PROFIT MAXIMIZING SHIPMENT CONSOLIDATION WITH UNCERTAIN SHIPMENT REQUESTS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ERGİN ÖZTÜRK

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
INDUSTRIAL ENGINEERING

AUGUST 2022

Approval of the thesis:

PROFIT MAXIMIZING SHIPMENT CONSOLIDATION WITH UNCERTAIN SHIPMENT REQUESTS

submitted by **ERGİN ÖZTÜRK** in partial fulfillment of the requirements for the degree of **Master of Science in Industrial Engineering, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** ________________

Prof. Dr. Esra Karasakal
Head of the Department, **Industrial Engineering** ________________

Assoc. Prof. Dr. Seçil Savaşaneril Tüfekci
Supervisor, **Industrial Engineering, METU** ________________

Assist. Prof. Banu Yüksel Özkaya
Co-Supervisor, **Industrial Eng., Hacettepe University** ________________

Examining Committee Members:

Prof. Dr. Pelin Bayındır
Industrial Engineering, METU ________________

Assoc. Prof. Dr. Seçil Savaşaneril Tüfekci
Industrial Engineering, METU ________________

Assist. Prof. Banu Yüksel Özkaya
Industrial Engineering, Hacettepe University ________________

Prof. Dr. Serhan Duran
Industrial Engineering, METU ________________

Assist. Prof. Önder Bulut
Industrial Engineering, Yaşar University ________________

Date: 31.08.2022

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name Last name : Ergin Öztürk

Signature :

ABSTRACT

PROFIT MAXIMIZING SHIPMENT CONSOLIDATION WITH UNCERTAIN SHIPMENT REQUESTS

Öztürk, Ergin
Master of Science, Industrial Engineering
Supervisor: Assoc. Prof. Dr. Seçil Savaşaneril Tüfekci

Co-Supervisor: Assist. Prof. Banu Yüksel Özkaya

August 2022, 90 pages

In this study, a profit maximizing shipment consolidation problem is under consideration. There are multiple shippers characterized by uncertain shipment requests, who consolidate their orders and make dispatch decisions jointly, in order to maximize total profit. The problem is modeled as a continuous-time Markov Decision Process. For the two-shipper setting the structure of the optimal policy is characterized under certain conditions. For the multiple-shipper setting obtaining the optimal policy is difficult due to the curse of dimensionality. Heuristic policies are proposed and performance of the policies are evaluated.

Keywords: Shipment Consolidation, Markov Decision Process, Heuristics, Policy Iteration Method

ÖZ

RASSAL GÖNDERİ TALEPLERİ ALTINDA KÂRI EN ÇOKLAYAN GÖNDERİ ORTAKLAŞTIRILMASI

Öztürk, Ergin
Yüksek Lisans, Endüstri Mühendisliği
Tez Yöneticisi: Doç. Dr. Seçil Savaşaneril Tüfekci
Ortak Tez Yöneticisi: Dr. Öğr. Üy. Banu Yüksel Özkaya

Ağustos 2022, 90 sayfa

Bu çalışmada kârı en çoklayan gönderi ortaklaşması probleminin değerlendirilmesi amaçlanmaktadır. Gönderi talep zamanları rassal olarak gerçekleşen birden fazla gönderici, toplam kârı ençoklamak amacıyla gönderilerini ortak bir araçta birleştirip gönderi kararını birlikte verir. Problem, Sürekli-Zaman Markov Karar Süreci olarak modellenmiştir. İki göndericili durumda en iyileştirilmiş politika belirli koşullarla karakterize edilir. Daha fazla göndericili durumlarda en iyileştirilmiş politikayı elde etmek büyüyen problem boyutundan dolayı daha zordur. Bu durum için sezgisel politikalar öne sürülüp bu politikaların performansları değerlendirilmektedir.

Anahtar Kelimeler: Gönderi Ortaklaşması, Markov Karar Süreci, Sezgisel Yöntemler, Politika İyileştirme

To my family and friends

ACKNOWLEDGMENTS

I would like to thank my supervisors, Assoc. Prof. Dr. Seçil Savaşaneril Tüfekci and Assist. Prof. Banu Yüksel Özkaya for their continuous guidance and support.

I would like to thank each faculty member of METU IE Department for providing me the required knowledge during my BS and MS years of study.

I would like to express gratitude to my family and friends who have been around during the times of need.

I thank TÜBİTAK BİDEB for providing scholarship under 2210-A program during my years of MS study.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Logistic activities are significant cost items for any business that ships products to its customers. For the outgoing products, outsourcing the logistic activities within the supply chain to third-party logistic providers (3PL) is becoming increasingly common across various industries (Marasco, 2008). Decision of not outsourcing logistic activities to 3PL would require significant investment and additional management effort which is not desirable for most businesses. 3PL companies function as a hub receiving shipments from multiple shippers. Different shipments with the same destination can be consolidated to be sent together.

Shipment consolidation is the action of collecting a number of shipments to utilize the capacity of the freight vehicle to achieve cost savings compared to shipping the loads individually. Utilizing the capacity of the freight vehicle more lowers the freight cost per shipment due to freight cost being shared by more shipments. Ergun et al. (2007) states that repositioning a freight vehicle is very expensive and considering a significant portion of repositioning movements are empty, it was estimated as of 2007 that empty-load related losses are in billions of dollars for the US market only. Considering the increasing desire of companies for better efficiency, shipment consolidation becomes more important.

A single shipper may not benefit the cost efficiency obtained through capacity utilization to the highest extent. This is due to earlier shipment requests waiting a large amount of time until the vehicle is dispatched. When there are delivery deadlines and/or shipment quality depreciation over time, large waiting times for shipments are not desirable. Multiple shippers from the same origin that are sending shipments to the same destination can consolidate their shipments in the same freight vehicle to tackle long waiting times. Collaborating shippers will get to benefit through economies of scale. Logistics service providers tend to offer better prices for collaborating shippers as it decreases the empty movements, and it provides more repetitive jobs for the drivers (Ergun et al., 2007). Such benefits are shared by the collaborating shippers in forms of additional profitability or efficiency.

Besides the cost efficiency achieved by utilizing the capacity, decreasing the CO_2 emission levels caused by transportation industry is an emerging motivation due to increasing concern regarding global warming. An example study is Pan et al. (2013) that considered a transportation pooling problem for retail supply chains in France with the objective of minimizing the CO_2 emissions. It is suggested that pooling the product flows using road and rail transport modes could lead to 52% reduction in CO_2 emissions. Increasing public awareness on the carbon footprint of the products is likely to encourage firms to consider lowering their emission levels to present an environmentally friendly public image.

In this study, we consider a shipment consolidation setting where multiple shippers with same origin and destination collaborate in consolidating their loads on the same freight vehicle. It is aimed to maximize the benefits of the shippers obtained from collaboration by determining load acceptance and vehicle dispatching levels. We characterize the optimal policy structure and propose heuristic policies that are applied to large problem instances. Performance of these heuristics are evaluated and compared with each other and the optimal policy, whenever it was obtainable.

In Chapter 2, we review the literature on related problems and position this study regarding its similarities and differences to other studies. In Chapter 3, we describe the Markov Decision Process (MDP) model for this problem, and we study the optimal policy structure based on the MDP formulation. In Chapter 4, we use the findings on the optimal policy structure and use them for developing heuristic policies to propose solutions for large sized problems. In Chapter 5, we assess the performances of the proposed heuristics. In Chapter 6, we discuss the main findings.

CHAPTER 2

LITERATURE REVIEW

Shipment consolidation (SC) problems aim to optimize the cost tradeoff between fixed cost of dispatching the vehicle with consolidated shipments and cost of holding these shipments until the time of dispatch. This tradeoff also exists in a typical inventory replenishment problem as well with inventory holding costs and fixed cost of order placement.

Suppose a problem environment with constant deterministic arrivals of demand (in an inventory replenishment problem) or shipments (in a SC problem). In a SC cycle, inventory level increases until the dispatch point which is the end of the cycle. Hence, maximum inventory level is observed at the dispatch point. In an inventory replenishment cycle, inventory level begins from the maximum and depletes until the end of the cycle. Decision variable in both problems is the maximum inventory level in a cycle that begins (or ends) with 0 inventory. In calculation of average cost per time unit for both problem settings, fixed cost incurs once per cycle and expected holding cost per time is calculated using the average inventory level in a cycle.

Due to this observation, it can be said that SC problems and inventory replenishment problems study an equivalent decision environment. Extensive literature for both is available with different problem settings such as single/multiple types of loads, deterministic/stochastic arrivals of demand or shipments, inventory capacity limitation etc.

Threshold policies are common in both SC and inventory replenishment literature. Heuristic policies in this study are proposed by coming up with such thresholds and conducting policy improvement procedures over threshold policies.

2.1 Shipment Consolidation (SC) Literature

In this section, papers studying SC problems are reviewed.

Çetinkaya and Bookbinder (2003) studied quantity-based and time-based threshold policies for the dispatch decision for consolidation cycles with the aim of minimizing expected cost per unit time. There is a single type of shipment. Quantity-based threshold dispatches once the accumulated quantity reaches the threshold while time-based threshold dispatches once the time since the beginning of the cycle reaches the threshold. Demand for shipments arrive according to a Poisson process. Problem is modelled as a renewal process. Optimal values for both thresholds are characterized analytically. The setting where a logistics provider offers a quantity discount after a threshold is also studied. Optimal threshold policies are characterized for this setting.

Papadaki and Powell (2007) studied the monotonicity of the value function for MDP's with multidimensional state-space. Monotonicity observations were illustrated on the batch dispatch model where multiple types of products exist and the vehicle has limited capacity. Available decisions are to wait for the next arrival or dispatch an amount from each type and the aim is to minimize expected total discounted cost. Under this setting, it is shown that value function is partially nondecreasing for states with increasing number of loads.

Mutlu et al. (2010) have studied time and quantity (TQ) based policy and compared to the time-based and quantity-based threshold policies within a shipment consolidation scheme with a single type of shipment that arrives according to a Poisson process. TQ-based policy uses both thresholds and chooses to dispatch whenever one of the time-based or quantity-based thresholds are exceeded. The objective is to minimize the expected long run average cost. It is shown that quantity-based policy provides minimum cost. Whenever there is an explicitly stated time limitation for a shipment consolidation, it is suggested that TQ-based policy can be useful and the expressions to find optimal quantity threshold for a given time limit is provided.

Ülkü (2012) has studied shipment consolidation cycle length decision for a multiple shipment type environment with constant deterministic arrival rates. Two optimization models are developed for this problem. The first model aims to maximize the cost difference between employing an immediate shipment policy that ships each period versus the shipment consolidation policy. In the second model, difference between the emission levels of each policy is also considered within the objective. Due to proposed models being non-linear and integer programs, optimal solutions were found through an enumeration procedure. Along with the optimization models, numerical study is made to observe the sensitivity of the optimal solutions with respect to problem parameters.

Yılmaz and Savaşaneril (2012) studied a shipment consolidation problem with multiple shippers consolidating their loads at different dispatch locations. Loads arrive according to a Poisson process. Problem is modeled as an MDP with states being number of loads from each shipper at each dispatch location and decisions being which location the shipper arrivals will be assigned and when to dispatch at each location. Each shipper obtains utility once their loads are dispatched. Following the formulation for the total utility maximization (coalition), shipper strategies are proposed to model the shipper behavior in the absence of coalition. Multiple saving allocation schemes are proposed that would motivate the shippers to collaborate. Computational study is conducted to observe the performances of the proposed shipper strategies and sensitivity of the model parameters.

Lai et al. (2016) have studied quantity-based and time-based threshold policies for the dispatch decision for consolidation cycles with the aim of minimizing expected cost per unit time. Multiple suppliers exist in the environment providing different types of shipments each having different Poisson arrival rates and inventory holding costs. Suppliers consolidate their loads at the warehouse of a third-party logistics provider. Threshold values for quantity and time are found for both dispatch policies using a procedure similar to finding the economic order quantity (EOQ). Cost allocation in cooperative inventory consolidation game is also studied. A rule for proportional cost allocation is proposed. It is found that time-based policy works better with this rule while the performance of the quantity-based policy in a cooperative structure is highly dependent on the number of suppliers and ratio of fixed cost versus holding cost.

Satır et al. (2018) studied a shipment consolidation environment with two types of shipments where one of them being expedited. Shipments arrive according to a Poisson process. Problem is modeled as an MDP where states are the number of loads from each shipment type that are waiting. Decisions are dispatching (all of the accumulated load or a portion) or waiting at a decision epoch. Dispatching incurs a one-time fixed cost while holding costs that differ for each shipment type incur for the loads as they wait. Objective is to minimize the total discounted cost. Optimal threshold policies for capacitated and uncapacitated settings that are quantity-based are observed. In light of these observations, a solution procedure is proposed to be applied to larger sized problems. Through simulation, performance of the policy obtained through proposed solution procedure is found to be better than performances of time-based threshold policies applied by two logistics providers.

2.2 Inventory Replenishment Literature

Among the papers reviewed in this section, only Çetinkaya and Lee (2002) study a problem setting with a single product. Remaining papers are Joint Replenishment Problems (JRP) that study replenishment decisions of multiple types of products.

Çetinkaya and Lee (2002) consider an inventory management problem of a warehouse facing deterministic demand for a product at a constant rate. Demand can be backordered at the warehouse, until a predetermined amount, Q_c, is consolidated. Each time Q_c accumulates, a shipment is made to the customers and inventory at the warehouse is depleted by Q_c, Replenishment of inventory as well as dispatching the Q_c amount, each has a fixed cost. There is a unit cost of keeping inventory, as well as a unit cost of backordering customer demand. Aim is to determine the optimal dispatch and replenishment quantities. Under the restriction that there is a constant dispatch cycle, an EOQ-like optimal dispatch cycle length is found. Under the same restriction, replenishment cycle length is taken as an integer multiple of shipment consolidation cycle length. The problem is also studied under a capacitated setting. An enumerative solution procedure is proposed for this problem, searching the integers for shipment consolidation cycles, and finding the best combination. Main finding is that unequal dispatch cycle lengths of dispatch and replenishment cycles results in lower cost under both capacitated and uncapacitated settings.

Porras & Dekker (2006) have studied joint replenishment setting with minimum order quantities under a deterministic demand rate. An analytical procedure is proposed for finding the base cycle length and its integer multiples for individual stock item types that minimize sum of ordering and inventory keeping costs. Numerical analysis is conducted using data from a real-life case to assess the performance of the proposed solution method. A separate numerical analysis is conducted to observe the sensitivity of the problem to varying parameters.

Moon et al. (2006) has studied a joint replenishment setting with a capital investment restriction that would limit the maximum inventory levels. Replenishment period lengths of individual stock item types are determined as an integer multiple of the base period length that aim to minimize total cost per unit time. The RAND algorithm that previously existed in the literature is modified to be applied to resource constrained joint replenishment problem. A genetic algorithm is also proposed. Numerical analysis is conducted to compare the performances of these algorithms. It is suggested that modified RAND algorithm performed better than the genetic algorithm. However, genetic algorithm has extension ability that makes it useful for constrained joint replenishment problems.

Tanrıkulu et al. (2010) studied a stochastic joint-replenishment problem where the inventory system is modeled as a continuous-time Markov chain. Demand of individual items arrive according to a Poisson process where individual item's demands are independent. Objective is to minimize expected cost per unit time. Optimal policy is found through an enumerative procedure. The (s, Q) policy is proposed where total order size Q is placed and allocated to each individual item whenever inventory level of one of the units drop to s. It is suggested that described (s, Q) policy outperforms the (Q, S) policy where an amount Q is ordered whenever inventory position is below $S - Q$ especially when lead times are small and backordering costs are large. Optimal values were found through a search procedure and performance comparisons were made in numerical analysis.

In Fung et al.'s (2010) study, a joint replenishment environment with compound Poisson demand and positive lead times is considered while having service level constraints to be met. Periodic order-up-to level (T, S) policy is mathematically modeled for multi-product case. The objective is to determine the optimal period lengths and order-up-to levels for individual products that minimizes the expected cost. Heuristic methods are proposed for the solution of this problem that find local optimum values through an enumeration procedure. It is suggested that (T, S) policy outperforms the well-studied can-order policy (s, c, S) especially when lead times are significant.

Salameh et al. (2014) studied joint replenishment problem under a constant deterministic demand rate where two products can be partial substitutes of each other. When inventory level of a product is zero, the other product observes an additional demand rate that is a portion of the finished product's demand. Minimum cost solution procedure by solving non-linear programs is proposed to find the optimal order quantities for both products within this setting. Numerical study is conducted to compare the performance of substitutable product case with no substitution case. It was shown that when product substitution is available, significant cost savings could be achieved.

Feng et al. (2015) studied a multi-product inventory control system where demand arrivals follow a Poisson process. Problem is formulated as an MDP. Demand in a period may consist of multiple products of multiple types. Hence it is observed by a joint pdf of product types within the subset of all types. Optimal inventory control policy is observed for smaller experiments while heuristic procedures are proposed for larger problems. The (s, c, d, S) policy is where a replenishment order is triggered once inventory order of a product type drops to s, types with inventory level less than c $(s < c < d < S)$ are ordered up to a level between d and S. This way, excess inventory holding costs are balanced compared to pure order up to level of S. This is shown by numerical examples and algorithms to find (s, c, d, S) values are presented.

Kouki et al. (2016) considered the replenishment decision for a multi-item inventory control system. Demand follows a Poisson distribution for each item. Items have lifetimes distributed exponentially. Inventory system is modeled as a Markov chain where states denote the inventory level. For zero lead time setting, optimal (s, c, S) values that minimize the expected total cost are characterized for the replenishment policy using a decomposition approach. For positive lead time setting, a heuristic procedure is proposed based on the findings from zero lead time setting. Numerical study is conducted to observe the performance of the proposed heuristic as well as observing sensitivity of the model for varying parameters.

Braglia et al. (2017) has studied a joint-replenishment problem for a supplier dealing with a family of products. Unlike most replenishment studies, lead times and the fixed cost of ordering are controllable. Incentive to decrease the lead times and the fixed cost is balanced by costs associated to these decisions. Only information available regarding the demand distribution is its mean and variance. Objective is to determine the replenishment cycle lengths for each individual product as an integer multiple of the global cycle length as well as the lead times and fixed cost. Under this setting, an algorithm to find the optimal replenishment policy was proposed which was inefficient for large problems. Two heuristic methods are proposed, and numerical study is conducted to compare the performances of proposed algorithms. It was suggested that the heuristic methods could be of use for large size problems.

In the study of Muriel et.al. (2022), joint-replenishment problem is considered in a constant demand environment. Instead of accounting for fixed cost of orders of individual items, a minimum order quantity (MOQ) is used. Multi-item replenishment cycle length that minimizes the cost per unit time is proposed. It is suggested that ordering a constant amount whenever inventory level is zero is suboptimal in MOQ environment. Numerical experiments are conducted to compare the performance of these policies as well as observing the sensitivity of the model to changing parameters.

Creemers & Boute (2022) studied the joint replenishment problem with a base case first and then its extensions with positive lead times, backorders, and compound demands. A Markov chain is developed for exact evaluation of the policies studied based on costs they incur. It is shown that exact optimal policy is only marginally better than can-order policy (s, c, S) with optimal parameters. Using the exact evaluation method, the optimal policy parameters for the can-order policy can be found.

Noblesse et al. (2022) have studied a two-product joint replenishment production-inventory model employing can-order policy (s, c, S) as a markov decision process. Demand arrives according to a compound Poisson process. Lead times for orders are positive and endogenous as they are affected by the order amount. Steady-state distribution of the system is characterized. Using enumerative methods, policy parameters that minimize the costs are found. Scenarios where joint replenishment with can order policy performs better than independent replenishment (s, S) is discussed.

2.3 Dynamic Stochastic Knapsack Literature

Dynamic Stochastic Knapsack Problem (DSKP) is the problem of accepting or rejecting arriving items. Accepted items consume a resource (capacity) and rewards are obtained for the item. Rejecting items can incur a penalty cost that represents the loss of goodwill from the sender. Aim is to maximize the sum of rewards minus the sum of costs.

