C}^\infty f(z) dz \leq \int_{\lceil \frac{t}{C} \rceil C - t + iC}^\infty f(z) dz \leq \int_{iC}^\infty f(z) dz.$$

In addition,

$$\sum_{i=0}^\infty \int_{iC}^\infty f(z) dz = \sum_{i=1}^\infty \int_{(i-1)C}^{iC} i \cdot f(z) dz \leq \int_0^\infty \frac{z}{C} f(z) dz + 1 = \frac{\mu Z}{C} + 1,$$

and

$$\sum_{i=0}^\infty \int_{(i+1)C}^\infty f(z) dz = \sum_{i=1}^\infty \int_{(i-1)C}^{iC} (i-1) \cdot f(z) dz \geq \int_0^\infty \frac{z}{C} f(z) dz - 1 = \frac{\mu Z}{C} - 1$$

together imply

$$\frac{\mu Z}{C} - 1 \leq \sum_{i=0}^{\infty} \int_{\lceil \frac{t}{C} \rceil C - t + iC}^{\infty} f(z) dz \leq \frac{\mu Z}{C} + 1. \quad (4.32)$$

Examine the difference between the costs of two choices, we have

$$u_k^1(t) - u_k^2(t) = wt + E(V_{k-2}(Z)) - E(V_{k-1}(Z)) + \Delta \cdot \sum_{i=0}^{\infty} \int_{\lceil \frac{t}{C} \rceil C - t + iC}^{\infty} f(z) dz.$$

Since the term $[E(V_{k-1}(Z_0)) - E(V_k(Z))]$ is constant, $\sum_{i=0}^{\infty} \int_{\lceil \frac{t}{C} \rceil C - t + iC}^{\infty} f(z) dz$ is bounded according to (4.32), and wt is strictly increasing, $u_k^2(t) - u_k^1(t)$ is positive when t is sufficiently large. In other words, TU_k exists. On the other hand, Since $u_k^1(0) = u_k^2(0)$ and

$$\lim_{t \rightarrow 0^+} u_k^2(t) = K_D + \Delta + E(V_{k-1}(Z)) > E(V_{k-1}(Z)) = u_k^1(0^+),$$

TL_k exists such that for $0 < t \leq TL_k$, $u_k^1(t) \leq u_k^2(t)$. The proof is complete. ■

Proposition 7 specifies the optimal dispatch decision for the consolidated load that is less than TL_n or greater than TU_n . To characterize the optimal decision for the load located within TL_n and TU_n , let $S_C^{n,1}$ be the greatest value such that whenever $t \leq S_C^{n,1}$, $u_n^1(t) \leq u_n^2(t)$. Then we can locate $S_C^{n,1}$ in a unique interval $((m-1)C, mC]$ for some integer m . By Lemma 7, there exist a sequence of numbers $S_C^{n,1}, S_C^{n,2}, \dots, S_C^{n,M}$, such that

$$S_C^{n,1} \leq \left\lceil \frac{S_C^{n,1}}{C} \right\rceil C < S_C^{n,2} \leq \left\lceil \frac{S_C^{n,2}}{C} \right\rceil C < \dots < S_C^{n,M},$$

and the optimal dispatch decision of period n can be expressed as

$$a_n^*(t) = \begin{cases} 0 \text{ (consolidate the demands),} & t \leq S_C^{n,1} \\ t \text{ (dispatch the demands),} & S_C^{n,i} < t \leq \left\lceil \frac{S_C^{n,i}}{C} \right\rceil C, i = 1, \dots, M-1 \\ 0 \text{ (consolidate the demands),} & \left\lceil \frac{S_C^{n,i}}{C} \right\rceil C < t \leq S_C^{n,i+1}, i = 1, \dots, M-1 \\ t \text{ (dispatch the demands),} & t \geq S_C^{n,M}. \end{cases} \quad (4.33)$$

M is finite because of the existence of TU_n . By the definition of $S_C^{n,1}$, $S_C^{n,1} \geq TL_n$. Similarly, $S_C^{n,M} \leq TU_n$. From expression (4.33), the optimal decision shifts between “consolidate the demands” and “dispatch the demands” in between TL_n and TU_n .

IV.5.3. Single Period Problem

A single period problem is examined in order to find the strict upper bound and lower bound for parameters M , $S_C^{1,1}$ and $S_C^{1,M}$. In the proof of Proposition 7, we have shown that when $t \leq \frac{K_D \cdot P(Z>0)}{w}$, $u_1^1(t) \leq u_1^2(t)$, and when $t \geq \frac{K_D \cdot P(Z>0) + \Delta}{w}$, $u_1^1(t) \geq u_1^2(t)$. Thus we need to examine the values of $u_1^1(t)$ and $u_1^2(t)$ for $t \in \left(\frac{K_D \cdot P(Z>0)}{w}, \frac{K_D \cdot P(Z>0) + \Delta}{w} \right)$.

Recall that at $t = mC$ where m is a positive integer, $E\left(\left\lceil \frac{t+Z}{C} \right\rceil\right) = \left\lceil \frac{t}{C} \right\rceil + E\left(\left\lceil \frac{Z}{C} \right\rceil\right)$. Since $t > \frac{K_D \cdot P(Z>0)}{w}$,

$$K_D + wt + E\left(\left\lceil \frac{t+Z}{C} \right\rceil\right) \cdot \Delta > K_D + K_D \cdot P(Z > 0) + \left\lceil \frac{t}{C} \right\rceil \Delta + E\left(\left\lceil \frac{Z}{C} \right\rceil\right) \Delta.$$

Equivalently, $u_1^2(t) < u_1^1(t)$. Hence, it is optimal to dispatch the consolidated load.

In fact, the weight of a full-truck consolidated load (in the form of mC) within the range $\left(\frac{K_D \cdot P(Z>0)}{w}, \frac{K_D \cdot P(Z>0) + \Delta}{w} \right)$ always belongs to the set below:

$$\mathcal{L} = \left\{ \left(\left\lceil \frac{K_D \cdot P(Z > 0)}{wC} \right\rceil + 1 \right) C, \dots, \left(\left\lceil \frac{K_D \cdot P(Z > 0) + \Delta}{wC} \right\rceil - 1 \right) C \right\}.$$

If

$$\left\lceil \frac{K_D \cdot P(Z > 0) + \Delta}{wC} \right\rceil - 1 < \left\lfloor \frac{K_D \cdot P(Z > 0)}{wC} \right\rfloor + 1, \quad (4.34)$$

there is no full-truck consolidated load in this range. Therefore, the optimal policy is actually a threshold policy, and $M = 1$.

If (4.34) does not hold, the range $\left(\frac{K_D \cdot P(Z > 0)}{w}, \frac{K_D \cdot P(Z > 0) + \Delta}{w} \right)$ can be divided into several pieces by the points of \mathcal{L} . (If there are I points in \mathcal{L} , then this range can be divided into $I + 1$ pieces). On each piece, the transportation cost are the same, i.e., $\left\lceil \frac{t}{C} \right\rceil$ is constant. As a result, $u_1^2(t)$ is constant. Since $u_1^1(t)$ is continuous and strictly increasing, $u_1^1(t) = u_1^2(t)$ has at most 1 solution on each piece.

Consider one piece. Denote t_l and t_r as the end points where $t_l < t_r$. From the above analysis, whether the end point t_r is in the form of mC or $\frac{K_D \cdot P(Z > 0) + \Delta}{w}$, we always have $u_1^1(t_r) \geq u_1^2(t_r)$. And at the end point t_l , then

- if $t_l = \frac{K_D \cdot P(Z > 0)}{w}$,

$$u_1^1(t_l) = K_D \cdot P(Z > 0) + K_D + E \left(\left\lceil \frac{t_l + Z}{C} \right\rceil \right) \Delta.$$

and

$$u_1^2(t_l) = K_D + \left\lceil \frac{t_l}{C} \right\rceil \Delta + K_D \cdot P(Z > 0) + E \left(\left\lceil \frac{Z}{C} \right\rceil \right) \Delta,$$

Thus, $u_1^1(t_l) \leq u_1^2(t_l)$. It follows that there exists a critical value that divides this piece into two parts. On the first part, $u_1^1(t) \leq u_1^2(t)$. On the second part, $u_1^1(t) > u_1^2(t)$.

- if $t_l = mC$ for some integer m then $u_1^2(t_l^+) = u_1^2(t_l) + \Delta = u_1^2(t_l + C)$. Explicitly,

$$\begin{aligned} u_1^2(t_l) &= K_D + \left\lceil \frac{t_l}{C} \right\rceil \Delta + K_D \cdot P(Z > 0) + E \left(\left\lceil \frac{Z}{C} \right\rceil \right) \Delta \\ &= K_D + m\Delta + K_D \cdot P(Z > 0) + E \left(\left\lceil \frac{Z}{C} \right\rceil \right) \Delta, \end{aligned}$$

and

$$u_1^2(t_l^+) = K_D + m\Delta + K_D \cdot P(Z > 0) + \Delta + E\left(\left\lceil \frac{Z}{C} \right\rceil\right) \Delta.$$

On the other hand, $u_1^1(t)$ is continuous and

$$u_1^1(t_l^+) = u_1^1(t_l) = wt_l + K_D + E\left(\left\lceil \frac{t_l + Z}{C} \right\rceil\right) \Delta = wt_l + K_D + m\Delta + E\left(\left\lceil \frac{Z}{C} \right\rceil\right) \Delta.$$

Since $t_l \in \left(\frac{K_D \cdot P(Z > 0)}{w}, \frac{K_D \cdot P(Z > 0) + \Delta}{w}\right)$,

$$K_D \cdot P(Z > 0) < wt_l < K_D \cdot P(Z > 0) + \Delta.$$

Thus,

$$u_1^2(t_l) < u_1^1(t_l) = u_1^1(t_l^+) < u_1^2(t_l^+).$$

This implies that on a piece $(t_l, t_r]$, $u_1^2(t)$ is constant, $u_1^1(t)$ is increasing, and $u_1^1(t) = u_1^2(t)$ has exactly one solution. Therefore

$$M = \left\lceil \frac{K_D \cdot P(Z > 0) + \Delta}{wC} \right\rceil - \left\lfloor \frac{K_D \cdot P(Z > 0)}{wC} \right\rfloor,$$

$$\frac{K_D \cdot P(Z > 0)}{w} < S_C^{1,1} \leq \left(\left\lfloor \frac{K_D \cdot P(Z > 0)}{wC} \right\rfloor + 1 \right) C,$$

and

$$\left(\left\lceil \frac{K_D \cdot P(Z > 0) + \Delta}{wC} \right\rceil - 1 \right) C < S_C^{1,M} \leq \frac{K_D \cdot P(Z > 0) + \Delta}{w}.$$

Especially when the demand in each period is uniformly distributed over $[0, b]$, it can be seen that

$$u_n^1(t) = wt + \frac{1}{b} \int_0^b V_{n-1}(t+x) dz.$$

Taking the first derivative of $u_n^2(t)$, we have

$$\frac{du_n^1(t)}{dt} = w + \frac{1}{b} \cdot \frac{d}{dt} \int_0^b V_{n-1}(t+z) dz = w + \frac{1}{b} \cdot (V_{n-1}(t+b) - V_{n-1}(t)).$$

It can be demonstrated that,

$$w + \frac{\Delta}{b} \cdot \left\lfloor \frac{b}{C} \right\rfloor \leq w + \frac{1}{b} \cdot \left(V_{n-1}(t + \left\lfloor \frac{b}{C} \right\rfloor C) - V_{n-1}(t) \right) \leq \frac{du_n^1(t)}{dt} \leq w + \frac{\Delta}{b} \cdot \left\lceil \frac{b}{C} \right\rceil.$$

Then the bounds of M_n can be determined as follows: If there exist some integers m_1 and m_2 such that

$$\frac{\Delta}{C} \left(1 + \frac{1}{m_1}\right) \leq w + \frac{\Delta}{b} \left\lfloor \frac{b}{C} \right\rfloor < \frac{\Delta}{C} \left(1 + \frac{1}{m_1 - 1}\right) \quad (4.35)$$

and

$$\frac{\Delta}{C} \left(1 + \frac{1}{m_2}\right) \leq w + \frac{\Delta}{b} \left\lceil \frac{b}{C} \right\rceil < \frac{\Delta}{C} \left(1 + \frac{1}{m_2 - 1}\right), \quad (4.36)$$

then

$$M_n^{max} = m_1 + 1, \text{ and } M_n^{min} = m_2.$$

Reorganize (4.35), we have

$$\frac{1}{m_1} \leq \frac{wC}{\Delta} + \frac{C}{b} \left\lfloor \frac{b}{C} \right\rfloor - 1 < \frac{1}{m_1 - 1},$$

and hence

$$\left(\frac{wC}{\Delta} + \frac{C}{b} \left\lfloor \frac{b}{C} \right\rfloor - 1 \right)^{-1} \leq m_1 < \left(\frac{wC}{\Delta} + \frac{C}{b} \left\lfloor \frac{b}{C} \right\rfloor - 1 \right)^{-1} + 1.$$

Therefore,

$$M_n^{max} = \left\lceil \left(\frac{wC}{\Delta} + \frac{C}{b} \left\lfloor \frac{b}{C} \right\rfloor - 1 \right)^{-1} \right\rceil + 1.$$

Similarly,

$$M_n^{min} = \left\lfloor \left(\frac{wC}{\Delta} + \frac{C}{b} \left\lceil \frac{b}{C} \right\rceil - 1 \right)^{-1} \right\rfloor.$$

IV.6. Summary

The focus of this chapter is on the theoretical analysis of the exact optimal policies for outbound shipment consolidation. By using a stochastic dynamic programming

approach, we study a periodic-review consolidation problem where a single collection depot serves multiple retailers located in a given market area. Four different options including the private carriage without cargo capacity and cost, single-truck transportation with cargo capacity and fixed cost, common carriage, and multi-truck fleet with cargo capacity are considered for the transportation.

After examining the properties of the value functions of the stochastic dynamic models, we demonstrate that the structure of the optimal policies for the private fleet transportation without cargo capacity, is in the form of a state-dependent threshold policy. Specifically, when the total weight of consolidated orders is observed, there exists a critical amount such that if the observed total weight is higher than this amount, it is optimal to dispatch all the waiting demands; otherwise, it is optimal to continue to consolidate the orders. For the capacitated private fleet transportation model, the optimal policy is also a threshold policy. For the common carriage model, the exact optimal policy is difficult to identify due to the special structure of transportation cost. With different parameter settings, the optimal policy can be an immediate dispatch policy (no matter how many orders are consolidated, it is always optimal to satisfy them immediately in each period), or a threshold policy (there exists a threshold value so that whenever the consolidated weight exceeds the threshold, it is optimal to release all waiting orders; otherwise, continue to consolidate), or a more complicated and named as (S_L, S_U) policy (when the consolidated weight is less than S_L or greater than S_U , dispatch the load; otherwise, consolidate). We provide sufficient conditions under which these policies are optimal. For the multi-truck model, the exact optimal policy is also complicated. And some preliminary results are provided.

CHAPTER V

THE VENDOR'S OPTIMAL STOCK REPLENISHMENT AND SHIPMENT SCHEDULING POLICY UNDER TEMPORAL SHIPMENT CONSOLIDATION

In Chapter III and Chapter IV, we examine the optimal inventory replenishment decision and the optimal consolidation dispatch schedule separately. Numerous theoretical results and practical experiences demonstrate that the efficient management of a supply chain system requires coordination between inventory control and transportation scheduling. Such integration is particularly implementable in a Vendor-Managed inventory (VMI) system.

VMI practices have been increasingly popular over the past decade following their widespread implementation by major manufacturers, such as Proctor and Gamble, and mass-retailers, such as Wal-Mart. In this chapter, we consider a joint stock replenishment and shipment scheduling problem applicable in the context of a VMI contract under stochastic demand. Çetinkaya and Lee (2000) are the first to introduce the problem of interest in the current research while focusing on a practical—but clearly suboptimal—policy. Our objective on the other hand is to identify the structural properties of the vendor's optimal joint policies and, to the best of our knowledge, the current research is the first attempt to this end.

More specifically, we examine a single-product, stochastic demand, periodic-review, two-echelon inventory model for a vendor who makes inbound stock replenishment and outbound shipment scheduling decisions, simultaneously, in the VMI setting of interest. This setting is further characterized by the vendor's flexibility to consolidate smaller orders over time from a group of retailers located in a given geographical region to realize transportation scale economies.

The retailers are willing to wait to receive their orders at an additional expense for the vendor to include retailer waiting (order delay) and inventory holding costs due to shipment consolidation. It is worth noting that such a shipment consolidation practice is known as temporal consolidation which makes practical sense if it can offer acceptable customer service without excessive order delay and economical sense when immediate order deliveries are expensive.

We assume that the vendor operates a private truck with ample capacity to transport the merchandise from its supplier (manufacturer) to her warehouse as well as from the warehouse to the retailers' locations. As a result, both the inbound replenishment cost and the outbound transportation cost are composed of a fixed and a linear term. This type of cost function has been used very often in the literature because of its simple structure representing the economies of scale in production, procurement, and transportation. Considering the vendor's inbound replenishment, inventory holding, outbound transportation, and customer waiting costs, we propose a stochastic dynamic programming approach for the purpose of computing the optimal joint policy specifying the optimal (i) inbound replenishment and (ii) outbound dispatch quantities.

