Spare Parts Inventory Management for Substitute Consumer Durable Products Under Uncertainties

SPARE PARTS INVENTORY MANAGEMENT FOR SUBSTITUTE CONSUMER DURABLE PRODUCTS UNDER UNCERTAINTIES

BY

SHUAI ZHANG, M.Eng., M.Sc., B.Sc., B.Econ.

A Thesis

Submitted to the DeGroote School of Business

and School of Graduate Studies

of McMaster University

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

Ph.D. Dissertation (2022)

DeGroote School of Business, McMaster University

Hamilton, Ontario, Canada

TITLE:	Spare Parts Management for Substitute Products Under Uncertainties
AUTHOR:	Shuai Zhang DeGroote School of Business McMaster University, Hamilton, Canada
SUPERVISOR:	Dr. Kai Huang
COMMITTEE MEMBERS:	Dr. Kai Huang Dr. Elkafi Hassini Dr. Yufei Yuan
NUMBER OF PAGES:	xii, 188

Abstract

In this thesis, we investigate mathematical models based on adaptive robust optimization (ARO) and stochastic programming (SP) for the spare parts inventory management under various uncertainties for the substitute consumer durable products offered in an assortment. In the first part of the thesis, we present a comprehensive literature review of spare parts inventory management and 142 papers are surveyed and classified. In the second part, we consider a multi-period spare parts inventory system providing spare parts for the substitute products in an assortment and aim to develop the spare parts inventory policies when the assortment is given in advance and there are uncertainties in the failure rates of both products and spare parts. We formulate a multi-stage adaptive mixed-integer robust optimization model and improve the partition-and-bound method to solve it. In the third part, we consider a multi-period dynamic assortment planning problem for an original equipment manufacturer (OEM) who launches and sells the substitute product variants through an online platform. To handle the uncertainties embedded in the customer preferences estimation, a multi-stage stochastic programming model is proposed and a branch-and-price (B&P) algorithm is designed based on the block-angular structure of the model. In the last part of this thesis, we study an assortment planning problem from the product lifecycle perspective and intend to simultaneously determine the assortment decisions, spare parts inventory policies, and returned products remanufacturing decisions when there are uncertainties in the failure rates of products and spare parts, and the return rates of the used products. The aim of this study is to explore the impacts on product assortment decisions brought by implementing the last-time buy (LTB) and remanufacturing strategies to supply the spare parts over the warranty periods.

To my loving parents.

Acknowledgements

I would like to express my sincerest gratitude to all those who have helped and supported me throughout my Ph.D. study.

First and foremost, I am extraordinarily grateful to my supervisor, Dr. Kai Huang. In last five years, he provided me with the valuable guidance, support, inspiration, and encouragement and showed me what features a great researcher should possess. He has guided me in selecting the thesis topic and patiently aided me to develop my research skills. Working under his supervision has been both a great honour and a pleasure to me. His advice regarding to research and career development has been a precious fortune that I would carry with in my life. I feel incredibly blessed to have him as my supervisor and could not have imagined having a better supervisor for my Ph.D. study.

Secondly, I would like to send my deepest gratitude to my supervisory committee members Dr. Elkafi Hasini and Dr. Yufei Yuan, for their valuable advice to my study and insightful comments and stimulating suggestions to this thesis. I would also like to thank Dr. Yun Zhou for kindly taking the role of comprehensive examination committee member. My gratitude also goes to Dr. Parkash Abad and Dr. Mahmut Parlar, who gave me high quality Ph.D. courses in the first two years.

Finally but the most importantly, I owe much to my parents Ruibo Zhang and Fengzhen Liu for their unending love and encouragement. Without them, I would not have made it through. I would also like to greatly appreciate my friends, especially Xiangchao Xu, Xiang Li, Zhihao Cao, and Alex Wang, for supporting and cheering me on.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation

Spare parts are stock items used in maintenance activities to keep equipment in operating conditions (Kennedy et al., 2002). The spare parts inventory management is critical because the cost of spare parts accounts for a big share of the equipment's lifecycle cost: The value of spare parts annually consumed by a machinery, which might have a lifetime of around 30 years, amounts to near 2.5% of the original purchasing price (Hu et al., 2018). The non-availability of spare parts may induce great financial losses to equipment owners when they have failed equipment required for repairs. In some industries where the manufacturers provide after-sales services, good spare parts inventory management can improve customer satisfaction by reducing equipment downtime and increasing equipment reliability (Jin and Tian, 2012). Furthermore, spare parts often have an obsolescence problem, which usually happens when an equipment enters the end of its lifecycle. Overstocked spare parts become obsolete when there is no demand for them so that they must be discarded at a quite low value. To tackle the trade-offs between overstocking and understocking spare parts, more actions related to the joint optimization of the maintenance and inventory operations are required. In summary, spare parts inventory management plays an important role in achieving the desired equipment availability at a minimum economic cost.

As the concept of supply chain sustainability has been growing considerably, spare part inventory management is also given more consideration with their role in supporting sustainability. In past decades, the original equipment manufacturers (OEMs) are more likely to advocate for a culture of planned obsolescence: By designing their products to be short-lived and hard to repair, they can seize more revenues because customers are forced to purchase more new products when the old ones are not functioning properly. However, this culture contributes to wasting more natural resources and energy, generating more greenhouse gases, and further escalating global warming. For example, the carbon emissions of producing an iPhone 12 account for nearly 80 percent of the total emissions during its lifecycle (Apple Inc., 2020). In the United States, a motion known as “right to repair" has been calling for legislation that requires companies make their parts, tools, and information available to consumers and repair shops (Rosa-Quino, 2020). The motivation of this motion is curbing that culture because the longer the product lifecycle is, the fewer unnecessary product purchases will be, and finally the lower pollution the production processes will generate. For the OEMs, they may suffer from the decreased sales of new products from this motion, but they can obtain revenues by expanding their after-sales services. Moreover, advocating sustainability practice can demonstrate manufacturers embrace corporate social responsibility (CSR) to customers, contribute to a positive brand image, and reinforce their corporate reputation (Ukko et al., 2019; López-Pérez et al., 2017; Aguilera-Caracuel and Guerrero-Villegas, 2018). In this context, to reach the balance between sustainability and profitability, the OEMs need more spare parts for repairing the faulty products and an efficient spare parts inventory management system is necessary.

However, the management of spare parts faces several difficulties. Firstly, the intermittent demand patterns are common among spare parts and difficult to predict. This point is extremely hard to solve for the consumer durable products which are usually offered in the form of product assortment. Nowadays, OEMs tend to use product segmentation strategy in which multiple substitute products belonging to one category are provided to customers from different groups. This product category is normally referred to as the product assortment. Although such a strategy potentially can increase the total revenue by attracting more customers but it adds more complexity to the spare parts inventory management. This is because the spare parts are more likely to be simultaneously used by two or more products in the assortment. In this case, the demand for spare parts is hard to predict due to its dependency on the various quantities of different products in the assortment sold to the customers. Secondly, the number and variety of spare parts are usually very large. It is difficult to identify an appropriate strategy for each spare part type. Thirdly, inventory decisions have to reduce both the penalties of excess stocks and the costs of equipment downtime incurred by the spare parts shortages. Last but not least, the consumption of spare parts is closely related to the equipment usage, damage, and maintenance (Hu et al., 2018). To develop the proper policies for managing spare parts inventory, the OEMs normally borrow the power from the historical data and statistical techniques to predict the spare parts demand. However, such estimations usually are coupled with errors/uncertainties and the OEMs have to take those uncertainties into consideration.

The goal of this thesis is to provide general modelling frameworks for computing optimal policies for the spare parts inventory management under various uncertainties in a multi-period planning horizon, especially for the substitute consumer durable products offered in an assortment. In other words, our main focus is to incorporate the spare parts inventory management with assortment planning. Specifically, we study this problem under three different scenarios:

(I) Develop the spare parts inventory policies when the product assortment is given in advance and there are uncertainties in the failure rates of both the products and the spare parts.

(II) Develop the dynamic product assortment decisions and spare parts inventory policies when there are uncertainties in the customer preferences to the products in the assortment.

(III) Develop the product assortment decisions, spare parts inventory policies, and returned product remanufacturing decisions when there are uncertainties in the failure rates of both the products and the spare parts and the return rates of the used products.

To handle the uncertainties, robust optimization (RO) and stochastic programming (SP) approaches are commonly used in the literature. RO is usually applied to the situations where limited distributional information is provided, and focuses on the worst-case scenario. On the other hand, SP assumes the full distributional knowledge of uncertain parameters and optimizes the expected performance. We are interested in RO and SP not only because they are mathematical programming-based modelling approaches that contribute to the robustness of solutions against data uncertainty, but also because they have the potential to incorporate multiple sources of uncertainty into model development.

1.2 Contributions and Organization of the Thesis

In this thesis, we mainly focus on integrating the spare parts inventory management with assortment planning under various uncertainties to address the research gaps in the literature.

In Chapter 2, we review current studies on the spare parts inventory management. Our review has the following highlights. Firstly, we focus on analyzing the supply chain structure of different inventory networks for managing spare parts. Secondly, current literature are classified based on three analytics techniques, i.e., descriptive analytics, predictive analytics, and prescriptive analytics. Thirdly, several research gaps in this field are identified and discussed from the perspectives of consumer durable goods, inventory network structure and policies, reverse logistics, spare parts demand pattern modelling, and big data analytics.

In Chapter 3, we focus on managing the spare parts inventory of a product assortment which includes several substitute products over multiple time periods under the uncertainties in the failure rates of both products and spare parts. We develop a multi-stage adaptive robust optimization model in which the demand is determined by the multinomial logit (MNL) model of consumer choice over the substitutes in an assortment. The main contributions of this chapter are multi-fold. Firstly, we propose a model considering managing the spare parts inventory of multiple substitute consumer products in an assortment. The spare parts demand induced by the users of these products is estimated based on the MNL model. Our purpose is to jointly manage the spare parts for these products. To our best knowledge, there is no study on this problem so far. This problem is complicated because some spare parts may be commonly used by several products in the assortment while some may be uniquely used by one product.

In Chapter 4, we consider an OEM who produces a dynamic assortment of products and sells them through online platform over a selling season with multiple periods under the uncertainties of customer preferences. The contributions of this chapter are as follows. First of all, to our best knowledge, this multi-period dynamic assortment planning problem with a blended setup of uncertain customer preferences and component stocking was unexplored in the literature. This problem models the situation faced by many OEMs who produce and sell product assortments through the online platforms and are able to utilize the historical data to estimate the customer preferences over the selling season. Secondly, a branch-and-price (B&P) algorithm is designed to solve the proposed multi-stage stochastic programming model. Through extensive numerical experiments, the complexity of this problem is illustrated and the performance of the proposed algorithm is validated. The advantage of dynamic assortment planning, i.e., dynamically changing the assortment at different periods based on the estimated customer preferences, is also highlighted in the numerical experiments.

In Chapter 5, we study an assortment planning problem from the perspective of the product lifecycle and aim to integrate the warranty service operations into the strategic assortment planning decisions for the OEMs. In this problem, we consider the lifecycle costs of the products when making strategic product assortment planning decisions. To be specific, the expected costs related to the warranty services for assortment products during the end-of-life (EOL) phase are included in the decision-making. Furthermore, we consider both components last-time buy (LTB) and remanufacturing to be used as the supply sources of the spare parts inventory during the EOL phase. To the best of our knowledge, this setting is novel in both the literature of assortment planning and those of spare parts inventory management for the products in the EOL phase. Through the numerical experiments, we explore the advantages of joint optimization on the assortment planning decisions and the spare parts procurement and remanufacturing decisions compared to the separate optimizations on those decisions. Afterwards, we discuss the impacts of the uncertainty levels of those two uncertainties on the expected total profits of the OEM.

Finally, the major contents and contributions of this thesis are summarized in Chapter 6.