Stochasticity of the problem is due to the uncertainty of the item arrival times and possibly the rewards. Problem is said to be dynamic when parameters become known at the times of the item arrivals (Range et al., 2018). In the dynamic case, inventory holding costs can be included in the models as time becomes an aspect in decision making with sequential item arrivals. These problems are related with capacity allocation problems with uncertain arrivals in various environments, including transportation problems as in this thesis. Accepting or rejecting random and sequential arrivals with a capacity limitation are common actions in this thesis and DSKP literature.

Kleywegt and Papastavrou (1998) has proposed a generic DSKP model and suggested possible applications of this model to multiple industries. In the description for application of the model to a transportation environment, loads (transportation requests) from different senders arrive in uncertain times. Arrivals are modeled as a Poisson process. Arriving items have random rewards that become known at the arrival. It is assumed that arrivals have the same size (capacity consumption). Decision at each load arrival is to accept or reject the load where a penalty cost exists for rejection. There is a waiting cost per unit per time. Problem is modeled as a continuous-time MDP. Dispatch decision is embedded as a stopping state in the Markov chain. The optimal solution to this problem is characterized as a threshold policy based on the reward and remaining capacity. Calculation method for optimal time of stopping is provided.

Kleywegt and Papastavrou (2001) have extended their study in 1998 where item sizes are random and become known at arrival along with reward parameters. It was found that optimal policy is characterized using thresholds in this scenario as well.

Lu (2018) has studied a DSKP where an initial inventory is depleted over time with arriving demands of random quantity and prices. Demand price and quantity are due to a discrete probability distribution. Once demand arrives and corresponding quantity and price becomes known, seller decides to accept or reject the demand. Structural properties of the value function are studied for unit-arrival case and generalizations are made for the random arrival quantity case. Time and price-based threshold heuristic policies are proposed.

2.4 Rationing Literature

Rationing problems aim to determine the rations of resources allocated to satisfy different classes of demand. Objectives could be cost minimization or profit/benefit maximization. A variety of decisions are studied depending on options in the formulation such as lost sales and backordering. Unlike the knapsack problems, capacity limitation for the stock level may not exist in rationing problems.

Rationing problems that study sequential demand arrivals with uncertain rates are related with this thesis. Such problems are commonly modeled as MDP's with state definition being the stock level at a given time.

Ha (1997a) has studied a production system with a single type of product being produced to stock with a finite rate of production. Unit demand from different customer classes arrive according to a Poisson process. Arriving demand may be accepted or rejected. Rejecting a demand incurs a penalty cost. Production can be continued or stopped. Problem is modeled as an MDP. Aim is to minimize the system cost where cost is incurred for holding inventory. It is shown that a base-stock level is optimal for production decision and thresholds for demand classes exist for accepting (when the threshold is exceeded) or rejecting the demand arrivals.

Ha (1997b) has studied the modified setting of Ha (1997a) where backorders are allowed instead of rejecting demand arrivals. The setting with two demand classes is studied where it is costlier to backorder demands from the first class. It is shown that base-stock level for production decision still holds. Sensitivity of the rationing level is characterized for changing number of backordered demands in place.

Véricourt et al. (2002) has generalized the problem of Ha (1997b) for n classes of demand instead of two. An algorithm to find optimal rationing levels as well as the base-stock level for production is proposed.

Bulut & Fadıloğlu (2011) studied another modification of Ha (1997a) where there can be multiple servers for production instead of one and this number can be increased. It is shown that base-stock policy is optimal for production and threshold policies are optimal for rationing decisions.

2.5 Contribution of This Work

In this section, contribution of this work in terms of the studied problem environment and solution approach is discussed.

Threshold policies are proposed in a number of studies in the reviewed literature for each of the four problem types. In this study, threshold policies for the MDP are proposed following the structural observations on the optimal policy that has shown monotonous behavior. To propose policies for the MDP, a constant arrival rate model is formulated that made use of the monotonicity observations.

In formulating a constant arrival rate model in this study, a global cycle length was determined which is common in JRP models. Process of determining this cycle length follows an EOQ-like solution procedure where average profit per time is maximized. Note that among the reviewed literature for inventory replenishment studies, a profit maximization objective does not exist. Loads of individual shipment types were accepted in a proportion of the global cycle length and individual cycle lengths were embedded into the formulation. Most JRP literature considers an uncapacitated problem environment. In this study, maximum shipments that can be dispatched is limited by the vehicle capacity. Formulations of the optimization models were made accordingly.

In this study, a shipment consolidation setting is studied where there are loads arriving from different shippers each having a rate of arrival. Each shipper gains a unit revenue from each load shipped, which is obtained when the consolidated loads are dispatched. Unit revenue can be interpreted as the sum of additional benefits obtained per unit when the shipper engages in the collaborative shipment consolidation scheme.

In mathematically modeling a shipment consolidation problem, approaches such as representing waiting time related issues as cost incurred per shipment per unit time are used. This is a common approach in the literature where shipments with tighter due dates or shipments that have more importance have a large cost per unit time. This way of modeling the costs discourages long waiting times for such shipments. There is a fixed cost of dispatching the freight vehicle.

It is aimed to maximize the profit of the coalition, that is the sum of profits obtained by individual shippers. In other words, aim is to balance the fixed cost and inventory holding costs with existence of revenue. In this setting, arriving shipments with large costs incurred per unit per time may be rejected. It is assumed that such shipments are sent separately where they do not obtain the additional benefits from the SC scheme which are defined as the unit revenue.

This study differs from the SC literature on different levels. The SC problem in place is modeled under a stochastic shipment arrival setting. Solution approach is based on the MDP formulation and heuristics proposed that make use of the observations on the structure of the optimal policy for the MDP. Main focus on other SC studies were determining the dispatch point that minimizes the costs. In this study, each load brings revenue on dispatch making the objective function a profit maximization. Only other study with a profit maximization objective is Yılmaz and Savaşaneril (2012) that focused on collaborative strategies of the shippers. Yılmaz and Savaşaneril have also used an MDP formulation but considered cross-assignment of the arriving shipments besides the dispatch decision.

In the study of Satır et al. (2018) which also modeled the problem as an MDP, dispatch action may dispatch a portion of the consolidated loads while in this study, the loads are consolidated within the vehicle and all of them will be shipped when the vehicle is dispatched. While Satır et al. have proposed a linear staircase heuristic policy based on the quantity of loads at hand, heuristics in this study propose thresholds based on the accumulated inventory holding cost up to given states.

Moreover, heuristics proposed in this study can be applied to a setting with any number of different shippers.

Besides determining dispatch levels, rejecting the arriving loads is an available option which is also observed in optimal policies in the computational study. Thresholds for rejection action for each type of shipment are proposed along with dispatch threshold in the heuristic methods that were proposed to be used in large sized problems.

Summary of problem setting and solution approach aspects of this study and reviewed literature is presented in Tables 2.1 to 2.3 for the reviewed literature.

Note that abbreviations for the objective function are for cost minimization (CM) and profit maximization (PM). Uncapacitated problem settings are denoted by (U) while capacitated ones are denoted by (C). Stochastic distribution of demand or arrival (depending on the problem type studied) is abbreviated as (S) while deterministic setting is abbreviated as (D).

Table 2.1 Papers studying inventory replenishment problems

Past work	Type of problem	Number of load types	Objective function	Partial or full shipments	Vehicle capacity	Demand or arrival distribution	Solution method
Çetinkaya & Lee (2002)	Rep.	Single	MC	Partial allowed	Both studied	D	Analytical
Porras & Dekker (2006)	JRP	Many	MC	(U)	U	D	Analytical
Moon & Cha (2006)	JRP	Many	MC	Full	C	D	Analytical, heuristic
Tanrıkulu et.al. (2010)	JRP	Many	MC	(U)	U	S (Poisson)	Heuristics
Fung et.al. (2010)	JRP	Many	MC	(U)	U	S (Poisson)	Heuristics
Salameh et.al. (2014)	JRP	Many	MC	(U)	U	D	Non-linear program
Feng et.al. (2015)	JRP	Many	MC	(U)	U	S (Poisson)	MDP, Analytical, Heuristic
Kouki et.al. (2016)	JRP	Many	MC	(U)	U	S (Poisson)	MDP, Analytical, Heuristic
Braglia et.al. (2017)	JRP	Many	MC	(U)	U	S (Distribution free)	Analytical, heuristic
Muriel et.al. (2022)	JRP	Many	MC	(U)	U	D	Analytical
Creemers & Boute (2022)	JRP	Many	MC	(U)	U	S (Poisson)	Analytical, heuristic
Noblesse et.al. (2022)	JRP	Many	MC	(U)	U	S (Poisson)	Analytical, Heuristic
This study	SC	Many	PM	Partial allowed	C	S (Poisson)	MDP, heuristic

Table 2.2 Papers studying shipment consolidation problems

Past work	Type of problem	Number of load types	Objective function	Partial or full shipments	Vehicle capacity	Demand or arrival Distribution	Solution method
Çetinkaya & Bookbinder (2003)	SC	Single	MC	(U)	U	S (Poisson process)	Analytical
Papadaki & Powell (2007)	SC	Many	MC	Partial allowed	C	S (Poisson process)	Analytical, MDP
Mutlu et.al. (2010)	SC	Single	MC	(U)	U	S (Poisson process)	Analytical
Ülkü (2012)	SC	Many	MC	Partial allowed	C	D	Analytical
Yılmaz & Savaşaneril (2012)	SC	Many	PM	Partial allowed	C	S (Poisson process)	MDP
Lai et.al. (2016)	SC	Many	MC	(U)	U	S (Poisson process)	Analytical
Satır et.al. (2018)	SC	Many	MC	Partial allowed	C	S (Poisson process)	MDP, heuristic
This study	SC	Many	PM	Partial allowed	C	S (Poisson process)	MDP, heuristic

Table 2.3 Papers studying DSKP and Rationing (RT) problems

Past work	Type of problem	Number of load types	Objective function	Partial or full shipments	Vehicle capacity	Demand or arrival Distribution	Solution method
Kleywegt & Papastavrou (1998)	DSKP	Many	PM	Partial allowed	C	S (Poisson)	MDP
Kleywegt & Papastavrou (2001)	DSKP	Many	PM	Partial allowed	C	S (Poisson)	MDP
Lu (2018)	DSKP	Single	PM	Partial allowed	C	S (Poisson)	Heuristics
Ha (1997a)	RT	Single	CM	(U)	U	S (Poisson)	MDP
Ha (1997b)	RT	Two	CM	(U)	U	S (Poisson)	MDP
Véricourt et al. (2002)	RT	Many	CM	(U)	U	S (Poisson)	MDP, Heuristics
Bulut & Fadıloğlu (2011)	RT	Single	CM	(U)	U	S (Poisson)	MDP
This study	SC	Many	PM	Partial allowed	C	S (Poisson process)	MDP, heuristic

CHAPTER 3

THE MODEL

In the shipment consolidation setting, there are N different shippers consolidating their shipments in a freight vehicle that can carry up to K loads. Loads from each shipper $i \in \{1, \dots, N\}$ arrive one by one where interarrival times are uncertain. In the literature, uncertain interarrival times are commonly modeled using the exponential distribution (see Yılmaz & Savaşaneril, 2012). Number of arrivals within a time interval follows a Poisson distribution when interarrival times are exponentially distributed. Hence, load arrival process is modeled as a Poisson process with rate λ_i units per time unit.

The vehicle dispatches instantaneously when a "sufficient" amount of load accumulates or when total accumulated loads reach K. When the vehicle is dispatched, all of the accumulated loads are dispatched. Whenever total number of accumulated loads reach K with the new arrival, vehicle dispatches. Note that dispatching a full vehicle may not be an optimal decision as it is forced in this formulation. Arriving loads are directly assigned to the vehicle, not to a warehouse as in Satır et al. (2018) where it was possible to dispatch a portion of available loads. Not considering this option can also cause suboptimality in this formulation. Whenever a vehicle dispatches, a new empty one is assumed immediately available.

It is assumed that shippers obtain benefits for each load shipped when participating in the shipment consolidation setting. Each shipper has a revenue per unit load R_i that represent these benefits. There is a unit holding (waiting) cost per unit per unit time, c_i for each load. Holding cost represents the due dates and/or depreciation factor for the loads. Shipments that have to be dispatched with less waiting time have a large value of holding cost. Dispatching the freight vehicle to the destination has fuel and labor costs. Sum of these costs is represented as fixed cost of dispatching the vehicle denoted by A. The process continues throughout an infinite horizon. Problem parameters are summarized as below.

λ_i: Arrival rate of shipper i in $\frac{units}{time}$

R_i: Revenue per unit load shipped for shipper i in $\frac{monetary\ unit}{unit}$

c_i: Inventory holding cost per unit load per unit time for shipper i in $\frac{monetary\ unit}{unit \times time}$

K: Vehicle capacity in units

A: Fixed cost of dispatching the vehicle in monetary units

In this setting, the aim is to maximize the profit per unit time of N shippers. The profit maximization problem is modeled as a continuous-time Markov Decision Process (MDP) with average reward criterion. Note that all of the described parameters have positive values.

States of the Markov chain are the number of available loads from each shipper in the vehicle denoted by $\vec{x} = (x_1, \dots, x_N)$. Due to Markovian property, time index is dropped, and the states are defined as the number of loads waiting for each shipper. Due to memoryless property, the decisions are made only upon customer arrivals. This is an equivalent definition without loss of optimality. Furthermore, in the MDP, an equivalent discrete-time MDP is constructed by defining the duration of a stage as average time between consecutive arrivals. Lump and continuously accruing costs are defined accordingly (see Serfozo, 1979).

Since the vehicle has to be dispatched every time its capacity is full, the set of feasible states is $\mathbb{S} = \{\vec{x}: \sum_i x_i \leq K - 1\}$. State transitions occur when a new load arrives or when the vehicle is dispatched.

For each state $\vec{x}$, decisions are made regarding the load arrivals and vehicle dispatches. Two types of actions are possible regarding an arriving load from shipper i that are accepting or rejecting the load. When load of shipper i is accepted, x_i in the current state is increased by one. If rejected, current state does not change.

Other two types of actions exist to decide whether the vehicle should wait for the next unit load to arrive or be dispatched with consolidated loads. Wait option incurs the holding costs of the loads consolidated on current state while dispatch option dispatches the consolidated loads earning revenue from each unit load incurring the fixed cost of dispatching and holding costs for consolidated loads (excluding the last arriving load).

The action set consists of combinations of load related and vehicle related decisions. Available actions are Accept&Wait (AW) that adds the arriving load to the vehicle and waits for the next arrival; Reject&Wait (RW) that dismisses the arriving load and waits for the next arrival with the same load composition and Accept&Dispatch (AD) that adds the arriving load to the vehicle and immediately (without incurring additional holding cost for the arriving load) dispatches the vehicle and earns revenue from consolidated loads incurring the fixed cost of dispatching and holding costs. Note that Reject&Dispatch decision is suboptimal as dispatch decision does not incur any additional unit cost for the last arriving load. Reject&Dispatch would not bring any additional revenue for the last arriving load and incurs holding cost for an additional period.

Action for the unit load arrival from shipper i at state $\vec{x}$ is denoted by $a_{i,\vec{x}} \in \{AW, RW, AD\}$. Action vector $\vec{a}_{\vec{x}} = (a_{1,\vec{x}}, \dots, a_{N,\vec{x}}) \in \mathbb{A} = \{AW, RW, AD\}^N$ denotes actions for each shipper's arrival at state $\vec{x}$. Since unit loads of each shipper have different contributions to revenue and holding cost, it is useful to treat their arrivals and actions differently.

The optimality equation of the MDP is as follows.

$$v(\vec{x}) + g = -\frac{\sum_i x_i c_i}{\Lambda}$$

$$+\sum_i \frac{\lambda_i}{\Lambda} max\{v(\vec{x}+e_i), v(\vec{x}), \sum_j R_j x_j + R_i - A + v(\vec{0})\} \quad \forall \vec{x} \in \mathbb{S} \qquad (3.1)$$

In (3.1), $v(\vec{x})$ denotes the bias function for state $\vec{x}$ and g denotes the gain, i.e., profit obtained per state transition (Note since state $\vec{0}$ is reachable under all possible policies, the MDP is well-defined and under optimal policy g and $v(\vec{x})$ exist and are finite). Bias of a state $\vec{x}$ is defined as the total expected reward gained when initial state is $\vec{x}$ relative to a benchmark state, say $\vec{0}$. Parameter $\Lambda = \sum_i \lambda_i$ is the arrival rate of a new load from an arbitrary shipper and $\frac{1}{\Lambda}$ corresponds to the expected time between two consecutive load arrivals. Given that g denotes profit per transaction, profit per unit time is obtained as $\frac{g}{\Lambda}$.

Expected holding cost of spending a transition interval in state $\vec{x}$ is expressed as $-\frac{\sum_i x_i c_i}{\Lambda}$. Let e_i be a 1 by N vector with i^{th} entry equal to 1 and other entries equal to 0. Then, when action for i^{th} shipper's arrival is AW, the next state becomes $\vec{x} + e_i$. If RW action is in place for i^{th} shipper's arrival, next state is still $\vec{x}$. When action for i^{th} shipper's arrival is AD, next state becomes $\vec{0}$, that is a 1 by N vector consisting of zeros corresponding to an empty vehicle.

In contrast to AW and RW actions where no additional cost or revenue incurs except from the inventory holding cost, AD brings revenues from the consolidated loads and the newly arriving loads incurring the fixed dispatch cost besides the holding cost.

Let rew function denote the one stage expected reward for a given state-action pair $(\vec{x}, \vec{a}_{\vec{x}})$ expressed as follows.

$$rew(\vec{x}, \vec{a}_{\vec{x}}) = -\frac{\sum_i x_i c_i}{\Lambda} + \sum_i \delta_{\{a_{i,\vec{x}} = AD\}} \frac{\lambda_i}{\Lambda} \left(\sum_j x_j R_j + R_i - A \right)$$

Where δ is a binary function that is equal to one if the specified condition is satisfied.

One stage transition probabilities are expressed as follows.

$$P(\vec{x}' | \vec{x}, a_{i,\vec{x}}) = \begin{cases} \dfrac{\lambda_i}{\Lambda} & \begin{array}{c} if\ \vec{x}' = \vec{x} + e_i\ and\ a_{i,\vec{x}} = AW \\ or\ \vec{x}' = \vec{x}\ and\ a_{i,\vec{x}} = RW \\ or\ \vec{x}' = \vec{0}\ and\ a_{i,\vec{x}} = AD \end{array} \\ 0 & otherwise \end{cases}$$

Hence, probability that the next arrival being from shipper i is expressed as $\frac{\lambda_i}{\Lambda}$.

3.1 Optimal Policy Structure for the Two-Shipper Uncapacitated Setting

In this section, optimal policy structure is studied to characterize monotonous behavior of optimal actions, if any exists. Such observations are helpful in terms of approximating the optimal policy for large size problems that are costly to solve in computation time. For the simplicity of the analysis the structure of the optimal policy is characterized under discounted reward criterion. We limit the analysis to the two-shipper setting.

Define $V_\beta(i)$ as the expected discounted total reward then the chain starts at state i and when per stage discount factor is β.

Theorem (Ross, 1983, p.95)

If there exists a finite value, say $H < \infty$ such that

$\left|V_\beta(i) - V_\beta(0)\right| < H, \quad \forall \beta, i$ then

- A bounded function $v(i)$ and a constant g exist satisfying the optimality equation

$$v(i) + g = \max_a \left(R(i,a) + \sum_j p_{i,j}(a) v(j) \right)$$

Where $R(i,a)$ is the one-stage reward in state i and action a is taken. Transition probability from state i to j under action a is denoted by $p_{i,j}$. Note that this is a maximization formulation.

- For $\beta_n \to 1$, $v(i) = \lim_{n\to\infty} \left(V_{\beta_n}(i) - V_{\beta_n}(0) \right)$ where n denotes the n^{th} transition. Note that $\beta_n \to 1$ corresponds to no discount.
- $\lim_{n\to\infty} (1 - \beta\,) V_\beta(0) = g$

■

This theorem indicates that the structure of the value function under discounted reward criteria is preserved under average reward criteria. We continue our analysis under discounted reward criteria.