Our first main result identifies the structure of the optimal policy for a single-period problem for an arbitrary period (i.e., not necessarily the last period), provided that the cost-to-go function of the next period belongs to a specific family of functions characterized explicitly. The optimal policy is basically a zoned, state-dependent threshold policy. We characterize the optimal policy based on the difference between the two state variables, i.e., the consolidated load waiting to be released minus the on-hand inventory. This quantity is instrumental for our analysis and represents the consolidated load excess of on-hand inventory when it is positive, and it stands for the inventory excess of consolidated load when it is negative. We call this quantity the

“excess position” in the remaining part of this chapter. “Load excess” and “inventory excess” are also respectively used to denote the absolute values of the positive or negative excess position when necessary. Depending on the excess position, the two-dimensional state space can be divided into three zones, and hence, the optimal policy characterizing the vendor’s optimal inbound replenishment and outbound dispatch quantities can be described as follows: On each zone, the optimal inbound stock replenishment and outbound shipment scheduling decisions can be specified by a threshold policy. As a result, in each period the optimal decisions are based on the following four options: (1) do not replenish and do not dispatch; (2) do not replenish and dispatch the entire consolidated load; (3) do not replenish and dispatch the entire on-hand inventory; (4) replenish an amount so that after dispatching the entire consolidated load, the remaining inventory level is equal to a critical target value.

It is worth noting that when option (3) is the optimal decision for a given state, the on-hand inventory must be less than the consolidated load. After the decision is executed, the on-hand inventory level drops to zero, and the consolidated load remains positive. This phenomenon indicates that this joint policy does not have the clearing property. We show that when the replenishment quantity is given and fixed, the optimal dispatch decision is either to dispatch as much as possible or not to dispatch. Of course, if the vendor replenishes sufficient inventory to satisfy the entire consolidated load when an outbound shipment needs to be made, then the clearing property will be satisfied. However, due to the economies of scale in the inbound replenishment represented by the fixed replenishment cost, it is possible that the most economical replenishment quantity necessitating the clearance of the consolidated load is small. Then, replenishing a small quantity to clear the load may cost more than paying some waiting cost for delaying the fulfillment of part of the consolidated load without stock replenishment.

Our second main result provides a formal proof that the cost-to-go function of each period always belongs to the specific family of functions characterized, regardless of the parameter settings. It follows that in any period during a finite horizon problem, the optimal policy is in the form of a zoned, state-dependent threshold policy described above. For any specific period, the vendor should replenish her inventory if and only if the vendor's load excess exceeds a certain level and the on-hand inventory level is above a threshold value. The corresponding optimal replenishment quantity is equal to the summation of the load excess and a critical value. In all other situations, the lowest cost is achieved when there is no inventory replenishment at the vendor. For the cases when the size of consolidated load is small, or both the on-hand inventory level and the load excess are low, it is preferable to choose to be idle in that period, i.e., do not replenish and do not dispatch. When the load excess is lower than a quantity or there is inventory excess, if the on-hand inventory level is higher than a state-dependent threshold value, then it is optimal to dispatch as much as possible consolidated load.

We find that the optimal policy has the following interesting characteristics:

- From our common intuition, inbound replenishment is not required when the on-hand inventory level is high, and an outbound shipment needs to be made when the consolidated load is large. However, the optimal joint policy structure we characterized relies more on the difference between the two states (the excess position) instead of the values of the states themselves. For example, when both the on-hand inventory level and the size of the consolidated load are large, people would think to make an outbound dispatch without replenishing the inventory. But in fact, as far as the load excess (a positive excess position) exceeds a critical level, the vendor still needs to replenish her stock. On the

contrary, when both amounts are small, as far as their difference is also small, it is still possible to make an outbound dispatch without replenishment.

- In the literature, when the inbound replenishment cost includes a fixed term, usually the replenishment policy consists of one parameter denoting an order-up-to level. In our problem, when the outbound dispatch scheduling is integrated with the inbound replenishment, the optimal replenishment quantity actually varies case by case, and a common order-up-to level does not exist. However, whenever the inventory is replenished, after the consolidated load is dispatched, the remaining on-hand inventory is always equal to a critical value.
- When the vendor has a large amount of on-hand inventory that is sufficient to satisfy the consolidated load, it is not always optimal to dispatch the load. That is to say, in some situations, paying for some extra holding cost and order waiting cost is still more preferable to paying a transportation cost for dispatching the load. Due to the fixed transportation cost, when the size of the consolidated load is small, it is justified to wait until an economical dispatch quantity is accumulated in a later period.
- When the consolidated load exceeds the on-hand inventory level, and the vendor chooses to make an outbound dispatch, it is not necessary to satisfy all the waiting orders. Therefore, “clearing policy” may not be optimal for the joint inventory replenishment and outbound dispatch scheduling problem. The consolidated load is only cleared under two situations: on-hand inventory is more than the consolidated load, and the load excess is vast.

The remainder of this chapter is organized as follows: In Sublevel V.1, we discuss the details of the problem setting and present the problem formulation; in Sublevel

V.2, we examine the properties of the terminal cost, solve a “single-period” problem and propose a new joint inventory replenishment and shipment scheduling policy; in Sublevel V.3, we analyze the optimality of the new policy for multiple-period models; and in Sublevel V.4, we provide summary and recommend the focus for future research.

V.1. Notation and Problem Formulation

In our model, a single vendor receives orders from a group of retailers and replenishes her own inventory from an ample supplier over a discrete and finite time-horizon. In contrast to the traditional inventory systems where orders are satisfied as they arrive at the vendor, we assume that the vendor controls the retailer’s resupply under a VMI contract, thus, the vendor has the autonomy to consolidate small orders until a dispatch quantity that economizes shipping costs is accumulated. In this setting, the vendor has the authority to coordinate the inbound inventory replenishment and the outbound shipment scheduling to achieve the maximum cost savings. The cost involved in this model consists of inbound replenishment cost, holding cost for excess inventory items, outbound transportation cost, and waiting cost for delayed orders.

We consider the system over N periods. Adopting the standard dynamic programming approach, we index the periods in a backward order so that they occur over time in the order $N, N - 1, \dots, 0$, and period 0 is the end of the planning horizon. An illustration is given in Figure 18, and a summary of basic notation is provided below. However, notation is also introduced throughout this chapter when needed.

System Parameters:

N	length of the planning horizon
n	period index

Z_n	random quantity of retailer orders in period n (we use Z to denote a generic element with density and distribution functions $f(\cdot)$ and $F(\cdot)$)
A_R	fixed cost of inbound replenishment
A_D	fixed cost of outbound transportation
c_R	unit inbound replenishment cost
c_D	unit outbound transportation cost
h	inventory holding cost per unit per period
w	customer waiting cost per unit per period
$W(a)$	inbound replenishment cost for replenishing a units $W(a) = A_R \cdot I_{[a>0]} + c_R \cdot a$
$\tilde{C}(l)$	outbound transportation cost for dispatching l units $\tilde{C}(l) = A_D \cdot I_{[l>0]} + c_D \cdot l$

States:

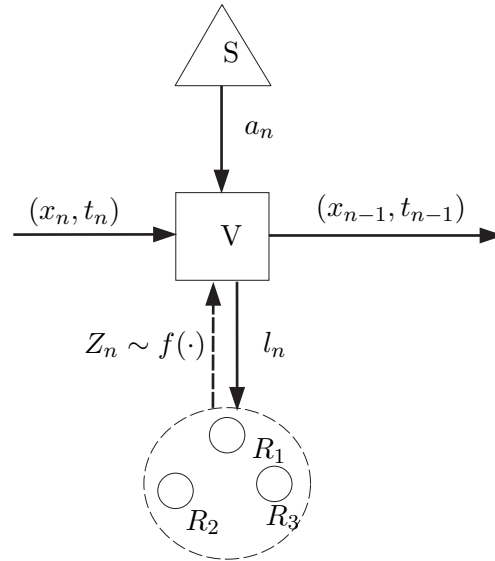
x_n	on-hand inventory level in period n , before the joint decisions are made
t_n	consolidated load waiting to be released in period n , before the joint decisions are made

Decisions:

$a_n^*(x, t)$	optimal inbound replenishment quantity in period n with x on-hand inventory and t consolidated load
$l_n^*(x, t)$	optimal outbound dispatch quantity in period n with x on-hand inventory and t consolidated load

Optimality Equation:

$V_n(x, t)$	optimal expected total cost from period n to the end with x on-hand inventory and t consolidated load (called the cost-to-go function of period n)
-------------	---

Figure 18: Problem Setting of the VMI System

In Figure 18, S , V and R_i represent the supplier, the vendor and the retailers, respectively. During period n , a combined retailer order with quantity Z_n is received at the vendor. Z_n 's are independent and identically distributed random variables described by the density function $f(\cdot)$. Whenever a retailer order is received, the vendor reviews her on-hand inventory level x_n and the consolidated load t_n . Based on x_n and t_n , the vendor makes decisions regarding the inbound replenishment quantity a_n and the outbound dispatch quantity l_n . At the end of period n , the remaining inventory x_{n-1} is carried to the next period (period $n-1$), and all unsatisfied retailer orders t_{n-1} in period n are consolidated for at least one more period. We assume that the vendor's replenishment is received instantaneously and can be dispatched to the retailer in the same time period. All retailer demands should be satisfied by the end of the planning horizon.

We denote $W(a)$ as the inbound replenishment cost for replenishing a ($a \geq 0$) units and denote $\tilde{C}(l)$ as the outbound transportation cost for dispatching l ($l \geq 0$) units. Since a private fleet is assumed to be used in both the inbound and outbound logistics, both the inbound replenishment cost and the outbound transportation cost are linear with a fixed cost, expressed as $W(a) = A_R \cdot I_{[a>0]} + c_R a$ and $\tilde{C}(l) = A_D \cdot I_{[l>0]} + c_D l$. Here, A_R and A_D are the fixed costs associated with the inbound replenishment and outbound shipment, respectively. In practice, A_R may include the cost for reviewing the inventory levels and the administrative cost for replenishing inventory. A_D may include the fixed cost of processing a dispatching command, the maintenance and usage cost of the truck, and even the salary paid to the truck driver. c_R and c_D are unit replenishment or transportation costs that are volume-related. The remaining inventory incurs inventory holding cost at h per unit per period, and the remaining consolidated load incurs waiting cost at w per unit per period. All transportation and inventory cost parameters are stationary.

Because we assume the joint decisions are made after the realization of the retailer demands and all demands should be satisfied by the end of the planning horizon, the unit transportation cost c_D actually does not affect the joint decisions. To simplify the analysis, we assume $c_D = 0$, therefore, $\tilde{C}(l)$ can be modified as $\tilde{C}(l) = A_D \cdot I_{[l>0]}$.

We can formulate this problem as a stochastic dynamic program. Since the dispatch quantity cannot exceed the available on-hand inventory level and also would not be more than the consolidated load, it follows that $0 \leq l_n \leq \min(x_n + a_n, t_n)$. Let $V_n(x, t)$ be the infimum of the expected total cost over periods $n, n-1, \dots, 0$ starting with on-hand inventory of x units and consolidated load of t units. Then $V_n(x, t)$ is expressed as follows:

$$V_n(x, t) = \begin{cases} \inf_{\substack{a \geq 0 \\ 0 \leq l \leq \min(x+a, t)}} \left\{ \begin{aligned} &W(a) + \tilde{C}(l) + h(x+a-l) + w(t-l) \\ &+ E(V_{n-1}(x+a-l, t-l+Z)) \end{aligned} \right\}, & n = 1, \dots, N, \\ \inf_{\substack{a \geq (t-x)^+ \\ l=t}} \left\{ W(a) + \tilde{C}(l) + h(x+a-l) \right\}, & n = 0. \end{cases} \quad (5.1)$$

Letting

$$v_n(a, l|x, t) = W(a) + \tilde{C}(l) + h(x+a-l) + w(t-l) + E(V_{n-1}(x+a-l, t-l+Z)) \quad (5.2)$$

denote the cost for replenishing a units of inventory and dispatching l units of consolidated orders in period n when the on-hand inventory level is x and the size of the consolidated load is t , we have

$$V_n(x, t) = \inf_{\substack{a \geq 0 \\ 0 \leq l \leq \min(x+a, t)}} v_n(a, l|x, t) \quad \text{for } n = 1, \dots, N.$$

To make the optimal inventory replenishment and dispatch scheduling decisions simultaneously, one method is to consider the cases of “do not dispatch” and “dispatch” separately, optimize each case individually and choose the one that gives a lower expected cost. For this purpose, two functions $u_n^1(x, t)$ and $u_n^2(x, t)$ are defined such that $u_n^1(x, t)$ represents the optimal cost-to-go function if no outbound shipment is going to be released in period n , and $u_n^2(x, t)$ represents the optimal cost-to-go function if a positive size of consolidated load will be dispatched. Specifically, $u_n^1(x, t)$ and $u_n^2(x, t)$ are represented as

$$u_n^1(x, t) = \inf_{\substack{a \geq 0 \\ l=0}} v_n(a, l|x, t), \quad x \geq 0, t \geq 0$$

$$u_n^2(x, t) = \inf_{\substack{a \geq 0 \\ 0 < l \leq \min(x+a, t)}} v_n(a, l|x, t), \quad x \geq 0, t > 0.$$

Here, $u_n^1(x, t)$ is defined on a two-dimensional space with $x \in [0, \infty)$ and $t \in$

$[0, \infty)$. For $u_n^2(x, t)$, when no retailer order is consolidated at the vendor, the outbound dispatch quantity can only take the value of zero, i.e., if $t = 0$, then $l = 0$. Thus, $u_n^2(x, t)$ is defined on a two-dimensional space with $x \in [0, \infty)$ and $t \in (0, \infty)$. Consequently, the optimality equation can be rewritten as

$$V_n(x, t) = \begin{cases} u_n^1(x, t), & \text{if } t = 0, \\ \min \{u_n^1(x, t), u_n^2(x, t)\}, & \text{if } t > 0. \end{cases} \quad (5.3)$$

By this formulation, the optimization problem in each period is decomposed into two subproblems associated with optimizing $u_n^1(x, t)$ and $u_n^2(x, t)$, respectively. In the remaining part of this chapter, these two subproblems are solved individually, and the optimal joint decisions are obtained.

V.2. A “Single-Period” Problem

According to (5.1), the optimal joint decisions in period n actually depend on the cost-to-go function of period $n - 1$, i.e., $V_{n-1}(x, t)$. However, due to the existence of the fixed costs in the inbound replenishment and outbound transportation, the cost-to-go function $V_{n-1}(x, t)$ is neither jointly convex in x and t nor monotonically increasing or decreasing in any dimension; hence, specific structures of the cost-to-go function need to be characterized.

In Sublevel V.2.1, we examine the properties of the terminal cost, i.e., $V_0(x, t)$. In Sublevel V.2.2, these properties are assumed for the cost-to-go function of an arbitrary period $n - 1$, and the optimal joint decisions for period n are identified.

V.2.1. Properties of the Terminal Cost

The terminal cost is the cost that occurs at the end of the planning horizon for handling the ending inventory and ending consolidated load. In this model, the terminal cost is $V_0(x, t)$ for x units of ending inventory and t units of consolidated load. Since all demand needs to be satisfied, the dispatch quantity $l = t$. Then we have

$$V_0(x, t) = \inf_{a \geq (t-x)^+} \left\{ W(a) + \tilde{C}(t) + h(x + a - t) \right\}.$$

With the term in the braces increasing in a , the optimal replenishment quantity is $a_0^*(x, t) = (t - x)^+$, and $V_0(x, t)$ can be explicitly written as follows:

$$\begin{aligned} V_0(x, t) &= W((t - x)^+) + \tilde{C}(t) + h(x - t + (t - x)^+) \\ &= \begin{cases} A_D \cdot I_{[t > 0]} + h(x - t), & \text{if } t \leq x, \\ A_R + c_R(t - x) + A_D, & \text{if } t > x. \end{cases} \end{aligned} \quad (5.4)$$

This equation shows that if the ending inventory is more than the consolidated load, the terminal cost is equal to the transportation cost for dispatching the entire consolidated load plus the holding cost for the excess inventory. Otherwise, the vendor needs to replenish her stock so that every retailer order can be fulfilled. Thus, the terminal cost is equal to the transportation cost plus the cost for replenishing extra units. The terminal cost function expressed in (5.4) possesses the following properties.