Chapter 2

Literature Review

2.1 Introduction

Spare parts are stock items used in maintenance activities to keep equipment in operating conditions (Kennedy et al., 2002). The spare parts inventory management is critical because the cost of spare parts accounts for a big share of the equipment's lifecycle cost: The value of spare parts annually consumed by a machinery, which might have a lifetime around 30 years, amounts to near 2.5% of the original purchasing price (Hu et al., 2018). The non-availability of spare parts may induce great financial losses to equipment owners when they have failed equipment required for repairs. In some industries where the manufacturers provide after-sales services, decent spare parts inventory management can improve customer satisfaction by reducing equipment downtime and increasing equipment reliability (Jin and Tian, 2012). Furthermore, spare parts often have an obsolescence problem, which usually happens when an equipment enters the end of lifecycle. Overstocking spare parts become obsolete when there is no demand for them so that they must be discarded at a quite low value. This causes wastes in resource to equipment owners or original equipment manufacturers (OEMs). To tackle the trade-offs between overstocking and understocking spare parts, more actions related to joint optimization in maintenance and inventory operations are required. In overall, spare parts inventory management plays an important role in achieving the desired equipment availability at a minimum economic cost. However, the management of spare parts faces several difficulties. Firstly, the intermittent demand patterns are common among spare parts and difficult to predict. Secondly, the number and variety of spare parts are usually very large. It is difficult to identify an appropriate strategy for each spare part type. Thirdly, inventory decisions have to reduce both the penalties of excess stock and the costs of equipment downtime incurred by the shortage of spare parts. Last but not least, the consumption of spare parts is closely related to the equipment usage, damage, and maintenance (Hu et al., 2018).

2.1.1 Motivations and Objectives

Over last decade, a significant number of studies on spare parts inventory management have been published to provide managerial insights to practitioners. Nevertheless, to our best knowledge, there is no literature review that organizes current literature on spare parts inventory management from the perspectives of supply chain management and supply chain analytics, even though such perspectives have been prevalent in contemporary business world.

There are seven literature reviews on this topic in the last thirty years. Cho and Parlar (1991) review the publications of optimal maintenance and replacement models for multi-unit systems and classify the surveyed literature into five categories based on the maintenance operations topic, but the spare parts inventory management is only one of the sub-topics discussed in their review. Guide and Srivastava (1997) present a review on the studies of repairable spare parts inventory management and the reviewed studies are grouped based on network structure (single versus multi-echelon), solution methodology, and solution types (exact versus approximate solutions). Kennedy et al. (2002) present the first review which completely focuses on the literature of spare parts inventory management and analyze the relationships between equipment maintenance and spare parts inventory. They identify the main issues in determining spare parts inventory and discuss future potential research directions. Their classification on literature is based on the types of implemented maintenance strategies including preventive maintenance and corrective maintenance. Paterson et al. (2011) give a review of the literature on lateral transshipments within an inventory system and point out that the reactive transshipments are normally used for managing spare parts inventory. The literature classification is made based on the characteristics of inventory system such as inventory item number, echelon number, ordering policy, inventory pooling implementation, and transshipments types. Van Horenbeek et al. (2013) present a literature review on joint maintenance and inventory optimization and the scope of their surveyed publications is limited to the models for non-repairable spare parts. Two classification schemes are proposed. One scheme is based on the following seven criteria, i.e., inventory policy, maintenance characteristic, delay, multi-echelon network, single-unit versus multi-unit system, objective function, and optimization technique. Another scheme is made under the topics of optimization and the studies are split into three groups, i.e., optimization of parameter, optimization of replenishment quantity, and design of reuse supply chain.

The most recent literature reviews are conducted by Basten and van Houtum (2014) and Hu et al. (2018) respectively. Basten and van Houtum (2014) focus on the studies related to managing spare parts inventories of technical systems. The classification in the review is made based on the characteristics of inventory network including network's service provider, number of echelon levels, availability of lateral or emergency transshipment, and so on. Hu et al. (2018) present a framework for Operational Research (OR) area in spare parts inventory management and analyze the literature on four critical aspects of OR in spare parts inventory management, i.e., spare parts classification, demand forecasting, inventory optimization, and supply chain system simulation.

Past literature reviews clearly show that spare parts inventory management is attracting increasing attentions from academia, especially in Operational Research and Management Science (OR/MS) area. Spare parts inventory management research origins as a sub-topic under the equipment maintenance studies and has become an individual research topic. The studied spare parts types are not only restricted to the spare parts used in capital intensive systems which require in-time maintenance and support, but also include the ones in consumer durable products whose maintenances are more flexible. In addition, some recent studies begin to study the problem under broader views in time dimension and network structure dimension, some intend to evaluate the impacts brought by different inventory operations on the lifecycle cost of product or the total cost of ownership, and some try to investigate the cooperative operations between different participants in spare parts supply chain.

The main differences between our literature review and the aforementioned reviews are as follows. Firstly, we focus on analyzing the supply chain structure of different inventory networks for managing spare parts. Secondly, current literature are classified based on three analytics techniques, i.e., descriptive analytics, predictive analytics, and prescriptive analytics. Thirdly, several research gaps in this field are identified and discussed from the perspectives of consumer durable goods, inventory network structure and policies, reverse logistics, spare parts demand pattern modelling, and big data analytics implementation.

2.1.2 Methodology

The database used for searching publications is ABI/INFORM Collection which is one of the most comprehensive business databases in OR/MS field. The initial search is conducted by searching *spare* and *inventory* as keywords in the abstracts of papers in peer reviewed publications during January 1, 2010 and January 1, 2020. After the initial search, 124 papers are identified in the review pool. To keep the literature review scientific and systematic, these papers are searched in Google Scholar and the new papers citing each paper in the pool are checked. The criterion to enrich the pool is: If one paper is published by a publication indexed in ABI/INFORM Collection database and contains the required keywords in the abstract, but cannot be found in the database, it will be included in the review pool. For example, the database indexes the papers in *European Journal of Operational Research* (EJOR) before February of 2017 but does not index the ones after. If a paper citing one of 124 papers is published in EJOR after February of 2017 and contains *spare* and *inventory* in the abstract, it is included in the pool. Overall, 142 papers are in the literature pool for reviewing.

In this review, all 142 papers are classified based on two different groups of perspectives. The first group of perspectives includes the characteristics of spare parts, products, inventory system, and supply chain while the second focuses on the characteristics of research methodologies and topics in the reviewed studies. These two groups of classification perspectives are discussed in Section 2.2 and Section 2.3 respectively. The purpose of this review is to identify the research gaps in spare parts inventory management field from the perspective of supply chain management.

2.2 Typology Based on Systematic Characteristics

In this section, the first typology based on systematic characteristics is presented. The reviewed literature are classified from three perspectives including product and spare part characteristics, spare parts supply chain characteristics, and spare parts inventory management characteristics.

2.2.1 Perspective of product and spare part characteristics

Five product and spare part characteristics including product system type, product lifcycle phase, spare part type, product system complexity, and performance measure are used to depict the studied product systems and spare parts in the literature. In the following contents, these characteristics are presented and discussed in detail. The corresponding classification results are shown in Table 2.1.

Product types		Product lifecycle phases		Spare part types		Product system complexity		Performance measures	
Capital goods	109	Initial phase	2	Repairable parts	41	Single-unit products	60	LCC measure	8
Consumer durable goods	18	Maturity phase	80	Non-repairable parts	83	Multi-unit products	76	TCO measure	2
Non-specific goods	15	End-of-life phase	9	Both	8			SCOR measure	114
		Whole lifecycle	11					Other measures	4

Table 2.1: Number of studies in different product and spare part characteristics

Product types

In the literature on spare parts inventory management, consumer durable goods and capital goods are two important products types which have long lifetime and require after-sales services (Rezapour et al., 2016). Consumer durable goods are referred as the products purchased by individual customers for consumption and not used for the production of another good. Examples of consumer durable goods include automobiles, household appliances, and consumer electronics. Capital goods are high value tangible assets used by a company as an input for producing other goods or services. In the literature, capital goods are also referred as capital assets or capital intensive assets. Examples of capital goods include computer networks, medical and defense systems, and aircraft.

There are several dissimilarities between consumer durable goods and capital goods. First, capital goods directly involves producing other goods or providing services to customers, while the consumer durable goods does not. Secondly, they are supported by distinct after-sales services. For consumer durable goods, warranties are usually provided as after-sales services of consumer durable goods whereas service contracts are provided by the OEM or third party maintenance company as that of capital goods. Discussions on both kinds of after-sales services are shown in the “After-sales services and maintenance strategies” part of Section 2.2.2. Thirdly, these two product have different spare parts demand patterns. The consumer durable goods are purchased by a large number of individuals. This leads to a large variation among the demands of customers because the product usage varies among customers. For capital goods, owners usually purchase a fleet of identical products in use and these products normally have same usage levels and work under same environment so that the variation in service demand of each product does not vary drastically.

The studies on the spare parts of capital goods prevail in this research area. As illustrated in the first column of Table 2.1, 109 out of 142 studies focus on the spare parts of capital goods and only 18 papers consider the spare parts of consumer durable goods. Meanwhile, we cannot identify the types of targeted products in 15 papers.

Product lifecycle phases

The product lifecycle consists of three phases including initial phase, maturity phase, and end-of-life (EOL) phase (Basten and van Houtum, 2014). The initial phase begins when a new product is launched onto the market. The product demand increases sharply during the initial phase and the number of products on market surges as well. In contrast, the product failure due to deterioration is very low, inducing the spare parts demands are not significant during this phase. After the initial phase, the maturity phase starts when the number of products on market gets stable. Such stability is caused not only by the declining product demands, but also by the increasing product failures which also incur increasing spare parts demands. The EOL phase usually starts when the manufacture of products and spare parts stops and ends when the last warranty or service contract period expires (Pourakbar et al., 2012). During this phase, the product sale ceases and the quantity of products on market starts to decrease. However, the products on market still need spare parts to recover functionality when they fail. Therefore, extra units of spare parts need to procure and stock at the beginning of this phase.

Glancing at the whole product lifecycle, spare parts demands follow products demands but with a time lag. Such a phenomena is referred as "lifecycle mismatch" (Solomon et al., 2000) in the literature. Moreover, such a mismatch indicates the spare parts demands are correlated with the quantity of products on market (Inderfurth and Mukherjee, 2008). Figure 2.1 illustrates the relationship between the products sales, spare parts demand, and the products on market over a product lifecycle.

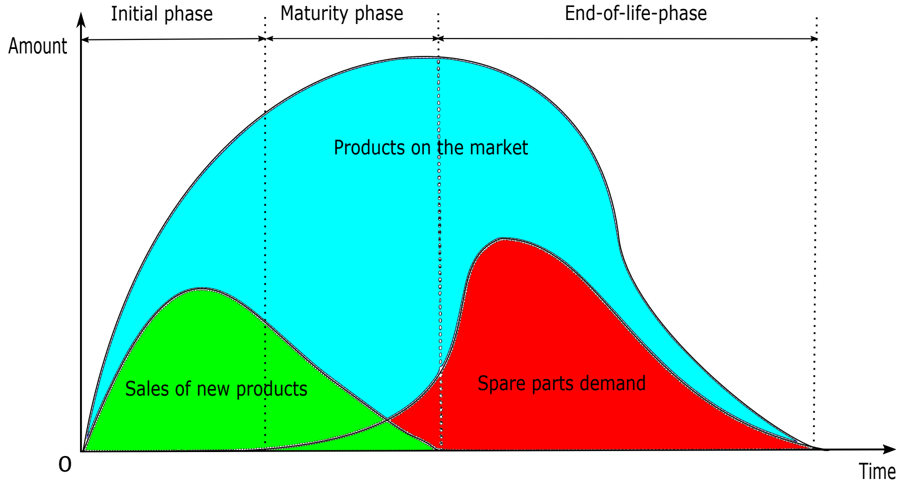

Figure 2.1: New products sales, spare parts demand, and products on market in the product lifecycle (adapted form Inderfurth and Mukherjee (2008))

As shown in the second column of Table 2.1, most of studies in spare parts inventory management literature focus on the inventory problems arising during the maturity phase when the quantity of products on market is stable. There are 9 studies focusing the final order problem or last-time-buy problem occurring at the EOL phase (Pourakbar et al., 2012; Inderfurth and Kleber, 2013; Nguyen et al., 2013; Hur et al., 2018; Li et al., 2018; Behfard et al., 2018; Frenk et al., 2019a,b; Shi, 2019). Moreover, several studies put their efforts on the spare parts inventory decision-makings over entire product lifecycle (Sahyouni et al., 2010; Öner et al., 2010; Liu and Tang, 2016; Duran and Afonso, 2019).