Under discounted reward criterion, one-stage expected reward is expressed for continuous time discount factor α as follows.

$$rew(\vec{x}, \vec{a}_{\vec{x}}) = -\frac{\sum_i x_i c_i}{\Lambda + \alpha} + \sum_i \delta_{\{a_{i,\vec{x}}=AD\}} \frac{\lambda_i}{\Lambda + \alpha}\left(\sum_j x_j R_j + R_i - A\right)$$

Here one may interpret per period discount factor β as $\frac{1}{1+\alpha}$. In other words, as $\alpha \to 0$, $\beta \to 1$ and the setting becomes undiscounted. To apply the discount, following modification is made over transition probabilities.

$$P_\alpha(\vec{x}'|\vec{x}, a_{i,\vec{x}}) = \begin{cases} \dfrac{\lambda_i}{\Lambda + \alpha} & \begin{array}{c} if\ \vec{x}' = \vec{x} + e_i\ and\ a_{i,\vec{x}} = AW \\ or\ \vec{x}' = \vec{x}\ and\ a_{i,\vec{x}} = RW \\ or\ \vec{x}' = \vec{0}\ and\ a_{i,\vec{x}} = AD \end{array} \\ 0 & otherwise \end{cases}$$

Resulting optimality function under total expected discounted reward criteria under discount factor α is as follows.

$$V(\vec{x}) = -\frac{\sum_i x_i c_i}{\Lambda + \alpha} + \sum\nolimits_i \frac{\lambda_i}{\Lambda + \alpha} max\left\{V(\vec{x} + e_i), V(\vec{x}), \sum\nolimits_j R_j x_j + R_i - A + V(\vec{0})\right\}$$

$$\forall \vec{x} \in \mathbb{S} \qquad (3.2)$$

In (3.2), $V(\vec{x})$ is the expected discounted total reward when initial state is $\vec{x}$ (For the simplicity of notation remove the discount index in $V(\vec{x})$ is removed).

Structure of the optimal policy is studied for the two-shipper setting to characterize any monotonous behavior over state-action pairs. Note that findings are not dependent on α and can be generalized to the average reward criterion.

Definition 1. Let $V(x_1, x_2)$ denote the optimal value function as defined in (3.2). $V_{\{a_1,a_2\}}(x_1, x_2)$ is defined as follows.

$$
\begin{aligned}
V_{\{a_1,a_2\}}(\vec{x}) = \quad & -\frac{c_1x_1 + c_2x_2}{\Lambda + \alpha} \\
& + \frac{\lambda_1}{\Lambda + \alpha}\big(V(\vec{x} + e_1)\delta_{\{a_1=AW\}} + V(\vec{x})\delta_{\{a_1=RW\}} \\
& \qquad\qquad + \big(R_1(x_1 + 1) + R_2x_2 - A + V(0,0)\big)\delta_{\{a_1=AD\}}\big) \\
& + \frac{\lambda_2}{\Lambda + \alpha}\big(V(\vec{x} + e_2)\delta_{\{a_2=AW\}} + V(\vec{x})\delta_{\{a_2=RW\}} \\
& \qquad\qquad + \big(R_1x_1 + R_2(x_2 + 1) - A + V(0,0)\big)\delta_{\{a_2=AD\}}\big)
\end{aligned}
$$

■

Lemma 3.1 states that the optimal value function increment between states $\vec{x}$ and $\vec{x} + e_i$ is less than R_i.

Lemma 3.1

$v(x_1 + 1, x_2) - v(x_1, x_2) \leq \frac{\Lambda R_1 - c_1}{\Lambda + \alpha} < R_1, \forall x_1, x_2$ (3.3)

$v(x_1, x_2 + 1) - v(x_1, x_2) \leq \frac{\Lambda R_2 - c_2}{\Lambda + \alpha} < R_2, \forall x_1, x_2$ (3.4)

Proof.

The proof is done by induction: It is assumed that the inequalities hold for the value functions when there are n transitions to go, the inequalities are shown to hold for the value functions when there are $n + 1$ transitions to go. For the sake of simplicity, the index for transition number is dropped.

Suppose the optimal actions at state $(x_1 + 1, x_2)$ are (AD, AW), (Accept and Dispatch when shipper 1 arrives, Accept and Wait when shipper 2 arrives). The following expression holds by the optimality equation in (3.2).

$$
\begin{aligned}
v(x_1 + 1, x_2) = \quad & \frac{-c_1(x_1 + 1) - c_2x_2}{\Lambda + \alpha} \\
& + \frac{\lambda_1}{\Lambda + \alpha}(R_1(x_1 + 2) + R_2x_2 + v(0,0) - A) \\
& + \frac{\lambda_2}{\Lambda + \alpha}v(x_1 + 1, x_2 + 1)
\end{aligned}
$$

Following relation holds by definition of $V(x_1, x_2)$.

$$v(x_1, x_2) \geq v_{\{AD,AW\}}(x_1, x_2)$$

Following relation holds for the differences in the value functions.

$$v(x_1 + 1, x_2) - v_{\{AD,AW\}}(x_1, x_2) \geq v(x_1 + 1, x_2) - v(x_1, x_2)$$

Left hand side of the above relation can be expressed as follows.

$$v(x_1 + 1, x_2) - v_{\{AD,AW\}}(x_1, x_2)$$

$$= \frac{-c_1}{\Lambda + \alpha} + \frac{\lambda_1}{\Lambda + \alpha}\big(R_1(x_1 + 2) + R_2 x_2 + v(0,0) - A - (R_1(x_1 + 1) + R_2 x_2 + v(0,0) - A)\big)$$

$$+ \frac{\lambda_2}{\Lambda + \alpha}\big(v(x_1 + 1, x_2 + 1) - v(x_1, x_2 + 1)\big)$$

$$= \frac{-c_1}{\Lambda + \alpha} + \frac{\lambda_1}{\Lambda + \alpha} R_1 + \frac{\lambda_2}{\Lambda + \alpha}\big(v(x_1 + 1, x_2 + 1) - v(x_1, x_2 + 1)\big)$$

By (3.3) (note that (3.4) is applicable when comparing states $\vec{x}$ and $\vec{x} + e_2$), this expression is less than $\frac{-c_1}{\Lambda+\alpha} + \frac{\lambda_1}{\Lambda+\alpha} R_1 + \frac{\lambda_2}{\Lambda+\alpha} R_1 = \frac{\Lambda R_1 - c_1}{\Lambda+\alpha}$ which is less than R_1. Thus, it can be concluded that below expressions hold true.

$$v(x_1 + 1, x_2) - v(x_1, x_2) \leq v(x_1 + 1, x_2) - v_{\{AD,AW\}}(x_1, x_2) < R_1$$

$$v(x_1 + 1, x_2) - v(x_1, x_2) < R_1$$

When optimal action is other than (AD, AW) under $(x_1 + 1, x_2)$ it is still possible to show that (3.3) holds (the proof follows similar lines and is skipped). For (3.4) the proof follows similar lines.

■

Lemma 3.2 states that the optimal value function increments are nondecreasing with respect to load count of a shipper given the other shippers' load count remains the same.

Lemma 3.2

$v(x_1+1,x_2)-v(x_1,x_2)\geq v(x_1,x_2)-v(x_1-1,x_2)$ where $x_1>0$. (3.5)

$v(x_1,x_2+1)-v(x_1,x_2)\geq v(x_1,x_2)-v(x_1,x_2-1)$ where $x_2>0$. (3.6)

Proof.

Following arrangement can be made for (3.5).

$$v(x_1+1,x_2)-v(x_1,x_2)\geq v(x_1,x_2)-v(x_1-1,x_2)$$

$$v(x_1+1,x_2)+v(x_1-1,x_2)\geq 2v(x_1,x_2)$$

Suppose the optimal action at state (x_1,x_2) is (AW,AD).

$v(x_1+1,x_2)+v(x_1-1,x_2)\geq v_{\{AW,AD\}}(x_1+1,x_2)+v_{\{AW,AD\}}(x_1-1,x_2)$

$=-2\left(\frac{c_1x_1+c_2x_2}{\Lambda+\alpha}\right)+\frac{\lambda_1}{\Lambda+\alpha}\big(v(x_1+2,x_2)+v(x_1,x_2)\big)$

$+\frac{\lambda_2}{\Lambda+\alpha}2(R_1x_1+R_2(x_2+1)+v(0,0)-A)$

$\geq\ -2\left(\frac{c_1x_1+c_2x_2}{\Lambda+\alpha}\right)+\frac{\lambda_1}{\Lambda+\alpha}\big(2v(x_1+1,x_2)\big)$

$+\frac{\lambda_2}{\Lambda+\alpha}2(R_1x_1+R_2(x_2+1)+v(0,0)-A)$

$=2v(x_1,x_2)$

Where the first inequality follows from definition of $V(x_1,x_2)$, the second inequality follows from (3.5) and the last equality follows from the assumption that under (x_1,x_2) optimal action is (AW,AD).

Different action pairs are omitted as same result can be attained using this approach. Note that (3.6) is the symmetrical case for increasing loads from shipper 2 while loads from shipper 1 are constant.

■

Lemma 3.3 states the supermodularity of the optimal value function.

Lemma 3.3

$$v(x_1+1, x_2+1) + v(x_1, x_2) \geq v(x_1+1, x_2) + v(x_1, x_2+1) \qquad (3.7)$$

Proof.

Following relation holds due to Definition 1.

$$v(x_1+1, x_2+1) + v(x_1, x_2) \geq v_{\{AD,AD\}}(x_1+1, x_2+1) + v_{\{RW,RW\}}(x_1, x_2)$$

Suppose the optimal action pairs for states $(x_1+1, x_2), (x_1, x_2+1)$ are (RW, RW) and (AD, AD) respectively. Then (3.7) holds if following relation holds.

$$v_{\{AD,AD\}}(x_1+1, x_2+1) + v_{\{RW,RW\}}(x_1, x_2) \geq v(x_1+1, x_2) + v(x_1, x_2+1)$$

Value functions are extended as follows.

$$v_{\{AD,AD\}}(x_1+1, x_2+1) + v_{\{RW,RW\}}(x_1, x_2)$$

$$= -\left(\frac{c_1(2x_1+1) + c_2(2x_2+1)}{\Lambda+\alpha}\right)$$

$$+\frac{\lambda_1}{\Lambda+\alpha}\big(R_1(x_1+2) + R_2(x_2+1) + v(0,0) - A + v(x_1, x_2)\big)$$

$$+\frac{\lambda_2}{\Lambda+\alpha}\big(R_1(x_1+1) + R_2(x_2+2) + v(0,0) - A + v(x_1, x_2)\big)$$

$$\geq v(x_1+1, x_2) + v(x_1, x_2+1)$$

$$= -\left(\frac{c_1(2x_1+1) + c_2(2x_2+1)}{\Lambda+\alpha}\right)$$

$$+\frac{\lambda_1}{\Lambda+\alpha}(v(x_1+1,x_2)+R_1(x_1+1)+R_2(x_2+1)+v(0,0)-A)$$

$$+\frac{\lambda_2}{\Lambda+\alpha}(v(x_1+1,x_2)+R_1x_1+R_2(x_2+2)+v(0,0)-A)$$

This expression can be arranged as follows.

$$\frac{\lambda_1}{\Lambda+\alpha}\big(R_1+v(x_1,x_2)-v(x_1+1,x_2)\big)$$

$$+\frac{\lambda_2}{\Lambda+\alpha}\big(R_1+v(x_1,x_2)-v(x_1+1,x_2)\big)\geq 0$$

$$R_1\geq v(x_1+1,x_2)-v(x_1,x_2)$$

The inequality holds due to Lemma 3.1. We extend the analysis to the cases where we assume optimal action pairs for states (x_1+1,x_2) and (x_1,x_2+1) are (RW,RW) and (AW,AW) respectively and (AW,AW) and (AD,AD) respectively. Analysis under other possible optimal actions follow similar lines and are thus skipped.

Suppose the optimal action pairs for states $(x_1+1,x_2),(x_1,x_2+1)$ are $(RW,RW),(AW,AW)$ respectively. Again, the following relation holds due to Definition 1.

$$v(x_1+1,x_2+1)+v(x_1,x_2)\geq v_{\{AW,AW\}}(x_1+1,x_2+1)+v_{\{RW,RW\}}(x_1,x_2)$$

(3.7) holds if following relation holds.

$$v_{\{AW,AW\}}(x_1+1,x_2+1)+v_{\{RW,RW\}}(x_1,x_2)\geq v(x_1+1,x_2)+v(x_1,x_2+1) \tag{3.8}$$

Value functions are extended as follows.

$$v_{\{AW,AW\}}(x_1+1,x_2+1)+v_{\{RW,RW\}}(x_1,x_2)$$

$$=\frac{\lambda_1}{\Lambda+\alpha}\big(v(x_1+2,x_2+1)+v(x_1,x_2)\big)+\frac{\lambda_2}{\Lambda+\alpha}\big(v(x_1+1,x_2+2)+v(x_1,x_2)\big)$$

$$\geq v(x_1+1,x_2)+v(x_1,x_2+1)$$

$$=\frac{\lambda_1}{\Lambda+\alpha}\big(v(x_1+1,x_2)+v(x_1+1,x_2+1)\big)$$

$$+\frac{\lambda_2}{\Lambda+\alpha}\big(v(x_1+1,x_2)+v(x_1,x_2+2)\big)$$

This expression can be arranged as follows.

$$\frac{\lambda_1}{\Lambda+\alpha}\big(v(x_1+2,x_2+1)+v(x_1,x_2)-v(x_1+1,x_2)-v(x_1+1,x_2+1)\big)$$

$$+\frac{\lambda_2}{\Lambda+\alpha}\big(v(x_1+1,x_2+2)+v(x_1,x_2)-v(x_1+1,x_2)-v(x_1,x_2+2)\big)\geq 0$$

Note that the inequality holds due to Lemma 3.2 and induction assumption in (3.7). Hence, (3.8) holds.

Next, suppose the optimal action pairs for states $(x_1+1,x_2),(x_1,x_2+1)$ are $(AW,AW),(AD,AD)$ respectively. Following relations hold due to Definition 1.

$$v(x_1+1,x_2+1)+v(x_1,x_2)\geq v_{\{AD,AD\}}(x_1+1,x_2+1)+v_{\{AW,AW\}}(x_1,x_2)$$

(3.7) holds if following relation holds.

$$v_{\{AD,AD\}}(x_1+1,x_2+1)+v_{\{AW,AW\}}(x_1,x_2)\geq v(x_1+1,x_2)+v(x_1,x_2+1) \tag{3.9}$$

Value functions are extended as follows.

$$v_{\{AD,AD\}}(x_1+1,x_2+1)+v_{\{AW,AW\}}(x_1,x_2)$$

$$=\frac{\lambda_1}{\Lambda+\alpha}\big(R_1(x_1+2)+R_2(x_2+1)+v(0,0)-A+v(x_1+1,x_2)\big)$$

$$+\frac{\lambda_2}{\Lambda+\alpha}\big(R_1(x_1+1)+R_2(x_2+2)+v(0,0)-A+v(x_1,x_2+1)\big)$$

$$\geq (x_1+1,x_2)+v(x_1,x_2+1)$$

$$= \frac{\lambda_1}{\Lambda + \alpha}(v(x_1 + 2, x_2) + R_1(x_1 + 1) + R_2(x_2 + 1) + v(0,0) - A)$$

$$+ \frac{\lambda_2}{\Lambda + \alpha}(v(x_1 + 1, x_2 + 1) + R_1 x_1 + R_2(x_2 + 2) + v(0,0) - A)$$

This expression can be arranged as follows.

$$\frac{\lambda_1}{\Lambda + \alpha}\big(R_1 + v(x_1 + 1, x_2) - v(x_1 + 2, x_2)\big)$$

$$+ \frac{\lambda_2}{\Lambda + \alpha}\big(R_1 + v(x_1, x_2 + 1) - v(x_1 + 1, x_2 + 1)\big) \geq 0$$

By Lemma 3.1, the following expressions hold.

$$R_1 + v(x_1 + 1, x_2) - v(x_1 + 2, x_2) > 0$$

$$R_1 + v(x_1, x_2 + 1) - v(x_1 + 1, x_2 + 1) > 0$$

Thus, the inequality (3.9) holds, and it is concluded that the value function is supermodular.

■

Given these lemmas, action pairs are compared to observe monotonous behavior of optimal actions for increasing load levels. It is shown that value function increments are nondecreasing (Lemma 3.2) for increasing loads of one of the shippers and the increment has an upper bound (Lemma 3.1). Supermodularity of the value function (Lemma 3.3) is also shown. Based on these observations, a typical optimal action sequence in the optimal policy is described for states with increasing load levels.

3.1.1 AD vs AW Monotonicity for Increasing Loads

Here it is shown that when upon arrival of shipper i, $i = 1,2$, action AD is preferred over AW in a given state $\vec{x}$, then AD is preferred over AW in state $\vec{x} + e_j$, $j = 1,2$. If it is known that action AD is preferrable to (provides higher value) action AW for an arriving shipment from shipper 1 at (x_1, x_2), following expression holds from the optimality function definition.

$$v(0,0) + R_1(x_1 + 1) + R_2x_2 - A > v(x_1 + 1, x_2) \qquad (3.10)$$

Adding R_1 to both sides of (3.10):

$$v(0,0) + R_1(x_1 + 2) + R_2x_2 - A > v(x_1 + 1, x_2) + R_1$$

From Lemma 3.1, following relation can be observed.

$$v(x_1 + 2, x_2) - v(x_1 + 1, x_2) < R_1$$

Thus, it can be concluded that following relation holds

$$v(0,0) + R_1(x_1 + 2) + R_2x_2 - A > v(x_1 + 1, x_2) + R_1 > v(x_1 + 2, x_2)$$

which implies AD is a better option than AW at state $(x_1 + 1, x_2)$ upon an arrival from shipper 1.

Adding R_2 to both sides of (3.10), following expression holds.

$$v(0,0) + R_1(x_1 + 1) + R_2(x_2 + 1) - A > v(x_1 + 1, x_2) + R_2$$

From Lemma 3.2, following relation can be observed.

$$v(x_1 + 1, x_2 + 1) - v(x_1 + 1, x_2) < R_2$$

Thus, it can be concluded that following relation holds

$$v(0,0) + R_1(x_1 + 2) + R_2x_2 - A > v(x_1 + 1, x_2) + R_2 > v(x_1 + 1, x_2 + 1)$$

which implies AD being a better option than AW at state $(x_1, x_2 + 1)$ upon an arrival from shipper 1.

If AD is a better option than AW for one of the shippers at state (x_1, x_2), AD is a better option than AW for new arrivals from that shipper at states (x_1', x_2') where $x_1' \geq x_1$ and $x_2' \geq x_2$.

3.1.2 AW vs RW Monotonicity for Increasing Loads

Here it is shown that when upon arrival of shipper i, $i = 1,2$, action AW is preferred over RW in a given state $\vec{x}$, then AW is preferred over RW in state $\vec{x} + e_j$, $j = 1,2$.

If it is known that AW action is preferrable compared to RW for an arriving shipment from shipper 1 at (x_1, x_2), following relation holds from the optimality function definition.

$$v(x_1 + 1, x_2) > v(x_1, x_2) \qquad (3.11)$$

By Lemma 3.1 and (3.11), following relation is observed

$$v(x_1 + 2, x_2) - v(x_1 + 1, x_2) \geq v(x_1 + 1, x_2) - v(x_1, x_2) > 0$$

which implies AW is a better action than RW at state $(x_1 + 1, x_2)$ upon an arrival from shipper 1.

By (3.11) and Lemma 3.3, following relations hold

$$v(x_1 + 1, x_2) - v(x_1, x_2) > 0 \Rightarrow v(x_1 + 1, x_2 + 1) - v(x_1, x_2 + 1) > 0$$

$$v(x_1 + 1, x_2 + 1) > v(x_1, x_2 + 1)$$

which implies AW is a better option than RW at state $(x_1, x_2 + 1)$ upon an arrival from shipper 1.

Hence, it can be said that given action AW provides larger value compared to RW at state (x_1, x_2) for an arrival from a given shipper, at states (x_1', x_2') where $x_1' \geq x_1$ and $x_2' \geq x_2$, action AW will provide larger value compared to RW.