Property 6 *For any fixed $t \geq 0$, $V_0(x, t) \rightarrow \infty$ as $x \rightarrow \infty$. $E(V_0(x, Z)) + (c_R + h)x$ is continuous in x and has a minimizer S on $x \in [0, \infty)$. $E(V_0(x, t + Z))$ is also continuous on the two-dimensional state space of (x, t) for $x, t \in [0, \infty)$.*

Proof of Property 6: First, let $\gamma_{0,1}(x, t) = A_D \cdot I_{[t > 0]} + h(x - t)$ and $\gamma_{0,2}(x, t) =$

$A_R + c_R(t - x) + A_D$; then

$$V_0(x, t) = \begin{cases} \gamma_{0,1}(x, t), & \text{if } t \leq x, \\ \gamma_{0,2}(x, t), & \text{if } t > x. \end{cases}$$

For any fixed $t \geq 0$,

$$\lim_{x \rightarrow \infty} V_0(x, t) = \lim_{x \rightarrow \infty} \gamma_{0,1}(x, t) = \lim_{x \rightarrow \infty} (A_D \cdot I_{[t > 0]} + h(x - t)) = \infty.$$

Notice that for a fixed t , both $\gamma_{0,1}(x, t)$ and $\gamma_{0,2}(x, t)$ are continuous in x . Then,

$$E(V_0(x, Z)) = \int_0^\infty V_0(x, z)f(z)dz = \int_0^x \gamma_{0,1}(x, z)f(z)dz + \int_x^\infty \gamma_{0,2}(x, z)f(z)dz$$

is also continuous in x . Consequently, $E(V_0(x, Z)) + (c_R + h)x$ is continuous in x .

When $x \rightarrow \infty$ for a fixed t , since $V_0(x, t)$ goes to ∞ , $E(V_0(x, Z)) + (c_R + h)x$ also goes to ∞ . Combining this observation with the condition that $E(V_0(x, Z)) + (c_R + h)x$ is continuous on $x \in [0, \infty)$, there exists an $S \in [0, \infty)$ such that $E(V_0(S, Z)) + (c_R + h)S \leq E(V_0(x, Z)) + (c_R + h)x$, for any $x \geq 0$. In other words, S is a minimizer of $E(V_0(x, Z)) + (c_R + h)x$ on $[0, \infty)$. The continuity of $E(V_0(x, t + Z))$ in x and t can be proved similarly by writing the expectation as $E(V_0(x, t + Z)) = \int_0^{(x-t)^+} \gamma_{0,1}(x, t + z)f(z)dz + \int_{(x-t)^+}^\infty \gamma_{0,2}(x, t + z)f(z)dz$. ■

Property 6 says that the terminal cost approaches infinity when the on-hand inventory level goes to infinity. In addition, if the vendor clears the consolidated load in period 1 and needs to carry some on-hand inventory into the last period (period 0), the sum of the procurement cost for replenishing that inventory amount, the holding cost for carrying it for one period, and the expected terminal cost is minimized when the vendor carries S (i.e., the critical amount of inventory from period 1 to period 0).

Property 7 For any fixed $x \geq 0$, $t \geq 0$ and $\delta > 0$, we have $V_0(x, t) - V_0(x + \delta, t) \leq$

$$A_R + (c_R + h)\delta.$$

Proof of Property 7: For fixed $x \geq 0$, $t \geq 0$ and $\delta > 0$, we examine the value of $V_0(x + \delta, t) - V_0(x, t)$ in three cases.

Case 1: If $x < x + \delta < t$,

$$\begin{aligned} V_0(x + \delta, t) - V_0(x, t) &= A_R + c_R(t - x - \delta) + A_D - A_R - c_R(t - x) - A_D \\ &= -c_R\delta \geq -A_R - (c_R + h)\delta. \end{aligned}$$

Case 2: If $x < t \leq x + \delta$,

$$\begin{aligned} V_0(x + \delta, t) - V_0(x, t) &= A_D + h(x + \delta - t) - A_R - c_R(t - x) - A_D \\ &= -A_R - (c_R + h)(t - x) + h \cdot \delta \geq -A_R - (c_R + h)\delta + h \cdot \delta \geq -A_R - (c_R + h)\delta. \end{aligned}$$

Case 3: If $t \leq x < x + \delta$,

$$\begin{aligned} V_0(x + \delta, t) - V_0(x, t) &= A_D \cdot I_{[t>0]} + h(x + \delta - t) - A_D \cdot I_{[t>0]} - h(x - t) \\ &= h \cdot \delta \geq -A_R - (c_R + h)\delta. \end{aligned}$$

In all cases, $V_0(x + \delta, t) - V_0(x, t) \geq -A_R - (c_R + h)\delta$. Thus, Property 7 holds. \blacksquare

Property 7 says that the marginal cost for having δ units less on-hand inventory at the beginning of period 0 is at most equal to the cost for replenishing δ units in the previous period and carrying them to the current period.

Property 8 For any fixed Δ , $V_0(x, x + \Delta)$ is non-decreasing in x .

Proof of Property 8: For any fixed Δ , the terminal cost $V_0(x, x + \Delta)$ is given by:

$$V_0(x, x + \Delta) = \begin{cases} A_D \cdot I_{[x+\Delta>0]} - h \cdot \Delta, & \text{if } \Delta \leq 0, \\ A_R + c_R \cdot \Delta + A_D, & \text{if } \Delta > 0. \end{cases}$$

From the above equation, it is straightforward to observe that $V_0(x, x + \Delta)$ is always non-decreasing in x . ■

Property 8 says that if the excess position is fixed, the terminal cost is non-decreasing in the on-hand inventory level.

Property 9 For $t > 0$, $\frac{dV_0(0,t)}{dt} \geq c_R$.

Proof of Property 9: For any $t > 0$, $V_0(0, t) = A_R + c_R t + A_D$. Therefore, $\frac{dV_0(0,t)}{dt} = c_R \geq c_R$. ■

Property 9 says that if there is no on-hand inventory at the beginning of the last period, the marginal terminal cost with respect to the consolidated load exceeds the unit procurement cost.

V.2.2. Optimal Joint Decision for Period n

Replacing $V_0(x, t)$ with a general two-dimensional function $V(x, t)$, we define a family of functions such that each function in this family possesses Properties 6 through 9 described above.

Definition 4 For given parameters A_R , c_R and h , define a family \mathcal{V} of two dimensional functions, such that when a function $V(x, t) \in \mathcal{V}$, the following conditions are satisfied.

(A1) For any fixed $t \geq 0$, $V(x, t) \rightarrow \infty$ as $x \rightarrow \infty$. $E(V(x, Z)) + (c_R + h)x$ is continuous in x and has a minimizer S on $x \in [0, \infty)$. $E(V(x, t + Z))$ is continuous on the state space of (x, t) for $x, t \in [0, \infty)$.

(A2) For any fixed $x \geq 0$, $t \geq 0$ and $\delta > 0$, we have $V(x + \delta, t) - V(x, t) \geq -A_R - (c_R + h)\delta$.

(A3) For any fixed Δ , $V(x, x + \Delta)$ is non-decreasing in x .

(A4) For $t > 0$, $\frac{dV(0,t)}{dt} \geq c_R$.

For the purpose of this chapter, $V(x, t)$ represents the cost-to-go function given by equation 5.1. With this representation, Conditions (A1)–(A4) have the similar interpretations as Properties 6 through 9, and Conditions (A1)–(A4) will be used to solve for the optimal joint decisions. For period n ($n > 0$), the cost-to-go function is obtained via solving the following optimality equation:

$$V_n(x, t) = \inf_{\substack{a \geq 0 \\ 0 \leq l \leq \min(x+a, t)}} \left\{ \begin{array}{l} W(a) + \tilde{C}(l) + h(x + a - l) + w(t - l) \\ + E(V_{n-1}(x + a - l, t - l + Z)) \end{array} \right\}.$$

Therefore, the optimal joint decision in period n depends on the cost-to-go function of period $n - 1$. According to expression (5.3), we examine the optimal joint decisions for “do not dispatch” and “dispatch” cases separately and select the more preferable one, i.e., the one with a lower total expected cost.

Lemma 8 (Optimal joint decision for the “do not dispatch” case) *For the case where no outbound shipment is going to be released in period n , if $V_{n-1}(x, t)$ belongs to the family \mathcal{V} , then the optimal replenishment quantity in period n is zero. That is, the optimal joint decision is given by*

$$a_n^* = 0, \quad l_n^* = 0, \tag{5.5}$$

and the corresponding cost-to-go is

$$u_n^1(x, t) = v_n(0, 0|x, t) = hx + wt + E(V_{n-1}(x, t + Z)). \tag{5.6}$$

Proof of Lemma 8: In period n , when there are x ($x \geq 0$) units of on-hand inventory and t units of consolidated load waiting to be released, the minimum achievable cost

for consolidating all orders to the next period is

$$u_n^1(x, t) = \inf_{\substack{a \geq 0 \\ l=0}} v_n(a, l|x, t) = \inf_{a \geq 0} \{W(a) + h(x + a) + wt + E(V_{n-1}(x + a, t + Z))\}.$$

Since $V_{n-1}(x, t) \in \mathcal{V}$, according to Condition (A2) and the linearity of the expectation operator, for $\forall a > 0, x \geq 0$ and $t \geq 0$, we have $E(V_{n-1}(x + a, t + Z)) - E(V_{n-1}(x, t + Z)) \geq -A_R - (c_R + h)a$. Thus, the cost for ordering a positive quantity a and dispatching nothing is

$$\begin{aligned} v_n(a, 0|x, t) &= A_R + c_R a + h(x + a) + wt + E(V_{n-1}(x + a, t + Z)) \\ &\geq A_R + c_R a + h(x + a) + wt + E(V_{n-1}(x, t + Z)) - A_R - (c_R + h)a \\ &= hx + wt + E(V_{n-1}(x, t + Z)) = v_n(0, 0|x, t). \end{aligned}$$

This implies that when no consolidated order is going to be dispatched, it is never optimal to replenish the vendor's inventory. Therefore, (5.5) and (5.6) hold and the proof is complete. \blacksquare

When no shipment is going to be released in period n , since replenishing the stock without dispatching a shipment will incur extra inventory holding cost compared to replenishing that amount in a later period, it is optimal to be idle. Note that when there is no consolidated order ($t = 0$), the vendor surely chooses not to dispatch. By Lemma 8, the optimal replenishment quantity is also zero, therefore, the optimal joint decision is “do not replenish and do not dispatch”, and the cost-to-go function is given by $V_n(x, 0) = v_n(0, 0|x, 0)$. For $t > 0$, the optimal joint decision requires the comparison of $u_n^1(x, t)$ and $u_n^2(x, t)$. Hence, our next step is to examine $u_n^2(x, t)$ to find the optimal inbound replenishment and outbound dispatch quantities if the vendor must dispatch a positive size of consolidated load. It is worth mentioning that $u_n^2(x, t)$ is defined on $x \in [0, \infty)$ and $t \in (0, \infty)$.

Lemma 9 (Optimal joint decision for the “dispatch” case) For the case where a positive size of consolidated load must be dispatched in period n , i.e., $t > 0$, if $V_{n-1}(x, t)$ belongs to \mathcal{V} , then there exist two nonnegative parameters S_n and Δ_n^* that define the joint policy. S_n is the minimizer of $E(V_{n-1}(x, Z)) + (c_R + h)x$ on $x \in [0, \infty)$, and Δ_n^* is the unique solution to the following equation:

$$A_R + c_R \Delta + (c_R + h)S_n + A_D + E(V_{n-1}(S_n, Z)) = A_D + w\Delta + E(V_{n-1}(0, \Delta + Z)), \quad (5.7)$$

for $\Delta \in [0, \infty)$. Then, the optimal joint decision for period n is

$$\begin{cases} a_n^* = 0, l_n^* = t, & \text{if } x > 0, t - x \leq 0, \\ a_n^* = 0, l_n^* = x, & \text{if } x > 0, 0 < t - x \leq \Delta_n^*, \\ a_n^* = t - x + S_n, l_n^* = t, & \text{if } x = 0 \text{ or } t - x > \Delta_n^*, \end{cases} \quad (5.8)$$

and the corresponding cost-to-go is

$$u_n^2(x, t) = \begin{cases} v_n(0, t|x, t) & \text{if } x > 0, t - x \leq 0, \\ = A_D + h(x - t) + E(V_{n-1}(x - t, Z)), \\ v_n(0, x|x, t) & \text{if } x > 0, 0 < t - x \leq \Delta_n^*, \\ = A_D + w(t - x) + E(V_{n-1}(0, t - x + Z)), \\ v_n(t - x + S_n, t|x, t) & \text{if } x = 0 \text{ or } t - x > \Delta_n^*. \\ = A_R + A_D + c_R(t - x) + (c_R + h)S_n + E(V_{n-1}(S_n, Z)), \end{cases} \quad (5.9)$$

Proof of Lemma 9: By the definition of $u_n^2(x, t)$,

$$u_n^2(x, t) = \inf_{\substack{a \geq 0 \\ 0 < l \leq \min(x+a, t)}} v_n(a, l|x, t)$$

$$= \inf_{\substack{a \geq 0 \\ 0 < l \leq \min(x+a, t)}} \left\{ \begin{array}{l} W(a) + \tilde{C}(l) + h(x+a-l) + w(t-l) \\ + E(V_{n-1}(x+a-l, t-l+Z)) \end{array} \right\}.$$

Given states (x, t) , when z is the realized demand and the replenishment quantity a is fixed, $V_{n-1}(x+a-l, t-l+z)$ is non-increasing in l according to Condition (A3). Then, its expectation, $E(V_{n-1}(x+a-l, t-l+Z))$, over Z is also non-increasing in l . Furthermore, if the dispatch quantity l is positive, then $v_n(a, l|x, t)$ given by equation (5.2) is strictly decreasing in $l \in (0, \min(x+a, t)]$ provided that a is fixed. In other words, if the replenishment quantity a is chosen, the optimal dispatch quantity should be as much as possible, i.e., $l = \min(x+a, t)$ and $u_n^2(x, t) = \inf_{l=\min(x+a, t)}^{a \geq 0} v_n(a, l|x, t)$. In order to analyze the cost function $u_n^2(x, t)$, we divide the two-dimensional state space into two subspaces based on the excess position, i.e., $t-x$. The optimal joint decision is determined for each subspace.

Subspace 1: If $t-x \leq 0$, the on-hand inventory is at least as much as the consolidated load. Then, the optimal dispatch quantity l is equal to $\min(x+a, t) = t > 0$ regardless of the value of a . Thus,

$$\begin{aligned} & u_n^2(x, t) \\ &= \inf_{\substack{a \geq 0 \\ l=t}} \left\{ W(a) + \tilde{C}(l) + h(x+a-l) + w(t-l) + E(V_{n-1}(x+a-l, t-l+Z)) \right\} \\ &= \inf_{a \geq 0} \left\{ A_R \cdot I_{[a>0]} + c_R a + A_D + h(x+a-t) + E(V_{n-1}(x+a-t, Z)) \right\}. \end{aligned}$$

For any positive replenishment quantity a , Condition (A2) implies that

$$\begin{aligned} v_n(a, t|x, t) &= A_R + c_R a + A_D + h(x+a-t) + E(V_{n-1}(x+a-t, Z)) \\ &\geq A_D + h(x-t) + E(V_{n-1}(x-t, Z)) = v_n(0, t|x, t). \end{aligned}$$

Consequently, if a shipment must be released when the on-hand inventory is sufficient to satisfy the consolidated load, the least cost-to-go is attained when the vendor dispatches the entire consolidated load without replenishing her inventory. In summary, when $t - x \leq 0$ (since $x \geq t > 0$), the optimal joint decision for the “dispatch” case is given by $a_n^* = 0$ and $l_n^* = t$ with

$$u_n^2(x, t) = v_n(0, t|x, t) = A_D + h(x - t) + E(V_{n-1}(x - t, Z)). \quad (5.10)$$

Subspace 2: If $t - x > 0$, the on-hand inventory is not sufficient to clear the consolidated load. We consider the following three cases for the replenishment quantity a : 1) $a \geq t - x$ (replenish more than the load excess), 2) $0 < a < t - x$ (replenish less than the load excess), and 3) $a = 0$ (do not replenish). The cost for each case is examined and the one with the lowest cost is chosen to be optimal. Mathematically,

$$u_n^2(x, t) = \min \{u_n^2(x, t|a \geq t - x), u_n^2(x, t|0 < a < t - x), u_n^2(x, t|a = 0)\}, \quad (5.11)$$

where $u_n^2(x, t|\text{Region of } a) = \inf_{l=\min(x+a,t)}^{\text{Region of } a} v_n(a, l|x, t)$.

Case 1: When $a \geq t - x$, the available inventory after the vendor’s replenishment is more than the consolidated load. As a result, the dispatch quantity $l^* = \min(x + a, t) = t$, and

$$\begin{aligned} u_n^2(x, t|a \geq t - x) \\ = \inf_{a \geq t-x > 0} \{A_R + c_R a + A_D + h(x + a - t) + E(V_{n-1}(x + a - t, Z))\}. \end{aligned}$$

Since $V_{n-1}(x, t) \in \mathcal{V}$, according to Condition (A1), there exists $S_n \geq 0$ that minimizes $E(V_{n-1}(x, Z)) + (c_R + h)x$. Letting $x + a^* - t = S_n$, we have the optimal replenishment quantity $a^* = t - x + S_n$ and

$$u_n^2(x, t|a \geq t - x)$$

$$= v_n(t - x + S_n, t|x, t) = A_R + c_R(t - x + S_n) + A_D + hS_n + E(V_{n-1}(S_n, Z)). \quad (5.12)$$

(5.12) implies that when the on-hand inventory is not sufficient to satisfy the consolidated load, and the replenishment quantity a is required to be greater than $t - x$, a should be equal to $t - x + S_n$. Following this result, the inventory level after the vendor's replenishment is $t + S_n$, and the ending inventory level of period n is S_n due to the reason that the optimal dispatch quantity is t .