The rest 40 studies cannot be classified because they do not clearly indicate the lifecycle phases they focus on.

Spare part types

The spare part studied in the reviewed literature mainly include two types, the repairable spare parts and the non-repairable spare parts.

Different types of spare parts possess dissimilar characteristics. For instance, some parts are critical such that the malfunctions can lead to product breakdown whilst other parts are not critical, i.e., the part malfunctions do not hinder product function for a short time. To avoid product downtime, more critical spare parts are needed to be stocked to prepare for replacements once failures occur. Therefore, the management operations need to be decided based on the types of spare parts.

In the literature on spare parts inventory management, the studies normally lack uniformly defined definition for spare part types. For the spare parts of capital goods, there is one clear definition given by Arts (2014). Based on maintenance strategies, they introduce three types of spare parts, i.e., rotables, repairables, and consumables. However, this classification may not be applicable to consumer durable goods. In the following part of this subsection, five characteristics of parts including critical level, specialization level, value, demand pattern, and supply source will be discussed.

Parts criticality The criticality of a part is relevant to the consequences triggered by its failure in the process if no replacement is available (Huiskonen, 2001). Based on the critical level, spare parts can be categorized into two types, i.e., critical spare parts and non-critical spare parts. A spare part is deemed as critical when its failure causes product breakdown (Öner et al., 2013). On the contrary, it is non-critical if its failure does not lead to a product failure.

Parts specificity The specificity of a part refers to the specialization level indicating that the part is specifically tailored for and used by a particular type of product. This kind of products is usually customized for certain customers and provides unique functions to meet customer's specific demands and requirements. In the review, this part type is referred as parts with high specificity level. As a contrast, the parts which are widely used by many products are standard parts and referred as parts with low specificity level.

The supply chain characteristics of high specificity parts differ from those of low specificity parts. Due to high demand volumes and economies of scale, low specificity parts are normally supplied and stocked by many suppliers, who are willing to cooperate with manufacturers. Therefore, low specificity parts usually have high availability. On the contrary, high specificity parts possess low demand volumes and have less suppliers. The suppliers are unwilling to stock spare parts because of high obsolescence risks which may lead to low spare parts availability. In this case, manufacturer has to make advance spare parts orders and stock them to meet the demands.

Part specialization level may change during different phases of product lifecycle. For a single part, one can find supply source more easily at maturity phase than at EOL phase. This means the specificity level of parts is dynamic rather than static. For instance, some systems are faced up with the EOL decisions regarding final order placement and spare parts inventory control because the acquisition of parts is no longer guaranteed at EOL phase (Pourakbar et al., 2012).

Table 4.1 introduces an example of automobile parts classification based on criticality and specificity levels. For automobile parts, the specificity level is closely related to the

		Criticality	
		Non-critical	Critical
Specificity	Low	Mufflers	Wheels
	High	Bumpers	Engines

Table 2.2: Examples of automobile parts

features of automobile such as exterior shapes, interior decorates, and engine powers. These features normally vary among the models of different makes and sometimes even among the models of the same make. The criticality level of automobile parts is measured by the consequences of part failures. For instance, the damage in bumpers usually does not hinder the vehicle functions, therefore bumper is one of non-critical part. On the other hand, different car models have dissimilar bumpers because their exterior shape and body dimensions are different. The replacement of a damaged bumper needs to be fulfilled by the dealer or the OEM who sells or produces that automobile model. Therefore, bumpers are classified as a non-critical part with high specificity.

Value The value of spare part is another vital factor influencing the structure of spare parts supply chain. High value spare parts are usually not favoured by neither manufacturers nor suppliers to stock, because holding such spare parts requires high investments in procurement and inventory. Therefore, more collaborative work needs to be done between suppliers and manufacturers to seek other alternatives rather than holding stock to satisfy the demands. For low value spare parts, a trade-off between the stocking and procuring decisions has to be considered by decision makers. If a large procurement quantity is ordered, the inventory level will be so high that corresponding inventory cost increases. Otherwise, the administrative cost might increase because more orders are placed, leading to the increases in ordering and transportation cost. To sum up, replenishment arrangements have to be efficient so that the inventory cost and administrative cost can reach a balance proportion to the value of spare parts.

Repairable v.s. Non-repairable parts The repairable parts are the items which are replaced by new ones and then sent to repair when they fail (Arts, 2014). After being repaired, the items are restocked as ready-for-use units (RFU's) which can be used to replace faulty items in the future. The non-repairable parts are the items which cannot be repaired after replacements and are usually referred as consumables because they are discarded after replacements and the inventories are replenished by procurement from suppliers.

The inventory management of repairable parts are more complex than that of consumables. For repairable parts, supply sources not only are limited to suppliers, but also include repair workshops. The workshop operations are highly relevant to the spare parts inventory management because inventory policies are significantly affected by the factors including repair capacity, repair time, etc. In the literature, the workshop is considered as an important and unique section in the inventory network. Especially, the joint repair shop scheduling and spare parts inventory management problem arises when the capability of repair shop is limited. As shown in the third column of Table 2.1, 41 papers focus on repairable parts only and 3 papers consider both the repairable and non-repairable parts. More than half of the reviewed literature focus on non-repairable parts. There are 10 papers cannot be identified because they do not explicitly show the types of spare parts considered.

Product system complexity

Base on system complexity, the product systems in the literature are specified into two types including single-unit system and multi-unit system. The single-unit system is the product system which has or is assumed to have single critical parts. As a contrast, the multi-unit system has more than one critical parts.

The studies on both kinds of product systems are quite abundant. Among the literature, 60 papers focus on single-unit systems while 75 papers on multi-unit systems. One interesting observation is most multi-unit system studies assume that the failures of dissimilar parts are independent with each other. There are only two studies, Moharana and Sarmah (2016) and Liu and Tang (2016), adopt the dependent failures assumption. In the real world, it is common to see that the failure of one part might induce the failure of others. Therefore, independent failure assumption is not always reasonable though it simplifies the studied problems.

The fourth column in Table 2.1 reveals that there are 60 reviewed studies on single-unit product systems while 70 studies on multi-unit product systems. Note that six papers cannot be classified because they do no involve this issue in their studies.

Performance measures

In the literature, various performance measures are used to evaluate the outcomes brought by different inventory policies, maintenance schedules, etc. Theses performance measures are adopted by inventory network owners to manage the spare parts inventory of different products. In the following context, three different measures including lifecycle cost, total cost of ownership, and supply chain performance metrics are introduced and the reviewed literature are classified based on these measures.

Lifecycle cost The general definition of lifecycle cost (LCC) is referred as the summation of all cost components to manufacturers, users, society during the product lifetime. The LLC can be decomposed into categories in a cost breakdown structure as shown in Table 2.3 which is adapted and developed from Asiedu and Gu (1998). Table 2.3 reveals that OEM, users, and society are faced up with four kinds of cost categories, i.e., design costs, production costs, usage costs, and disposal or recycling costs. The new element we add to the cost breakdown structure proposed by Asiedu and Gu (1998) is that OEM may also have disposal or recycling costs if the reverse logistics (RL) is implemented. Nowadays, the concept of RL is getting popular in various industries such as electronics, automotive, and consumer appliances. The RL involves the operations related to the return of damaged, unsold, end-of-life products along with handling, consolidation, remanufacturing and disposal (Diabat et al., 2015). Therefore, the RL operations induce

	OEM cost	Users cost	Society cost
Design	Product development Market Recognition		
Production	Materials Salaries, wages etc. Facilities Energy		Wastes Pollution Health damages
Usage	Inventory Transportation Wastes Breakage after-sales services	Inventory Transportation Energy Materials Maintenance	Packaging Wastes Pollution Health damages
Disposal or recycling	Disposal Recycling Remanufacturing	Disposal Recycling	Wastes Disposal Pollution Health damage

Table 2.3: Lifecycle stages and costs

new cost categories which have to be considered when conducting lifecycle cost analysis for the OEM. Detailed discussions on the relationship between RL and spare parts inventory management will be shown in Section 2.4.2.

When analyzing the LCC of a product, it is important to identify the listed cost categories based on the analysis subjects. To be specific, the LCC may contain different cost categories when the LCC analysis is carried for OEM, users, or society respectively, because they are only interested in the costs categories that they can control. For example, Öner et al. (2010) study a problem in which an OEM is providing performance based contract (PBC) to the customers who purchase a set of equipment. The LCC considered in their study includes design costs, production costs, and usage costs including spare parts inventory costs, repair costs, and downtime costs. In general, the LCC is a widely used economic measure to be considered when making decisions at the design phase of products and relevant services. When using LCC in the studies, appropriate cost categories should be included in the LCC analysis based on the analysis subjects and their interests.

Total cost of ownership Total cost of ownership (TCO) is a concept which helps customers to understand the true cost of buying a particular good or service from a particular supplier (Ellram, 1995). This concept can be applied to any products to assist purchase decision-makings. Except product price, the TCO normally contains the costs incurred by product maintenance, downtime, disposal, and other activities happened during the periods starting from making purchasing decisions until disposing the products.

It should be highlighted that the maintenance related cost may account for a huge portion (70% - 80%) of the TCO for some complicated technical products used in industries such as military, medical, and power generation. In this case, spare parts inventory management is important due to its critical role in maintenance operations. The studies adopting TCO as the performance measure for evaluating spare parts inventory system generally seek to find optimal inventory policy to minimize the TCO.

Supply chain performance metrics Supply chain performance metrics are widely used to measure supply chain performances in the literature. One of the most popular supply chain performance metrics is Supply Chain Operations Reference (SCOR) model. This model advocates measuring the performance of supply chain from four perspectives, i.e., lead time metrics, cost metrics, service/quality metrics, and assets metrics (Pfohl and Ester, 1999). In this review, SCOR performance metrics are applied to classify spare parts inventory literature. The example metrics of the mentioned four kinds of metrics are shown in Table 2.4.

The number of reviewed studies using different performance measures are listed in the last column of Table 2.1. It is worth mentioning that multiple performance measures can be used to evaluate system performance simultaneously. For example, an OEM who provides performance-based contracts to customers has to not only consider the cost metrics of maintenance and inventory system, but also ensure the requested service metrics in contract to be satisfied (e.g. product availability level). In this case, a study may use

Metrics	Example
Lead time metrics	Spare part replenishment lead time, repair lead time for repair operation, etc.
Cost metrics	Inventory holding cost, repair cost for repairable spare parts, shipment cost, etc.
Service/quality metrics	Product's availability level, inventory system service level, etc.
Assets metrics	Spare parts inventory level, repair center capacity, etc.

Table 2.4: Different kinds of supply chain performance metrics

the total cost of managing spare parts as an objective and consider the agreed availability level as the constraints when establishing mathematical models. Consequently, we have to consider the number of usages of different performance measures in the reviewed literature. From the last column of Table 2.1, it can be easily concluded that supply chain performance metrics are dominantly used in the literature. In contrast, there are 8 studies using the LCC measure and 2 studies using the TCO measure respectively. This phenomena reveals that the studies so far lack a product lifecycle view. There are four studies using other measures such as company profits of selling products and providing warranty services (Ahiska et al., 2017; Rezapour et al., 2016), profits of using spare parts inventory pooling among different companies (Zhao et al., 2019), and warranty cost (Li et al., 2018).

2.2.2 Perspective of spare parts supply chain characteristics

The spare parts supply chain refers to the network of firms and facilities involving in transforming raw materials to spare parts and distributing these spare parts. The supply chain characteristics reveal the stakeholders, structure, operation mode, and flow paths of the spare parts supply chain and they include network ownership, number of echelons, lead times, lateral transshipments and emergency shipments, and after-sales services and maintenance strategies. The obtained results are shown in Table 2.5.