3.1.3 AD vs RW Monotonicity for Increasing Loads

If it is known that action AD is preferrable to action RW for an arriving shipment from shipper 1 at (x_1, x_2) for (x_1, x_2) being nonnegative integers, following relation holds from the optimality function definition.

$$v(0,0) + R_1(x_1 + 1) + R_2 x_2 - A > v(x_1, x_2) \tag{3.12}$$

Adding R_1 to both sides of (3.12):

$$v(0,0) + R_1(x_1 + 2) + R_2 x_2 - A > v(x_1, x_2) + R_1$$

From Lemma 3.1, following relation can be observed.

$$v(x_1 + 1, x_2) - v(x_1, x_2) < R_1$$

Thus, it can be concluded that following relation holds

$$v(0,0) + R_1(x_1 + 2) + R_2 x_2 - A > v(x_1, x_2) + R_1 > v(x_1 + 1, x_2)$$

which implies AD is a better option than RW at state $(x_1 + 1, x_2)$ upon an arrival from shipper 1.

Adding R_2 to both sides of (3.12):

$$v(0,0) + R_1(x_1 + 1) + R_2(x_2 + 1) - A > v(x_1, x_2) + R_2$$

From Lemma 3.1, following relation can be observed.

$$v(x_1, x_2 + 1) - v(x_1, x_2) < R_2$$

Thus, it can be concluded that following relation holds

$$v(0,0) + R_1(x_1 + 1) + R_2(x_2 + 1) - A > v(x_1, x_2) + R_2 > v(x_1, x_2 + 1)$$

which implies AD is a better option than RW at state $(x_1, x_2 + 1)$ upon an arrival from shipper 1.

Hence, it can be said that if AD is a better option than RW at a state (x_1, x_2) for one of the shippers, AD is a better option than RW for new arrivals from that shipper at states (x_1', x_2') where $x_1' \geq x_1$ and $x_2' \geq x_2$.

■

Given the monotonicity findings for pairs of actions, typical action sequence observed in the optimal policy is $RW(if\ exists) \rightarrow AW \rightarrow AD$ when observing states for increasing number of loads.

It can be said that once it is optimal to choose AW for arrivals from a shipper at state $\vec{x}$, RW will not be the optimal action for same shipper's arrivals for states $\vec{x}'$ satisfying $x_i' \geq x_i$ for every shipper i.

Same conclusion can be made for AD action compared to both AW and RW. Once it is optimal to choose AD for arrivals from a shipper at state $\vec{x}$, AW or RW will not be optimal actions for same shipper's arrivals for states $\vec{x}'$ satisfying $x_i' \geq x_i$ for every shipper i.

Example.

Example optimal policy for a 2-shipper instance including all three actions is provided in Figure 3.1. Parameter values for this problem are as given below.

$\vec{\lambda} = (5.9, 6.9)$

$\vec{R} = (2.5, 11.7)$

$\vec{c} = (6.5, 9.2)$

$K = 10$

$A = 15$

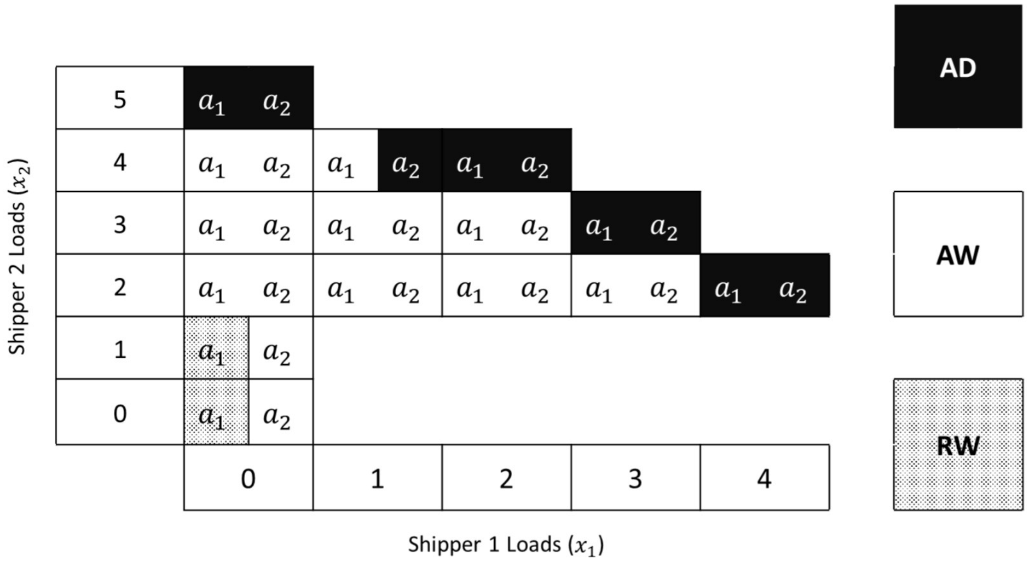

Figure 3.1 Example 2-Shipper optimal policy

Note that states that are not reached due to AD and RW decisions are not shown in Figure 3.1.

These findings gave the insight that thresholds for accepting the loads of different shippers and a threshold for dispatching the vehicle with consolidated loads could be developed to propose heuristic policies. These policies are discussed in the next chapter.

CHAPTER 4

HEURISTIC POLICIES

Due to curse of dimensionality, optimal policies cannot be obtained in reasonable computing times for problem instances with more than three shippers as state space grows. To handle this issue, heuristic policies are proposed based on the observations on the structure of the optimal policy.

First one is an EOQ-based heuristic that uses the solution obtained for a constant arrival rate model (abbreviated as CM) of the problem into developing a threshold policy for the MDP which is called the Workload Threshold Heuristic (WLH). Two other heuristics are also developed based on the CM model, in which the aim is to modify actions in certain states, with an effort to obtain an improved performance. First of these is the policy iteration-based look ahead policy One-Step Policy Improvement Heuristic (PI). Second one is the Whole-State Policy Improvement Heuristic (WPI). Each of these policies modify the WLH policy differently. Finally, a naïve policy is proposed as a benchmark Full-Capacity Dispatch Heuristic (FC) which updates the WLH policy only considering the vehicle capacity.

4.1 Constant Arrival Rate Heuristic

This is a heuristic that initially obtains a solution for a model where discreteness of load arrivals and stochasticity no longer holds. In other words, the loads are assumed to arrive deterministically at a constant rate over time, as in an EOQ model. Let τ denote the time between two dispatch decisions, i.e., the cycle length in this problem. To solve the problem, global dispatch cycle length τ and individual cycle lengths for each shipper that are proportions ($p_i \in [0,1]$) of τ.

Using the cycle lengths and problem parameters, thresholds are found to be used to propose policies for the MDP.

4.1.1 The Model Under a Constant Arrival Rate

The shipment consolidation cycle can be considered as a reverse inventory replenishment cycle where instead of depleting inventory over time, it increases up to a point where the consolidated loads are dispatched, and zero-inventory is reached again.

In shipment consolidation model, maximum inventory is reached at the end of the cycle. Under constant arrival rate, quantities can be continuous, and they increase linearly from the start to the end of a cycle, during τ units of time. Since rejecting arrivals is also an available option, this is integrated to the model by shippers having individual cycle lengths that are a proportion $p_i \in [0,1]$ of the global cycle. Note that this representation is also consistent with the optimal policy under MDP especially because of the observation on the optimal policy structure where shippers are rejected at the beginning of dispatch cycles (states with small number of loads in them). Once large enough states are reached, (later in a dispatch cycle) shippers that were rejected before can also be accepted. This structure is also valid in the constant arrival rate model (CM) by admitting loads from shippers for the last $p_i\tau$ time units of the cycle.

A model where N shippers with constant continuous arrival rates λ_i and unit revenue R_i as well as unit inventory holding costs per time c_i is developed. As in the MDP model, the vehicle has a limited capacity K and fixed cost of dispatching the vehicle is A. As stated earlier, this model relaxes the discreteness and randomness of the arrivals in the MDP model. Objective is to determine the decisions that maximize average total profit per unit time. These decisions are the dispatch cycle length τ and acceptance proportion of τ that is, p_i for every shipper i.

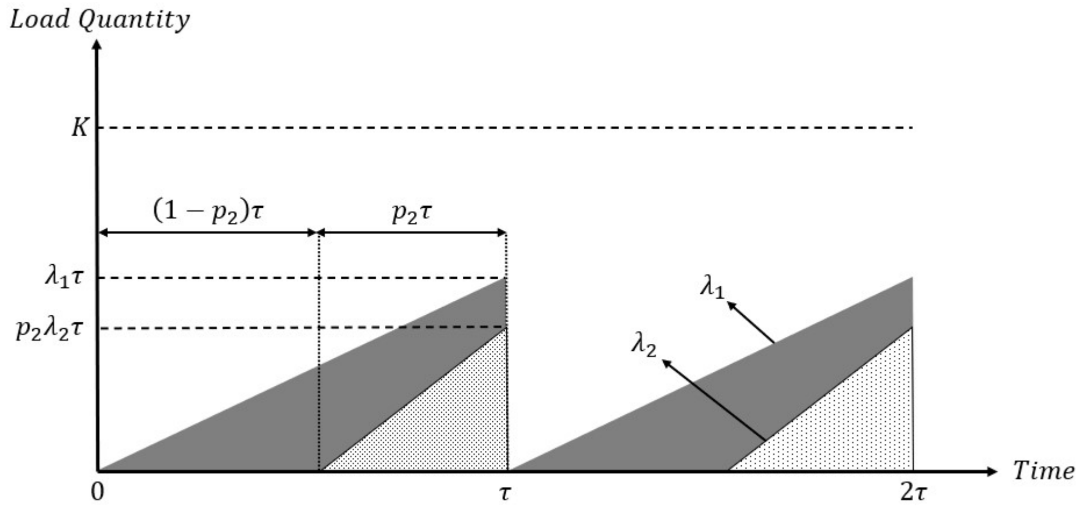

Figure 4.1 Representation of CM with two shippers

In Figure 4.1, a graphical representation of the accumulated loads over time for the CM with two shippers is presented. Note that first shipper is a full shipper ($p_1 = 1$) and second shipper is a partial shipper ($0 < p_2 < 1$) whose loads were rejected at the beginning of each cycle until its acceptance threshold is reached.

Below is the formulation of the problem

$$\text{(CM) } max \qquad \pi(\vec{p},\tau) = -\frac{A}{\tau} + \sum_{i=1}^{N}\left(R_i p_i \lambda_i - \frac{c_i \lambda_i p_i^2 \tau}{2}\right) \tag{4.1}$$

$$\sum_i p_i \tau \lambda_i \leq K \tag{4.2}$$

$$0 \leq p_i \leq 1$$

$$\tau \geq 0$$

where $\vec{p} = (p_1, \dots, p_N)$ is the vector of acceptance proportions (equivalently fraction of time spent in cycle) for each shipper. In the following subsections 4.1.1.1 and 4.1.1.2, optimal solution of CM is studied for uncapacitated and capacitated settings.

4.1.1.1 Uncapacitated Setting

Optimal solution in the absence of the capacity constraint can be determined by finding the optimal p_i^* for a given τ and then finding the optimal τ^*.

Step 1: Determining $\boldsymbol{p_i^*}$ given a cycle length $\boldsymbol{\tau}$

Treating τ as a constant in the profit function π in (4.1) and ignoring the upper bound on p_i for the moment, first order condition (FOC) yields the following expression.

$$\frac{\partial \pi(\vec{p})}{\partial p_i} = R_i\lambda_i - c_i\lambda_i p_i\tau = 0$$

Since second derivative w.r.t. p_i of the profit function with constant τ yields $-c_i\lambda_i\tau$ which is negative, the function is concave with respect to p_i and has a global maximum at p_i satisfying the FOC. Then, optimal p_i^* can be expressed as follows considering its boundaries.

$$R_i\lambda_i = c_i\lambda_i p_i\tau \Rightarrow p_i^* = min\left\{\frac{R_i}{c_i\tau}, 1\right\}$$

This relation ensures p_i^* is a positive value less than or equal to one since problem parameters are positive. Additionally, for every finite τ value, p_i cannot be zero for the uncapacitated setting. Note that, for a given τ, the problem is decomposable with respect to index i. Thus, optimal value of p_i can be determined only considering the terms with index i in (4.1), independent of the other terms.

Shippers with $p_i = 1$ are named as full shippers whose loads are always accepted until the dispatch point is reached. Shippers with $0 < p_i < 1$ are partial shippers whose loads are rejected for the initial $1 - p_i$ proportion of the dispatch cycle beyond which accepted until the dispatch point. A shipper with $p_i = 0$ is a null shipper and its loads are never accepted to the vehicle. Note, in the uncapacitated setting, for a given i, none of the shippers are null shippers under optimality.

Step 2: Determining optimal cycle length τ^*

For $i \in I = \{1, \dots, N\}$ where I denotes the set of shippers, let $F(\tau)$ and $P(\tau)$ be sets of full and partial shippers for a given τ respectively expressed as below.

$$i \in F(\tau) \quad if \quad \frac{R_i}{c_i} \geq \tau$$

$$i \in P(\tau) \quad if \quad \frac{R_i}{c_i} < \tau$$

Plugging in the expression for optimal p_i^* values, the profit function can be expressed as follows in terms of τ only.

$$\pi(\tau) = -\frac{A}{\tau} + \sum_{i \in F(\tau)} \left(R_i \lambda_i - \frac{c_i \lambda_i \tau}{2} \right) + \sum_{i \in P(\tau)} \frac{R_i^2 \lambda_i}{2 c_i \tau}$$

This is a piecewise function as its expression changes for different τ values having different $F(\tau)$ and $P(\tau)$ members. Let $k \in I$ be an arbitrary shipper and $F(\tau)$ and $P(\tau)$ be defined as stated earlier for $i \in I \setminus \{k\}$. Following expressions hold for left and right limits of the profit function at $\tau = \frac{R_k}{c_k}$.

$$\lim_{\tau \to \left(\frac{R_k}{c_k}\right)^-} \pi(\tau) = -\frac{A}{\tau} + \sum_{i \in F(\tau)} \left(R_i \lambda_i - \frac{c_i \lambda_i \tau}{2} \right) + \sum_{i \in P(\tau)} \frac{R_i^2 \lambda_i}{2 c_i \tau} + R_k \lambda_k - \frac{c_k \lambda_k \tau}{2}$$

$$\lim_{\tau \to \left(\frac{R_k}{c_k}\right)^+} \pi(\tau) = -\frac{A}{\tau} + \sum_{i \in F(\tau)} \left(R_i \lambda_i - \frac{c_i \lambda_i \tau}{2} \right) + \sum_{i \in P(\tau)} \frac{R_i^2 \lambda_i}{2 c_i \tau} + \frac{R_k^2 \lambda_k}{2 c_k \tau}$$

Plugging $\tau = \frac{R_k}{c_k}$ in both expressions, the following relation holds.

$$\lim_{\tau \to \left(\frac{R_k}{c_k}\right)^-} \pi(\tau) - \lim_{\tau \to \left(\frac{R_k}{c_k}\right)^+} \pi(\tau) = R_k \lambda_k - \frac{R_k \lambda_k}{2} - \frac{R_k \lambda_k}{2} = 0$$

Since both left and right limits at $\tau = \frac{R_k}{c_k}$ have the same value, $\pi(\tau)$ is continuous at points where members of $F(\tau)$ and $P(\tau)$ change.

Left and right limits of the derivative of $\pi(\tau)$ at $\tau = \frac{R_k}{c_k}$ are observed to check for differentiability.

$$\lim_{\tau \to \left(\frac{R_k}{c_k}\right)^-} \frac{d\pi(\tau)}{d\tau} = \frac{1}{\tau^2}\left(A - \sum_{i \in P(\tau)} \frac{R_i^2 \lambda_i}{2c_i}\right) - \sum_{i \in F(\tau)} \frac{c_i \lambda_i}{2} - \frac{c_k \lambda_k}{2}$$

$$\lim_{\tau \to \left(\frac{R_k}{c_k}\right)^+} \frac{d\pi(\tau)}{d\tau} = \frac{1}{\tau^2}\left(A - \sum_{i \in P(\tau)} \frac{R_i^2 \lambda_i}{2c_i} - \frac{R_k^2 \lambda_k}{2c_k}\right) - \sum_{i \in F(\tau)} \frac{c_i \lambda_i}{2}$$

Plugging $\tau = \frac{R_k}{c_k}$ in both expressions, following relation holds.

$$\lim_{\tau \to \left(\frac{R_k}{c_k}\right)^-} \frac{d\pi(\tau)}{d\tau} - \lim_{\tau \to \left(\frac{R_k}{c_k}\right)^+} \frac{d\pi(\tau)}{d\tau} = -\frac{c_k \lambda_k}{2} + \frac{c_k \lambda_k}{2} = 0$$

Since both left and right limits of the derivative of $\pi(\tau)$ at $\tau = \frac{R_k}{c_k}$ have the same value, $\pi(\tau)$ is differentiable at points where members of $F(\tau)$ and $P(\tau)$ change.

Let shippers $i \in I$ be ordered in increasing $\frac{R_i}{c_i}$ and let $k \in I$ be the shipper that satisfies following relations.

$$\sum_{i=1}^{k-1} \frac{R_i^2 \lambda_i}{2c_i} < A$$

$$\sum_{i=1}^{k} \frac{R_i^2 \lambda_i}{2c_i} \geq A$$

Second derivative of the profit function as follows.

$$\frac{d^2\pi(\tau)}{d\tau^2} = \frac{2}{\tau^3}\left(\sum_{i \in P(\tau)} \frac{R_i^2 \lambda_i}{2c_i} - A\right)$$

Let $\tau_A = \frac{R_k}{c_k}$. Due to the sign of the second derivative, $\pi(\tau)$ is concave for $\tau < \tau_A$ and convex for $\tau \geq \tau_A$. Following observations can be made for the first derivative using definition of τ_A:

$$\frac{d\pi(\tau)}{d\tau} = \frac{1}{\tau^2}\left(A - \sum_{i\in P(\tau)} \frac{R_i^2 \lambda_i}{2c_i}\right) - \sum_{i\in F(\tau)} \frac{c_i \lambda_i}{2} < 0 \quad where\ \tau \geq \tau_A$$

$$\lim_{\tau\to\infty} \frac{d\pi(\tau)}{d\tau} = \frac{1}{\tau^2}\left(A - \sum_{i\in I} \frac{R_i^2 \lambda_i}{2c_i}\right) = 0$$

Hence, profit is decreasing for $\tau > \tau_A$ and $\pi(\tau)$ asymptotically approaches 0 as τ approaches ∞. Hence, there is a global maximum at $\tau^* < \tau_A$ which is the value that makes the first derivative of $\pi(\tau)$ equal to zero. This τ^* is expressed as follows.

$$\tau^* = \sqrt{\frac{2A - \sum_{i\in P(\tau)} \frac{R_i^2 \lambda_i}{c_i}}{\sum_{i\in F(\tau)} c_i \lambda_i}}$$

Note that it is not straightforward to find this value since full and partial shippers depend on τ. A search procedure for finding τ^* is proposed following Lemma 4.1. According to these findings, an example graph for $\pi(\tau)$ can be seen in Figure 4.2.