Case 2: When $0 < a < t - x$, since $0 \leq x < x + a < t$, $l = \min(x + a, t) = x + a$ and

$$u_n^2(x, t|0 < a < t - x) = \inf_{0 < a < t - x} v_n(a, x + a|x, t),$$

where $v_n(a, x + a|x, t) = A_R + c_R a + A_D + w(t - x - a) + E(V_{n-1}(0, t - x - a + Z))$.

Taking the first partial derivative of $v_n(a, x + a|x, t)$ with respect to a , we have

$$\begin{aligned} \frac{\partial v_n(a, x + a|x, t)}{\partial a} &= \frac{\partial}{\partial a} [A_R + c_R a + A_D + w(t - x - a) + E(V_{n-1}(0, t - x - a + Z))] \\ &= c_R - w + \frac{\partial}{\partial a} E(V_{n-1}(0, t - x - a + Z)) \\ &\leq c_R - w - c_R = -w < 0. \end{aligned}$$

The “ \leq ” inequality holds due to Condition (A4). $\frac{\partial v_n(a, x + a|x, t)}{\partial a} < 0$ means that for a fixed x and t , the function $v_n(a, x + a|x, t)$ is decreasing in a on $(0, t - x)$. In other words, when the on-hand inventory level and the size of the consolidated load are observed, if the vendor chooses to replenish her inventory with a quantity that is less than the load excess $(t - x)$, then the larger is the replenishment quantity the lower is the expected total cost-to-go. Hence,

$$\begin{aligned} u_n^2(x, t|0 < a < t - x) &= \inf_{0 < a < t - x} v_n(a, x + a|x, t) \geq v_n(t - x, t|x, t) \\ &\geq v_n(t - x + S_n, t|x, t) = u_n^2(x, t|a \geq t - x). \end{aligned}$$

(since $t - x + S_n$ minimizes $v_n(a, t|x, t)$ for $a \geq t - x$ (Case 1))

Therefore, $u_n^2(x, t|0 < a < t - x)$ has a higher expected cost-to-go than $u_n^2(x, t|a \geq t - x)$ does. More specifically, if the consolidated load is more than the on-hand inventory, and the vendor decides to replenish her inventory, then it is never optimal to replenish less than the load excess.

Case 3: When $a = 0$, since we define $u_n^2(x, t)$ as the minimum cost for dispatching a positive size of consolidated load, the action $\{a = 0\}$ is infeasible when there is no on-hand inventory, i.e., $x = 0$. When $x > 0$, if the replenishment quantity $a^* = 0$, the dispatch quantity $l^* = \min(x + a, t) = x > 0$, i.e.,

$$u_n^2(x, t|a = 0) = \inf_{\substack{a=0 \\ l=x}} v_n(a, l|x, t) = v_n(0, x|x, t) = A_D + w(t - x) + E(V_{n-1}(0, t - x + Z)).$$

To sum up, for the “dispatch” case on Subspace 2, when $x = 0$, since $a = 0$ is infeasible and $u_n^2(x, t|0 < a < t - x)$ is dominated by $u_n^2(x, t|a \geq t - x)$ according to the analysis in Case 2, the optimal decision is to replenish $t + S_n$ units and dispatch t units. When $x > 0$, the optimal joint decisions are either to replenish $t - x + S_n$ units and dispatch t units, or to replenish nothing and dispatch x units. By (5.11), $u_n^2(x, t)$ on Subspace 2 is expressed as:

$$u_n^2(x, t) = \begin{cases} v_n(t - x + S_n, t|x, t), & \text{if } x = 0, \\ \min\{v_n(t - x + S_n, t|x, t), v_n(0, x|x, t)\}, & \text{if } x > 0, t - x > 0. \end{cases} \quad (5.13)$$

To determine the explicit expression of $u_n^2(x, t)$, we denote

$$\begin{aligned} g_1(x, t) &= v_n(t - x + S_n, t|x, t) \\ &= A_R + c_R(t - x + S_n) + A_D + hS_n + E(V_{n-1}(S_n, Z)), \\ g_2(x, t) &= v_n(0, x|x, t) = A_D + w(t - x) + E(V_{n-1}(0, t - x + Z)). \end{aligned}$$

If we define $\Delta := t - x$ where $\Delta > 0$ and replace t with $x + \Delta$ in both $g_1(x, t)$ and $g_2(x, t)$, we obtain

$$\begin{aligned} g_1(x, x + \Delta) &= A_R + c_R \Delta + (c_R + h)S_n + A_D + E(V_{n-1}(S_n, Z)), \text{ and} \\ g_2(x, x + \Delta) &= A_D + w\Delta + E(V_{n-1}(0, \Delta + Z)). \end{aligned}$$

Taking partial derivatives with respect to Δ of $g_1(x, x + \Delta)$ and $g_2(x, x + \Delta)$, we have

$$\frac{\partial g_1(x, x + \Delta)}{\partial \Delta} = c_R, \text{ and } \frac{\partial g_2(x, x + \Delta)}{\partial \Delta} = w + \frac{\partial}{\partial \Delta} E(V_{n-1}(0, \Delta + Z)) \geq w + c_R.$$

The last inequality holds due to Condition (A4). Since $\lim_{\Delta \rightarrow \infty} \frac{\partial [g_2(x, x + \Delta) - g_1(x, x + \Delta)]}{\partial \Delta} \geq w$, it follows that $\lim_{\Delta \rightarrow \infty} g_2(x, x + \Delta) > \lim_{\Delta \rightarrow \infty} g_1(x, x + \Delta)$. In addition, observe that

$$\begin{aligned} \lim_{\Delta \downarrow 0} g_1(x, x + \Delta) &= A_R + (c_R + h)S_n + A_D + E(V_{n-1}(S_n, Z)), \text{ and} \\ \lim_{\Delta \downarrow 0} g_2(x, x + \Delta) &= A_D + E(V_{n-1}(0, Z)). \end{aligned}$$

If $S_n = 0$, obviously, $\lim_{\Delta \downarrow 0} g_1(x, x + \Delta) > \lim_{\Delta \downarrow 0} g_2(x, x + \Delta)$. If $S_n > 0$, according to Condition (A2) and the linearity of the expectation operator, $E(V_{n-1}(S_n, Z)) - E(V_{n-1}(0, Z)) \geq -A_R - (c_R + h)S_n$, and hence, $\lim_{\Delta \downarrow 0} g_1(x, x + \Delta) \geq \lim_{\Delta \downarrow 0} g_2(x, x + \Delta)$. Comparing $g_1(x, x + \Delta)$ and $g_2(x, x + \Delta)$ for fixed x , we have two possible situations:

I. If $\lim_{\Delta \downarrow 0} g_1(x, x + \Delta) > \lim_{\Delta \downarrow 0} g_2(x, x + \Delta)$, since $\frac{\partial g_1(x, x + \Delta)}{\partial \Delta} < \frac{\partial g_2(x, x + \Delta)}{\partial \Delta}$ and $\lim_{\Delta \rightarrow \infty} g_1(x, x + \Delta) < \lim_{\Delta \rightarrow \infty} g_2(x, x + \Delta)$, there exists a unique real number $\Delta_n^* > 0$ such that when $0 < \Delta \leq \Delta_n^*$, $g_1(x, x + \Delta) \geq g_2(x, x + \Delta)$, and when $\Delta > \Delta_n^*$,

$g_1(x, x + \Delta) < g_2(x, x + \Delta)$. Therefore,

$$u_n^2(x, x + \Delta) = \min\{g_1(x, x + \Delta), g_2(x, x + \Delta)\} = \begin{cases} g_2(x, x + \Delta), & \text{if } 0 < \Delta \leq \Delta_n^*, \\ g_1(x, x + \Delta), & \text{if } \Delta > \Delta_n^*, \end{cases}$$

II. If $\lim_{\Delta \downarrow 0} g_1(x, x + \Delta) = \lim_{\Delta \downarrow 0} g_2(x, x + \Delta)$, since $\frac{\partial g_1(x, x + \Delta)}{\partial \Delta} < \frac{\partial g_2(x, x + \Delta)}{\partial \Delta}$ for any Δ , letting $\Delta_n^* = 0$, we have $g_1(x, x + \Delta) \leq g_2(x, x + \Delta)$ for any $\Delta > \Delta_n^*$. Equivalently,

$$u_n^2(x, x + \Delta) = \min\{g_1(x, x + \Delta), g_2(x, x + \Delta)\} = g_1(x, x + \Delta).$$

Thus, unique $\Delta_n^* \geq 0$ exists and satisfies $g_1(x, x + \Delta_n^*) = g_2(x, x + \Delta_n^*)$, i.e.,

$$A_R + c_R \Delta_n^* + (c_R + h)S_n + A_D + E(V_{n-1}(S_n, Z)) = A_D + w \Delta_n^* + E(V_{n-1}(0, \Delta_n^* + Z)).$$

Therefore, $u_n^2(x, t)$ can be explicitly presented as

$$u_n^2(x, t) = \begin{cases} v_n(t - x + S_n, t|x, t), & \text{if } x = 0, \\ v_n(0, x|x, t), & \text{if } x > 0, 0 < t - x \leq \Delta_n^*, \\ v_n(t - x + S_n, t|x, t), & \text{if } x > 0, t - x > \Delta_n^*. \end{cases} \quad (5.14)$$

Summarizing (5.10) of Subspace 1 and (5.14) of Subspace 2, we obtain $u_n^2(x, t)$

$$u_n^2(x, t) = \begin{cases} v_n(t - x + S_n, t|x, t), & \text{if } x = 0, \\ v_n(0, t|x, t), & \text{if } x > 0, t - x \leq 0, \\ v_n(0, x|x, t), & \text{if } x > 0, 0 < t - x \leq \Delta_n^*, \\ v_n(t - x + S_n, t|x, t), & \text{if } x > 0, t - x > \Delta_n^*. \end{cases} \quad (5.15)$$

Reorganizing (5.15), we obtain (5.9). Note that when $\Delta_n^* = 0$, $0 < t - x \leq \Delta_n^*$ does not result in a valid range for $t - x$, and hence, the third situation in (5.15) (equivalently, the second case in (5.9)) can be ignored. The corresponding optimal

joint decisions are presented in (5.8) accordingly. ■

Lemma 9 provides the expressions for the optimal joint decisions and the corresponding cost-to-go functions under the condition that a positive size of consolidated load must be dispatched in period n . Depending on the excess position, the joint decisions have three options: (1) do not replenish and dispatch the entire consolidated load when the on-hand inventory is sufficient; (2) do not replenish and dispatch all the on-hand inventory when the load excess is in low volume; (3) replenish a quantity that is equal to the sum of the load excess and a critical value when the load excess exceeds a certain level. When the parameter $\Delta_n^* = 0$, option (2) cannot be optimal, then the optimal dispatch quantity $l^* = t$ implies that the vendor should always clear the consolidated load if an outbound dispatch needs to be made.

With the optimal decisions for “do not dispatch” and “dispatch” cases known, we can obtain the optimal joint decisions for $t > 0$ by comparing their corresponding expected costs (5.6) and (5.9).

Theorem 6 (Optimal joint decision) *If $V_{n-1}(x, t) \in \mathcal{V}$, then there exist two nonnegative parameters S_n and Δ_n^* , and a state-dependent parameter H_n^{t-x} , such that the optimal joint decision for period n is given by:*

$$\left\{ \begin{array}{ll} a_n^* = 0, l_n^* = 0, & \text{if } x < H_n^{t-x}, \\ a_n^* = 0, l_n^* = t, & \text{if } x \geq H_n^{t-x}, t > 0, t - x \leq 0, \\ a_n^* = 0, l_n^* = x, & \text{if } x \geq H_n^{t-x}, 0 < t - x \leq \Delta_n^*, \\ a_n^* = t - x + S_n, l_n^* = t, & \text{if } x \geq H_n^{t-x}, t - x > \Delta_n^*. \end{array} \right. \quad (5.16)$$

The cost-to-go function for period n ($n > 0$) is

$$V_n(x, t) = \begin{cases} v_n(0, 0|x, t) & \text{if } x < H_n^{t-x}, \\ = hx + wt + E(V_{n-1}(x, t + Z)), & \\ v_n(0, t|x, t) & \text{if } x \geq H_n^{t-x}, t > 0, t - x \leq 0, \\ = A_D + h(x - t) + E(V_{n-1}(x - t, Z)), & \\ v_n(0, x|x, t) & \text{if } x \geq H_n^{t-x}, 0 < t - x \leq \Delta_n^*, \\ = A_D + w(t - x) + E(V_{n-1}(0, t - x + Z)), & \\ v_n(t - x + S_n, t|x, t) & \text{if } x \geq H_n^{t-x}, t - x > \Delta_n^*. \\ = A_R + c_R(t - x + S_n) + A_D + hS_n + E(V_{n-1}(S_n, Z)), & \end{cases} \quad (5.17)$$

S_n and Δ_n^* are defined in Lemma 9, and H_n^{t-x} is continuous in $t-x$. In addition, there exists $\Delta_n' > \Delta_n^*$ such that when $t - x \geq \Delta_n'$, $H_n^{t-x} = 0$.

Proof of Theorem 6: According to Lemma 8, when there is no retailer order consolidated at the vendor ($t = 0$), we have $V_n(x, t) = v_n(0, 0|x, t)$. When $t > 0$, we need to compare the costs associated with the optimal decisions for “do not dispatch” and “dispatch” options. By Lemma 8 and Lemma 9, the costs for these two options can be expressed as (5.6) and (5.9), respectively. For convenience, we set $t := x + \Delta$ and rewrite equations (5.6) and (5.9) for $x \geq 0$ and $x + \Delta > 0$ as follows:

$$u_n^1(x, x + \Delta) = v_n(0, 0|x, x + \Delta) = (h + w)x + w\Delta + E(V_{n-1}(x, x + \Delta + Z)), \quad (5.18)$$

$$u_n^2(x, x + \Delta) = \begin{cases} v_n(0, x + \Delta | x, x + \Delta) & \text{if } x > 0, -x < \Delta \leq 0, \\ = A_D - h\Delta + E(V_{n-1}(-\Delta, Z)), & \\ v_n(0, x | x, x + \Delta) & \text{if } x > 0, 0 < \Delta \leq \Delta_n^*, \\ = A_D + w\Delta + E(V_{n-1}(0, \Delta + Z)), & \\ v_n(\Delta + S_n, x + \Delta | x, x + \Delta) & \text{if } x = 0 \text{ or } \Delta > \Delta_n^*. \\ = A_R + A_D + c_R\Delta + (c_R + h)S_n + E(V_{n-1}(S_n, Z)), & \end{cases} \quad (5.19)$$

Since $V_{n-1}(x, x + \Delta) \in \mathcal{V}$ and the expectation operator is linear, according to Condition (A3), $E(V_{n-1}(x, x + \Delta + Z))$ is non-decreasing in x for any fixed Δ , hence, $u_n^1(x, x + \Delta)$ is strictly increasing in x for any fixed Δ .

For $u_n^2(x, x + \Delta)$ on $x \geq 0$ and $\Delta > -x$, $\{x = 0\} \cup \{\Delta > \Delta_n^*\} = \{x = 0, 0 < \Delta \leq \Delta_n^*\} \cup \{x \geq 0, \Delta > \Delta_n^*\}$. Therefore, (5.19) can be rewritten as

$$u_n^2(x, x + \Delta) = \begin{cases} v_n(0, x + \Delta | x, x + \Delta) & \text{if } x > 0, -x < \Delta \leq 0, \\ = A_D - h\Delta + E(V_{n-1}(-\Delta, Z)), & \\ v_n(0, x | x, x + \Delta) & \text{if } x > 0, 0 < \Delta \leq \Delta_n^*, \\ = A_D + w\Delta + E(V_{n-1}(0, \Delta + Z)), & \\ v_n(\Delta + S_n, x + \Delta | x, x + \Delta) & \text{if } x \geq 0, \Delta > \Delta_n^*, \\ = A_R + A_D + c_R\Delta + (c_R + h)S_n + E(V_{n-1}(S_n, Z)), & \\ v_n(\Delta + S_n, x + \Delta | x, x + \Delta) & \text{if } x = 0, 0 < \Delta \leq \Delta_n^*. \\ = A_R + A_D + c_R\Delta + (c_R + h)S_n + E(V_{n-1}(S_n, Z)), & \end{cases} \quad (5.20)$$

When $\Delta_n^* = 0$, we have $\{0 < \Delta \leq \Delta_n^*\} = \emptyset$, then the second and the last cases of (5.20) do not exist. Although the expressions of $u_n^2(x, x + \Delta)$ for $\Delta_n^* = 0$ and

$\Delta_n^* > 0$ are different, the optimal policies for the two cases follow the same argument. Therefore, it is sufficient to examine $u_n^2(x, x + \Delta)$ in the form of (5.20).