Network ownership		Number of echelons		Lead time		Lateral transshipments and emergency shipments		Maintenance strategies	
User network	49	Single-echelon	90	Zero lead time	21	Lateral transshipments	5	Preventive maintenance	31
OEM network	55	Multi-echelon	32	Deterministic lead time	59	Emergency shipments	17	Corrective maintenance	86
Third party network	28			Stochastic lead time	38	Both	2	Both	20

Table 2.5: Number of studies in different spare parts supply chain characteristics

Network ownership

The spare parts supply chain supports the maintenance activities for the products in use. Based on the ownership, the spare parts supply chain can be identified as three types including the supply chain network of the OEM or system integrator (the firm bringing component subsystems into a whole and ensuring those subsystems function together) who provides after-sales services to the product systems they have sold, the network of users who maintain the product systems that they use (Basten and van Houtum, 2014), and the network of third-parties who are outsourced by the users for maintaining the product systems in use. In the following context, these three types of networks are referred as OEM network, user network, and third-party network respectively.

Different types of networks represent the different patterns of maintenance activities. User network is a traditional network for maintaining the products in use and still quite popular in the industries such as military and transportation. In user network, users take over the maintenance activities, aiming to avoid product downtime. Therefore, users have to make a trade-off between the spare parts inventory cost, repairman labor cost, products downtime cost and so on.

As the product structure and maintenance complexity increase, the OEM network starts to prevail in many industries such as high-tech industries because the OEM generally owns more knowledge and techniques and thus can handle the complexity of product system more easily than users. In addition, the prevalence of lean management programs is another driver of this trend. The philosophy of lean management requires the users have smaller buffers for disturbances and also high system availability. In this context, rather than reserving the teams and resources to maintain the product systems in use, users prefer outsourcing the maintenance to the OEM or a third-party company.

In OEM network, OEM establishes facilities and resources to provide after-sales services to users through service contracts such as warranties and service contracts which will be further discussed in the part of "After-sales services and maintenance strategies" in Section 2.2.2. Nowadays, OEM is increasingly willing to provide after-sale services due to following reasons. Firstly, providing such services induces a competitive advantage to other OEMs since customers increasingly require high quality after-sales services; secondly, OEM can earn profits by selling such services, which is at least as high as that by selling products (Oliva and Kallenberg, 2003).

In some cases, a third-party company takes the role of OEM in network to provide after-sales services to users to earn profits and such network is referred as the third-party network. In some ways, this network is similar to the OEM network. For example, both OEM and third-party company own and manage the facilities and resources to support the after-sales services. However, there are some differences between the OEM network and the third-party network. First, in the OEM network, the OEM is able to obtain the feedback from after-sales services to improve products through the designing-for-maintenance or designing-for-lifecycle-cost approach (Basten and van Houtum, 2014). On the contrary, third party companies do not benefit from this when they perform such services. Second, in third-party network, after-sales services usually are provided to a much bigger user groups who use the products from different OEMs while in the OEM network, they are only provided to the products from a particular OEM.

In the first column of Table 2.5, the number of studies adopting networks with various ownership are summarized. Compared to the studies using user networks or OEM networks, the studies using third-party networks are less. In addition, there are 10 papers whose network types cannot be identified because this information is not provided.

Number of echelons

The number of echelons is one of the most important characteristics in the inventory management literature, because it reflects the inventory system structure. The single echelon inventory system contains single stock point in one particular region or several local stock points in different regions. Each local stock point is responsible for serving the demand in the region where it locates and there is no interaction between any two local stock points in different regions. In this context, ordering and replenishing decisions are made at each stock point independently.

The multi-echelon inventory system have more than one echelon, i.e., national and regional stock points are utilized. The national warehouse orders from suppliers, stocks inventories, and replenishes the regional stock points which are responsible for meeting corresponding demands. Spare parts inventories are located at every stock points in the inventory system. In this system, systematic ordering and replenishing decisions need to be made based on inventory and demand information at each stock point.

The second column in Table 2.5 shows that a huge part of reviewed studies focus on the single-echelon inventory system for managing spare parts. Among 142 reviewed papers, only 32 consider the spare parts inventory management problems with multi-echelon inventory systems. There are 20 papers cannot be classified because the structures of inventory systems are not clearly indicated.

Lead time

Lead time is another important spare parts supply chain characteristic and is usually referred as the period begins with an order placement and ends with the receipt of corresponding order. In the literature, different assumptions regarding order lead time are made including zero lead time, deterministic lead time, and stochastic lead time. In addition, the lead time concept varies among different spare part types. For non-repairable spare parts, the lead time mainly includes the transportation time, while the repair time of faulty repairable parts is also included in lead time.

Note that different inventory operations may involved different lead times. For example, if emergency replenishment is allowed in the studied inventory system, the lead time of emergency procedure is much shorter than that of normal procedure but the emergency one usually costs more. Therefore, one has to make a trade-off between placing normal or emergency orders and reducing cost of product downtime. As shown in the third column in Table 2.5, 21 studies do not consider lead times, while 97 studies do: 59 papers with deterministic lead times and 38 with stochastic ones. In addition, 24 studies cannot be identified because they do not show this characteristic.

Lateral transshipments and emergency shipments

Lateral transshipments and emergency shipments are widely used to improve inventory system performance in practice. The reviewed literature is classified based on if lateral transshipments or emergency shipments are used.

Lateral transshipments Lateral transshipments are defined as the stock movements between different stock points within the same echelon in an inventory system (Basten and van Houtum, 2014). Such stock movements are executed under inventory pooling strategy, referred as the arrangements in which different stock points share their inventories (Wong et al., 2007). In some studies, inventory pooling and lateral transshipments are treated as interchangeable strategies (Wong et al., 2006). However, we would like distinguish inventory pooling from lateral transshipments because the first is a inventory strategy while the latter is one type of shipping operations. Even though inventory pooling strategy is realized through lateral transshipments to move stocks between different stock points, they are not same in concept. In the following context, a brief introduction to different types of lateral transshipments are presented at first. Afterwards, lateral transshipment strategy used in the spare parts inventory management literature will be discussed.

Proactive lateral transshipment v.s. reactive lateral transshipments

In the inventory management literature, one key criterion to classify the studies is based on the timing when transshipment decisions are made. Specifically, the literature considering lateral transshipments can be categorized into two main streams, the studies with proactive transshipments and the studies with reactive transshipments (Paterson et al., 2011). In proactive transshipments, all lateral transshipments are scheduled in advance and all stock points are redistributed simultaneously. In contrast, reactive transshipments refer to the stock movements from one stock point which has sufficient on-hand stock to another which faces a stock-out situation at anytime.

Partial pooling v.s. complete pooling

Another way to classify the literature regarding lateral transshipments is based on the extent to which the stock at one point can be used in transshipments to other stock points. If all inventory at one stock point can be used, then we call such transshipments as complete pooling. In contrast, if only a part of the inventory at each stock points can be shared with other points and the rest inventory is reserved for covering future local demands, then we call such transshipments as partial pooling.

In the literature, reactive transshipments using complete pooling is widely used because the transshipment costs is usually significantly less than the sum of stock holding cost and the shortage cost when the demands cannot be fulfilled immediately (Paterson et al., 2011). In fact, we indeed find that all studies involving lateral transshipments in spare parts inventory system use reactive transshipments with complete pooling. Therefore, in this review, we restrict our attentions to the reactive transshipments with complete pooling and will refer it as transshipments for convenience.

Emergency shipments Except lateral transshipments, emergency shipments are also widely used when on-hand spare parts stocks are not sufficient to support in-time maintenance activities. In this case, a normal order will be too late such that the downtime of product system will incur a huge cost. Unlike lateral transshipments which redistribute stocks among stock points within the same echelon, emergency transshipments is referred to the emergency and unplanned stock movements between the stock points at different echelons, e.g. the stock-out points and the inventory supply source (central warehouse or suppliers).

As shown in fourth column of Table 2.5, the discussion on implementing lateral transshipments and emergency shipments is a big gap in spare parts inventory management literature. There are only 5 studies considering lateral transshipments, 17 studies considering emergency shipments and 2 studies considering both. The rest 118 papers usually ignore or are irrelevant to this issue.

After-sales services and maintenance strategies

Product after-sales services are highly relevant to spare parts inventory management. Such services are regulated by the agreed contracts between product owners and maintenance providers, and are realized by different maintenance strategies. In the following content, two types of after-sales services, warranty and service contracts, are discussed and several maintenance strategies are illustrated.

After-sales service contracts As discussed previously, the after-sales services of consumer durable goods are provided through the warranty contracts, which request OEM to fix or replace faulty products under some conditions. Warranty contracts are important strategic profiles in modern manufacturers because they can improve company image, generate business and profits, and foster brand (Martinez et al., 2007).

The after-sales services of capital goods are often provided through service contracts which includes two main types including Material Based contract (MBC) and Performance Based contract (PBC) (Mirzahosseinian et al., 2016). MBC used to be the most common mechanism to provide maintenance services in the industry. In MBC, each time a service task is completed, OEM is compensated for the service cost. Lately, PBC has emerged as a new service mode redefining the acquisition, operation, and maintenance of capital equipment. Under PBC, service provider is compensated based on the system performance. Compared to MBC, PBC has advantages in motivating OEM to improve product quality and reliability during design and manufacturing phase so as to reduced failures and repair costs of the sold products.

Maintenance strategies Maintenance strategies determine the timing when products are maintained or parts are replaced. Preventive maintenance and corrective maintenance are two most discussed maintenance strategies in the literature.

Preventive maintenance strategy intends to maintain products or replace parts before the failure occurs to avoid product breakdowns. One important assumption in preventive maintenance is that the product state which subjects to stochastic failure is always known with certainty (Cho and Parlar, 1991). Based on the known state, preventive maintenance activities can be scheduled in the planning periods and induce planned demands for spare parts. Preventive maintenance strategy can be divided into usage based maintenance and condition based maintenance. Usage based maintenance is implemented when the usage of parts reaches a target threshold. A good example of usage based maintenance is the automobiles are usually maintained after a certain mileage. Condition based maintenance is implemented when the condition of part reaches certain states. The condition of one part can be either checked periodically via inspections or monitored continuously via sensors. However, preventive maintenance cannot eliminate part failures because parts might fail before the replacements or maintenance. In addition, some types of spare parts, such as electronic parts, do not wear such that states of parts cannot be used to decide the maintenance schedule. In this case, corrective maintenance strategies are used when the uncertain part failures occur, which would incur unplanned demands for spare parts. For example, the automobile engine may fail even though scheduled preventive maintenance are implemented.

The last column In Table 2.5 shows the numbers of studies under preventive and corrective maintenance are summarized. It is clear that a huge number of studies focusing on corrective maintenance while 31 studies focusing on preventive maintenance. There are 20 papers consider the spare parts inventory problem under both preventive and corrective maintenance. Five papers cannot be identified based on this topic.

2.2.3 Perspective of spare parts inventory management characteristics

The characteristics of spare parts inventory management discussed in this review include inventory policy, number of inventory units, and inventory supply source.

Inventory policy

During different phases of product lifecycle, various policies can be implemented to control spare parts inventory. In the following context, main types of continuous and periodic review policies used during the initial and maturity phases, and final order policy used in the EOL phase are discussed in detail.

Continuous review policy Under continuous review policy, spare parts inventory level is inspected in a continuous manner. New spare parts are ordered when inventory level falls below a certain level. Based on how the order quantity is decided, one can identify two types of continuous review policies including (s, S) policy and (q, r) policy.

In (s, S) policy, a new order is placed to make inventory level reach the order-up-to level (S) once the level falls below reorder point (s). The ordered quantity is equal to the difference between the order-up-to level and the inventory level at the ordering time. $(S-1, S)$ policy (which is also referred as one-for-one replenishment policy or base stock policy) is a special case of (s, S) when the reorder point takes the value of order-up-to level minus 1. $(S-1, S)$ policy is widely used in the repairable spare parts inventory management studies because the one-for-one replenishment mode can mimic the repair process. For example, when a product is under repair, inducing one unit demand for a repairable spare part, a new spare part is used for replacing the faulty one, which is sent to repair and will return as one unit of inventory after repair. Unlike (s, S) policy whose order size is uncertain, (q, r) policy has a fixed order size (q) and the order is placed when inventory level is no higher than reorder point (r).