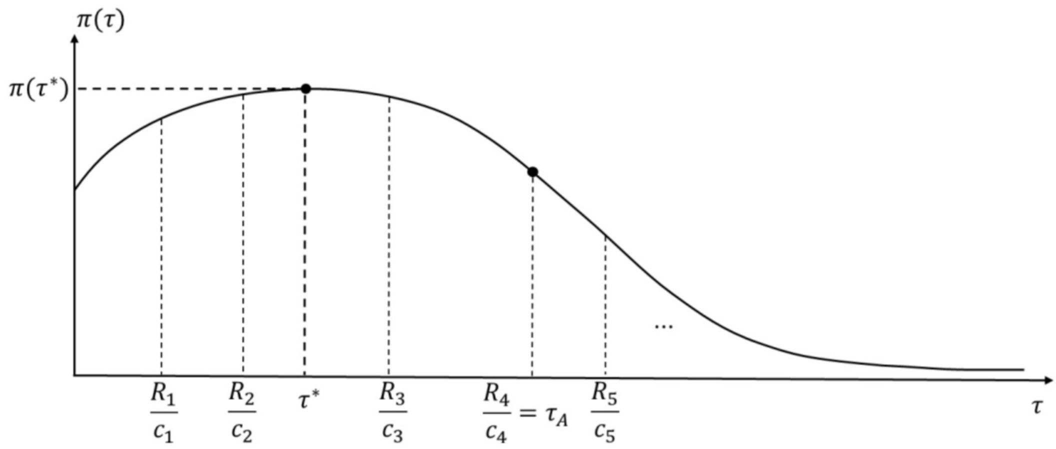

Figure 4.2 Example graph of $\pi(\tau)$ versus τ for a hypothetical setting

Definition 2. Let an interval for τ be called as the i^{th} interval, a convex subset of $\left[0, \frac{R_N}{c_N}\right]$, if $\frac{R_{i-1}}{c_{i-1}} \leq \tau < \frac{R_i}{c_i}$. Note that $\frac{R_0}{c_0} = 0$ is used in this definition and shippers $i \in I$ are ordered in increasing $\frac{R_i}{c_i}$.

Let profit function of an interval be the function where full and partial shippers are selected according to the corresponding τ value within the given interval e.g., $\pi_n(\tau)$ is the profit function where $F(\tau) = \{n, \dots, N\}$ and $P(\tau) = \{1, \dots, n-1\}$. By Definition 2, $\pi_n(\tau)$ is expressed as follows.

$$\pi_n(\tau) = -\frac{A}{\tau} + \sum_{i \geq n} \left(R_i \lambda_i - \frac{c_i \lambda_i \tau}{2} \right) + \sum_{i<n} \frac{R_i^2 \lambda_i}{2 c_i \tau}$$

Let τ_i^* be the optimal cycle length found for the profit function arranged for full and partial members if given τ is in i^{th} interval. For instance, in a setting with 5 total shippers, τ_3^* would be expressed as follows.

$$\tau_3^* = \sqrt{\frac{2A - \sum_{i \in \{1,2\}} \frac{R_i^2 \lambda_i}{c_i}}{\sum_{i \in \{3,4,5\}} c_i \lambda_i}}$$

Note that $\tau_i^* < \tau_j^*$ when $i < j$ by this definition.

Before describing the approach to find τ^*, the Lemma 4.1 is introduced.

Lemma 4.1

Let optimal τ^* lie in the interval, $\frac{R_{j-1}}{c_{j-1}} \leq \tau^* < \frac{R_j}{c_j}$.

For the optimal τ_i^*, following relation holds.

$$\tau_i^* \leq \tau^*, \quad \forall i$$

Proof.

This lemma holds if,

$$\frac{d\pi_j(\tau)}{d\tau} \geq \frac{d\pi_n(\tau)}{d\tau}, \quad \forall\tau \tag{4.3}$$

holds where $\pi_n(\tau)$ is the profit function for n^{th} interval and $\pi_j(\tau)$ is the profit function for optimal interval j. If the optimal τ^* is in n^{th} interval ($j = n$), $\pi_n(\tau) = \pi_j(\tau)$ will hold. (4.3) holds if the profit function of n^{th} interval is maximized at a smaller τ compared to the optimal τ^*. Derivatives of $\pi_n(\tau)$ and $\pi_j(\tau)$ are written as follows.

$$\frac{d\pi_n(\tau)}{d\tau} = \frac{1}{\tau^2}\left(A - \sum_{i=1}^{n-1}\frac{R_i^2\lambda_i}{2c_i}\right) - \sum_{i=n}^{N}\frac{c_i\lambda_i}{2}$$

$$\frac{d\pi_j(\tau)}{d\tau} = \frac{1}{\tau^2}\left(A - \sum_{i=1}^{j-1}\frac{R_i^2\lambda_i}{2c_i}\right) - \sum_{i=j}^{N}\frac{c_i\lambda_i}{2}$$

For $n < j$, difference of derivatives of $\pi_j(\tau)$ and $\pi_n(\tau)$ are written as follows.

$$\frac{d\pi_j(\tau)}{d\tau} - \frac{d\pi_n(\tau)}{d\tau} = -\frac{1}{\tau^2}\sum_{i=n}^{j-1}\frac{R_i^2\lambda_i}{2c_i} + \sum_{i=n}^{j-1}\frac{c_i\lambda_i}{2}$$

$$\frac{d\pi_j(\tau)}{d\tau} - \frac{d\pi_n(\tau)}{d\tau} = \sum_{i=n}^{j-1}\left(\frac{c_i\lambda_i}{2} - \frac{R_i^2}{\tau^2 c_i^2}\frac{c_i\lambda_i}{2}\right) \tag{4.4}$$

Let $\omega_i = \frac{R_i^2}{\tau^2 c_i^2}$ for every i. Note that ω_i decreases for increasing τ. Then, (4.4) can be written as follows.

$$\frac{d\pi_j(\tau)}{d\tau} - \frac{d\pi_n(\tau)}{d\tau} = \sum_{i=n}^{j-1}\frac{c_i\lambda_i}{2}(1 - \omega_i) \tag{4.5}$$

Plugging $\tau = \frac{R_n}{c_n}$ in ω_i for each i, following expression is obtained.

$$\omega_i = \frac{c_n^2 R_i^2}{c_i^2 R_n^2}, \quad \forall i$$

Since $i \in I$ are ordered as described in Definition 2, $\omega_i \leq 1$ for $i \geq n$. Note that the smallest value of τ for $i \geq n$ is $\tau = \frac{R_n}{c_n}$. Hence by (4.5), following relation holds when $n < j$.

$$\frac{d\pi_j(\tau)}{d\tau} - \frac{d\pi_n(\tau)}{d\tau} \geq 0$$

For $n > j$, difference of derivatives of $\pi_j(\tau)$ and $\pi_n(\tau)$ are written as follows.

$$\frac{d\pi_j(\tau)}{d\tau} - \frac{d\pi_n(\tau)}{d\tau} = \frac{1}{\tau^2}\sum_{i=j}^{n-1} \frac{R_i^2 \lambda_i}{2c_i} - \sum_{i=j}^{n-1} \frac{c_i \lambda_i}{2}$$

$$\frac{d\pi_j(\tau)}{d\tau} - \frac{d\pi_n(\tau)}{d\tau} = \sum_{i=j}^{n-1} \left(\frac{R_i^2}{\tau^2 c_i^2} \frac{c_i \lambda_i}{2} - \frac{c_i \lambda_i}{2} \right)$$

This expression can be written using the earlier definition of ω_i as follows. Note that ω_i increases for decreasing τ.

$$\frac{d\pi_j(\tau)}{d\tau} - \frac{d\pi_n(\tau)}{d\tau} = \sum_{i=n}^{j-1} \frac{c_i \lambda_i}{2} (\omega_i - 1) \qquad (4.6)$$

Plugging $\tau = \frac{R_n}{c_n}$ in ω_i for each i, following expression is obtained.

$$\omega_i = \frac{c_n^2 R_i^2}{c_i^2 R_n^2}, \quad \forall i$$

Since $i \in I$ are ordered as described in Definition 2, $\omega_i \geq 1$ for $i \leq n$. Note that the largest value of τ for $i \leq n$ is $\tau = \frac{R_n}{c_n}$. Hence by (4.6), following relation holds when $n > j$.

$$\frac{d\pi_j(\tau)}{d\tau} - \frac{d\pi_n(\tau)}{d\tau} \geq 0$$

Since $n = j$ leads to $\pi_n(\tau) = \pi_j(\tau)$,

$$\frac{d\pi_j(\tau)}{d\tau} - \frac{d\pi_n(\tau)}{d\tau} \geq 0$$

holds for every n and the proof is complete.

■

Knowing that $\pi(\tau)$ is concave for $\tau < \tau_A$ from earlier observations, it can be concluded that there is only one interval i where $\frac{R_{i-1}}{c_{i-1}} \leq \tau_i^* < \frac{R_i}{c_i}$ which will be the global maximum of the profit function τ^*. Lemma 4.1 guarantees that a search procedure over the optimal τ_i^* values for intervals i will not exceed the global maximum τ^*. Knowing that $\tau_i^* < \tau_j^*$ when $i < j$, following search algorithm will converge to the optimal τ^*.

Algorithm 1.

1. Order shippers $i \in I$ in increasing order of $\frac{R_i}{c_i}$ and set $j = 1$
2. If $\sum_{i=1}^{N} \frac{R_i^2 \lambda_i}{2c_i} < A$, then a positive profit solution does not exist and τ^* approaches infinity. Set $p_i^* = 0$ for every shipper and stop. Otherwise proceed to step 3.
3. Find $\tau_j^* = \sqrt{\frac{2A - \sum_{i=0}^{j-1} \frac{R_i^2 \lambda_i}{c_i}}{\sum_{i=j}^{N} c_i \lambda_i}}$
4. Find k satisfying $\frac{R_{k-1}}{c_{k-1}} \leq \tau_j^* < \frac{R_k}{c_k}$

   If $k = j$, stop. $\tau^* = \tau_j^*$, otherwise set $j = k$ and go to 3.

Once τ^* is determined, optimal p_i^*'s can be found using the expression $p_i^* = min\left\{\frac{R_i}{c_i \tau^*}, 1\right\}$.

After obtaining the solution of the CM for uncapacitated setting through Algorithm 1, let variable $Q_i^* = \lambda_i p_i^* \tau^*$ be the expected dispatch quantity for shipper i at the end of the cycle. Then, expected number of loads accumulated right before the vehicle is dispatched can be calculated as $\sum_i Q_i^*$. If $\sum_i {Q_i}^*$ is less than or equal to K, then the optimal solution can be used for the proposed heuristic in section 4.1.2. The heuristic in 4.1.2 uses p_i and τ to obtain a policy to be used in the MDP with thresholds for accepting and dispatching the loads.

If $\sum_i Q_i^*$ exceeds K, then we discuss the possible computational approach in 4.1.1.2 below.

4.1.1.2 Capacitated Setting

Capacity constraint which is Eq (4.2) in (CM) is binding if $\sum_i {Q_i}^*$ in the uncapacitated setting exceeds K. In this case, a solution can be obtained using Karush-Kuhn-Tucker (KKT) optimality conditions. Theorem 9 in Winston (2003, Section 11.9) states the conditions for a solution to be optimal in a maximization problem given as follows.

max $\quad f\big(x = (x_1, \dots, x_n)\big)$

subject to $\quad g_i(x) \le b_i \quad \forall i \in \{1, \dots, m\}$

For a solution $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)$ to be optimal, following expressions must hold with multipliers $\bar{\lambda}_1, \dots, \bar{\lambda}_m$.

$$\frac{\partial f(\bar{x})}{\partial x_j} - \sum_{i=1}^{m} \bar{\lambda}_i \frac{\partial g_i(\bar{x})}{\partial x_j} = 0 \quad \forall j \in \{1, \dots, n\}$$

$$\bar{\lambda}_i\big(b_i - g_i(\bar{x})\big) = 0 \quad \forall i \in \{1, \dots, m\}$$

$$\bar{\lambda}_i \ge 0 \quad \forall i \in \{1, \dots, m\}$$

Let g and g' be the left-hand sides of the constraints in CM where g_i are the $p_i \le 1$ constraints and the capacity constraint while g_i' are the non-negativity constraints. Let the dual variables corresponding to g_i constraints be θ and dual variables corresponding to g_i' constraints be θ'. For the general model with N shippers, KKT conditions are expressed for the following modification as the constraints have to be less or equal to type.

$$(max) \quad \pi(\vec{p},\tau) = -\tfrac{A}{\tau} + \textstyle\sum_{i=1}^{N}\left(R_i\lambda_i p_i - \tfrac{c_i\lambda_i\tau p_i^2}{2}\right) \ \textit{subject to}$$

$$(g_i): p_i \leq 1 \qquad \forall i$$

$$(g_{N+1}): \sum_{i=1}^{N} p_i\lambda_i\tau \leq K$$

$$(g_i'): -p_i \leq 0 \qquad \forall i$$

$$(g_{N+1}'): -\tau \leq 0$$

For this formulation, following expressions have to hold for a solution to be optimal based on the given theorem.

$$\frac{\partial\pi(\vec{p},\tau)}{\partial p_i} - \sum_{j=1}^{N}\theta_j\frac{\partial p_j}{\partial p_i} - \theta_{N+1}\frac{\partial(\sum_{i=1}^{N}p_i\lambda_i\tau)}{\partial p_i} - \sum_{j=1}^{N}\theta_j'\frac{\partial(-p_j)}{\partial p_i} - \theta_{N+1}'\frac{\partial(-\tau)}{\partial p_i} = 0$$

$\forall i \in \{1, \dots, N\}$

$$\frac{\partial\pi(\vec{p},\tau)}{\partial \tau} - \sum_{j=1}^{N}\theta_j\frac{\partial p_j}{\partial \tau} - \theta_{N+1}\frac{\partial(\sum_{i=1}^{N}p_i\lambda_i\tau)}{\partial \tau} - \sum_{j=1}^{N}\theta_j'\frac{\partial(-p_j)}{\partial \tau} - \theta_{N+1}'\frac{\partial(-\tau)}{\partial \tau} = 0$$

$$\theta_i(1-p_i) = 0 \qquad \forall i \in \{1, \dots, N\}$$

$$\theta_{N+1}\left(K - \sum_{i=1}^{N}p_i\lambda_i\tau\right) = 0$$

$$\theta_i' p_i = 0 \qquad \forall i \in \{1, \dots, N\}$$

$$\theta_{N+1}'\tau = 0$$

$$\theta_i \geq 0, \quad \theta_i' \geq 0 \qquad \forall i \in \{1, \dots, N+1\}$$

These expressions can then be simplified to obtain the conditions for N shipper case as follows.

- $R_i\lambda_i - c_i\lambda_i p_i\tau - \theta_i - \theta_{N+1}\lambda_i\tau + \theta_i' = 0 \qquad \forall i \in \{1,\dots,N\}$
- $\frac{A}{\tau^2} - \sum_{i=1}^{N} \frac{c_i\lambda_i p_i^2}{2} - \theta_{N+1}\sum_{i=1}^{N} p_i\lambda_i + \theta_{N+1}' = 0$
- $\theta_i(1-p_i) = 0 \qquad \forall i \in \{1,\dots,N\}$
- $\theta_i' p_i = 0 \qquad \forall i \in \{1,\dots,N\}$
- $\theta_{N+1}(K - \sum_{i=1}^{N} p_i\lambda_i\tau) = 0$
- $\theta_{N+1}'\tau = 0$
- $\theta_i \geq 0, \quad \theta_i' \geq 0 \qquad \forall i \in \{1,\dots,N+1\}$

Dual variables θ_i for $i \in \{1,\dots,N\}$ are positive when corresponding shipper i is a full shipper that is, $p_i = 1$. θ_i' for $i \in \{1,\dots,N\}$ are positive when corresponding shipper i is always rejected that is, $p_i = 0$. Note that θ_{N+1} is positive when capacity constraint is binding. Also note that θ_{N+1}' is always 0 for a feasible solution with a positive τ.

A special case with two shippers was intended to be studied with the aim of generalizing its solution to the model with N shippers. For the problem with $N = 2$ shippers, KKT conditions can be written in open form as follows. Note that $\theta_3 > 0$ always holds for solutions where capacity constraint is binding in 2-shipper case.

$$R_1\lambda_1 - c_1\lambda_1 p_1\tau - \theta_1 - \theta_3\lambda_1\tau + \theta_1' = 0 \tag{4.7}$$

$$R_2\lambda_2 - c_2\lambda_2 p_2\tau - \theta_2 - \theta_3\lambda_2\tau + \theta_2' = 0 \tag{4.8}$$

$$\frac{A}{\tau^2} - \frac{c_1\lambda_1 p_1^2}{2} - \frac{c_2\lambda_2 p_2^2}{2} - \theta_3(p_1\lambda_1 + p_2\lambda_2) + \theta_3' = 0 \tag{4.9}$$

$$\theta_3(K - p_1\lambda_1\tau - p_2\lambda_2\tau) = 0 \tag{4.10}$$

$$\theta_i(1-p_i) = 0 \qquad \forall i \in \{1,2\} \tag{4.11}$$

$$\theta_i' p_i = 0 \qquad \forall i \in \{1,2\} \tag{4.12}$$

$$\theta_3'\tau = 0 \tag{4.13}$$

$$\theta_i \geq 0, \quad \theta_i' \geq 0 \qquad \forall i \in \{1,2,3\}$$

Motivation of observing this special case is to come up with parametric conditions that would make it possible to determine the resulting p_i's and τ, without resorting to computational methods. There are several possible cases where different constraints are binding leading to different multipliers being positive such as existence of a zero-shipper (say shipper i) making $\theta_i' > 0$ or a full shipper making $\theta_i > 0$.

Note that for any solution where capacity constraint is binding, θ_3 is positive by (4.10) and for a positive τ value, θ_3' is always zero by (4.13).

In the following, we only present three possible cases of the analysis. Due to the complexity of the analysis, the remaining cases are left as future work.

Case 1: $\boldsymbol{p_1, p_2 = 1}$

For this case, following conditions must hold.

$$\theta_1, \theta_2, \theta_3 > 0$$

$$\theta_1', \theta_2' = 0$$

Value of τ is obtained by (4.10):

$$\tau = \frac{K}{\lambda_1 + \lambda_2}$$

Conditions for $\theta_1, \theta_2, \theta_3$ are obtained by solving equations (4.7), (4.8), and (4.9) together using the expression for τ.

$$\theta_1 = R_1\lambda_1 - \frac{A\lambda_1}{K} - \frac{K}{2}\left(c_1 - \frac{\lambda_2(c_1\lambda_2 + c_2\lambda_1)}{(\lambda_1 + \lambda_2)^2}\right) > 0$$

$$\theta_2 = R_2\lambda_2 - \frac{A\lambda_2}{K} - \frac{K}{2}\left(c_2 - \frac{\lambda_1(c_1\lambda_2 + c_2\lambda_1)}{(\lambda_1 + \lambda_2)^2}\right) > 0$$

$$\theta_3 = \frac{A(\lambda_1+\lambda_2)}{K^2} - \frac{1}{2}\left(c_1\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right) + c_2\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)\right) > 0$$

Case 2: $\boldsymbol{p_1 \in (0,1), p_2 = 1}$

Note that $p_2 \in (0,1)$, $p_1 = 1$ is the symmetrical case and the indexes of the given expressions can be reversed to find the corresponding conditions. Following conditions must hold for this case.

$$\theta_2, \theta_3 > 0$$

$$\theta_1, \theta_1', \theta_2' = 0$$

Value of τ is obtained by (4.10).

$$\tau = \frac{K}{\lambda_1 p_1 + \lambda_2}$$

Condition for partial shipper $p_1 \in (0,1)$, condition for θ_2 corresponding to the full shipper, and condition for θ_3 are obtained by solving equations (4.7), (4.8), and (4.9) together using the expression for τ.

$$0 < p_1 = \frac{-\lambda_2 + K\sqrt{\dfrac{\lambda_2(c_1\lambda_2 + c_2\lambda_1)}{c_1K^2 - 2R_1\lambda_1K + 2A\lambda_1}}}{\lambda_1} < 1$$

$$\theta_2 = \frac{KR_1\lambda_2 - 2A\lambda_2 + KR_2\lambda_2}{K} + \frac{\left(\lambda_2 - K\sqrt{\dfrac{\lambda_2(c_1\lambda_2 + c_2\lambda_1)}{c_1K^2 - 2R_1\lambda_1K + 2A\lambda_1}}\right)(c_1K^2 - 2R_1\lambda_1K + 2A\lambda_1)}{\lambda_1 K} > 0$$

$$\theta_3 = \frac{c_1\lambda_2}{\lambda_1} + \left(R_1 - \frac{c_1K}{\lambda_1}\right)\sqrt{\frac{\lambda_2(c_1\lambda_2 + c_2\lambda_1)}{c_1K^2 - 2R_1\lambda_1K + 2A\lambda_1}} > 0$$

Case 3: $\boldsymbol{p_1 = 0, p_2 = 1}$

Note that $p_2 = 0$, $p_1 = 1$ is the symmetrical case and the indexes of the given expressions can be reversed to find the corresponding conditions.

$$\theta_1', \theta_2, \theta_3 > 0$$

$$\theta_1, \theta_2' = 0$$

Value of τ is obtained by (4.10).

$$\tau = \frac{K}{\lambda_2}$$

Conditions of θ_1' for the null shipper, θ_2 for the full shipper, and condition for θ_3 are obtained by solving equations (4.7), (4.8), and (4.9) together using the expression for τ.

$$\theta_1' = \frac{A\lambda_1}{K} - \frac{c_2\lambda_1K}{2\lambda_2} - R_1\lambda_1 > 0$$

$$\theta_2 = R_2\lambda_2 - \frac{c_2K}{2} - \frac{A\lambda_2}{K} > 0$$

$$\theta_3 = \frac{A\lambda_2}{K^2} - \frac{c_2}{2} > 0$$

■

Due to complexity of the equation system under cases where $0 < p_i < 1$ for both shippers, and $p_1 = 0, 0 < p_2 < 1$, analysis of these cases is left as future work.