Let us compare $u_n^1(x, x + \Delta)$ and $u_n^2(x, x + \Delta)$ for four cases discussed in (5.20). For the first three cases, when Δ is fixed, $u_n^2(x, x + \Delta)$ is constant in x . Since $u_n^1(x, x + \Delta)$ is strictly increasing in x for any fixed Δ , there exists $H_n^\Delta \geq 0$, such that if $x < H_n^\Delta$, $u_n^1(x, x + \Delta) < u_n^2(x, x + \Delta)$. If $x \geq H_n^\Delta$, $u_n^1(x, x + \Delta) \geq u_n^2(x, x + \Delta)$. That is,

$$V_n(x, x + \Delta) = \min \{u_n^1(x, x + \Delta), u_n^2(x, x + \Delta)\} = \begin{cases} u_n^1(x, x + \Delta), & \text{if } x < H_n^\Delta, \\ u_n^2(x, x + \Delta), & \text{if } x \geq H_n^\Delta. \end{cases} \quad (5.21)$$

For the last case of (5.20), we have $x = 0$ and $0 < \Delta \leq \Delta_n^*$. This case is only valid when $\Delta_n^* > 0$. We observe that $u_n^1(0, \Delta) = w\Delta + E(V_{n-1}(0, \Delta + Z))$, and $u_n^2(0, \Delta) = A_R + A_D + c_R\Delta + (c_R + h)S_n + E(V_{n-1}(S_n, Z))$. By (5.7), for any $0 < \Delta \leq \Delta_n^*$, $u_n^1(0, \Delta) < u_n^1(0, \Delta) + A_D < u_n^2(0, \Delta)$. Thus,

$$V_n(0, \Delta) = u_n^1(0, \Delta) = v_n(0, 0|0, \Delta). \quad (5.22)$$

When we substitute (5.18) and (5.20) into (5.21) and (5.22), the optimal cost-to-go function for $x + \Delta > 0$ is written as follows:

$$V_n(x, x + \Delta) = \begin{cases} 1).v_n(0, 0|x, x + \Delta), & \text{if } x < H_n^\Delta, x > 0, -x < \Delta \leq 0, \\ 2).v_n(0, 0|x, x + \Delta), & \text{if } x < H_n^\Delta, x > 0, 0 < \Delta \leq \Delta_n^*, \\ 3).v_n(0, 0|x, x + \Delta), & \text{if } x < H_n^\Delta, x \geq 0, \Delta > \Delta_n^*, \\ 4).v_n(0, x + \Delta|x, x + \Delta), & \text{if } x \geq H_n^\Delta, x > 0, -x < \Delta \leq 0, \\ 5).v_n(0, x|x, x + \Delta), & \text{if } x \geq H_n^\Delta, x > 0, 0 < \Delta \leq \Delta_n^*, \\ 6).v_n(\Delta + S_n, x + \Delta|x, x + \Delta), & \text{if } x \geq H_n^\Delta, x \geq 0, \Delta > \Delta_n^*, \\ 7).v_n(0, 0|0, \Delta), & \text{if } x = 0, 0 < \Delta \leq \Delta_n^*. \end{cases} \quad (5.23)$$

The range of H_n^Δ can be examined by analyzing the following three cases, respectively:

Case 1: When $x > 0$ and $-x < \Delta \leq 0$, for a fixed Δ , $\lim_{x \downarrow -\Delta} u_n^1(x, x + \Delta) = -h\Delta + E(V_{n-1}(-\Delta, Z))$, and $\lim_{x \downarrow -\Delta} u_n^2(x, x + \Delta) = A_D - h\Delta + E(V_{n-1}(-\Delta, Z))$. Thus, $\lim_{x \downarrow -\Delta} u_n^1(x, x + \Delta) < \lim_{x \downarrow -\Delta} u_n^2(x, x + \Delta)$. Due to the observations that $u_n^1(x, x + \Delta)$ is strictly increasing in x for a fixed Δ and $u_n^2(x, x + \Delta)$ is constant, $H_n^\Delta > -\Delta$.

Case 2: When $x > 0$ and $0 < \Delta \leq \Delta_n^*$, for a fixed Δ , $\lim_{x \downarrow 0} u_n^1(x, x + \Delta) = w\Delta + E(V_{n-1}(0, \Delta + Z))$, and $\lim_{x \downarrow 0} u_n^2(x, x + \Delta) = A_D + w\Delta + E(V_{n-1}(0, \Delta + Z))$. Thus, $\lim_{x \downarrow 0} u_n^1(x, x + \Delta) < \lim_{x \downarrow 0} u_n^2(x, x + \Delta)$. Consequently, $H_n^\Delta > 0$.

Case 3: When $x \geq 0$ and $\Delta > \Delta_n^*$, we observe that $u_n^1(0, \Delta) = w\Delta + E(V_{n-1}(0, \Delta + Z))$, and $u_n^2(0, \Delta) = A_R + A_D + c_R\Delta + (c_R + h)S_n + E(V_{n-1}(S_n, Z))$. Since Δ_n^* is defined by (5.7),

$$\lim_{\Delta \downarrow \Delta_n^*} A_R + c_R\Delta + (c_R + h)S_n + A_D + E(V_{n-1}(S_n, Z)) = \lim_{\Delta \downarrow \Delta_n^*} A_D + w\Delta + E(V_{n-1}(0, \Delta + Z)).$$

Equivalently, $\lim_{\Delta \downarrow \Delta_n^*} u_n^2(0, \Delta) = \lim_{\Delta \downarrow \Delta_n^*} A_D + u_n^1(0, \Delta)$. It follows that

$$\lim_{\Delta \downarrow \Delta_n^*} u_n^2(0, \Delta) > \lim_{\Delta \downarrow \Delta_n^*} u_n^1(0, \Delta).$$

However, by Condition (A4), $u_n^1(0, \Delta)$ increases faster in Δ than $u_n^2(0, \Delta)$ does. That means $u_n^1(0, \Delta)$ will finally surpass $u_n^2(0, \Delta)$. Hence, there exists $\Delta'_n > \Delta_n^*$, such that when $\Delta_n^* < \Delta < \Delta'_n$, $u_n^1(0, \Delta) < u_n^2(0, \Delta)$ and $H_n^\Delta > 0$. When $\Delta \geq \Delta'_n$, $u_n^1(0, \Delta) \geq u_n^2(0, \Delta)$ and $H_n^\Delta = 0$.

To sum up, $H_n^\Delta = 0$ if $\Delta \geq \Delta'_n$, and $H_n^\Delta > 0$ otherwise. Thus,

$$\{x \geq H_n^\Delta, x > 0, -x < \Delta \leq 0\} = \{x \geq H_n^\Delta, -x < \Delta \leq 0\}, \quad (5.24)$$

$$\{x \geq H_n^\Delta, x > 0, 0 < \Delta \leq \Delta_n^*\} = \{x \geq H_n^\Delta, 0 < \Delta \leq \Delta_n^*\}, \quad (5.25)$$

$$\{x \geq H_n^\Delta, x \geq 0, \Delta > \Delta_n^*\} = \{x \geq H_n^\Delta, \Delta > \Delta_n^*\}. \quad (5.26)$$

In addition, when $x > 0$ and $0 < \Delta \leq \Delta_n^*$, since $H_n^\Delta > 0$, $\{x = 0, 0 < \Delta \leq \Delta_n^*\} = \{x < H_n^\Delta, x = 0, 0 < \Delta \leq \Delta_n^*\}$, and hence, cases 1), 2), 3) and 7) in (5.23) can be combined together as case “ $x < H_n^\Delta, -x < \Delta$ ”. In addition, substituting (5.24), (5.25) and (5.26) into cases 4), 5) and 6) in (5.23), we obtain the expression of $V_n(x, x + \Delta)$ for $x + \Delta > 0$ as

$$V_n(x, x + \Delta) = \begin{cases} v_n(0, 0|x, x + \Delta), & \text{if } x < H_n^\Delta, -x < \Delta, \\ v_n(0, x + \Delta|x, x + \Delta), & \text{if } x \geq H_n^\Delta, -x < \Delta \leq 0, \\ v_n(0, x|x, x + \Delta), & \text{if } x \geq H_n^\Delta, 0 < \Delta \leq \Delta_n^*, \\ v_n(\Delta + S_n, x + \Delta|x, x + \Delta), & \text{if } x \geq H_n^\Delta, \Delta > \Delta_n^*. \end{cases}$$

For a fixed $\Delta \leq 0$, since $H_n^\Delta > -\Delta$, when x approaches $-\Delta$ from the right, it is true that $x < H_n^\Delta, -x < \Delta$. Then, the limit of the cost-to-go function is $\lim_{x \downarrow -\Delta} V_n(x, x + \Delta) = \lim_{x \downarrow -\Delta} u_n^1(x, x + \Delta) = v_n(0, 0|x, x + \Delta)$. Note that when

$x + \Delta = 0$, for any $x \geq 0$, $V_n(x, x + \Delta) = v_n(0, 0|x, x + \Delta)$. Thus, the cost-to-go function for $x \geq 0$, $x + \Delta \geq 0$ can be summarized as

$$V_n(x, x + \Delta) = \begin{cases} u_n^1(x, x + \Delta) = v_n(0, 0|x, x + \Delta), & \text{if } x < H_n^\Delta, \\ u_n^2(x, x + \Delta) = v_n(0, x + \Delta|x, x + \Delta), & \text{if } x \geq H_n^\Delta, -x < \Delta \leq 0, \\ u_n^2(x, x + \Delta) = v_n(0, x|x, x + \Delta), & \text{if } x \geq H_n^\Delta, 0 < \Delta \leq \Delta_n^*, \\ u_n^2(x, x + \Delta) = v_n(\Delta + S_n, x + \Delta|x, x + \Delta), & \text{if } x \geq H_n^\Delta, \Delta > \Delta_n^*, \end{cases} \quad (5.27)$$

and the optimal joint decisions are represented as

$$\begin{cases} a_n^* = 0, l_n^* = 0, & \text{if } x < H_n^\Delta, \\ a_n^* = 0, l_n^* = x + \Delta, & \text{if } x \geq H_n^\Delta, -x < \Delta \leq 0, \\ a_n^* = 0, l_n^* = x, & \text{if } x \geq H_n^\Delta, 0 < \Delta \leq \Delta_n^*, \\ a_n^* = \Delta + S_n, l_n^* = x + \Delta, & \text{if } x \geq H_n^\Delta, \Delta > \Delta_n^*. \end{cases} \quad (5.28)$$

Substituting Δ with $t - x$ in the above expressions, we get the optimal decisions as (5.16), and the optimality equation as (5.17).

In addition, since $E(V_{n-1}(x, t + Z))$ is continuous on the state space of (x, t) based on the hypothesis that $V_{n-1}(x, t) \in \mathcal{V}$ and Condition (A1), $u_n^1(x, t)$ given by (5.6) is also continuous in x and t . By (5.9), it is obvious that $u_n^2(x, t)$ is continuous on the subspace of $x \geq t > 0$. On the subspace of $t > x > 0$, $u_n^2(x, t)$ is defined as the minimum of two continuous functions: $v_n(0, x|x, t)$ and $v_n(t - x + S_n, t|x, t)$, and hence, $u_n^2(x, t)$ is also continuous. We still need to check $u_n^2(x, t)$ on the boundary $t = x$. (1) When $\Delta_n^* > 0$, $\lim_{x \downarrow t} u_n^2(x, t) = \lim_{x \downarrow t} v_n(0, t|x, t) = v_n(0, t|t, t) = \lim_{x \uparrow t} v_n(0, x|x, t) = \lim_{x \uparrow t} u_n^2(x, t)$. (2) When $\Delta_n^* = 0$,

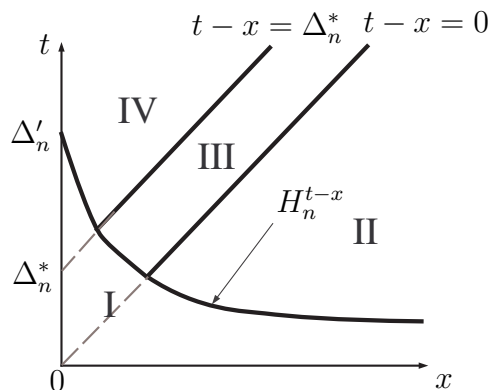
$$\lim_{x \downarrow t} u_n^2(x, t) = \lim_{x \downarrow t} v_n(0, t|x, t) = A_D + E(V_{n-1}(0, Z)),$$

$$\lim_{x \uparrow t} u_n^2(x, t) = \lim_{x \uparrow t} v_n(t - x + S_n, t | x, t) = A_R + A_D + (c_R + h)S_n + E(V_{n-1}(S_n, Z)).$$

By (5.7), the right hand sides of the above two equations are equal when $\Delta_n^* = 0$, and hence, $\lim_{x \downarrow t} u_n^2(x, t) = \lim_{x \uparrow t} u_n^2(x, t)$. It follows that $u_n^2(x, t)$ is also continuous when $t = x$.

Considering that H_n^{t-x} is essentially the intersection of the two continuous functions, i.e., $u_n^1(x, t)$ and $u_n^2(x, t)$, we obtain that H_n^{t-x} is continuous in $t - x$. ■

Figure 19: Illustration of the Optimal Joint Policy



Theorem 6 identifies the structure of the optimal policy for a single-period problem, provided that the cost-to-go function of the next period belongs to the family of functions \mathcal{V} in Definition 4. The optimal policy expressed in (5.16) is basically a zoned, state-dependent threshold policy illustrated by Figure 19. In this policy, the domain of $x \geq 0$ and $t \geq 0$ is divided into three zones by two lines: $t - x = 0$ and $t - x = \Delta_n^*$. In the zone where $t - x \leq 0$, there is inventory excess of consolidated load; in the zone where $0 < t - x \leq \Delta_n^*$, there is a small amount of load excess of on-hand inventory; and in the zone where $t - x > \Delta_n^*$, the load excess is vast. In each zone, the optimal policy is a threshold policy defined by a state-dependent parameter

H_n^{t-x} . More specifically, in each zone when $x < H_n^{t-x}$, it is optimal not to replenish and not to dispatch; otherwise, a positive quantity of consolidated load should be dispatched, and both the optimal dispatch and replenishment quantities depend on which zone the state (x, t) is in. By Theorem 6, we can divide the state space of (x, t) into four regions as depicted by: I, II, III and IV in Figure 19. Note that when $\Delta_n^* = 0$, the state space is divided into two zones by $t - x = 0$, and three regions as: I, II, IV. The optimal joint inventory replenishment and shipment scheduling decisions of the vendor can be characterized as follows:

- Region I: do not replenish and do not dispatch;
- Region II: do not replenish and dispatch the entire consolidated load;
- Region III: do not replenish and dispatch the entire on-hand inventory;
- Region IV: replenish by ordering $t - x + S_n$ units and dispatch the entire consolidated load.

V.3. Optimal Joint Policy for Multi-Period Problems

To ensure that the zoned, state-dependent threshold policy described in the previous sublevel is optimal in period n , the cost-to-go function of the next period (period $n-1$) should be in the class of functions \mathcal{V} , which seems quite restrictive and complicated to examine. However, a careful analysis reveals that for a finite horizon problem, the cost-to-go function of each period always belongs to \mathcal{V} , regardless of the parameter settings.

Theorem 7 *For a finite horizon problem, the optimal joint policy for each period can be described in the form of (5.16).*

Proof of Theorem 7: It is sufficient to show that $V_n(x, t) \in \mathcal{V}$ for any $n = 0, \dots, N - 1$. We prove this by induction.

- When $n = 0$, according to Properties 1-4, $V_0(x, t) \in \mathcal{V}$ as shown in Sublevel V.2.1.
- Now suppose that $V_{k-1}(x, t) \in \mathcal{V}$ for $x \geq 0$ and $t \geq 0$. We need to show that $V_k(x, t) \in \mathcal{V}$. By Theorem 6,

$$V_k(x, t) = \begin{cases} v_k(0, 0|x, t) & \text{if } x < H_k^{t-x}, \\ = hx + wt + E(V_{k-1}(x, t + Z)), & \\ v_k(0, t|x, t) & \text{if } x \geq H_k^{t-x}, t > 0, t - x \leq 0, \\ = A_D + h(x - t) + E(V_{k-1}(x - t, Z)), & \\ v_k(0, x|x, t) & \text{if } x \geq H_k^{t-x}, 0 < t - x \leq \Delta_k^*, \\ = A_D + w(t - x) + E(V_{k-1}(0, t - x + Z)), & \\ v_k(t - x + S_k, t|x, t) & \text{if } x \geq H_k^{t-x}, t - x > \Delta_k^*. \\ = A_R + c_R(t - x + S_k) + A_D + hS_k + E(V_{k-1}(S_k, Z)), & \end{cases}$$

Condition (A1):

Define $\gamma_{k,1}(x, t) = V_k(x, t)$ on $t \leq x$ subspace and $\gamma_{k,2}(x, t) = V_k(x, t)$ on $t > x$ subspace. By the definitions of H_k^{t-x} and Δ_k^* , it can be verified that $\gamma_{k,1}(x, t)$ and $\gamma_{k,2}(x, t)$ are continuous on their respective (x, t) domains. Thus, using the same approach in the proof of Property 1, $E(V_k(x, Z)) + (c_R + h)x$ is continuous in x and $E(V_k(x, t + Z))$ is continuous in x and t . When t is fixed, as x goes to infinity, the limit value of $\gamma_{k,1}(x, t)$ needs to be examined. Note that for a sufficiently large x , $\gamma_{k,1}(x, t) = v_k(0, t|x, t) = A_D + h(x - t) + E(V_{k-1}(x - t, Z))$.