Periodic review policy Under periodic review policy, replenish orders are placed at the beginning of each order cycle. The most widely used periodic review policy in the literature is (R, S) policy, in which the order is placed at the start of every fixed ordering cycle (R) to make the inventory level reach the order-up-to level (S), which is decided based on the predicted demand during next cycle and lead time. The ordering cycle is usually predetermined by decision makers as a fraction of year, or a certain amount of weeks or months, based on their experiences and preferences. More important, the implement of continuous review policy or periodic review policy is mainly decided by the decision makers based on how they manage spare parts inventory in practice. In addition, it is possible to adopt both types of inventory policies in an particular inventory network. For example, in some studies, different inventory policies are used in different echelons of a multi-echelon inventory network (e.g. Topan and Bayindir, 2012).

Final order policy The final order for spare parts is placed to satisfy the demand during EOL phase of product lifecycle Pourakbar et al. (2012). In the literature, it is also referred as EOL inventory problem, end of production problem (EOP), or final buy problem (FBP). The key to solve this problem is how to decide the optimal final order quantity. If too many spare parts were ordered, inventory holder would take huge obsolescence and disposal risks and pay for a high inventory holding cost at EOL phase.

On the contrary, if the final order quantity was not enough to cover the spare parts demand, the OEM or maintenance provider would not fulfill the service contracts or warranty obligations, faced up with fiscal penalties or damages in customer satisfaction or brand image.

Table 2.6 shows the number of studies on managing spare parts inventory though different policies. One should note that a study may involve different inventory policies. The studies of Topan and Bayindir (2012), Bacchetti et al. (2013), Çapar (2013), Inderfurth and Kleber (2013), Panagiotidou (2014), Hu et al. (2017), and Duran and Afonso (2019) consider more than one inventory policies to manage spare parts inventory and these studies are counted in more than one study group in Table 2.6. In addition, 2 papers adopt other inventory policies and they are not included in the table. Liu and Tang (2016) propose a base cumulative order size (BCOS) policy and Sahba et al. (2018) propose a Multilevel rationing (MR) policy for spare parts inventory management. There are 28 papers which cannot be identified and classified based on the inventory policy.

Initial and maturity phases				EOL phase
Continuous review policy			Periodic review policy	Final order policy
(s, S)	$(S-1, S)$	(q, r)	(R, S)	
12	51	12	36	7

Table 2.6: The number of studies using different inventory policies

Number of inventory items

From the perspective of inventory item numbers, the reviewed literature can be categorized into two groups: the studies with single-item inventory and the studies with multi-item inventory. In the first group of studies, only one type of spare parts is considered while the second group of studies consider more than one spare part types.

It is worth mentioning that the single-item and multi-item inventory classification is highly relevant to the approach in which how the target service level of inventory system is defined. Basically, there are two approaches including item approach and system approach. In the item approach, each individual part is defined with a target service level. As a result, ordering decisions are independently made for each part. In the system approach, a target service level is defined for the demand weighted average of all the performance measures over all parts (Topan et al., 2017).

Among the literature, the studies on multi-item spare parts inventory system are prevalent, with 85 papers focusing on this characteristic. Additionally, there are 44 studies on single-item inventory system. Even though single-item system is not popular in practice, corresponding studies could still be valuable if more factors affecting spare parts inventory management are considered. In addition, there are 13 studies cannot be classified because they do not cover this issue.

Inventory supply sources

The last characteristics regarding inventory system is the supply source. Both consumable and repairable spare parts are normally procured from external suppliers or manufactured by OEMs. However, it is possible that repairable spare parts are supplied from repair shops through fixing faulty ones. The identification of supply source is important because different supply modes represent various emergency levels of orders. For example, when the inventory are depleted and orders cannot be backlogged, ordering new spare parts from external suppliers are more preferable if lead time is short because waiting for repairing a faulty part possibly cost more (e.g., penalties on long customer waiting time beyond the agreed level in the service contracts).

From Table 2.7, it can be concluded that the supply sources in most studies are suppliers and repair shops.There are only 8 studies have some other interesting supply modes. In the study of Rezapour et al. (2016), repair shop is responsible for fixing repairables and suppliers are sending orders of non-repairables. Ahiska et al. (2017) study a problem in which OEM produces and repairs or remanufactures spare parts to support the after-sales services of products sold. Togwe et al. (2019) focus on a case in which a portion of spare parts inventory comes from additive manufacturing while the rest is ordered from suppliers. In the study of Song and Yang (2015), spare parts are supplied from three sources, repair shop, supplier, and cannibalization (an operation attempting to achieve a maximum number of operative equipment by interchanging components in failed ones). Dreyfuss et al. (2018) consider both cannibalization and supplier as inventory supply sources.

Suppliers	OEM production	Repair shop	Others
52	5	72	5

Table 2.7: Number of studies with different supply sources

2.3 Typology Based on Research Methodologies and Topics

In Section 2.2, the reviewed studies are categorized based on the physic characteristics of the studied products, spare parts, and supply chain networks. These characteristics are used to help readers understand the backgrounds and contents of the studies. In this section, another typology is introduced to classify the studies based on research methodologies and topics. Corresponding analysis is carried out to check current research status, identify research gaps, and find potential topics. This typology is constructed based on the supply chain analytics and the research topics in the studies.

2.3.1 Perspective of research analytics

Spare parts supply chain refers to the network of firms and facilities involving in transforming raw materials to spare parts and distributing spare parts to customers. The studies on spare parts inventory management should be carried out under the big picture of spare parts supply chain. In this subsection, the studies of spare parts inventory management are classified based on their research methodologies. We borrow the concept of supply chain analytics to identify different research methodologies. Supply chain analytics can be further divided into three different types including descriptive, predictive, and prescriptive analytics. It is worth mentioning that one study may use multiple analytics.

Descriptive analytics

Descriptive analytics answer the question of what is happening by utilizing and deriving the information from significant amount of data (Souza, 2014). Spare parts inventory management studies adopting descriptive analytics mainly focus on particular cases, conduct simulations, and perform performance analysis based on the collected historical data. The research issues are typically related to spare parts supply chain network structure design, spare parts classification, and new technology adoption in a particular industry. The numbers of research works adopting descriptive analytics in different forms is shown in Table 4.2.

Case study	Simulation	Performance analysis
49	19	7

Table 2.8: Number of studies with different descriptive analytics methods

The table shows that 49 studies have real-life cases relevant to spare parts inventory management problems. There are 19 papers using simulation models to mimic the real-world spare parts inventory management processes. The purpose of using simulation models as a descriptive method is to reveal the value of implementing certain policies when managing spare parts inventory. Using such simulation experiments is preferable because the cost of testing policies in real-life management operations could be high. The performance analysis are used in 7 papers for testing or comparing different spare parts demand classification or forecasting methods.

Predictive analytics

Predictive analytics answer the question of what will be happening by utilizing historical data in prediction techniques. In the spare parts inventory management literature, the studies adopting predictive analytics normally address the issues of how to identify different spare parts demand patterns and how to correctly forecast the demands.

Demand patterns The spare parts demands are induced by the failures and replacements of parts in use during the product lifecycle and normally in intermittent and lumpy demand patterns (Boylan and Syntetos, 2010). Such demand patterns cause troubles in the demand forecasting, thus leading to difficulties in choosing inventory policies.

In the literature, various techniques are used to depict the intermittent and lumpy spare part demands based on historical data. Traditionally, spare part demands are often modelled by stochastic processes including stationary and non-stationary processes. Stationary stochastic process assumes the demand possesses same probability distribution in different time periods. This assumption holds when the product installed base is large and constant over planning horizon. The most widely used stochastic demand pattern is Poisson process which assumes that the duration between any two consecutive failures are independent random variables with identical exponential distributions. On the contrary, non-stationary pattern assumes that demand possesses different probability distributions in distinctive periods. This assumption holds when the number of target products changes over planning horizon. One typical non-stationary demand pattern is non-homogeneous Poisson process.

The selection between stationary and non-stationary demand processes is affected by the product lifecycle phase on which the study focuses. If the study only targets on the products in maturity phase, stationary demand process is preferred because the number of products in use remains at a stable level thus leading to a steady demand rate. However, the demands are not stable neither at the initial phase nor the EOL phase. At initial phase, the number of products in use is building up quickly, inducing an increasing demand rate. On the contrary, the demand has a decreasing rate in the EOL phase due to the shrinking number of products in use. Therefore, non-stationary demand process should be adopted to capture the changes in demands over different phases of product lifecycle.

Another popular approach used for depicting the demand patterns is to assume the demand follows a general or certain probability distribution which is determined by the information retrieved from historical data. An advantage of this approach is that nearly all types of demand patterns can be modelled based on historical observations of spare parts demands. However, such general probability distributions may bring more complexities into the optimization problems such that the problem is hard to solve and structure results may not be guaranteed. Apart from the studies assuming stochastic demands, a few studies assumes deterministic spare part demands. In these studies, more efforts are put on determining the optimal inventory policy. However, such deterministic models could not capture the stochastic nature of part failure process.

The studies using different spare parts demand patterns are listed in Table 2.9. It is clear that a majority of the reviewed studies adopt stationary demand patterns. The prevalence of stationary demand pattern reflects that spare parts inventory management literature is developed from and heavily impacted by the equipment maintenance literature in which the stationary demand patterns are widely accepted as a basic assumption. However, in recent years, more and more studies start to focus on non-stationary demand patterns when the spare parts inventory management issues are considered under different lifecycle phases. In addition, 20 papers assume the spare parts demand patterns follow general probability distributions. 24 studies cannot be identified based on how the demand patterns are modelled.

Single demand pattern v.s Multiple demand patterns As mentioned previously, various demand patterns are used for forecasting spare parts demands and different spare

Stochastic			Deterministic
Stationary	Non-stationary	General distribution	
74	10	20	4

Table 2.9: Number of studies with different demand patterns

parts types may possess different demand patterns. Therefore, based on the number of spare parts demand patterns used, the reviewed studies using prescriptive analytics can be classified into two groups, one with single demand pattern and the other with multiple demand patterns.

The studies with single demand pattern usually consider one type of spare part (could be multiple identical spare parts). However, in the studies with multiple demand patterns, multiple non-identical spare parts belonging to different types are considered and each spare part possesses its individual demand pattern. In addition, for the complex systems with multiple parts, one spare part demand may be affected by the demands of other spare parts, because such systems are often subject to multiple statistically dependent failure processes. In this case, dependencies between spare part failures have to be considered if the failure of one spare part may lead to the failures of others. Only two papers, Moharana and Sarmah (2016) and Liu and Tang (2016), consider the dependency among different spare parts demand patterns while others assume different spare parts demands are independent. The number of studies with single demand pattern and multiple demand patterns are shown in Table 2.10 respectively.

Multiple demand patterns		Single demand pattern
Dependency considered	Dependency not considered	
2	40	66

Table 2.10: The number of studies with single demand pattern and multiple demand patterns

Prescriptive analytics

Prescriptive analytics answer the question of what should be happening. Based on descriptive and predictive analytics, prescriptive analytics utilize optimization models to provide decision makers with recommendations. Many studies adopt prescriptive analytics to investigate the inventory policies for managing spare parts in different industries to assist decision makers achieving favourable inventory system performance.

In this section, 86 studies using prescriptive analytic methods are examined. Among these studies, various optimization models are proposed to solve the problems on spare parts inventory management. To provide readers a clear and direct view on optimization models in the studies with prescriptive analytics, we divide the studies from three aspects, i.e., model settings, optimization model types, and solution methods.

Model settings Based on model settings, the mentioned 86 papers can be classified into two groups, one with stochastic settings and the other with deterministic settings. The key characteristic to distinguish studies with stochastic settings from those with deterministic settings is how the spare parts demands are described. To be specific, if a model assumes that spare part demand is stochastically distributed or follows certain stochastic process, then we say that this model uses a stochastic model setting.

The numbers of studies using different model settings are illustrated in Table 2.11. It is clear that the studies adopting stochastic settings dominate: 81 out of 86 papers use stochastic settings in their optimization models while only 5 papers use deterministic settings. From this results, we can see that stochastic settings are widely accepted in this research field. This phenomena reveals that studies in this filed aim to manage spare parts inventory under the uncertainties embedded in spare parts demands.