Characterization of optimal policy under capacitated setting when the number of shippers is more than 2 is not further pursued due to the complexity of the problem. Optimal solution is obtained via off-the-shelf solvers.

4.1.2 Workload Threshold Heuristic (WLH)

WLH is a heuristic that uses p_i's and τ obtained from CM to propose a policy for the MDP. It is shown that under optimal policy in uncapacitated stochastic unit load problem, there exist thresholds, in terms of the loads in the vehicle, that define when to accept a load of an arriving customer and when to reject the load. On the other hand, dispatch threshold determines whether the current state is eligible to be dispatched at the time of next arrival. Thus, thresholds indicate upon arrival of each shipper which action is to be taken in a given state. Even though existence of such thresholds were not shown to be optimal under capacitated setting in MDP, it is conjectured that similar threshold structure exists. In WLH heuristics, these thresholds of MDP are derived from the optimal policy of CM.

Inventory holding cost accumulated at the end of the cycle (cycle of length τ units of time) for CM is calculated to be used as the dispatch workload cost threshold $wl_{dispatch}$. This cost level is used as the threshold cost level that would trigger the dispatch of the vehicle in MDP.

Suppose $Q_i, p_i, \forall i$ and τ values are obtained to optimality in CM. Then,

$$wl_{dispatch} = \sum_{i \in \{1,2,\dots,N\}} \frac{Q_i c_i p_i \tau}{2} = \sum_{i \in \{1,2,\dots,N\}} \frac{\lambda_i c_i p_i^2 \tau^2}{2} = \sum_{i \in \{1,2,\dots,N\}} \frac{Q_i^2 c_i}{2\lambda_i}$$

This threshold is used in the MDP by comparing cost weighted workload $wp_{\vec{x}}$ of each state to $wl_{dispatch}$. Cost weighted workload at state $\vec{x}$ is calculated as follows:

$$wp_{\vec{x}} = \sum_{\forall i \in \{1,2,\dots,N\}} \frac{x_i(x_i+1)c_i}{2\lambda_i}$$

Suppose $p_i < 1$ for shipper i under CM. This implies loads of the shipper are not accepted until $(1-p_i)\tau$ time units into a cycle. At the time point where loads of shipper i are started to be accepted, the expected accumulated holding cost level is calculated for the shippers with $p_j > p_i$.

$$wl_i = \sum_{j \in \{1,2,\dots,N\}} \delta_{\{p_j > p_i\}} \frac{\lambda_j c_j \left(p_j \tau - p_i \tau\right)^2}{2} \qquad \forall i$$

Note that equivalently, once this threshold is exceeded, corresponding shipper's loads can be accepted. Note that a full shipper m ($p_m = 1$) will have $wl_m = 0$ while a null shipper i (with $p_i = 0$), has $wl_{dispatch} = wl_i$. Smaller values of wl_i indicate that the shipper is more prioritized.

In Figure 4.3, calculation method of wl_i is visualized. This figure represents a 2-shipper case where $p_1 = 1$ is the full shipper and $0 < p_2 < 1$ is the partial shipper. Loads from shipper 2 are accepted after $(1 - p_2)\tau$ units of time within a cycle. At time $(1 - p_2)\tau$, there are $\lambda_1(1 - p_2)\tau$ shipments from shipper 1 in the system. Hence, inventory holding cost incurred until time $(1 - p_2)\tau$ which is wl_2 is found as follows.

$$wl_2 = \frac{1}{2} c_1 \lambda_1 (1 - p_2)\tau(1 - p_2)\tau = \frac{c_1 \lambda_1 (1 - p_2)^2 \tau^2}{2}$$

Note that if another shipper exists in the system satisfying $p_3 > p_2$, area of another triangle multiplied by c_3 will be added to wl_2.

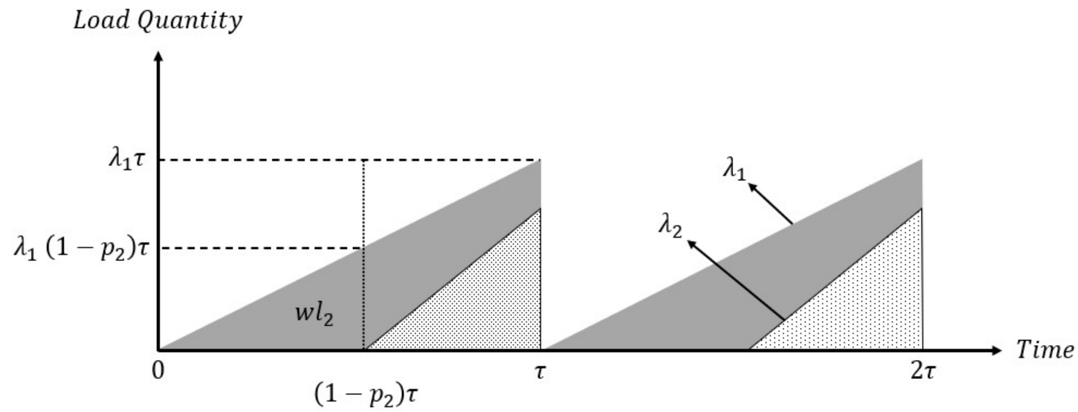

Figure 4.3 Representation of workload threshold for acceptance measure

Acceptance threshold is used in the MDP by comparing cost weighted workload $wp_{i,\vec{x}}$ of each shipper at each state to wl_i. Cost weighted workload for acceptance of i^{th} shipper at state $\vec{x}$ is calculated as follows.

$$wp_{i,\vec{x}} = \sum_{j\in\{1,2,\dots,N\}} \delta\{wl_j < wl_i \ and \ p_j = 1\} \frac{x_j(x_j+1)c_j}{2\lambda_j} \qquad \forall i$$

This way, a partial shipper is rejected until enough full shippers are accepted. Note that cost weighted workload calculation is limited to full shippers due to the observation that $wl_j < wl_i$ condition alone leads to longer rejection periods that result in less profitable solutions. Following these calculations, WLH policy is executed using procedure explained below.

For an arrival from i^{th} shipper at any state $\vec{x}$ where $\sum_i x_i \le K-1$:

- If $wp_{i,\vec{x}} < wl_i$ assign action $a_{i,\vec{x}}^{WLH} = RW$
- Else if $wp_{i,\vec{x}} \ge wl_i$ and ($wp_{\vec{x}} \ge wl_{dispatch}$ or $\sum_{x_j\in\vec{x}} x_j = K-1$), $a_{i,\vec{x}}^{WLH} = AD$
- Otherwise $a_{i,\vec{x}}^{WLH} = AW$

Actions are assigned for every shipper's arrival for every state. Note that any shipper i satisfying $wp_{i,\vec{x}} \ge wl_i$ will have AD action for its arrivals if $wp_{\vec{x}} \ge wl_{dispatch}$ or $\sum_{x_j\in\vec{x}} x_j = K-1$. This disallows AD and AW actions for different shipper's arrivals in the same state in a WLH policy.

Another possible threshold method that can be utilized is a quantity-based threshold that uses wl_i and $wl_{dispatch}$ in terms of load quantities. Computational results have shown that using cost weighted thresholds is better in terms of expected profit.

4.2 One-Step Policy Improvement Heuristic (PI)

Following observations after solving numerical examples, it was seen that WLH policy underestimates the dispatch points (by dispatching in states with less loads) compared to the optimal policy. One-step policy improvement procedure (PI) over WLH is applied to obtain an improved policy.

Tijms (1995, p.193) describes the three-step policy iteration algorithm as follows.

Step 0: Choose an initial policy, say R.

Step 1: Find the solution (v^R, g_R) for policy R solving the following system of equations.

$v_i^R + g_R = c_i(R_i) + \sum_{j \in I} p_{i,j}(R_i) v_j^R$, $\forall i \in I$ where I is the state space.

$v_0 = 0$ where state $0 \in I$ is an arbitrary reference state.

Note that $c_i(R_i)$ in this notation refers to one-stage expected cost of employing policy R in state i and this is a minimization setting.

Step 2: Choose best actions for each state to obtain improved policy R'.

$$a^{R'}(i) = \underset{a \in A(i)}{argmin} \left\{ c_i(a) - g_R + \sum_{j \in I} p_{i,j}(a) v_j^R \right\}, \quad \forall i \in I$$

Step 3: If $R' = R$, stop with improved policy being R'. Otherwise set $R = R'$ and go to Step 1.

In the PI heuristic, starting with the initial policy WLH, Step 1 and Step 2 of the policy iteration algorithm is executed. Step 2 is executed over a limited subset of states (to be defined as $\mathbb{S}_{(AD)}$) to obtain an improved policy. This limited subset is defined as follows. The WLH is conjectured to result in a policy that takes AD decisions in states with smaller number of loads compared to the optimal solution (due to the computational observations). Since it is the states with AD action that needs improvement, only those AD states under the WLH policy are considered. For those states, it is checked whether AW improves the value, if feasible. The feasible states are the ones where number of loads in the vehicle are at most $K-2$. Let $\mathbb{S}_{(AD)}$ denote the specified subset of states.

$$\mathbb{S}_{(AD)} = \left\{ \vec{x}: \sum_j x_j < K-1 \wedge \sum_j \delta_{\{a_{j,\vec{x}}=AD\}} > 0 \right\}$$

At Step 0, WLH policy is obtained as the initial policy to be used in PI.

PI is only applied to the states in $\mathbb{S}_{(AD)}$. Due to the property of the WLH policy that it can allow a mixture (both actions exist in given states for different arrivals) of AD and RW decisions in a given state (but not AD and AW in the same state), following expression is applicable for the value equations of states in $\mathbb{S}_{(AD)}$. Note that term $v(\vec{0})$ is omitted since its value is close to 0.

$$v(\vec{x}) + g = -\frac{\sum_i x_i c_i}{\Lambda} + \sum_i \left(\frac{\lambda_i v(\vec{x}) \delta_{\{a_{i,\vec{x}}=RW\}}}{\Lambda} + \frac{\lambda_i \left(\sum_j R_j x_j + R_i - A\right) \delta_{\{a_{i,\vec{x}}=AD\}}}{\Lambda} \right)$$

This expression is arranged for $v(\vec{x})$ as follows.

$$v(\vec{x}) = \frac{\sum_i \lambda_i \delta_{\{a_{i,\vec{x}}=AD\}} \left(\sum_j R_j x_j + R_i - A\right)}{\Lambda - \sum_i \lambda_i \delta_{\{a_{i,\vec{x}}=RW\}}} - \frac{\sum_{\forall i} x_i c_i}{\Lambda - \sum_i \lambda_i \delta_{\{a_{i,\vec{x}}=RW\}}} - \frac{g\Lambda}{\Lambda - \sum_i \lambda_i \delta_{\{a_{i,\vec{x}}=RW\}}} \quad (4.14)$$

In these calculations, an approximate value of g is used which is the average profit per unit time value found for the constant demand rate model used in finding WLH policy. Expression for finding the g value used is as follows.

$$g = \frac{\pi(\vec{p}, \tau)}{\Lambda}$$

Where $\pi(\vec{p}, \tau)$ is the CM objective function from Eq. (4.1).

Policy improvement step is made over the desired subset of states as follows.

$$a_{i,\vec{x}}^{PI} = \underset{a_{i,\vec{x}}^{PI} \in \{AW, RW, AD\}}{argmax} \left\{ v(\vec{x} + e_i), v(\vec{x}), \sum_{j \in \{1,\ldots,N\}} R_j x_j + R_i - A \right\} \quad (4.15)$$

$$\forall i \in \{1, \dots, N\}, \forall \vec{x} \in \mathbb{S}_{(AD)}$$

State action pairs obtained as $a_{i,\vec{x}}^{PI}$ form the PI heuristic policy. Note that there is no requirement to solve a system of equations since g is estimated.

4.3 Whole-State Policy Improvement Heuristic (WPI)

WPI is another modification to the WLH that aims to increase the number of states in which AW action is taken. This heuristic adjusts the WLH using the following procedure:

1. Let for shipper i, $x_{i,max}^{WLH}$ be defined as the maximum total load value in which AW action is taken upon arrival of shipper i under WLH.

$$x_{i,max}^{WLH} = max \left\{ \sum_{j \in \{1,..,N\}} x_j : a_{i,\vec{x}}^{WLH} = AW \right\}$$

2. For each shipper i, identify all states that have a total number of loads less than $x_{i,max}^{WLH}$. For shipper i, consider the actions taken at those states under WLH. (Note those actions would be RW, AW and AD). Identify the border states for shipper i, those that satisfy the relation $a_{i,\vec{x}}^{WLH} = AD$ and $a_{i,\vec{x}-e_i}^{WLH} = AW$.
3. Assign $a_{i,\vec{x}}^{WPI} = a_{i,\vec{x}}^{WLH}$ for states that are not border states.
4. Let $\mathfrak{B}$ denote the set of border states under WLH.
5. For $\vec{x} \in \mathfrak{B}$, calculate

$$cond = -\frac{\sum_j c_j x_j}{\Lambda} - \frac{\sum_j \lambda_j c_j}{\Lambda^2} + \frac{\sum_j R_j \lambda_j^2 + \sum_j \sum_{k \in \{1,\dots,N\} \setminus \{j\}} \lambda_j \lambda_k (R_j + R_k)}{\Lambda^2} \quad (4.16)$$

Evaluating $v(\vec{x})$ assuming that AD action is taken for all shipper arrivals:

$$v_{AD}(\vec{x}) = g - \frac{\sum_j c_j x_j}{\Lambda} + \sum_j R_j x_j + \frac{\sum_j \lambda_j R_j}{\Lambda} - A + v(\vec{0})$$

Suppose at given state $\vec{x}$, upon any shipper arrival AW decision is taken (regardless of the type of the shipper), and in the following state $(\vec{x} + e_i)$, upon any arrival AD decision is taken. Let AW' denote these successive actions. Then the value function would be expressed as follows.

$$v_{AW'}(\vec{x}) = g - \frac{2\sum_j c_j x_j}{\Lambda} - \frac{\sum_j \lambda_j c_j}{\Lambda^2} + \sum_j R_j x_j + \frac{\sum_i \sum_j \lambda_i \lambda_j (R_i + R_j)}{\Lambda^2} - A + v(\vec{0})$$

Comparing the value functions under these two approaches: upon any shipper arrival take AD vs. upon any shipper arrival take AW and then in the following state, upon any arrival take AD:

$$\begin{aligned}
v_{AW'}(\vec{x}) - v_{AD}(\vec{x}) = {} & -\frac{\sum_j c_j x_j}{\Lambda} - \frac{\sum_j \lambda_j c_j}{\Lambda^2} + \frac{\sum_i \sum_j \lambda_i \lambda_j (R_i + R_j) - \Lambda \sum_j \lambda_j R_j}{\Lambda^2} \\
= {} & -\frac{\sum_j c_j x_j}{\Lambda} - \frac{\sum_j \lambda_j c_j}{\Lambda^2} + \frac{\sum_i \sum_j \lambda_i \lambda_j (R_i + R_j) - \sum_i \sum_j \lambda_i \lambda_j R_j}{\Lambda^2} \\
= {} & -\frac{\sum_j c_j x_j}{\Lambda} - \frac{\sum_j \lambda_j c_j}{\Lambda^2} + \frac{\sum_i \sum_j \lambda_i \lambda_j R_i}{\Lambda^2} \\
= {} & -\frac{\sum_j c_j x_j}{\Lambda} - \frac{\sum_j \lambda_j c_j}{\Lambda^2} \\
& + \frac{\sum_j \lambda_j^2 R_j + \sum_j \sum_{k \in \{1,\dots,N\} \setminus \{j\}} \lambda_j \lambda_k (R_j + R_k)}{\Lambda^2}
\end{aligned}$$

We name the expression on RHS as $cond$. Note that $cond$ is a value independent of the arriving shipper.

- If $cond \geq 0$, for all shipper i arrival, assign $a_{i,\vec{x}}^{WPI} = AW$. Update the set of border states by removing $\vec{x}$, and for all i adding $\vec{x} + e_i$ to $\mathfrak{B}$.
- Otherwise, do not change $a_{i,\vec{x}}^{WPI}$, remove $\vec{x}$ from $\mathfrak{B}$.

Repeat Step 5 until $\mathfrak{B}$ becomes empty.

■

Following this procedure, $a_{i,\vec{x}}^{WPI}$ actions are obtained for every state for each shipper i arrival.

Note that the number of states that possibly assume different actions under WPI compared to WLH is bounded by $x_{i,max}$ for each arrival i. The reason for a bounded search is the empirical observations which have shown that WPI policy tends to choose AD at states with larger number of loads compared optimal policy in the absence of this bound. In other words, an unbounded search leads to WPI overusing the capacity compared to the optimal policy.

4.4 Full-Capacity Dispatch Heuristic (FC)

This heuristic forces the vehicle capacity to be fully utilized in every dispatch cycle while keeping the RW decisions (if any exists) as they are in the WLH policy.

FC can also be interpreted as a method that extends the AW decisions in WLH policy to states with larger number of loads. In contrast to the heuristics discussed earlier, problem parameters are not utilized. Instead, AD decision for every shipper's arrival within states $\vec{x} \in \mathbb{S}_{(AD)}$ is replaced by AW decision.

CHAPTER 5

COMPUTATIONAL RESULTS

In this chapter, details and the results of computational experiments are presented. These experiments are conducted to evaluate and to compare the performance of the proposed policies.

Recall that problem size is sensitive to the vehicle capacity K and number of shippers N as number of possible states increase for increasing N and K. Along with different values of K and N, experiment instances are generated according to the details given below.

Experiments are run for $N = \{2,3,4,5,6,12\}$, in total six different number of shipper types are tested. Under each number of shipper types, 100 random instances are generated. Distributions to randomize the parameters are chosen in a way that would allow existence of fully and/or partially rejected shippers in the optimal solution as well as both capacitated and uncapacitated policies. Distributions for the parameters are given as follows. Note that \$ is the monetary unit.

$\lambda_i \sim Uniform(0,10)$ in $\frac{units}{time}$

$R_i \sim Uniform(3,13)$ in $\frac{\$}{unit}$

$c_i \sim Uniform(2,7)$ in $\frac{\$}{unit \times time}$

$A \sim Uniform(12,32)$ in \$

Capacity variable K is dynamically adjusted for different values of N. This is to ensure a stable utilization level as N changes, as more shippers in the system lead to more frequent arrivals and requirement for more capacity for the system to behave as desired. To be able to observe a combination of capacitated and uncapacitated runs for each N setting, following lower and upper limits are considered for the range of possible values that K can take. This range is based on the capacity level observations over the uncapacitated CM solution for the experimental instances created by the given parameter distribution.

$K \sim Discrete\ Uniform\left(\left\lfloor\frac{3.5N+8}{3}\right\rfloor, \lceil 3.5N+8 \rceil\right)$ in $units$

Instances are created using the random number generator in MATLAB. Each instance is a distinct parameter combination according to the distributions given above in addition to the number of shippers N. Exact evaluation of policies was intended to be made through solving the linear programming formulation (see Puterman, 1994) of the MDP using CPLEX solver. GAMS 23.9 is used as the programming language. For 2-shipper case, policies could be evaluated within seconds. Runs with 3-shippers could be evaluated within a few minutes at most (depending on K). Hence, exact values of the objective function and the heuristic policies are obtained on GAMS for each run instance with $N = \{2,3\}$.