According to Condition (A1) for $V_{k-1}(x, t)$, as x goes to infinity, $V_{k-1}(x, t)$ goes to infinity, and hence, so do $\gamma_{k,1}(x, t)$ and $V_k(x, t)$. Thus, $E(V_k(x, Z)) + (c_R + h)x$ also goes to infinity as x goes to infinity, and on $x \in [0, \infty)$, there exists a value S_{k+1} that minimizes $E(V_k(x, Z)) + (c_R + h)x$.

Condition (A2):

To make the analysis easier, we introduce a new function $u_k'^2(x, t)$ for $x, t \geq 0$ and $k = 1, \dots, N$, which is given by:

$$u_k'^2(x, t) = \begin{cases} v_k(0, 0|x, t) + A_D, & \text{if } x = 0, 0 < t - x \leq \Delta_k^*, \\ v_k(0, 0|x, t) + A_D, & \text{if } t = 0, \\ u_k^2(x, t), & \text{otherwise.} \end{cases}$$

We can rewrite $u_k'^2(x, t)$ as follows:

$$u_k'^2(x, t) = \begin{cases} v_k(0, t|x, t) + A_D \cdot I_{[t=0]} & \text{if } t - x \leq 0, \\ = A_D + h(x - t) + E(V_{k-1}(x - t, Z)), & \\ v_k(0, x|x, t) + A_D \cdot I_{[x=0]} & \text{if } 0 < t - x \leq \Delta_k^*, \\ = A_D + w(t - x) + E(V_{k-1}(0, t - x + Z)), & \\ v_k(t - x + S_k, t|x, t) & \text{if } t - x > \Delta_k^*. \\ = A_R + A_D + c_R(t - x) + (c_R + h)S_k + E(V_{k-1}(S_k, Z)), & \end{cases}$$

It can be easily verified that the cost-to-go function $V_k(x, t)$ computed by equation (5.3) is equal to $\min \{u_k^1(x, t), u_k'^2(x, t)\}$, i.e., $V_k(x, t) = \min \{u_k^1(x, t), u_k'^2(x, t)\}$, for $x, t \geq 0$, where $u_k^1(x, t)$ is defined by equation (5.6). Therefore, we examine the function $V_k(x, t)$ by analyzing the properties of $u_k^1(x, t)$ and $u_k'^2(x, t)$.

First, for any fixed $x \geq 0$, $t \geq 0$ and $\delta > 0$,

$$\begin{aligned}
& u_k^1(x + \delta, t) - u_k^1(x, t) \\
&= h(x + \delta) + wt + E(V_{k-1}(x + \delta, t + Z)) - hx - wt - E(V_{k-1}(x, t + Z)) \\
&= h\delta + E[(V_{k-1}(x + \delta, t + Z) - V_{k-1}(x, t + Z))] \\
&\quad (\text{by the induction hypothesis that } V_{k-1}(x, t) \in \mathcal{V} \text{ and Condition (A2)}) \\
&\geq h\delta + (-A_R - (c_R + h)\delta) \geq -A_R - (c_R + h)\delta.
\end{aligned}$$

Second, for any fixed $x \geq 0$, $t \geq 0$ and $\delta > 0$, we show that $u_k'^2(x + \delta, t) - u_k'^2(x, t) \geq -A_R - (c_R + h)\delta$ by examining the following six cases:

Case 1: When $t - x - \delta < t - x \leq 0$,

$$\begin{aligned}
& u_k'^2(x + \delta, t) - u_k'^2(x, t) \\
&= v_k(0, t|x + \delta, t) - v_k(0, t|x, t) \\
&= h\delta + E(V_{k-1}(x + \delta - t, Z)) - E(V_{k-1}(x - t, Z)) \geq -A_R - (c_R + h)\delta.
\end{aligned}$$

Case 2: When $0 < t - x - \delta < t - x \leq \Delta_k^*$, since $V_{k-1}(x, t) \in \mathcal{V}$, Condition (A4) is satisfied. Then for any $t > 0$, $\frac{dV_{k-1}(0, t)}{dt} \geq c_R$. By the *mean value theorem*, $V_{k-1}(0, t - x - \delta + z) = V_{k-1}(0, z) + \frac{dV_{k-1}(0, \theta)}{d\theta}(t - x - \delta)$ where $z < \theta < t - x - \delta + z$. Since z represents a realized demand, $z \geq 0$, then $\theta > 0$. By Condition (A4), we have $\frac{dV_{k-1}(0, \theta)}{d\theta} \geq c_R$. Therefore, $V_{k-1}(0, t - x - \delta + z) \geq V_{k-1}(0, z) + c_R(t - x - \delta)$, and hence,

$$E(V_{k-1}(0, t - x - \delta + Z)) \geq E(V_{k-1}(0, Z)) + c_R(t - x - \delta). \quad (5.29)$$

By the definition of Δ_k^* in equation (5.7), $v_k(0, x|x, t) \leq v_k(t - x + S_k, t|x, t)$ when $0 < t - x \leq \Delta_k^*$. Equivalently, $A_D + w(t - x) + E(V_{k-1}(0, t - x + Z)) \leq$

$A_R + c_R(t - x + S_k) + A_D + hS_k + E(V_{k-1}(S_k, Z))$, which implies that

$$E(V_{k-1}(0, t - x + Z)) \leq A_R + c_R(t - x + S_k) + hS_k + E(V_{k-1}(S_k, Z)) - w(t - x). \quad (5.30)$$

Therefore,

$$\begin{aligned} u_k'^2(x + \delta, t) - u_k'^2(x, t) &= v_k(0, x + \delta | x + \delta, t) - v_k(0, x | x, t) \\ &= -w\delta + E(V_{k-1}(0, t - x - \delta + Z)) - E(V_{k-1}(0, t - x + Z)) \\ &\geq -w\delta + E(V_{k-1}(0, Z)) + c_R(t - x - \delta) - [A_R + c_R(t - x + S_k) \\ &\quad + hS_k + E(V_{k-1}(S_k, Z)) - w(t - x)] \quad (\text{by (5.29) and (5.30)}) \\ &\geq -w\delta + c_R(t - x - \delta) - A_R - c_R(t - x) + w(t - x) \\ &\quad \left(\begin{array}{l} \text{since } S_k \text{ minimizes } (c_R + h)x + E(V_{k-1}(x, Z)) \text{ on } [0, \infty), \\ E(V_{k-1}(0, Z)) \geq (c_R + h)S_k + E(V_{k-1}(S_k, Z)) \end{array} \right) \\ &= -A_R - c_R\delta + (t - x - \delta)w \geq -A_R - c_R\delta \geq -A_R - (c_R + h)\delta. \end{aligned}$$

Case 3: When $\Delta_k^* < t - x - \delta < t - x$,

$$\begin{aligned} u_k'^2(x + \delta, t) - u_k'^2(x, t) &= v_k(t - x - \delta + S_k, t | x + \delta, t) - v_k(t - x + S_k, t | x, t) \\ &= -c_R\delta \geq -A_R - (c_R + h)\delta. \end{aligned}$$

Case 4: When $t - x - \delta \leq 0$ and $0 < t - x \leq \Delta_k^*$, since S_k is the minimizer of $E(V_{k-1}(x, Z)) + (c_R + h)x$,

$$E(V_{k-1}(x + \delta - t, Z)) \geq E(V_{k-1}(S_k, Z)) + (c_R + h)(S_k - x - \delta + t). \quad (5.31)$$

In addition, by the definition of Δ_k^* in equation (5.7), when $0 < t - x \leq \Delta_k^*$,

$v_k(0, x|x, t) \leq v_k(t - x + S_k, t|x, t)$, which is equivalent to

$$\begin{aligned} A_D + w(t - x) + E(V_{k-1}(0, t - x + Z)) \\ \leq A_R + c_R(t - x + S_k) + A_D + hS_k + E(V_{k-1}(S_k, Z)). \end{aligned}$$

Reordering the terms, we have

$$E(V_{k-1}(0, t - x + Z)) \leq A_R + c_R(t - x + S_k) + hS_k + E(V_{k-1}(S_k, Z)) - w(t - x). \quad (5.32)$$

Hence,

$$\begin{aligned} u_k'^2(x + \delta, t) - u_k'^2(x, t) &= v_k(0, t|x + \delta, t) - v_k(0, x|x, t) \\ &= A_D + h(x + \delta - t) + E(V_{k-1}(x + \delta - t, Z)) \\ &\quad - [A_D + w(t - x) + E(V_{k-1}(0, t - x + Z))] \\ &\geq A_D + h(x + \delta - t) + E(V_{k-1}(S_k, Z)) + (c_R + h)(S_k - x - \delta + t) \\ &\quad - [A_D + A_R + c_R(t - x + S_k) + hS_k + E(V_{k-1}(S_k, Z))] \\ &\quad (\text{ by (5.31) and (5.32) }) \\ &= -A_R - c_R\delta \geq -A_R - (c_R + h)\delta. \end{aligned}$$

Case 5: When $t - x - \delta \leq 0$ and $t - x > \Delta_k^*$, since S_k is the minimizer of $E(V_{k-1}(x, Z)) + (c_R + h)x$, (5.31) still holds. Thus,

$$\begin{aligned} u_k'^2(x + \delta, t) - u_k'^2(x, t) &= v_k(0, t|x + \delta, t) - v_k(t - x + S_k, t|x, t) \\ &= A_D + h(x + \delta - t) + E(V_{k-1}(x + \delta - t, Z)) \\ &\quad - [A_R + A_D + c_R(t - x) + (c_R + h)S_k + E(V_{k-1}(S_k, Z))] \\ &\geq -A_R - c_R \cdot \delta (\text{ by (5.31) }) \geq -A_R - (c_R + h)\delta. \end{aligned}$$

Case 6: When $0 < t - x - \delta \leq \Delta_k^*$ and $t - x > \Delta_k^*$, (5.29) still holds. Thus,

$$\begin{aligned}
& u_k'^2(x + \delta, t) - u_k'^2(x, t) = v_k(0, x + \delta | x + \delta, t) - v_k(t - x + S_k, t | x, t) \\
& = A_D + w(t - x - \delta) + E(V_{k-1}(0, t - x - \delta + Z)) \\
& \quad - [A_R + A_D + c_R(t - x) + (c_R + h)S_k + E(V_{k-1}(S_k, Z))] \\
& \geq E(V_{k-1}(0, Z)) + c_R(t - x - \delta) - A_R - c_R(t - x + S_k) \\
& \quad - hS_k - E(V_{k-1}(S_k, Z)) + w(t - x - \delta) \text{ (by (5.29))} \\
& \geq -A_R - c_R\delta + w(t - x - \delta) \geq -A_R - (c_R + h)\delta. \\
& \quad \left(\begin{array}{l} \text{since } S_k \text{ minimizes } (c_R + h)x + E(V_{k-1}(x, Z)) \text{ on } [0, \infty), \\ E(V_{k-1}(0, Z)) \geq (c_R + h)S_k + E(V_{k-1}(S_k, Z)) \end{array} \right)
\end{aligned}$$

Thus, for all cases $u_k'^2(x + \delta, t) - u_k'^2(x, t) \geq -A_R - (c_R + h)\delta$.

To sum up, since $V_k(x, t) = \min \{u_k^1(x, t), u_k'^2(x, t)\}$, if $V_k(x + \delta, t) = u_k^1(x + \delta, t)$ and $V_k(x, t) = u_k^1(x, t)$, or $V_k(x + \delta, t) = u_k'^2(x + \delta, t)$ and $V_k(x, t) = u_k'^2(x, t)$, then obviously, $V_k(x + \delta, t) - V_k(x, t) \geq -A_R - (c_R + h)\delta$. If $V_k(x + \delta, t) = u_k'^2(x + \delta, t)$ and $V_k(x, t) = u_k^1(x, t)$, this implies $u_k^1(x, t) \leq u_k'^2(x, t)$, hence, $V_k(x + \delta, t) - V_k(x, t) \geq u_k'^2(x + \delta, t) - u_k'^2(x, t) \geq -A_R - (c_R + h)\delta$. The same logic can be applied to the case where $V_k(x + \delta, t) = u_k^1(x + \delta, t)$ and $V_k(x, t) = u_k'^2(x, t)$. Thus, $V_k(x + \delta, t) - V_k(x, t) \geq -A_R - (c_R + h)\delta$.

Condition (A3):

To prove the validation of this condition, we need to write $u_k^1(x, t)$ and $u_k'^2(x, t)$ in terms of x and Δ , i.e., $u_k^1(x, x + \Delta) = v_k(0, 0 | x, x + \Delta) = (h + w)x + w\Delta +$

$E(V_{k-1}(x, x + \Delta + Z))$, and

$$u_k^2(x, x + \Delta) = \begin{cases} v_k(0, x + \Delta | x, x + \Delta) + A_D \cdot I_{[x+\Delta=0]} & \text{if } \Delta \leq 0, \\ = A_D - h\Delta + E(V_{k-1}(-\Delta, Z)), \\ v_k(0, x | x, x + \Delta) + A_D \cdot I_{[x=0]} & \text{if } 0 < \Delta \leq \Delta_k^*, \\ = A_D + w\Delta + E(V_{k-1}(0, \Delta + Z)), \\ v_k(\Delta + S_k, x + \Delta | x, x + \Delta) & \text{if } \Delta > \Delta_k^*. \\ = A_R + A_D + c_R\Delta + (c_R + h)S_k + E(V_{k-1}(S_k, Z)), \end{cases}$$

Obviously, for a fixed Δ , $u_k^1(x, x + \Delta)$ is increasing in x , and $u_k^2(x, x + \Delta)$ is constant in x . Hence, $V_k(x, x + \Delta) = \min \{u_k^1(x, x + \Delta), u_k^2(x, x + \Delta)\}$ is non-decreasing in x for a fixed Δ .

Condition (A4):

When $x = 0$, $V_k(0, t)$ is either equal to $v_k(0, 0 | 0, t)$ or $v_k(t + S_k, t | 0, t)$. If $V_k(0, t) = v_k(0, 0 | 0, t) = wt + E(V_{k-1}(0, t + Z))$, then $\frac{dV_k(0, t)}{dt} = w + \frac{dE(V_{k-1}(0, t + Z))}{dt} \geq w + c_R$. Otherwise, $V_k(0, t) = v_k(t + S_k, t | 0, t) = A_R + A_D + c_R(t + S_k) + hS_k + E(V_{k-1}(S_k, Z))$, and $\frac{dV_k(0, t)}{dt} = c_R$. Thus, $\frac{dV_k(0, t)}{dt} \geq c_R$.

Since Conditions (A1)–(A4) also hold for $V_k(x, t)$, $V_n(x, t) \in \mathcal{V}$ for any $n = 0, \dots, N - 1$. ■

Theorem 7 establishes that in any period during a finite horizon problem, the optimal joint inbound inventory replenishment and outbound shipment scheduling policy is in the form of a zoned, state-dependent threshold policy. For any period n , the vendor should replenish her inventory if and only if the excess position $(t - x)$ exceeds a certain level (Δ_n^*) and the on-hand inventory level x is above a threshold value (H_n^{t-x}) . The corresponding optimal replenishment quantity is equal to the sum

of the excess position and a critical value S_n . In all other situations, the lowest cost is achieved when there is no inventory replenishment. For the cases when the size of the consolidated load is small, or both the on-hand inventory level and the load excess are in low volumes, it is more preferable to choose to be idle in that period, i.e., do not replenish and do not dispatch. When there is inventory excess or mild load excess, if the on-hand inventory level is higher than a state-dependent threshold value, then it is optimal to dispatch as many consolidated orders as possible.