Stochastic setting	Deterministic setting
81	5

Table 2.11: Number of studies with different model settings

Optimization model types Based on the types of optimization models used, the reviewed studies can be classified into the studies using stochastic models and the studies using deterministic models. It should be highlighted that a model with stochastic setting is not necessarily a stochastic optimization model. As mentioned in previous part, the reviewed studies usually adopt stochastic model settings. However, the studied problems can be modelled either as deterministic models or stochastic optimization models. In this section, the classification focuses on optimization models types rather than model settings. For example, Arts (2017) considers a repairable spare parts inventory problem with a stochastic setting, i.e., the demand for each part is a Markov modulated Poisson process. However, he formulates a non-linear non-convex integer programming. In this case, we classify this paper as a study with a deterministic optimization model under a stochastic setting.

Main types of stochastic optimization models include stochastic dynamic programming (Markov decision process) models, stochastic programming models, and robust optimization models. On the other hand, deterministic optimization models consist of linear programming models, non-liner programming models, and mixed integer (linear and nonlinear) programming models. Based on this classification scheme, the mentioned studies are classified into these two streams as shown in Table 2.12. In the table, it is clear that the studies using deterministic optimization models account for a majority. There are 22 studies using stochastic optimization models for studying spare parts inventory management problems. Among them, 19 studies use stochastic dynamic programming models as modelling technique, 2 studies using stochastic programming models, and 1 study using robust optimization model. The diversity in modelling techniques can facilitate the development of this research area and bring more new innovative topics. Therefore, we call on more studies to propose various stochastic optimization models to enrich this research area.

Stochastic optimization model			Deterministic Optimization model
Stochastic dynamic programming	stochastic programming	Robust optimization	
19	2	1	64

Table 2.12: Numbers of studies with different types of optimization models

Solution methods Faced up with different optimization problems with various degrees of complexity, the reviewed studies propose different methods solve the problems. The reviewed studies are further divided based on four types of solution methods, i.e., exact methods, approximation methods, heuristic methods, and simulation methods.

The exact methods are able to find optimal solutions to optimization problems. Analytic methods and large scale optimization methods such as branch-and-bound, Lagrangian relaxation, Benders decomposition and so on are considered as exact methods because these methods can theoretically find optimal or close-to-optimal solutions with guaranteed upper and lower bounds. The approximation methods are adopted when the performance measure in the studied problem is hard to be evaluated. This type of solution methods is widely used to solve stochastic optimization models. The heuristic methods are used when the problems are high in complexity. Heuristics usually are simple and easy to be implemented but optimal solutions cannot be guaranteed. Simulation methods are defined as using simulation experiments to find the best input variable values among all possibilities without explicitly evaluating each possibility (Carson and Maria, 1997). The simulation optimization methods indicate the impacts of changes in the interested variables on system performance but its advantages diminish when a large complex system is considered.

The reviewed studies adopting different solution methods are presented in Table 2.13. It reveals that exact methods and heuristic methods are two most widely used methods to obtain solutions in the reviewed literature. Note that more than one solution methods are proposed in some studies. In this case, the studies with multiple solution methods are included in multiple categories.

Exact methods	Approximation methods	Heuristic methods	Simulation methods
46	6	33	4

Table 2.13: Number of studies with different solution methods

2.3.2 Perspective of existing research topics

The state-of-the-art study topics in the reviewed literature are quite abundant because spare parts inventory management process involves various cooperative business activities to handle a huge amount of spare parts among different participants in a supply chain network. In this section, the reviewed studies are further classified based on four research topics including spare parts inventory system, joint maintenance and inventory optimization, supply chain network design or supply chain policy performance evaluation, and spare parts classification and demand forecasting.

Spare parts inventory problem

The studies focusing on spare parts inventory problem seek to find optimal policies managing inventories so that certain economic objectives are achieved while service performance is guaranteed. This problem is important to solve because the policies governing final product inventories are not applicable to manage spare parts inventories (Kennedy et al., 2002). In addition, repairable spare parts inventory problem is more complex than non-repairable spare parts inventory problem because repair operations and resources have to be taken into consideration.

Joint maintenance and inventory optimization

Spare parts inventories are held for satisfying the demands which are generated by the corrective and preventive maintenance. Due to the high interconnections between maintenance and spare parts inventory management, the companies who are responsible for both operations should consider and optimize these two problems simultaneously (Van Horenbeek et al., 2013). In the studies with this topic, efforts are made to investigate the performance of different combinations of inventory policies and maintenance strategies. It is worth mentioning that there are several studies considering the service contracts design for after-sales services provider (e.g., Mirzahosseinian et al. (2016); Mo et al. (2016); Li et al. (2018); Zhao et al. (2019)). In this case, inventory and maintenance decisions are made simultaneously to minimize contract cost.

Supply chain network design or supply chain policy performance evaluation

The studies relevant to network design or policy performance evaluation mainly introduce new operations or technologies to manage spare parts inventory and evaluate the resulting performance. There are several interesting topics such as reuse spare parts supply chain design (Abdallah et al., 2012; Diabat et al., 2015), spare parts additive manufacturing (Togwe et al., 2019; Zhao et al., 2019), repair shop design (Turan et al., 2018), etc. These studies provide theoretical supports to implement new techniques and operations in practice.

Spare parts classification and demand forecast

As mentioned previously, spare parts demand is hard to forecast because the demand pattern is usually intermittent and lumpy. In addition, different types of spare parts may possess different demand patterns. The purpose of classification and demand forecast is to identify spare parts classes, which is used in forecasting spare parts demands, and finally help deciding inventory policies.

In the reviewed literature, the studies considering spare parts classification or demand forecast to improve spare parts inventory management are abundant. There are 10 papers focusing on developing various multiple-criteria classification schemes for spare parts inventory control in different industries. It is noticeable that classification schemes help implement proper inventory policy to spare parts of different classes. For example, Bacchetti et al. (2013) propose a hierarchical multiple-criteria classification method to manage spare parts inventory of household appliances. Different forecasting methods and inventory policies are used to manage 12 spare parts classes.

There are 18 papers on spare parts demand forecast and many methods including Croston's method, Syntetos–Boylan Approximation, Single Exponential Smoothing (SES), and bootstrapping methods are proposed. To integrate spare parts demand forecast and inventory management, forecast methods should be evaluated through inventory management metrics instead of performance metrics which are directly related to the forecasting results (e.g., mean squared error) (van Wingerden et al., 2014). For instance, Zhu et al. (2017) use extreme value theory to forecast spare parts demand and the method are evaluated by inventory performance. In addition, forecast methods should be able to utilize more information from maintenance activities, which is highly relevant to spare parts demand. Van der Auweraer and Boute (2019) combine the failure behaviour of parts and the maintenance plan of equipment to predict spare parts demand.

Table 2.14 presents the numbers of studies with different topics. It can be concluded that the mentioned four study topics cover a big portion of related literature. The two most studied topics are "Joint maintenance and inventory optimization" and "Spare parts inventory problem", followed by "Net work design or policy performance evaluation" and "Spare parts classification and demand forecast". In addition, there are 11 papers cannot be categorized into the mentioned topics. Among them, 7 papers are review papers and the rest papers focus several topics such as repair kit problem (Bijvank et al., 2010), product reliability design problem (Öner et al., 2010), costs/benefits allocation among companies using cooperative pooling inventory (Karsten and Basten, 2014; Guajardo and Rönnqvist, 2015; Wang and Yue, 2015).

To further reveal the developing trend in study topics, a histogram in which the

Spare parts inventory problem	Joint maintenance and inventory optimization	Network design or policy performance evaluation	Spare parts classification and demand forecast	Others
33	2	24	28	11

Table 2.14: The number of studies with different study topics

cumulative numbers of studies on each topic are presented in a chronological order as shown in Figure 2.2. The figure clearly shows the studies on joint maintenance and inventory optimization dominate at the beginning several years. Entering the year 2017, joint maintenance and inventory optimization studies start to loss the dominating status because the studies on other topics, especially studies on spare parts inventory problems have been growing rapidly. This trend indicates that the studies on spare parts inventory management start to grow as an independent research field apart from the equipment maintenance field. More and more researchers begin to focus on this problem and bring interesting viewpoints and studies to enrich this filed.

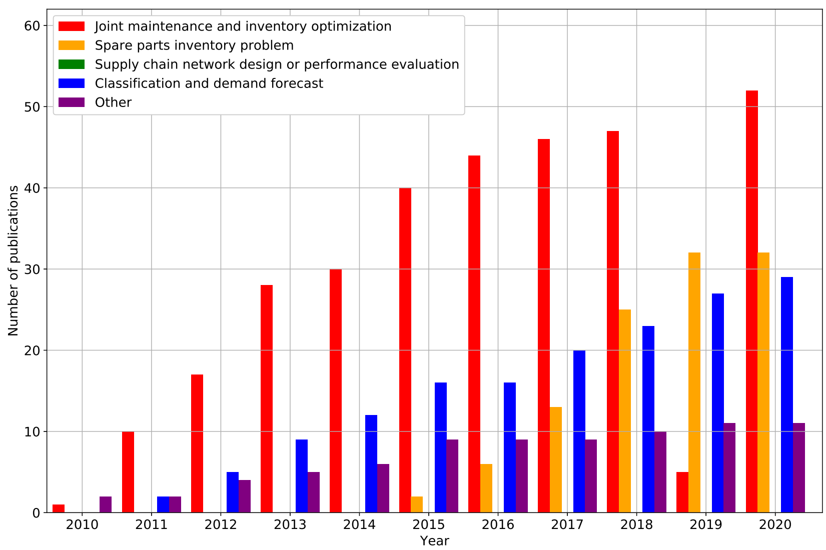

Figure 2.2: Cumulative number of studies with different topics

In this section, the reviewed studies are classified based on the perspectives of supply chain analytics techniques and research topics. The main results of this section are summarized as follows. First, the reviewed studies are classified based on analytics techniques. Afterwards, the topics discussed in the research field of spare parts inventory management are identified. In the next section, research gaps and future research directions are illustrated based on the classification analysis in this review.

2.4 Conclusions and Future Research

In this section, a general view of the developing trend in spare parts inventory management research is provided at first. The research gaps are then discussed based on our observations and analysis in previous content, followed by suggested extensions on current research topics to fill in the gaps. A concluding remark is presented at the end.

2.4.1 Developing trends

To clearly present the developing trend of spare parts inventory management research, the reviewed studies are grouped according to the publication year. Afterwards, the number of studies in each year are aggregated and chronically illustrated in Figure 2.3. From Figure 2.3, we can see that studies on spare parts inventory management experience

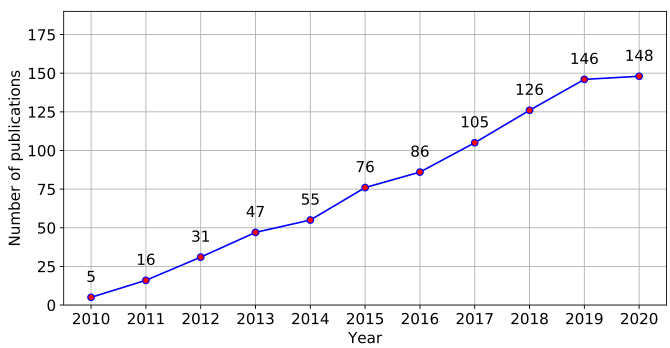

Figure 2.3: Cumulative number of studies in spare parts inventory management

a steady growth in last decade. This trend indicates that there is an increasing interest in spare parts inventory management and many questions are explored by researchers.

2.4.2 Research gaps and extensions

Based on the classifications made in previous sections, we identify the research gaps in the following aspects, i.e., consumer durable goods, supply chain network structure and policies, reverse logistics, spare parts demand pattern modelling, and big data analytics implementation. In the rest of this section, these gaps are elaborated one by one in detail. All the discussed research gaps are summarized in Table 2.15.