For $N = \{2, 3\}$, optimal policies are obtained through the exact approach. For $N \geq 4$, run times are long due to large problem size. For instance, in 4-shipper case, desired distribution of K leads to average capacity of an instance with $N = 4$ to be 14. However, due to GAMS run time exceeding an hour for a single instance with $K = 14$ without a result, optimal policy is not evaluated for $N \geq 4$. Simulation model for evaluating the heuristics is developed and implemented in MATLAB.

Note that comparing the simulated results with exact results for run instances with $N = 2$, it was observed that their difference was not significant. Further discussion on this comparison is available in Section 5.1.

5.1 Simulation Model

In this section, details of the simulation model and the implementation procedure is discussed.

Set of simulated policies are $Pol \in \{WLH, PI, WPI, FC\}$ for $N \geq 4$. Note that Opt is the abbreviation for the optimal policy, results of which was obtained only for $N = \{2,3\}$ via exact evaluation.

Constant arrival rate model (CM) has to be run for each instance to generate the heuristic policies. Resulting CM profit is also included in the comparisons of policy performances.

Each instance is replicated 10 times in simulation. Let $\pi_{Pol}(j, r)$ denote the resulting average profit per unit time $j \in \{1, \dots, 100\}$ being the index for instances and $r \in \{1, \dots, 10\}$ being the index for the replications and subscript Pol being the policy employed to obtain the given result.

Let $\bar{\pi}_{Pol}(j)$ defined as the average profit per time value obtained for the policy Pol. Values of $\bar{\pi}_{Opt}(j)$ and $\bar{\pi}_{CM}(j)$ as well as evaluations of heuristic policies $\bar{\pi}_{Pol}(j)$ of instances where $N = \{2,3\}$ correspond to the exact evaluations obtained through GAMS. These values are not obtained through simulation. Thus, they do not have any replications. When $\bar{\pi}_{Pol}(j)$ is obtained by simulation, it is calculated as the mean of $\pi_{Pol}(j, r)$ over 10 replications.

To have the simulated results significantly close to the exact evaluation, it is desired to choose a replication length that is long enough. Length of each replication is determined according to a precision threshold based on the coefficient of variation. Coefficient of variation of a run instance is defined as follows.

$$CV_j = \frac{\sigma_{\pi_{Pol}(j)}}{\bar{\pi}_{Pol}(j)}$$

Here, $\sigma_{\pi_{Pol}(j,r)}$ is the standard deviation of results for 10 replications in j^{th} instance, and $\bar{\pi}_{Pol}(j)$ is the mean of the simulated profit values for the j^{th} instance over 10 replications.

Number of transitions (equivalently, arrivals) for a simulated instance is determined such that coefficient of variation is sufficiently low. Specifically, it is ensured that for all given N, for at least 97% of the instances $CV_j < 0.005$. This condition is satisfied for each simulation run when the number of transitions is set to 1,000,000.

Simulation of each replication begins from state $\vec{0}$ and transitions occur following load arrivals. Number of load arrivals in a replication is 1,000,000. Interarrival times were assumed constant which is equal to the mean interarrival time $\frac{1}{\Lambda}$. Thus, the continuous-time Markov chain is treated as a discrete-time Markov chain. Under this setting, average inventory holding costs are incurred per transition.

Common seeds are used for creating the random arrivals to eliminate the additional variance in the performance comparisons of different heuristic policies. In particular, common shipper arrival times and common shipper types along with the same number of arrivals are used in the simulation of heuristic policies.

Validity of the simulation results are checked as policies were simulated for 2-shipper and 3-shipper settings as well. Exact evaluations obtained from GAMS are compared with simulated results using paired t-test. Differences of simulated results and exact results did not differ significantly, as the relative difference between the simulated and the exact profit values was at most 0.17% among all instances.

5.1.1 Simulation Procedure

At the beginning of the simulation, τ and $\vec{p}$ values are obtained solving the CM for the parameters of the given run instance. Solution for CM is obtained using Algorithm 1.

When $\sum_i Q_i^*$ exceeds K for the solution obtained by Algorithm 1, result is obtained by CONOPT solver in GAMS. If $\sum_i Q_i^*$ does not exceed K, Algorithm 1 solution is used. Solution time for CM is momentary when using Algorithm 1. It takes several additional seconds to obtain the solution from GAMS when required for any N in the experiments. Resulting $\pi(\vec{p}, \tau)$ is kept as the CM profit for the given run instance.

When the CM result $\bar{\pi}_{CM}(j) = \pi(\vec{p}, \tau) = 0$, run instance j is skipped with $\bar{\pi}_{Pol}(j) = 0$ being the average profit.

Every interval beginning at $\vec{0}$ and ending at the next dispatch (equivalently the next arrival at $\vec{0}$) is called a cycle. To evaluate the average profit of the run instance for the simulated heuristic, number of transitions as well as the total reward obtained until the next arrival at $\vec{0}$ are kept.

Let NS_k be the number of steps since the last arrival at $\vec{0}$ and TR_k be the total reward since the last arrival at $\vec{0}$ for k^{th} cycle. Let $\vec{x}(n)$ be the state after n^{th} arrival where $n \in \{1, \dots, 10^6\}$ is the set of arrivals in the replication.

Step 0. Simulation is initialized with variables provided as follows where $k = 1$ denotes the first cycle.

$$k = 1, \qquad n = 0, \qquad \vec{x}(n) = \vec{0}, \quad NS_1 = 0, \quad TR_1 = 0$$

Step 1. Shipper of an arrival is determined according to the arrival rates λ_i of each shipper. A random variable is created for each arrival, say $r \sim Uniform(0,1)$. That arrival belongs to shipper j which satisfies $\frac{\sum_{i=0}^{j-1} \lambda_i}{\Lambda} < r \leq \frac{\sum_{i=0}^{j} \lambda_i}{\Lambda}$. Note that $\lambda_0 = 0$ is used in this expression.

Step 2. For the current state $\vec{x}(n)$ and arriving load type i, WLH action $a_{i,\vec{x}(k)}^{WLH}$ is determined using the workload of the state and thresholds calculated using τ and $\vec{p}$ as described in section 4.1.2.

Step 3. According to the simulated heuristic policy, say Pol, the details of how the current action is obtained are presented below.

- $Pol = PI$:

  If current state $\vec{x}(n)$ is not in $\mathbb{S}_{(AD)}$ as defined in Section 4.2, $a_{i,\vec{x}(n)}^{Pol} = a_{i,\vec{x}(n)}^{WLH}$.

If current state is in $\mathbb{S}_{(AD)}$, $v(\vec{x}(n))$ and $v(\vec{x}(n)+e_i)$ are calculated using Eq. (4.14). Recall that g is estimated using CM objective by $g=\frac{\pi(\vec{p},\tau)}{\Lambda}$. PI action for arrival of shipper i in state $\vec{x}$ is determined as in Eq. (4.15).

- $Pol = WPI$:

  If $a_{i,\vec{x}(n)}^{WLH} = AD$ and $\sum_{j\in\{1,..,N\}} x_j(n) < x_{i,max}^{WLH}$, check if $cond \geq 0$ as in Eq. (4.16).

  When $cond \geq 0$, set $a_{i,\vec{x}(n)}^{Pol} = AW$. When $cond < 0$, $a_{i,\vec{x}(n)}^{Pol} = AD$.

  Otherwise $a_{i,\vec{x}(n)}^{Pol} = a_{i,\vec{x}(n)}^{WLH}$.

  Note that WLH simulation is done prior to WPI to obtain $x_{i,max}^{WLH}$ for each simulation instance.

- $Pol = FC$:

  If $a_{i,\vec{x}(n)}^{WLH} = AD$ and $\sum_{j\in\{1,..,N\}} x_j\,(n) < K-1$,

  Set $a_{i,\vec{x}(n)}^{Pol} = AW$.

  Otherwise $a_{i,\vec{x}(n)}^{Pol} = a_{i,\vec{x}(n)}^{WLH}$.

Step 4. For simulated heuristic Pol, once the action at the current state is determined, following updates are made.

- $TR_k = TR_k - \frac{\sum_i x_i c_i}{\Lambda} + \delta_{\left\{a_{i,\vec{x}(n)}^{Pol}=AD\right\}}\left(\sum_j R_j x_j + R_i - A\right)$
- $NS_k = NS_k + 1$
- If $a_{i,\vec{x}(n)}^{Pol} = AD$, then $\vec{x}(n+1) = \vec{0}$

  If $a_{i,\vec{x}(n)}^{Pol} = RW$, then $\vec{x}(n+1) = \vec{x}(n)$

  If $a_{i,\vec{x}(n)}^{Pol} = AW$, then $\vec{x}(n+1) = \vec{x}(n) + e_i$

If $\vec{x}(n+1) = \vec{0}$, following updates are made sequentially where Cy_k is the average profit obtained through cycle k.

- $Cy_k = \frac{TR_k \times \Lambda}{NS_k}$
- $TR_{k+1} = 0$

- $NS_{k+1} = 0$
- $k = k + 1$

Step 5. If $n < 10^6$, set $n = n + 1$ and return to Step 1.

If $n = 10^6$, calculate the average profit of the replication as follows.

- $k = k_{max}$
- $\pi_{Pol}(j, r) = \frac{\sum_{k=1}^{k_{max}} Cy_k NS_k}{\sum_{k=1}^{k_{max}} NS_k}$

■

Once 10 replications of a run instance are complete, mean of the results over 10 replications are found as follows.

$$\bar{\pi}_{Pol}(j) = \sum_{r=1}^{10} \frac{\pi_{Pol}(j, r)}{10}$$

5.2 Policy Performance Comparisons

In this section, the results of the performance comparisons of the heuristics are presented in detail. In the first approach, amount of average profit differences are reported in percentages. In this approach, comparisons are made to assess how profit differences behave.

In the second approach, a statistical test called the sign test based on the number of times a policy outperforms the other is conducted. Results of the sign test are presented to assess the overall performance of the heuristics in terms of the median profit differences.

5.2.1 Percentage Comparisons

The percentage difference between average profits per unit time of $Pol1$ and a benchmark policy, $Pol2$ is obtained using the Eq. (5.1). Note that J is the maximum number of instances.

$$\Delta_{Pol1,Pol2}(j) = \frac{\bar{\pi}_{Pol2}(j) - \bar{\pi}_{Pol1}(j)}{\bar{\pi}_{Pol2}(j)} \times 100, \qquad j \in \{1, \dots, J\} \tag{5.1}$$

Note that in obtaining $\Delta_{Pol1,Pol2}$, run instances where CM profit is 0 are discarded. Since heuristics are based on the CM result, none of them propose a policy for such instances. Also note that this result is consistent with the optimal policy since the proposed optimal decisions for an instance where $\bar{\pi}_{CM}(j) = 0$ were not accepting any loads e.g., a zero-profit solution.

Recall that, if the exact average profit per time value can be obtained (as in $N = \{2,3\}$ for each policy and every result for CM), $\bar{\pi}_{Pol}(j)$ is the corresponding average profit per time value in instance j of policy Pol. If the average profit can be obtained only through simulation, $\bar{\pi}_{Pol}(j)$ is the 10-replication average as explained in Section 5.1.

For each N shipper setting, Tables 5.1 - 5.6 present the results of percentage difference comparisons of policies. In these tables, percentiles of the comparison results are presented that give an insight on the distribution of the results. Mean of the percentage differences is also provided.

In Tables 5.1 and 5.2, presented values represent given percentiles or the mean of $\Delta_{Pol1,Opt}$ for $Pol1$ being the heuristic policy in the columns.

Table 5.1 $\Delta_{Pol1,Opt}$ for 2-Shipper runs

2 shippers	CM	WLH	PI	WPI	FC
min	1.34	0	0	0	0
25%	3.53	0.10	0	0	0
median	7.30	0.34	0.05	0.06	0.30
75%	12.4	1.07	0.32	0.63	3.03
max	68.2	22.2	22.2	22.2	39.5
mean	10.3	1.43	0.57	0.76	4.23

Table 5.2 $\Delta_{Pol1,Opt}$ for 3-Shipper runs

3 shippers	CM	WLH	PI	WPI	FC
min	0.95	0.07	0	0	0
25%	2.28	0.39	0	0	0
median	3.15	0.81	0.02	0.02	0.10
75%	5.66	2.42	0.11	0.20	2.78
max	65.6	75.1	5.57	4.56	48.0
mean	5.64	3.48	0.23	0.31	3.12

In Table 5.1 and Table 5.2, comparisons of heuristic policies versus the optimal policy are presented for $N = 2$ and $N = 3$. Results in both tables are the exact results obtained through GAMS. Recall that CM does not correspond to an MDP policy, but it is the difference comparison value for CM profit function that was obtained as in Eq. (4.1).

By results for run instances with $N = \{2,3\}$ shippers, it can be said that $\Delta_{PI,Opt}$ and $\Delta_{WPI,Opt}$ are similar with median percentage differences from optimal value of both policies being less than 0.06% in both tables. It can be inferred that in close to half of the instances, PI and WPI proposed the optimal policy. Median values for $\Delta_{WLH,Opt}$ and $\Delta_{FC,Opt}$ are less than 1%, but their 75 percentile values are both larger than 1%. $\Delta_{PI,Opt}$ and $\Delta_{WPI,Opt}$ have at most 0.63% difference at 75$^{\text{th}}$ percentile.

Regarding the mean values, PI has the smallest percentage difference versus Opt that is closely followed by WPI. WLH and FC performed worse than PI and WPI. However, it is not possible to suggest one of them has outperformed the other. It can be observed that CM provided worse average profit values compared to results from other heuristics.

In Tables 5.3 to 5.6, presented values represent given percentiles or the mean of $\Delta_{Pol1,PI}$ for $Pol1$ being the policy in the columns. PI is selected as the benchmark for $N \geq 4$ since PI resulted in better profits compared to other policies almost all the time.

Table 5.3 $\Delta_{Pol1,PI}$ for 4-Shipper runs

4 shippers	CM	WLH	WPI	FC
min	0.86	0.19	-0.6	-0.6
25%	1.75	0.60	0	0
median	2.34	1.25	0	0.02
75%	3.32	3.71	0.16	5.26
max	17.9	19.4	0.69	40.3
mean	3.12	3.14	0.09	3.36

Table 5.4 $\Delta_{Pol1,PI}$ for 5-Shipper runs

5 shippers	CM	WLH	WPI	FC
min	0.62	0.28	-0.1	0
25%	1.23	0.77	0	0
median	1.73	1.39	0.06	0.14
75%	2.38	3.39	0.13	2.78
max	8.55	14.7	4.36	56.0
mean	2.17	2.83	0.22	3.77

Table 5.5 $\Delta_{Pol1,PI}$ for 6-Shipper runs

6 shippers	CM	WLH	WPI	FC
min	0.7	0.38	0	0
25%	1.16	1.17	0.06	0
median	1.68	1.85	0.12	0.27
75%	2.16	3.95	0.22	2.98
max	5.82	26.9	5.14	34.0
mean	1.83	3.66	0.53	3.21

Table 5.6 $\Delta_{Pol1,PI}$ for 12-Shipper runs

12 shippers	CM	WLH	WPI	FC
min	0.38	1.08	0.05	0
25%	0.55	1.74	0.17	0.13
median	0.65	2.28	0.22	1.93
75%	0.79	3.53	0.38	5.08
max	2.45	21.6	7.75	22.7
mean	0.73	3.33	0.50	3.59

For the experimental instances with $N \geq 4$ (results in Tables 5.3 – 5.6), results for the optimal policy are not available due to computational complexity. Since PI is the heuristic with the best performance, it was selected as the benchmark policy.

It can be observed that WPI performs slightly worse than PI with median value of $\Delta_{WPI,PI}$ being at most 0.22% for $N = 12$ shippers. Mean value of $\Delta_{WPI,PI}$ was at most 0.53% for $N = 6$ shippers. Maximum percentage difference between WPI and PI is at most 7.75% for $N = 12$.

Average capacity of instances increases by N since distribution of K is a function of N. For smaller capacity runs, having a different action leads to a more significant difference between policies. This can be observed from comparisons with smaller N as maximum differences are larger compared to comparisons with larger N values. Also note that there is a significant increase for $\Delta_{Pol1,Pol2}$ values from 75 to 100 percentile (max) showing large differences are rare observations for every comparison in Tables 5.1 to 5.6.

For increasing N where $N = \{4,5,6,12\}$, it can be said that performance gap between PI and other heuristics increase. Only exception is the difference between CM and PI. Expected arrivals per unit time $\sum_{i=1}^{N} \lambda_i$ increases with N. This is due to distribution of λ_i is kept the same but as number of shippers increase, there are more shippers leading to more frequent arrivals. This makes to model behave similar to continuous arrivals, making results of the CM larger.

From the results of the percentage comparisons, it can be inferred that PI is the heuristic with the best performance. It is closely followed by WPI. WPI is followed by FC and WLH based on the mean and median of the percentage differences. However, FC and WLH had more extreme gaps either for comparisons versus Opt or PI. Maximum percentage differences of PI and WPI were not as extreme as FC and WLH. Although FC seemingly performs better than WLH based on the median difference, it could be far fetching to suggest that FC performs better than WLH especially since WLH performed better in terms of mean difference for $N =$

$\{2,4,5,12\}$. Further discussion is available in the following sections where heuristics are compared based on the sign test and sensitivity of the heuristic results were observed for varying capacity.

5.2.2 The Sign Test

In this section, results are statistically tested whether a policy outperforms the other. For this comparison. the sign test as described in Montgomery & Runger (2018) and Hines et. al. (2003) is adopted. This procedure tests the hypotheses to see if the median of paired samples is significantly different. Let the values in the samples be X_j^1 and X_j^2 where $j \in \{1, \dots, J\}$ is the set of sample entries. Median of these samples are $\tilde{\mu}_1$ and $\tilde{\mu}_2$. Tested hypotheses are as follows.

$$H_0: \tilde{\mu}_1 = \tilde{\mu}_2$$

$$H_1: \tilde{\mu}_1 \neq \tilde{\mu}_2$$

If medians $\tilde{\mu}_1$ and $\tilde{\mu}_2$ does not differ significantly (null hypothesis), then the sign of expression $X_j^1 - X_j^2$ has equal probability 0.5 of being negative or positive for any j. Two-sided test is done by first finding number of positive and negative $X_j^1 - X_j^2$ values, say n^+ and n^- respectively. Let $n^{test} = min\{n^+, n^-\}$. Note that it is assumed that there are no ties with $X_j^1 = X_j^2$. If ties exist, it is suggested to disregard them and test using remaining values. Test statistic $p = 2\sum_{i=0}^{n^{test}} \binom{J}{i} 0.5^J$ is the cumulative binomial probability of observing up to n^{test} successes multiplied by 2. If p is less than specified significance level α, null hypothesis is rejected e.g., sample medians are different from each other.

Let $\tilde{\mu}_{Pol1}$ and $\tilde{\mu}_{Pol2}$ denote the medians of samples $\bar{\pi}_{Pol1}$ and $\bar{\pi}_{Pol2}$ which are the average profits for $J = 100$ experimental instances of $Pol1$ and $Pol2$. In comparing the policies, null hypothesis is set to observe if $Pol1$ and $Pol2$ differ significantly in terms of median profit performance.

$$H_0: \tilde{\mu}_{Pol1} = \tilde{\mu}_{Pol2}$$

$$H_1: \tilde{\mu}_{Pol1} \neq \tilde{\mu}_{Pol2}$$

If positive and negative values of this difference are significantly different from half of the instances, null hypothesis is rejected. Note that different heuristics can produce exactly the same actions for each state and for each shipper's arrival. Due to exact evaluation of policies for $N = 2$ and same seed being used in simulation runs for $N \geq 3$, resulting average profit for different heuristics proposing exactly the same policy are equal. Hence, $\bar{\pi}_{Pol1}(j) = \bar{\pi}_{Pol2}(j)$ can be observed as a tied result. In conducting the sign test, instances with compared heuristics having the same average profit value are left out.