Figure 20: A Realization of the Process

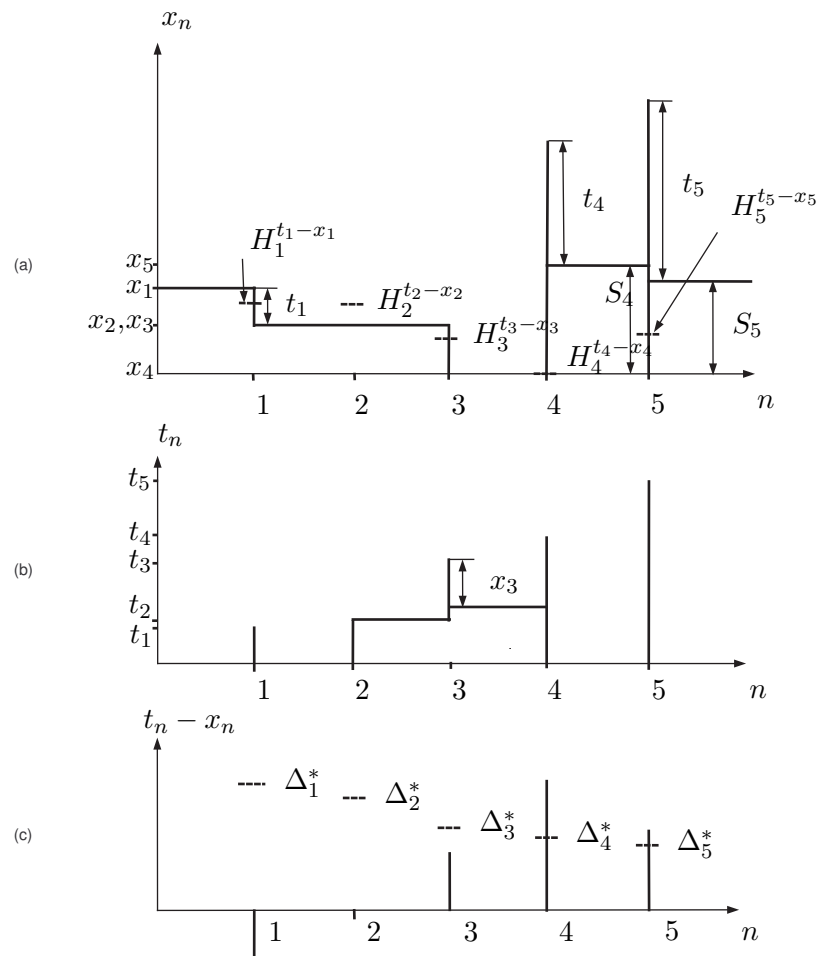


Figure 20 illustrates a realization of a VMI system running under the optimal policy for 5 periods. x_n and t_n denote the on-hand inventory level and the consolidated load in period n , respectively. x_n increases when the vendor replenishes her inventory, and t_n increases when the vendor receives an order (realized stochastic demand of a period). Both x_n and t_n decrease when a shipment is dispatched. Orders are received at the end of each period. The joint decision is made immediately after the order arrival at the end of each period.

$t_n - x_n$ is the excess position level observed right before the n -th decision is made. In period n , the optimal policy can be defined by parameters Δ_n^* , S_n and $H_n^{(\cdot)}$ where Δ_n and S_n are independent of the states x_n and t_n , and $H_n^{(\cdot)}$ is a function of $t_n - x_n$. For this example, in period 1, we observe that $t_1 - x_1 < 0$ (from part (c) of Figure 20) and $x_1 > H_1^{t_1 - x_1}$ (from part (a) of Figure 20). Hence, the optimal decision is to dispatch the entire consolidated load without replenishing the inventory. For period 2, $x_2 < H_2^{t_2 - x_2}$, which implies that the optimal decision is to postpone the dispatch until the next period. For period 3, a retailer order arrives to raise the consolidated load to t_3 such that $0 < t_3 - x_3 \leq \Delta_3^*$ (from part (c) of Figure 20). Since $x_3 \geq H_3^{t_3 - x_3}$, the optimal decision is to dispatch the entire on-hand inventory without replenishment. In period 4, the consolidated load is high enough such that $t_4 - x_4 > \Delta_4^*$ and $x_4 = 0 = H_4^{t_4 - x_4}$, and hence, the optimal decision is to replenish $t_4 - x_4 + S_4$ and dispatch t_4 . Thus, the remaining inventory level is exactly equal to S_4 . In period 5, a large order quantity t_5 arrives at the vendor so that $t_5 - x_5 > \Delta_5^*$ and $x_5 > H_5^{t_5 - x_5}$. It is optimal for the vendor to replenish $t_5 - x_5 + S_5$ and dispatch t_5 , and the remaining inventory level is S_5 .

V.4. Summary

Although the integration of inventory replenishment and shipment scheduling is investigated in the literature, all the existing research adopts a pre-defined (e.g., quantity-based or time-based) temporal shipment consolidation policy, and attempts to optimize for the underlying parameter values. Although these policies are practical, they might be suboptimal in the class of all feasible policies. This paper is the first to examine the optimal joint inventory replenishment and shipment consolidation policy in a multi-level supply chain system.

We formulate the problem as a discrete-time Markov decision process via a stochastic dynamic programming approach, and we examine the optimal joint policy for the case of private fleet transportation. To address the economies of scale in transportation, we include fixed inbound replenishment cost and fixed outbound transportation cost, respectively, for each echelon. We prove the optimality of a zoned, state-dependent threshold policy. More specifically, we show that depending on the difference between the two states (the on-hand inventory level and the size of the consolidated load), the two-dimensional state space can be divided into three zones. On each zone, the optimal joint replenishment and dispatch decisions can be determined by a threshold policy. As a result, in each period the optimal decision can only be chosen from the following four options: (1) do not replenish and do not dispatch; (2) do not replenish and dispatch the entire consolidated load; (3) do not replenish and dispatch the entire on-hand inventory; (4) replenish an amount so that after dispatching the entire consolidated load, the remaining inventory level is equal to a critical value. Due to the existence of the fixed replenishment cost, the optimal joint policy does not exhibit the clearing property.

CHAPTER VI

CONCLUSIONS AND FUTURE DIRECTIONS

This dissertation concentrates on coordinating the inventory and transportation decisions in stochastic dynamic demand supply chain systems. In contrast to the existing research that mainly aims at optimizing the cost over a set of feasible and predefined policies, we utilize stochastic dynamic programming techniques to determine the structures of the optimal inventory and/or shipment policies. In order to achieve our objectives, we investigate three classes of problems: (1) Pure inbound inventory model; (2) Pure outbound transportation models; and (3) Integrated inbound inventory and outbound transportation model.

Having roots in applied probability, optimization, inventory theory and optimal control theory, this dissertation makes several theoretical contributions in stochastic modeling and optimization. The contributions are discussed in detail in Sublevel VI.1. Future research directions are provided in Sublevel VI.2

VI.1. Contributions

In Chapter III, the pure inbound inventory model puts its emphasis on answering the questions of how often and in what quantities to replenish stock so that replenishment (including shipping cost), holding, and waiting costs are minimized. The studied model is a generalization of the classical stochastic dynamic inventory problem in that a multiple setup cost structure is included to explicitly represent a private fleet of trucks with finite capacity. We develop a new replenishment policy, called (Q, \vec{s}, \vec{S}) policy, which is more general than previously examined policies in the literature for the problem of interest. Sufficient conditions for the optimality of the new policy are provided. To this end, we introduce a new concept, non- (Δ, C) -decreasing, which is

analogous to the existing concept of non- K -decreasing. For the special case when the demand is uniformly distributed on $(0, mC)$ where m is an integer and C is the cargo capacity, we prove the optimality of the (Q, \vec{s}, \vec{S}) policy that in turn reduces to a simpler form. Our contributions are highlighted below:

- We defined new concepts of *non- (Δ, C) -decreasing* and *non- $(\Delta, C)_{\mathbb{N}}^K$ -decreasing*;
- We proposed a new replenishment policy: (Q, \vec{s}, \vec{S}) policy;
- We provided sufficient conditions under which the new policy is optimal; and
- We investigated a special case when the demand is uniformly distributed over and identifying the exact optimal policy which has a simple form characterized by parameters $(Q = s, s, S)$.

We note that the results of the special case are beneficial for generating easier-to-compute, approximate policies or for new products with no historical data of demand patterns so that demand is assumed to be uniformly distributed.

In Chapter IV, the pure outbound transportation model investigates the problem regarding when to dispatch and how large the dispatch quantity should be so that transportation scale economies are achieved while the timely service requirements are not sacrificed. Since transportation exhibits scale economies, savings are particularly realizable when the transportation decisions include a strategy for shipment consolidation, under which small shipments can be consolidated into a single larger load. Time-base, quantity-based, and hybrid consolidation policies are mostly implemented in practice. In the literature, Çetinkaya and Bookbinder (2003) compute the optimal parameter values of the time-based and quantity-based consolidation policies for private and common carriage cases, respectively. However, these two policies may not be optimal in all feasible policies. Hence, we model the same supply chain system

as a stochastic dynamic program and examine the optimal consolidation policies for different types of transportation costs: (1) private fleet transportation without cargo capacity constraint, (2) single-truck transportation with cargo capacity and fixed cost, (3) common carriage with quantity discount schedule, and (4) multi-truck transportation with cargo capacity. We characterize the structures of the optimal policies for cases (1) and (2). We show, theoretically that in these two cases, the optimal policy possesses a control limit property and the dispatch quantity is either zero or equal to the maximum possible value. Assuming the “clearing properties”, we investigate the optimality of some practical policies for case (3) and provide sufficient conditions under which they are optimal. We prove that under different parameter settings, the optimal policy for case (3) can be an immediate dispatch policy, or a threshold policy, or a more complicated (S_L, S_U) policy. The complexity of the exact optimal policy is illustrated numerically. In addition, preliminary results for case (4) are provided.

This research contributes to the periodic-review, stochastic, consolidation systems literature by explicitly considering transportation costs related to different industry practices. Analytically, we contribute by developing a new consolidation policy: (S_L, S_U) policy, and providing sufficient conditions under which the discussed practical policies are optimal.

The model that we consider in Chapter V optimizes the inbound inventory replenishment and outbound dispatch decisions simultaneously. We consider a two-echelon supply chain system applicable under a VMI contract where the vendor uses a private fleet for inbound replenishment and outbound shipments. Shipment consolidation strategy is again employed on the outbound logistics. There are some studies in the literature that analyze the same system, however, they all choose the inbound inventory and outbound transportation decisions in advance and concentrate on finding the policy parameter values that yield a minimal cost.

To characterize the exact optimal joint policy, we formulate the problem as a two-state, two-action stochastic dynamic program. We first identify the structure of the optimal policy for a single-period problem for an arbitrary period, provided that the value function of the next period belongs to a specific family of functions characterized explicitly in the paper. The exact optimal joint policies specifying the vendor's inbound replenishment and outbound dispatch quantities in successive periods is in the form of a zoned, state-dependent threshold policy. We then provide a formal proof that the value function of each period always belongs to the specific family of functions characterized, regardless of the parameter settings. It follows that in any period during a finite horizon problem, the optimal policy is a zoned, state-dependent threshold policy. Our contributions are highlighted below:

- Our research is the first to identify the structural properties of the vendor's exact optimal joint policies.
- The characterized *zoned, state-dependent threshold policy* is a new class of policies in multi-echelon stochastic inventory control theory.
- This research renders insights into ways that we design and implement supply chain applications, e.g., VMI and Third Party Logistics.

VI.2. Future Work

Several extensions of the presented work are possible.

1. **Infinite Horizon:** As noted earlier, our focus in this dissertation is on theoretical analysis of the optimal inventory and/or transportation policies with consideration of realistic transportation cost. All models investigated are over a finite horizon. Usually for a finite-period dynamic system, although the poli-

cies in different periods have the same structure, the parameters that define the policy can be very different from each other. However, for a supply chain system that needs to make frequent decisions over many time periods, it would be more convenient to have a stationary policy. Here, “stationary” means the policy parameters do not vary from period to period. Therefore, it is worthwhile to show the optimality of the characterized policies in the infinite horizon discounted and average cost cases.

2. **Solution Methods:** Although the analytical results regarding the structure of the optimal policies have a theoretical value, to strengthen the practical contribution of this research, solution methods or algorithms for efficient calculation of the policy values need to be developed. It is known that the computational limitations on stochastic dynamic programming have made it very difficult to find the optimal values of the policy parameters, even if it is a conceptually powerful technique. In addition, incorporation of time-windows, explicit consideration of general freight cost structures, and cargo capacity constraints increase the computational requirements for this class of problems. Therefore, designing a well-formed approximate algorithm is a possible future research avenue. When a solution method is available, the advantages of the optimal policy over those suboptimal, yet practical policies can also be evaluated.
3. **Generalized Transportation Cost in Integrated Model:** In Chapter V, the transportation cost is presented as the summation of a fixed setup cost and a linear variable cost. Clearly, this cost structure ignores the impact of transportation cost and capacity related to delivery of orders; thereby, also ignoring possible transportation scale economies achievable via optimization. An interesting generalization is to investigate the transportation costs for single

or multiple capacitated trucks, or common carriage transportation in both the inbound and outbound logistics of the integrated replenishment and shipment model.

4. **Multiple Items:** It would also be of interest to consider multiple products that have different demand distributions, procurement costs (hence, different inventory holding costs), and customer waiting costs. As mentioned previously, shipment consolidation can be applied to combine orders of the same item ordered by the customers at different time, or the orders of different items ordered at the same time (more accurately, during a sufficiently short period). In such a supply chain system, we need to consider the proper dispatch schedules to reduce the total cost of transportation and waiting.

REFERENCES

- Alp, O., N.K. Erkip, R. Güllü. 2003. Optimal lot-sizing/vehicle-dispatching policies under stochastic lead times and stepwise fixed costs. *Oper. Res.* **51**(1) 160–166.
- Anily, S., M. Tzur. 2005. Shipping multiple items by capacitated vehicles: An optimal dynamic programming approach. *Transp. Sci.* **39**(2) 233–248.
- Anily, S., M. Tzur. 2006. Algorithms for the multi-item multi-vehicles dynamic lot sizing problem. *Nav. Res. Logist.* **53** 157–169.
- Aucamp, D.C. 1982. Nonlinear freight costs in the EOQ problem. *Eur. J. Oper. Res.* **9** 61–63.
- Aucamp, D.C. 1984. A solution to the multiple set-up problem. *Int. J. Prod. Res.* **22**(4) 549–554.
- Axsäter, S. 2001. A note on stock replenishment and shipment scheduling for vendor-managed inventory systems. *Manage. Sci.* **47**(9) 1306–1310.
- Bagchi, P.K., F.W. Davis. 1988. Some insights into inbound freight consolidation. *Int. J. Phys. Distrib. Mater. Manage.* **18**(6) 27–33.
- Baita, F., W. Ukovich, R. Pesenti, D. Favaretto. 1998. Dynamic routing and inventory problems: a review. *Transp. Res.-A* **32**(8) 585–598.
- Ballou, R.H. 1992. *Business Logistics Management*. Prentice Hall, Englewood Cliffs, NJ.

- Bell, W.J., L.M. Dalberto, M.L. Fisher, A.J. Greenfield, R. Jaikumar, P. Kedia, R.G. Mack, P.J. Prutzman. 1983. Improving the distribution of industrial gases with an online computerized routing and scheduling optimizer. *Interfaces* **13**(6) 4–23.
- Ben-khedher, N., C.A. Yano. 1994. The multi-item joint replenishment problem with transportation and container effects. *Transp. Sci.* **28**(1) 37–54.
- Bookbinder, J.H., J.K. Higginson. 2002. Probabilistic modeling of freight consolidation by private carriage. *Transp. Res.-E* **38**(5) 305–318.
- Campbell, A., L. Clark, A. Kleywegt, M. Savelsberg. 1998. The inventory routing problem. T.G. Crainic and G. Laporte eds. *Fleet Management and Logistics*. Kluwer, Norwell, MA.
- Çalışkan Demirağ, O., F.Y. Chen, Y. Yang. 2011. Optimal lot-sizing for a periodic-review inventory system with quantity-dependent setup costs. *Submitted*, Black School of Business, Pennsylvania State University Erie, Erie, PA.
- Çetinkaya, S. 2005. Coordination of inventory and shipment consolidation decisions: A review of premises, models, and justification. Geunes, J. et al. eds. *Applications of Supply Chain Management and E-Commerce Research in Industry*. Kluwer, Norwell, MA.
- Çetinkaya, S., J.H. Bookbinder. 2003. Stochastic models for the dispatch of consolidated shipments. *Transp. Res.-B* **37**(8) 747–768.
- Çetinkaya, S., C.Y. Lee. 2000. Stock replenishment and shipment scheduling for vendor managed inventory systems. *Manage. Sci.* **46**(2) 217–232.
- Çetinkaya, S., C.Y. Lee. 2002. Optimal outbound dispatch policies: Modeling inventory and cargo capacity. *Nav. Res. Logist.* **49** 531–556.