Research gaps	Current studies	Future studies
Spare parts of consumer durable goods	• Few studies on the spare parts of durable consumer goods. • Lack lifecycle planning perspective. • Simple warranties are considered. • Lack integration of product marketing strategy and spare parts inventory management.	• Consider lifecycle planning perspective • Consider advanced warranty types of durable consumer goods. • Integrate product marketing strategies and spare parts inventory management.
Inventory network structure and policies	• Most of studies consider single-echelon structure and lack of implementation of transshipment strategies.	• Investigate and evaluate lateral or emergency transshipments in the multi-echelon inventory system.
Reverse logistics	• Reverse logistics has a close relationship with spare parts inventory management but few studies considering it. • Even though some studies incorporate reverse logistics in their problems, the considered reverse logistics has simple settings while the one in practice has more complicated settings.	• Combine various settings in structure and more inventory strategies with reverse logistics relevant to spare parts inventory management. • Evaluate the performance of spare parts reverse logistics through a product lifecycle perspective.
Spare parts demand pattern modelling	• Stationary stochastic process is mainly used to depict spare parts demand in the reviewed studies.	• Non-stationary stochastic process or general probability distribution should be used.
Big data analytics	• Few reviewed studies implement big data analytics (BA).	• Facilitating BA in demand forecasting, inventory system design, and system optimization.

Table 2.15: Summary of research gaps

Research gaps regarding consumer durable goods

The differences between consumer durable goods and capital goods are identified in previous contents. Based on these differences, we identify the following research gaps in managing spare parts of consumer durable goods and propose corresponding extensions.

Spare parts inventory problems of consumer durable goods As mentioned in Section 2.2.1, most of current literature target on managing spare parts of capital goods. The reason for this phenomena is two-fold. Firstly, the literature on spare parts inventory management origin from and are highly impacted by the literature on capital equipment maintenance scheduling. The capital equipment usually require high availability level. Therefore, spare parts inventory management plays a critical role in maintenance activities because sufficient spare parts inventory can effectively reduce equipment downtime and improve equipment availability. Secondly, spare parts of capital equipment usually are high value components with high specificity level. Therefore, the inventory management on such spare parts is able to create significant performance improvements in both economic and service perspectives.

However, there are very few studies focusing on the inventory problems of consumer durable goods. Unlike capital goods, consumer durable goods usually have a shorter lifecycle and a bigger consumer base. Inventory planning period of spare parts in consumer durable goods is shorter and spare parts demand is more variant. The spare parts inventory of consumer durable goods is normally held by an OEM or third-party service provider to fulfill the warranty contracts of consumer goods whilst that of capital goods is held to support maintenance services which are regulated by service contracts. Warranty contracts often do not have specific requirements on product performance while service contracts do. In this sense, the spare parts inventory management of consumer durable goods are more appropriate to be operated in a cost-minimizing manner. The mentioned differences indicate that the existing inventory policies for spare parts of capital goods are not likely to be optimal for that of consumer durable goods. Based on these differences, two research gaps are identified as follows.

The first research gap is that the studies on managing spare parts inventory of consumer durable goods lack the perspective of lifecycle planning. Only one study focus on the lifetime spare parts procurement problem for consumer durable goods (Sahyouni et al., 2010). Studies on consumer durable goods are more suitable to implement the perspective of lifecycle planning than those on capital goods due to the following reasons. First, the consumer durable goods normally have short lifecycles and the OEMs prefer rapid product development. In some industries such as electronics and telecommunications, it is common for an OEM to finish the production processes even before the products are available in retail outlets (Sahyouni et al., 2010). Consequently, the OEMs almost cannot procure spare parts to support maintenance operations during product warranty time after observing demand and subsequent return rates of defective products. Second, consumer durable goods usually have fast innovation clockspeed (Li et al., 2018). The OEMs launch new products quickly to keep the competitive advantages of their products, leading to multiple product generations are on market simultaneously. When an OEM switches its production to a new generation or a new model, suppliers incur high retooling costs to produce new parts. In this case, due to high setup cost, they are not willing to produce the parts in old generations or models in another production after the first production. Therefore, the spare parts of elder products may not be available when the production of these products ceases. In this case, OEM needs to develop a specific lifecycle spare parts managing plan for each product generation.

The second research gap is to manage spare parts inventory of consumer durable goods under warranty contracts. There are several existing studies focusing on this topic such as Pourakbar et al. (2012), Rezapour et al. (2016), Li et al. (2018), Frenk et al. (2019b), and Frenk et al. (2019a). However, these studies have various settings, focusing on different phases in product lifecycle and considering different types of spare parts (repairables or non-repairables), while only consider basic warranties. New studies on exploring the spare part inventory control under advance warranty contracts are needed. For example, extended warranties are provided by the OEMs for certain prices as value-added services and they actually generate revenues to the OEMs. Therefore, how to manage the spare parts inventory system to support such profit-generating contracts is a critical problem to maximize the profits. In addition, for some consumer durable goods, warranty periods usually are restricted in two-dimensions like the product age and usage (Shafiee and Chukova, 2013). A good example of two-dimensional warranty contracts is the automobile warranty. Studies intending to investigate how to manage spare parts inventory to support maintenance activities under two-dimensional warranty contracts will be more attractive.

Multiple substitutable products which share a common spare parts inventory system Nowadays, the OEMs are incline to use product segmentation strategy in which multiple substitutable models belonging to one product category are provided to customers from different groups. In the following context, we will refer the product category contains several models as product assortment. One real life example of product assortment is that one automobile manufacture sells several car models at different prices among various markets while these models may contain similar and different parts. In this case, the sales quantity of each car model is highly governed by customer preferences over the car assortment, resulting different market shares captured by different car models. For the spare parts which are used by different models, inventory management issues are more complicated because demand patterns of these spare parts are hard to model. The demands of these spare parts could be induced by car failures or accidents. The car failures numbers are determined by two factors including the quantity of cars on market and failure rate of each car model. First, the quantity of cars on market is different across all car models in the assortment. Second, in general, different car models prohibit various failure rates because of the differences in car design, manufacturing, and usage conditions. Consequently, the aggregated spare part demand should be estimated based on the mentioned factors. The illustrated example is not limited to the automobile industry and can be extended to other industries using product segmentation strategy.

Research gaps regarding inventory network structure and policies

Future studies could focus on the following aspects of inventory network structure and polices: Multi-echelon inventory system and lateral and emergency transshipments.

Multi-echelon inventory system As discussed in the "Number of echelons" part of Section 2.2.2, compared to the studies focusing on single-echelon inventory system, the studies considering multi-echelon inventory system are limited in number: There are 31 papers adopting multi-echelon inventory setting while 91 papers adopting single-echelon inventory system.

The multi-echelon inventory system for managing spare parts represents a more common scenario in practice, especially for the OEMs or third-party organizations who are responsible for maintaining the products over different geographic districts. In addition, other extensions regarding inventory policy are based on the multi-echelon structure. For example, the inventory pooling strategy can be implemented when considering a multi-echelon inventory system.

Lateral and emergency transshipment Lateral transshipment and emergency shipments are inventory strategies contributing to performance improvement of inventory system. Spare parts are held as inventory to support maintenance activities which are governed by service contracts. These contracts usually have specific requirements on the availability level of the maintained systems. Hence, lateral and emergency transshipment can be used to support timely maintenance service so as to decrease product downtime. However, the discussions in the "Lateral transshipments and emergency shipments" of Section 2.2.2 illustrate that only 24 of the reviewed studies consider lateral or emergency transshipment. Therefore, more studies can incorporate lateral or emergency transshipments into their inventory systems.

Research gaps regarding reverse logistics

As environmental concerns have been increasingly growing in recent years, the concept of supply chain sustainability is widely advocated in the industry and academia. Reverse logistics (RL) is one of the most popular topics in the literature on supply chain sustainability. RL involves operations related to the return of damaged, unsold, end-of-life products along with handling, consolidation, remanufacturing and disposal (Diabat et al., 2015). To be specific, RL is a process turning the inputs such as the used products, recycled materials, used parts to the outputs such as remanufactured products and spare parts (Pokharel and Mutha, 2009). There are four key subprocesses of RL, i.e., product acquisition, collection, inspection/sorting, and disposition (Agrawal et al., 2015). Product acquisition is the first step in RL process, referring to the acquisition of used products, components or materials from end users for further processing. After acquisition, returned items are collected in three ways: the OEM collects directly from the users, the OEM buys back the products collected from users by a retailer, and the OEM collects products through third-party logistics (Kumar and Putnam, 2008). The collected items are sent for inspection, sorting, and disposition. The inspection and sorting process is necessary because the conditions of returned items have to be inspected and evaluated and then they are sorted based on conditions. Afterwards, the disposition decisions for further processing will be made. There are five common disposition alternatives including reuse, repair, remanufacturing, recycling, and disposal. For detailed definitions of these disposition alternatives please refer to the literature review on RL presented by Agrawal et al. (2015).

Spare parts inventory management has a close relationship with RL because spare parts are one of the wanted outcomes of the RL system (Pokharel and Mutha, 2009). Based on their conditions, some of these spare parts are sent back to market as reused ones after inspection, cleaning and minor maintenance; some are used for repairing faulty products; some are sent to remanufacture new products; and some are recycled as raw materials. From this perspective, it is promising to explore the integration of spare parts inventory management and RL because how to manage the inventory of different spare part types in RL process is quite attractive and challenging.

Among the reviewed studies, there are several studies considering the RL. Abdallah et al. (2012) study an uncapacitated closed-loop location inventory model in which one type of returned products is collected and salvaged into spare parts which can be reused. Inderfurth and Kleber (2013) deal with a hybrid manufacturing/remanufacturing system for providing spare parts to support the after-sales service to the products in the EOL phase. Diabat et al. (2015) consider a single-echelon reverse supply chain where the returned products are remanufactured as spare parts and then sent back to retailers. In their problem, new items are produced either from manufacturing using externally supplied materials or from remanufacturing using returned items. Ahiska et al. (2017) propose heuristic inventory policies to control a manufacturing/remanufacturing system with downward product substitution. In addition, a downward substitution strategy is used, i.e., when a lower-value item stocks out, a higher-value item is substituted to meet the demand to reduce the stock-out cost. Shi (2019) introduces a spare parts inventory control problem for an OEM who remanufactures spare parts from returned products to meet warranty demand under part obsolescence.

The mentioned existing studies on RL in spare parts inventory management often have simple settings. More research on this topic can consider the extensions as follows. First of all, the supply chain structure in the existing studies is quite simple: Only Abdallah et al. (2012) consider a two-echelon inventory network while the rest studies consider single-echelon network. Second, all these studies only focus on one phase of product lifecycle, i.e., the changes in installed base is not considered as one impacting factor of returned products quantity. Third, various inventory strategies such as lateral and emergency transshipments are not considered even though these strategies may improve RL performance. For example, RL are common in the electronics supply chains such as cell phone, personal computer, etc. In these supply chains, returned products are disposed to remanufacture new products which can be either sold as new ones or used as replacements for the returned products. Considering the worldwide retailing network an OEM may possess, how to collect the returned products from the retailers over different geographical regions, then dispose collected returns and use spare parts or materials outcomes to remanufacture new products, and finally allocate the remanufactured products to those retailers is a relevant and challenging problem. In addition, lateral and emergency transshipments can be used to move the remanufactured products within the retailing network to decrease stock-out cases and customer waiting time. Last, it is interesting to see how to allocate inventory of different spare parts which are used in different disposition alternatives. For instance, as mentioned earlier, different spare parts in RL process may have different disposition alternatives. The optimal inventory policies need to be made for each type of spare parts so that corresponding costs are lowered.

Research gaps regarding spare parts demand pattern modelling

From the perspective of modelling spare parts demand patterns, attentions should be paid to implementation of non-stationary stochastic process and general probability distributions. As discussed in the "Predictive analytics" part of Section 2.3.1, most of the studies model the demand pattern of spare parts as stationary stochastic processes such as Poisson process by assuming that the demand pattern, as an random variable, follows a exponential probability distribution which is not changing over planning periods. This assumption makes sense when the installed base of product is in a steady-state condition and the product reliability is relatively mature (Jin et al., 2017). In other words, the stationary stochastic process assumption holds when the studied product is in the mature phase of its lifecycle. modelling spare parts demand patterns as a stationary stochastic processes can make the optimization model relatively simpler thus structural results might be achieved. However, demand pattern cannot be modelled as a stationary stochastic process when the planning horizon spans over several phases of the lifecyle or over the entire lifecycle, because product installed base is changing in different phases. Therefore, non-stationary stochastic processes or general probability distributions should be used for modelling the demand patterns of spare parts.