Let n^+, n^-, n^{eq}, n^{test} be defined as follows.

$$n^+ = \sum_{j=1}^{J} \delta_{\{\bar{\pi}_{Pol1}(j) - \bar{\pi}_{Pol2}(j) > 0\}}$$

$$n^- = \sum_{j=1}^{J} \delta_{\{\bar{\pi}_{Pol1}(j) - \bar{\pi}_{Pol2}(j) < 0\}}$$

$$n^{eq} = \sum_{j=1}^{J} \delta_{\{\bar{\pi}_{Pol1}(j) = \bar{\pi}_{Pol2}(j)\}}$$

$$n^{test} = min\{n^+, n^-\}$$

Test statistic is calculated as $p = 2\sum_{i=0}^{n^{test}} \binom{J'}{i} 0.5^{J'}$ where $J' = J - n^{eq}$ is the remaining sample size after tied entries are removed.

If $p < \alpha$ where α is the significance level for this two-sided test, null hypothesis is rejected. In this case, it is said that medians of the samples are not equal with α significance. Median profit of $Pol1$ is said to perform significantly better than median profit of $Pol2$ if $n^+ > n^-$ and $p < \alpha$. If $n^- > n^+$ and $p < \alpha$, median profit of $Pol2$ is better than median profit of $Pol1$. When $p \geq \alpha$, median profits of the tested policies do not differ significantly.

For every $N = \{2,3,4,5,6,12\}$ shipper setting, each heuristic pair is compared using the sign test with $\alpha = 0.05$. Results can be viewed from Tables 5.7 to 5.12. For an example comparison between $Pol1$ (rows) and $Pol2$ (columns), table cell provides the information $n^{+} \backslash n^{-}$ (n^{eq}) and the p-value.

Table 5.7 Sign test results for 2-Shipper runs

2 shippers	PI	WPI	FC
WLH	12\52 (36 eq.) $p = 0$	25\39 (36 eq.) $p = 0.06 > \alpha$	37\27 (36 eq.) $p = 0.17 > \alpha$
PI		25\12 (63 eq.) $p = 0.02$	38\0 (62 eq.) $p = 0$
WPI			29\0 (71 eq.) $p = 0$

From Table 5.7, it can be observed that median difference was insignificant in comparisons between $WLH - WPI$ and $WLH - FC$. Median profit for PI is significantly larger than medians of other policies. Instances where WLH profit are larger than PI were investigated since it was expected that PI would improve the WLH performance.

It was observed that these instances had a large gap (up to 66.7%) in favor of WLH between results of CM and WLH leading to inaccurate approximation of g which was used in deriving PI actions.

PI was evaluated again using $g = \frac{\pi(\vec{p},\tau)\times 1.0905}{\Lambda}$ after observing the average percentage gap between CM and WLH for $N = 2$ was 9.05%. With this adjustment in place, WLH profit was larger than PI for only a single instance. This instance had a gap of 21.3% between CM and WLH.

Due to CM performance improving with increasing N, approximation of g becomes more accurate for run instances with larger number of shippers.

Note that there are consider amount of equal profit observations between different heuristics. As number of shippers increase, such encounters become less common due to larger problem size making more state-action pairs available. In smaller problem instances in terms of N and K, these encounters are more common.

Table 5.8 Sign test results for 3-Shipper runs

3 shippers	PI	WPI	FC
WLH	2\98 $p = 0$	9\89 (2 eq.) $p = 0$	37\61 (2 eq.) $p = 0.01$
PI		36\23 (41 eq.) $p = 0.07 > \alpha$	53\6 (41 eq.) $p = 0$
WPI			36\0 (64 eq.) $p = 0$

From Table 5.8, it can be observed that median difference was insignificant only in comparison between $PI - WPI$ for run instances with $N = 3$ shippers. Notice that median difference between WLH and PI has become more apparent compared to results for $N = 2$.

Table 5.9 Sign test results for 4-Shipper runs

4 shippers	PI	WPI	FC
WLH	0\100 $p = 0$	5\95 $p = 0$	43\57 $p = 0.13 > \alpha$
PI		49\10 (41 eq.) $p = 0$	57\2 (41 eq.) $p = 0$
WPI			44\0 (56 eq.) $p = 0$

It can be observed from Table 5.9 that median difference was insignificant only in comparison between $WLH - FC$ for run instances with $N = 4$ shippers. Notice that median difference between PI and WPI has become more apparent compared to results for $N = \{2,3\}$.

Table 5.10 Sign test results for 5-Shipper runs

5 shippers	PI	WPI	FC
WLH	0\100 $p = 0$	0\100 $p = 0$	38\62 $p = 0.01$
PI		73\5 (22 eq.) $p = 0$	69\2 (29 eq.) $p = 0$
WPI			50\8 (42 eq.) $p = 0$

Table 5.11 Sign test results for 6-Shipper runs

6 shippers	PI	WPI	FC
WLH	0\100 $p = 0$	0\100 $p = 0$	39\61 $p = 0.02$
PI		87\2 (11 eq.) $p = 0$	74\2 (24 eq.) $p = 0$
WPI			53\26 (21 eq.) $p = 0$

Table 5.12 Sign test results for 12-Shipper runs

12 shippers	PI	WPI	FC
WLH	0\100 $p = 0$	0\100 $p = 0$	51\49 $p = 0.76 > \alpha$
PI		100\0 $p = 0$	89\3 (8 eq.) $p = 0$
WPI			72\27 (1 eq.) $p = 0$

From the results presented in Tables 5.10 to 5.12, it can be observed that PI has performed better than each heuristic in terms of the median profit value. WPI also performed better than FC and WLH. Comparisons for median profits of FC and WLH were either in favor of FC or the difference was insignificant.

Note that there are equal profit observations mainly between PI, WPI, FC. Also note that number of such observations decrease for increasing number of shippers as seen from Tables 5.7 to 5.12.

Results of the sign tests are consistent with percentage differences where it could be inferred that PI performed better than other heuristics. Comparisons of median profits through the sign test for PI either resulted in an insignificant median difference (at $N = 3$ versus WPI) or resulted in favor of PI. PI is followed by WPI which only performed worse than PI in terms of median profits. Comparisons of median profits between FC and WLH either resulted in favor of FC or median difference was insignificant. Due to insignificant difference being observed for multiple N values, FC performing better than WLH in terms of median profits is not a strong conclusion.

5.2.3 Heuristic Performance Sensitivity to Capacity

In previous sections, the whole of experimental instances for each N value is studied. In this section, it is aimed to observe how the heuristics perform under different capacity conditions.

In Figure 5.1, values of percentage difference comparisons for $N = 12$ are presented in a scatter plot for instances with increasing vehicle capacity K. Note that these values are from $\Delta_{Pol1,PI}$ where $Pol1 \in \{WLH, WPI, FC\}$ since PI is the benchmark policy for $N \geq 4$.

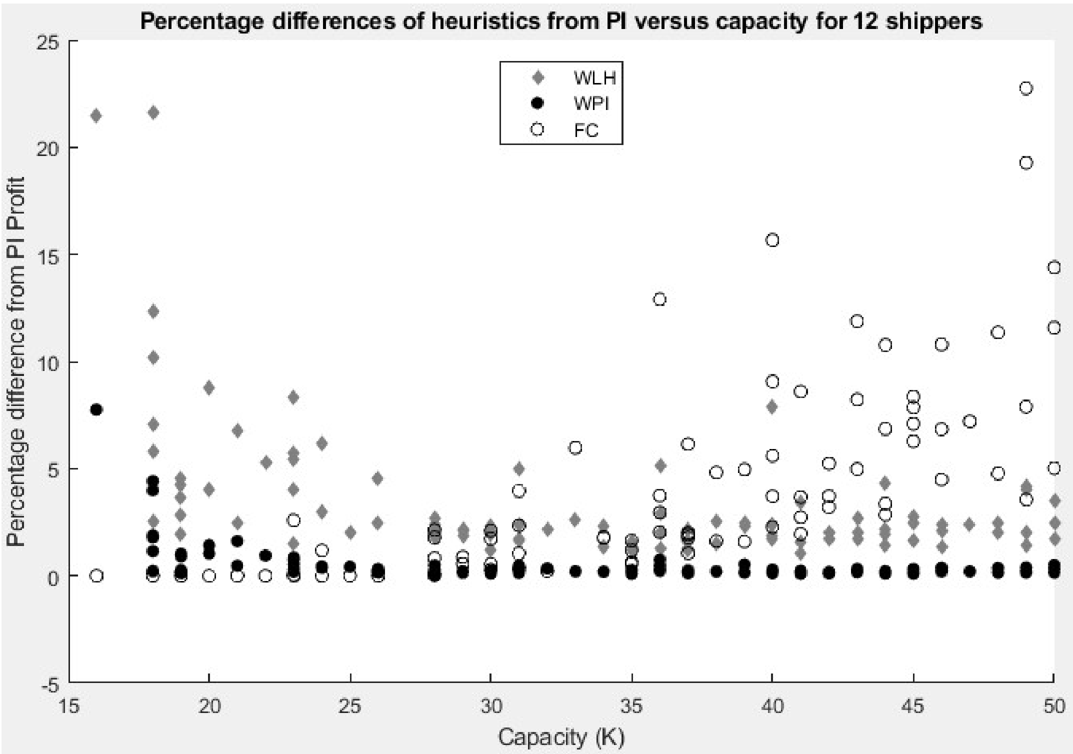

Figure 5.1 Percentage differences of heuristics with PI for 12 shippers versus vehicle capacity

From Figure 5.1, it can be observed that percentage differences of WLH and FC versus PI are highly sensitive to changing levels of K. For smaller capacity instances, WLH (diamonds) has more extreme results while for larger capacity instances, FC (empty circles) results are more extreme. WPI (filled circles) is more consistent except for run instances with very small capacity.

Following this observation, runs are grouped for capacity levels of instances $j \in \{1, \dots, J\}$. Small capacity group consists of j where $K_j \leq \overline{K}$ and large capacity group consists of j where $K_j > \overline{K}$. Note that $\overline{K} = \frac{\sum_j K_j}{J}$ is the mean of the capacity levels of instances. Mean values of the percentage comparisons are presented in Table 5.13 for each heuristic with instances being grouped for capacity levels.

Table 5.13 Percentage difference comparisons of low and high capacity runs versus PI for $N = 12$ shippers

$N = 12$	$K_j \leq \overline{K}$	$K_j > \overline{K}$
WLH	4.65	2.25
WPI	0.84	0.23
FC	0.66	5.98

From the results in Table 5.13, it is observed that WLH has a better average performance for larger capacity runs compared to FC. While FC has worse performance compared to other heuristics for larger capacity runs, it has better performance compared to WPI for smaller capacity runs.

This behavior can be explained intuitively since FC updates the policy obtained by WLH to dispatch the vehicle only when capacity is full. For large capacity runs, this leads to larger inventory holding costs as number of periods to wait increases.

Recall the observation on the WLH policy that it does not utilize capacity level as much as the optimal policy does. For lower capacity runs, it is more likely that the optimal policy is to fully utilize the capacity. Since WLH does not utilize capacity as much, it is expected that WLH performance will be worse for lower capacity instances.

WPI performance is worse for lower capacity instances, but it is still within 0.84% of the PI performance. Recall the observation made for WPI that when a bound for states to be improved ($x_{i,max}^{WLH}$ as defined in Section 4.3) does not exist, WPI tends to overuse the vehicle capacity. By results from instances with smaller capacity, it is apparent that this bound occasionally leads to not utilizing the capacity level as much as PI does.

Scatterplots as in Figure 5.1 are observed for other problem parameters (λ, c, R, A) to look for performance patterns of the heuristics depending on the values of the parameters. Heuristic profit results for these parameters did not follow a visible pattern that was observed for the K parameter as in Figure 5.1.

CHAPTER 6

CONCLUSIONS

In this study, a shipment consolidation problem is studied. In the studied problem setting, there are multiple shippers consolidating their shipments in a vehicle that has a limited capacity. Shipments arrive according to a Poisson process with arrival rates differing for each shipper. For each load, revenue is obtained once the vehicle is dispatched and inventory holding cost is incurred per load per unit time. There is a fixed cost for dispatching the vehicle. It is aimed to maximize the average profit obtained per unit time.

Problem is modeled as a continuous-time MDP with states being the number of shipments available from each shipper. At each decision epoch, actions related to accepting or rejecting the incoming arrival and dispatching the vehicle or waiting for the next shipment arrival are determined.

Optimal policy structure is observed to propose heuristic methods that would provide policies with comparable performance to the optimal policy for large-sized problems. Monotonicity of the optimal decisions were characterized for the 2-shipper setting. It was found that optimal sequence of actions is Reject & Wait (RW) (if exists), then Accept & Wait (AW), and lastly Accept & Dispatch (AD) for states with increasing number of loads.

Heuristic policies that make use of the findings from the observations on the optimal MDP policy for the 2-shipper setting. Constant arrival rate model (CM) is formulated to propose threshold policies for the MDP. First heuristic policy is WLH, which derives an MDP policy from the optimal values for consolidation cycle lengths for each shipper obtained by solving CM. Other heuristics which are called PI, WPI, FC, modify the WLH policy in different ways.

Performance of the proposed policies were evaluated in terms of the average profit per unit time. Computational experiments were made for different values for number of shippers $N \in \{1,2,3,4,5,6,12\}$. Optimal solutions of the MDP could be found for instances where $N = \{2,3\}$. Exact evaluations of the heuristic policies were made for $N = \{2,3\}$ as well. For instances with $N \geq 4$, heuristic policies were evaluated through simulation. For instances where $N \geq 4$, optimal policy was not evaluated due to long computation time.

Heuristics were compared based on the percentage differences of their profits and number of times each heuristic performed better than another. Among the alternative heuristics, PI provided better results compared to WLH, WPI, and FC. Although comparison versus the optimal policy could not be made for instances with more than 3 shippers, results for 2-shipper and 3-shipper instances showed the average percentage difference between PI and the optimal policy is found to be 0.57% for 2-shipper runs and 0.23% for 3-shipper runs.

Sensitivity of the heuristics to varying capacity level is also observed. It was seen that WLH performed better in instances with larger vehicle capacity compared to FC, which performed better than WLH in smaller capacity instances.

This study focused on maximizing the total profit of the system, not focusing on possible individual objectives of shippers. A possible extension to this study is considering collaborative strategies of shippers aiming to maximize their individual profits.

REFERENCES

Braglia, M., Castellano, D., & Song, D. (2017). Distribution-free approach for stochastic Joint-Replenishment Problem with backorders-lost sales mixtures, and controllable major ordering cost and lead times. Computers & Operations Research, 79, 161-173.

Bulut, Ö., & Fadiloğlu, M. M. (2011). Production control and stock rationing for a make-to-stock system with parallel production channels. IIE Transactions, 43(6), 432-450.

Cetinkaya, S., & Bookbinder, J. H. (2003). Stochastic models for the dispatch of consolidated shipments. Transportation Research Part B: Methodological, 37(8), 747-768.

Cetinkaya, S. & Lee, C.-Y. (2002), Optimal outbound dispatch policies: Modeling inventory and cargo capacity. Naval Research Logistics, 49: 531-556.

Creemers, S., & Boute, R. (2022). The joint replenishment problem: Optimal policy and exact evaluation method. European Journal of Operational Research.

Ergun, Ö., Kuyzu, G., & Savelsbergh, M. (2007). Shipper collaboration. Computers & Operations Research, 34(6), 1551-1560.

Feng, H., Wu, Q., Muthuraman, K., & Deshpande, V. (2015). Replenishment policies for multi-product stochastic inventory systems with correlated demand and joint-replenishment costs. Production and Operations Management, 24(4), 647-664.

Fung, R. Y., Ma, X., & Lau, H. C. (2010). (T, S) policy for coordinated inventory replenishment systems under compound Poisson demands. Production planning & control, 12(6), 575-583.

Ha, A. Y. (1997a). Inventory rationing in a make-to-stock production system with several demand classes and lost sales. Management Science, 43(8), 1093-1103.

Ha, A. Y. (1997b). Stock-rationing policy for a make-to-stock production system with two priority classes and backordering. Naval Research Logistics (NRL), 44(5), 457-472.

Hines, W. W., Montgomery, D. C., Goldsman, D. M., & Borror, C. M. (2003). Probability and Statistics in Engineering (4th ed.). Wiley.

Kleywegt, A. J., & Papastavrou, J. D. (1998). The dynamic and stochastic knapsack problem. Operations research, 46(1), 17-35.

Kouki, C., Babai, M. Z., Jemai, Z., & Minner, S. (2016). A coordinated multi-item inventory system for perishables with random lifetime. International Journal of Production Economics, 181, 226-237.

Lai, M., Xue, W., & Zhao, L. (2016). Cost allocation for cooperative inventory consolidation problems. Operations Research Letters, 44(6), 761-765.

Lu, Y. (2018). Heuristic policies for stochastic knapsack problem with time-varying random demand. Applied Stochastic Models in Business and Industry, 34(6), 1007-1016.

Marasco, A. (2008). Third-party logistics: A literature review. International Journal of production economics, 113(1), 127-147.

Montgomery, D. C., & Runger, G. C. (2018). Applied Statistics and Probability for Engineers (7th ed.). Wiley.

Moon, I. K., & Cha, B. C. (2006). The joint replenishment problem with resource restriction. European Journal of Operational Research, 173(1), 190-198.

Muriel, A., Chugh, T., & Prokle, M. (2022). Efficient algorithms for the joint replenishment problem with minimum order quantities. European Journal of Operational Research, 300(1), 137-150.

Mustafa Tanrikulu, M., Şen, A., & Alp, O. (2010). A joint replenishment policy with individual control and constant size orders. International Journal of Production Research, 48(14), 4253-4271.

Mutlu, F., Çetinkaya, S. I. L., & Bookbinder, J. H. (2010). An analytical model for computing the optimal time-and-quantity-based policy for consolidated shipments. IIE Transactions, 42(5), 367-377.

Noblesse, A. M., Sonenberg, N., Boute, R. N., Lambrecht, M. R., & Van Houdt, B. (2022). A joint replenishment production-inventory model as an MMAP [K]/PH [K]/1 queue. Stochastic Models, 1-26.

Pan, S., Ballot, E., & Fontane, F. (2013). The reduction of greenhouse gas emissions from freight transport by pooling supply chains. International journal of production economics, 143(1), 86-94.

Papadaki, K., & Powell, W. B. (2007). Monotonicity in multidimensional Markov decision processes for the batch dispatch problem. Operations research letters, 35(2), 267-272.

Porras, E., & Dekker, R. (2006). An efficient optimal solution method for the joint replenishment problem with minimum order quantities. European Journal of Operational Research, 174(3), 1595-1615.

Puterman, M. L. (1994). Markov Decision Processes: Discrete Stochastic Dynamic Programming (1st ed.). Wiley-Interscience.

Range, T. M., Kozlowski, D., & Petersen, N. C. (2018). A shortest-path-based approach for the stochastic knapsack problem with non-decreasing expected overfilling costs. Computers & Operations Research, 97, 111-124.

Ross, S. (1983). Introduction to Stochastic Dynamic Programming.

Salameh, M. K., Yassine, A. A., Maddah, B., & Ghaddar, L. (2014). Joint replenishment model with substitution. Applied Mathematical Modelling, 38(14), 3662-3671.

Satır, B., Erenay, F. S., & Bookbinder, J. H. (2018). Shipment consolidation with two demand classes: Rationing the dispatch capacity. European Journal of Operational Research, 270(1), 171-184.

Serfozo, R. F. (1979). An equivalence between continuous and discrete time Markov decision processes. Operations Research, 27(3), 616-620.

Tijms, H. C. (1995). Stochastic Models: An Algorithmic Approach (1st ed.). Wiley.

Ülkü, M. A. (2012). Dare to care: Shipment consolidation reduces not only costs, but also environmental damage. International Journal of Production Economics, 139(2), 438-446.

Véricourt, F., Karaesmen, F., & Dallery, Y. (2002). Optimal stock allocation for a capacitated supply system. Management Science, 48(11), 1486-1501.

Winston, W. L. (2003). Operations Research: Applications and Algorithms (4th ed.). Cengage Learning.

Yilmaz, O., & Savasaneril, S. (2012). Collaboration among small shippers in a transportation market. European journal of operational research, 218(2), 408-415.