- Çetinkaya, S., F. Mutlu, C.Y. Lee. 2006. A comparison of outbound dispatch policies for integrated inventory and transportation decisions. *Eur. J. Oper. Res.* **171** 1094–1112.
- Çetinkaya, S., M. Parlar. 2004. Computing a stationary base-stock policy for a finite horizon stochastic inventory problem with non-linear shortage costs. *Stoch. Anal. Appl.* **22**(3) 589–625.
- Çetinkaya, S., E. Tekin, C.Y. Lee. 2008. A stochastic model for joint inventory and outbound shipment decisions. *IIE Trans.* **40** 324–340.
- Chen, F. 2000. Optimal policies for multi-echelon inventory problems with batch ordering. *Oper. Res.* **48** 376–389.
- Chen, F., S. Ray, Y. Song. 2006. Optimal pricing and inventory control policy in periodic-review systems with fixed ordering cost and lost sales. *Nav. Res. Logist.* **53** 117–136.
- Chen, F., J.S. Song. 2001. Optimal policies for multi-echelon inventory problems with Markov-modulated demand. *Oper. Res.* **49** 226–234.
- Chen, F., T. Wang, T.Z. Xu. 2005. Integrated inventory replenishment and temporal shipment consolidation: A comparison of quantity-based and time-based models. *Ann. Oper. Res.* **135** 197–210.
- Chen, F., Y. Zheng. 1994. Lower bounds for multi-echelon stochastic inventory problems. *Manage. Sci.* **40** 1426–1443.
- Chen, H.D., D.W. Hearn, C.Y. Lee. 1994. A new dynamic programming algorithm for the single item capacitated dynamic lot size model. *J. Glob. Optim.* **4** 285–300.

- Cheung, K.L., H.L. Lee. 2002. The inventory benefit of shipment coordination and stock rebalancing in a supply chain. *Manage. Sci.* **48**(2) 300–306.
- Clark, A.J., H. Scarf. 1960. Optimal policies for a multi-echelon inventory problem. *Manage. Sci.* **6**(4) 475–490.
- Closs, D.J., R.L. Cook. 1987. Multi-stage transportation consolidation analysis using dynamic simulation. *Int. J. Phys. Distrib. Mater. Manage.* **17**(3) 28–45.
- Cooper, M.C. 1984. Cost and delivery time implications of freight consolidation and warehouse strategies. *Int. J. Phys. Distrib. Mater. Manage.* **14**(6) 47–67.
- Erengüç, S.S., N.C. Simpson, A.J. Vakharia. 1999. Integrated production/distribution planning in supply chains. *Eur. J. Oper. Res.* **115**(2) 219–236.
- Ertogral, K., M. Darwish, M. Ben-Daya. 2007. Production and shipment lot sizing in a vendor-buyer supply chain with transportation cost. *Eur. J. Oper. Res.* **176** 1592–1606.
- Federgruen, A., P. Zipkin. 1984. A combined vehicle routing and inventory allocation problem. *Oper. Res.* **32**(5) 1019–1037.
- Federgruen, A., P. Zipkin. 1984a. Computational issues in an infinite horizon, multi-echelon inventory model. *Oper. Res.* **32** 818–836.
- Federgruen, A., P. Zipkin. 1984b. Approximations of dynamic multi-location production and inventory problems. *Manage. Sci.* **30** 69–84.
- Florian, M., M. Klein. 1971. Deterministic production planning with concave costs and capacity constraints. *Manage. Sci.* **18**(1) 12–20.

- Gallego, G., O. Ozer. 2003. Optimal replenishment policies for multi-echelon inventory problems under advance demand information. *Manuf. Serv. Oper. Manage.* **5** 157–175.
- Gallego, G., O. Ozer. 2005. A new algorithm and a new heuristic for serial supply systems. *Oper. Res. Lett.* **33** 349–362.
- Gallego, G., A. Scheller-Wolf. 2000. Capacitated inventory problems with fixed order costs: Some optimal policy structure. *Eur. J. Oper. Res.* **126** 603–613.
- Gallego, G., L.B. Toktay. 2004. All-or-nothing ordering under a capacity constraint. *Oper. Res.* **52**(6) 1001–1002.
- Golden, B., A. Assad, R. Dahl. 1984. Analysis of a large scale vehicle routing problem with an inventory component. *Large Scale Syst.* **7** 181–190.
- Goyal, S.K. 1976. An integrated inventory model for a single supplier single customer problem. *Int. J. Prod. Res.* **15** 107–111.
- Goyal, S.K. 1988. A joint economic lot size model for purchaser and vendor: A comment. *Decis. Sci.* **19** 236–241.
- Goyal, S.K. 1995. A one-vendor multi-buyer integrated inventory model: A comment. *Eur. J. Oper. Res.* **82** 209–210.
- Goyal, S.K., F. Nebebe. 2000. Determination of economic production-shipment policy for a single-vendor-single-buyer system. *Eur. J. Oper. Res.* **121** 175–178.
- Guan, R.X., X.B. Zhao. 2010. On contracts for VMI program with continuous review (r,Q) policy. *Eur. J. Oper. Res.* **207**(2) 656–667.

- Gupta, Y.P., P.K. Bagchi. 1987. Inbound freight consolidation under just-in-time procurement: Application of clearing models. *J. Bus. Logist.* **8**(2) 74–94.
- Gurbuz, M.C., K. Moinzadeh, Y.P. Zhou. 2007. Replenishment strategies in inventory/distribution systems. *Manage. Sci.* **53**(2) 293–307.
- Higginson, J.K. 1993. Modeling shipper costs in physical distribution analysis. *Transp. Res.-A* **27A**(2) 113–124.
- Higginson, J.K., J.H. Bookbinder. 1994. Policy recommendations for a shipment consolidation program. *J. Bus. Logist.* **15**(1) 87–112.
- Higginson, J.K., J.H. Bookbinder. 1995. Markovian decision processes in shipment consolidation. *Transp. Sci.* **29**(3) 242–255.
- Hill, R.M. 1997. The single-vendor single-buyer integrated production inventory model with a generalized policy. *Eur. J. Oper. Res.* **97** 493–499.
- Hill, R.M. 1999. The optimal production and shipment policy for the single-vendor single-buyer integrated production inventory problem. *Int. J. Prod. Res.* **37** 2463–2475.
- Hoque, M.A., S.K. Goyal. 2000. An optimal policy for single-vendor single-buyer integrated production-inventory system with capacity constraint of transport equipment. *Int. J. Prod. Econ.* **65** 305–315.
- Hwang, H., D. Moon, S. Shinn. 1990. An EOQ model with quantity discounts for both purchasing price and freight cost. *Comput. Oper. Res.* **17**(1) 73–78.
- Hwang, H.C. 2009. Inventory replenishment and inbound shipment scheduling under a minimum replenishment policy. *Transp. Sci.* **43**(2) 244–264.

- Hwang, H.C. 2010. Economic lot-sizing for integrated production and transportation. *Oper. Res.* **58**(2) 428–444.
- Hwang, H.C., W. Jaruphongsa. 2006. Dynamic lot-sizing model with demand time windows and speculative cost structure. *Oper. Res. Lett.* **34** 251–256.
- Iglehart, D.L. 1963. Optimality of (s, S) policies in the infinite horizon dynamic inventory problem. *Manage. Sci.* **9**(2) 259–267.
- Iwaniec, K. 1979. An inventory model with full load ordering. *Manage. Sci.* **25**(4) 374–384.
- Jackson, G.C. 1981. Evaluating order consolidation strategies using simulation. *J. Bus. Logist.* **2**(2) 110–138.
- Jackson, G.C. 1985. A survey of freight consolidation practices. *J. Bus. Logist.* **6**(1) 13–34.
- Janakiraman, G., R.O. Roundy. 2004. Lost-sales problems with stochastic lead times: Convexity results for base-stock policies. *Oper. Res.* **52**(5) 795–803.
- Jaruphongsa, W., S. Çetinkaya, C. Y. Lee. 2005. A dynamic lot-sizing model with multi-mode replenishments: Polynomial algorithms for special cases with dual and multiple modes. *IIE Trans.* **37** 453–467.
- Jaruphongsa, W., S. Çetinkaya, C. Y. Lee. 2007. Outbound shipment mode considerations for integrated inventory and delivery lot-sizing decisions. *Oper. Res. Lett.* **35** 813–822.
- Jaruphongsa, W., C.Y. Lee. 2008. Dynamic lot-sizing problem with demand time windows and container-based transportation cost. *Optimi. Lett.* **2** 39–51.

- Jin, Y., A Muriel. 2009. Single-warehouse multi-retailer inventory systems with full truckload shipments. *Nav. Res. Logist.* **56** 450–464.
- Kaya, O., D. Kubali, L. Örmeci. 2010. Stochastic models for the coordinated production and shipment problem in a supply chain. *Submitted*, Department of Electrical and Electronics Engineering, Işık University, Istanbul, Turkey.
- Kleywegt, A.J., V.S. Nori, M.W.P. Savelsbergh. 2004. Dynamic programming approximations for a stochastic inventory routing problem. *Transp. Sci.* **38** 42–70.
- Lee, C.Y. 1986. The economic order quantity for freight discount costs. *IIE Trans.* **18**(3) 318–320.
- Lee, C.Y. 1989. A solution to the multiple set-up problem with dynamic demand. *IIE Trans.* **21**(3) 266–270.
- Lee, C.Y. 2004. Inventory replenishment model: Lot sizing versus just-in-time delivery. *Oper. Res. Lett.* **32** 581–590.
- Lee, C.Y., S. Çetinkaya, W. Jaruphongsa. 2003. A dynamic model for inventory lot sizing and outbound shipment scheduling at a third-party warehouse. *Oper. Res.* **51**(5) 735–747.
- Lee, C.Y., S. Çetinkaya, A.P.M. Wagelmans. 2001. A dynamic lot-sizing model with demand time windows. *Manage. Sci.* **47**(10) 1384–1395.
- Li, C.L., V.N. Hsu, W.Q. Xiao. 2004. Dynamic lot sizing with batch ordering and truckload discounts. *Oper. Res.* **52**(4) 639–654.
- Lippman, S.A. 1969a. Optimal inventory policy with multiple set-up costs. *Manage. Sci.* **16**(1) 118–138.

- Lippman, S.A. 1969b. Optimal inventory policy with subadditive ordering costs and stochastic demands. *SIAM J. Appl. Math.* **17**(3) 543–559.
- Lippman, S.A. 1971. Economic order quantities and multiple set-up costs. *Manage. Sci.* **18**(1) 39–47.
- Love, S.F. 1973. Bounded production and inventory models with piecewise concave costs. *Manage. Sci.* **20**(3) 313–318.
- Lu, L. 1995. A one-vendor multi-buyer integrated inventory model. *Eur. J. Oper. Res.* **81** 312–323.
- Masters, J.M. 1980. The effects of freight consolidation on customer service. *J. Bus. Logist.* **2**(1) 55–74.
- Mendoza, A., J.A. Ventura. 2008. Incorporating quantity discounts to the EOQ model with transportation costs. *Int. J. Prod. Econ.* **113** 754–765.
- Minkoff, A.S. 1993. A Markov decision model and decomposition heuristic for dynamic vehicle dispatching. *Oper. Res.* **41**(1) 77–90.
- Mutlu, F. 2006. The transporters impact on channel coordination and contractual agreements. Ph.D. dissertation, Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX.
- Mutlu, F., S. Çetinkaya. 2010. An integrated model for stock replenishment and shipment scheduling under common carrier dispatch costs. *Transp. Res.-E* **46** 844–854.
- Mutlu, F., S. Çetinkaya, H. Bookbinder. 2010. An analytical model for computing the optimal time-and-quantity-based policy for consolidated shipments. *IIE Trans.* **42** 367–377.

- Newbourne, M.J., C. Barrett. 1972. Freight consolidation and the shipper, parts 1 to 5. *Transp. Distrib. Manage.* **12**(2-6).
- Parlar, M., R. Rempala. 1992. Stochastic inventory problem with piecewise quadratic holding cost function containing a cost-free interval. *J. Optim. Theory Appl.* **75**(1) 133–153.
- Pollock, T. 1978. A management guide to LTL consolidation. *Traffic World*. April 3 29–35.
- Pooley, J., A.J. Stenger. 1992. Modeling and evaluating shipment consolidation in a logistics system. *J. Bus. Logist.* **13**(2) 153–174.
- Porteus, E.L. 1971. The optimality of generalized (s, S) policies. *Manage. Sci.* **17**(7) 411–426.
- Porteus, E.L. 1972. The optimality of generalized (s, S) policies under uniform demand densities. *Manage. Sci.* **18**(11) 644–646.
- Rieksts, B.Q., J.A. Ventura. 2010. Two-stage inventory models with a bi-modal transportation cost. *Comput. Oper. Res.* **37** 20–31.
- Savaseneril, S., N. Erkip. 2010. An analysis of manufacturer benefits under vendor-managed systems. *IIE Trans.* **42**(7) 455–477.
- Scarf, H. 1960. The optimality of (S, s) policies in the dynamic inventory problem. K.J. Arrow, S. Karlin, and P. Suppes eds. *Mathematical Methods in the Social Sciences*. Stanford University Press, Palo Alto, CA.
- Schal, M. 1976. On the optimality of (s, S) -policies in dynamic inventory models with finite horizon. *SIAM J. Appl. Math.* **30**(3) 528–537.

- Schwarz, L.B. 1973. A simple continuous review deterministic one-warehouse n-retailer inventory problem. *Manage. Sci.* **19** 555–566.
- Schwarz, L.B., J.E. Ward, X. Zhai. 2006. On the interactions between routing and inventory-management policies in a one-warehouse n-retailer distribution system. *Manuf. Serv. Oper. Manage.* **8**(3) 253–272.
- Sethi, S.P., F. Cheng. 1997. Optimality of (s, S) policies in inventory models with Markovian demand. *Oper. Res.* **45**(6) 931–939.
- Shaoxiang, C. 2004. The infinite horizon periodic review problem with setup costs and capacity constraints: A partial characterization of the optimal policy. *Oper. Res.* **52**(3) 409–421.
- Simchi-Levi, D., S. Kaminsky, E. Simchi-Levi. 2007. *Designing and Managing the Supply Chain*. 3rd ed. McGraw-Hill/Irwin, New York, NY.
- Simchi-Levi, D., D. Nelson, N. Mulani, J. Wright. 2008. “Crude calculations: Why high oil prices are upending the way companies should manage their supply chains,” *The Wall Street Journal*. September 22.
- Swoveland, C. 1975. A deterministic multi-period production planning model with piecewise concave production and holding-backorder costs. *Manage. Sci.* **21**(9) 1007–1013.
- Toptal, A. 2009. Replenishment decisions under an all-units discount schedule and stepwise freight costs. *Eur. J. Oper. Res.* **198** 504–510.
- Toptal, A., S. Çetinkaya. 2006. Contractual agreements for coordination and vendor-managed delivery under explicit transportation considerations. *Nav. Res. Logist.* **53** 397–417.

- Toptal, A., S. Çetinkaya. 2008. Quantifying the value of buyer-vendor coordination: Analytical and numerical results under different replenishment cost structures. *Eur. J. Oper. Res.* **187** 785–805.
- Toptal, A., S. Çetinkaya, C.Y. Lee. 2003. The buyer-vendor coordination problem: Modeling inbound and outbound cargo capacity and costs. *IIE Trans.* **35** 987–1002.
- van Houtum, G.J., A. Scheller-Wolf, J. Yi. 2007. Optimal control of serial inventory systems with fixed replenishment intervals. *Oper. Res.* **55**(4) 674–687.
- Veinott, A.F. 1966. On the optimality of (s, S) inventory policies: New conditions and a new proof. *SIAM J. Appl. Math.* **14**(5) 1067–1083.
- Veinott, A.F., H.M. Wagner. 1965. Computing optimal (s, S) inventory policies. *Manage. Sci.* **11**(5) 525–552.
- Wagner, H.M., T.M. Whitin. 1958. Dynamic version of the economic lot size model. *Manage. Sci.* **5**(1) 89–96.
- Wilson, R. 2008. *19th Annual State of Logistics Report*. Tech. rep., Council of Supply Chain Management Professionals, Lombard, IL.
- Zangwill, W.I. 1966. A deterministic multi-period production scheduling model with backlogging. *Manage. Sci.* **13**(1) 105–199.
- Zangwill, W.I. 1969. A backlogging model and a multi-echelon model of a dynamic lot size production system. *Manage. Sci.* **15**(9) 509–527.
- Zhao, Q.H., S.Y. Wang, K.K. Lai. 2007. A partition approach to the inventory/routing problem. *Eur. J. Oper. Res.* **177** 786–802.

Zhao, Q.H., S.Y. Wang, K.K. Lai, G.P. Xia. 2004. Model and algorithm of an inventory problem with the consideration of transportation cost. *Comput. Ind. Eng.* **46** 389–397.

Zheng, Y.S., A. Federgruen. 1991. Finding optimal (s, S) policies is about as simple as evaluating a single policy. *Oper. Res.* **39**(4) 654–665.

VITA

Liqing Zhang received her Bachelor of Science degree in automation from Tsinghua University, China in 2002 and Master of Science in control theory and control engineering from Tsinghua University in 2005. She joined the Department of Industrial and Systems Engineering at Texas A&M University in Fall 2005 for doctoral studies and graduated with her Ph.D. in August 2011. Her research interests are in supply chain management, inventory theory, stochastic modeling, and optimization.

Liqing Zhang can be reached at:

Liqing Zhang

c/o Sila Çetinkaya

Department of Industrial and Systems Engineering

Texas A&M University

College Station TX 77843-3131