As shown in Table 2.9 in Section 2.3.1, there are only 10 papers using non-stationary stochastic process and 20 papers using general probability distribution to capture the demand pattern respectively, compared to 74 papers using stationary stochastic process. Hence, more papers are called on to extend current literature by using non-stationary stochastic processes and general probability distributions to model spare part demand patterns.

Research gaps regarding implementing

Big data analytics (BA) is the study of practices, technologies and skills to evaluate operations and strategies to obtain insights and offer guidance to the business planning of an organization (Wang et al., 2016). Such evaluation is done toward product development, strategic management, and customer services, and so on, by utilizing evidence-based data, statistical and operations analysis, predictive modelling, forecasting, and optimization techniques (Chen et al., 2012). There is a clear gap in the reviewed literature that BA is barely used and this research area can be further enriched by implementing BA in several interesting aspects such as demand forecasting, inventory system design, and inventory system optimization.

Typical literature in spare parts inventory management mainly focus on managing at most hundred kinds of spare parts in inventory while it is common to see more than thousands of spare parts are held in practice. There is only one paper proposing a smart spare parts inventory management system in a semiconductor company for obtaining more information which can be used in BA (Zheng and Wu, 2017). The reason behind this phenomenon is that the computation capacity is limited when a huge number of spare parts are considered in the demand forecasting or optimization processes. Therefore, classic statistics or optimization methods are not applicable when facing up with a large amount of spare parts inventory to manage. In this case, BA will be useful when massive size of data are available. More research can be made toward to how to build smart system adopting BA to manage spare parts inventory.

2.4.3 Conclusions

In this review, 142 papers on spare parts inventory management have been surveyed and classified. Several research gaps have been identified for future research in this field. Our review has the following distinct features. Firstly, it provides a quick guide to a variety of classification schemes to the spare parts inventory management literature. Two different typologies are used for the literature classification. One typology classifies the literature based on systematic characteristics of spare parts inventory systems while the other typology classifies the literature based on research methodologies and topics. Secondly, this review presents a big picture on the spare parts supply chains to discuss the studies on spare parts inventory management. This big picture links the important aspects relevant to managing spare parts inventory system, such as product and spare part types, after-sales services, maintenance operations, inventory management strategies and policies, supply sources, demand patterns and so on. Thirdly, we classify the research methodology of each studies from the perspective of supply chain analytics. From this classification, current studies using descriptive, predictive and prescriptive analytics are identified. Finally, various research topics in the spare parts inventory management literature are summarized and corresponding research gaps and extensions are discussed.

The studies in this research field has been experiencing steady growth in the last decade. Researchers have put plenty of their interests on managing spare parts of capital goods, which are concerned with system availability of the capital goods. It is expected to see more studies on managing spare parts of durable consumer goods, which are faced up with different groups of consumers. In addition, diverse inventory policies implemented in the inventory network with different structures need to be considered in future research. New supply chain concepts such as reverse logistics can contribute to this field but have to be tested and evaluated through more studies. The most promising area for future research could be the study which combines supply chain inventory management problem with big data analytics. Big data analytics could contribute to many aspects such as spare parts demand forecast, inventory system optimization, and classification, when a large number of spare parts need to manage.

Chapter 3

Spare parts inventory management for substitute consumer products: An adaptive robust optimization approach

3.1 Introduction

Spare parts are stock items used in maintenance activities to keep equipment in operating conditions and extend their lifecycle (Kennedy et al., 2002). Spare parts inventory management is becoming more critical as manufacturers intend to provide after-sales services to customers as value-added services. Furthermore, spare parts also play an important role to the societies in the world when pursuing carbon neutral. In past decades, manufacturers are more likely to advocate for a culture of planned obsolescence: To design their products to be short-lived and hard to repair, they can seize more revenues because customers are forced to purchase more new products when their old ones are not functioning properly. However, this culture contributes to wasting more natural resources and energy, generating more greenhouse gases, and further escalating global warming. For example, the carbon emissions of producing an iPhone 12 account for nearly 80 percent of the total emissions during its lifecycle (Apple Inc., 2020). In the

United States, a motion known as “right to repair" has been calling for legislation that requires companies make their parts, tools, and information available to consumers and repair shops (Rosa-Quino, 2020). The motivation of this motion is that the longer the product lifecycle is, the fewer unnecessary product purchases are, and finally the lower pollution will the production processes generate. For the original equipment manufacturers (OEMs), they may suffer from the decreased sales of new products from this motion, but they can obtain revenues by expanding their after-sales services. Moreover, advocating sustainability practice can demonstrate manufacturers embrace corporate social responsibility (CSR) to customers, contribute to a positive brand image, and reinforce their corporate reputation (Ukko et al., 2019; López-Pérez et al., 2017; Aguilera-Caracuel and Guerrero-Villegas, 2018). In this context, to reach the balance between sustainability and profitability, the OEMs need more spare parts for repairing the faulty products and an efficient spare parts inventory management system is also necessary.

In this chapter, we focus on managing the spare parts inventory of a product assortment which includes several substitute products over multiple time periods. Product assortment refers to the variety of products and services that a company offers to the consumers and is an important marketing strategy advocated by practitioners and researchers (Simonson, 1999). In the assortment, products are alternatives to each other but have differentiation. In other words, these products have similar basic functions but different product characteristics such as price, quality, colour, and so on, to provide more choices to consumers with different preferences. Although the substitute products in the assortment are competing with each other, the availability of more products can lead to a higher customer utility. To accommodate the product assortments, the manufacturers may consider different strategies in their production process. For example, they may use part standardization strategy such that some parts are commonly used in the products to achieve the same basic product function, while some parts are uniquely used in certain products to achieve the differentiation.

Such product assortments with substitute products are quite common in the consumer durable goods market such as personal computer and automobile industries. For example, it is not rare that one model of an automobile make contain several different types. All types have same engines and transmission systems but different interior decorations, electronics, etc. In the literature of marketing and revenue management, the customer choices over the substitutes in an assortment are widely modelled through the multinomial logit (MNL) model and its variants (Gallego and Wang, 2014; Du et al., 2016). In this chapter, we develop a multi-stage adaptive robust optimization model in which the demand is determined by the MNL model of consumer choice over the substitutes in an assortment. Specifically, the market share of each product is estimated by the MNL model and then is multiplied by the market size to calculate the market sales in each planing period. Afterwards, the total on-market quantity of each product is calculated by aggregating the product sales over all periods. Finally, the product failure quantity is estimated based on the on-market quantity and failure rate. It should be noted that we assume the manufacturer does not know the probability distribution of the part failure rate but knows the rate falls into a certain range. In other words, the part failure quantity of each product in each period also falls in a range determined by the estimated product failure quantity and the range the part failure rate may fall in.

The main contributions of this study are multi-fold. Firstly, our proposed model considers managing the spare parts inventory of multiple substitute consumer products in an assortment. The spare parts demand induced by the users of these products is estimated based on the MNL model. Our purpose is to jointly manage the spare parts for these products. To our best knowledge, there is no study on this problem so far. This problem is complicated because some spare parts may be commonly used by several products in the assortment while some may be uniquely used by one product. Secondly, we manage the uncertainty inherent in this problem by an adaptive robust optimization model (ARO) with integer variables. In the literature, spare parts inventory management is normally studied by the stochastic inventory control (SIC) models (Zipkin, 2000; Axsäter, 2015). However, when the information on spare parts demand is imperfect, e.g., the probability distribution of demand is unknown, the SIC models cannot be used. In addition, the proposed problem is a complex problem with multiple spare parts managed over multiple periods, so that the SIC models may face computational challenges even the demand information is perfect. Therefore, an adaptive robust optimization approach is proposed in this chapter to deal with these challenges. Thirdly, we investigate the solution methodology for this ARO model. Our method is developed from the partition-and-bound method raised by Bertsimas and Dunning (2016). This method suffers from the "curse of dimension" when solving the large size problems. We improve this method through eliminating the redundant constraints in the model so that the model size is significantly decreased. Through extensive numerical analysis, our improved method is demonstrated to be able to solve the medium and large problem instances. Therefore, our study adds value to the literature on the multi-stage ARO problems because the application of the partition-and-bound method to solve the multi-stage ARO problems is scarce in the literature. The proposed method enriches current studies and may provoke more studies on this problem. Finally, we provide some managerial insights from conducting a sensitivity analysis to explore the impacts of spare parts purchase cost, product popularity, and product backorder cost on the inventory policy and total cost.

In the following, we first review the literature on spare parts inventory management and ARO problems respectively in Section 3.2. The model description is presented in Section 3.3. We present the improved partition-and-bound method to solve the proposed problem in Section 3.4. Then comprehensive numerical experiments are presented in Section 3.5. Finally, we conclude this chapter and discuss future research directions in Section 3.6. For the ease of reading, we will use "spare part/s" uniformly instead of "part/s" in the rest of this chapter.

3.2 Literature Review

In this section, the literature reviews on spare parts inventory management and multi-stage robust optimization are presented respectively.

3.2.1 Spare parts inventory management

In recent decades, spare parts inventory management has received increasing attentions from the researchers in operations research and management science. Studies on this topic stem from and are highly impacted by the literature on capital equipment maintenance scheduling because spare parts inventory management plays a critical role in maintenance activities: Sufficient spare parts inventory can effectively reduce equipment downtime and improve equipment availability. In addition, spare parts of capital equipment usually are high value components with high specificity level. Therefore, the inventory management on such spare parts is able to create significant performance improvements in both economic and service perspectives. In the literature of this field, there are abundant studies on managing spare parts inventory used for supporting preventive and corrective maintenance of capital equipment such as aircraft (Mirzahosseinian and Piplani, 2011; van Jaarsveld et al., 2015), technical systems (Öner et al., 2013; van Wijk et al., 2019), etc. Among these studies, multiple characteristics of spare parts inventory system are considered. Tiacci and Saetta (2011) examine the effectiveness of two lateral shipments approaches in reducing the mean supply delay of a non-repairable item in the capital equipment through the simulation model. Li and Ryan (2011) incorporate real-time condition monitoring information on spare parts into the proposed spare part inventory model for supporting the maintenance of capital machines. Mirzahosseinian and Piplani (2011) model a closed-loop spare parts inventory system under the performance-based contract for the Unmanned Aerial Vehicle (UAV). van Jaarsveld et al. (2015) study an multiple spare parts inventory control problem for an aircraft component repair shop where spare parts demand arrives as a stationary Poisson process and formulate an integer programming optimization problem. van Wijk et al. (2019) incorporate lateral transshipments with spare parts inventory system by proposing a periodic–review spare parts inventory model in which lateral transshipments are allowed between the stockpoints when the advanced technical systems break down. These mentioned studies clearly show that this research field is well developed.

However, as the marketing strategies adopted by manufacturers evolve rapidly, more focus is placed on providing after-sales services to enhance customer satisfactions and generate more profits. In this context, spare parts management is not only considered by the manufacturers of capital equipment but also those of consumer durable goods. However, in the literature, the studies on spare parts inventory management of consumer durable goods are scarce. Pourakbar et al. (2012) consider a spare parts inventory control problem for consumer electronics manufacturers in the final phase of the service life cycle when the spare parts production is terminated. The manufacturers have to decide the spare parts inventory level for supporting the repair operation until the service contract or warranty period expires. Inderfurth and Kleber (2013) work on a spare parts inventory system for the consumer durable products entering the end-of-production phase of their life cycle under after-sales contract. The inventory decisions they consider include spare parts final order quantity, remanufacturing quantity, and extra production quantity. To solve the proposed stochastic dynamic programming model, they develop a heuristic which can be applied to the problem with practical sizes. Frenk et al. (2019a) work on the spare parts end-of-life inventory problem that happens after the spare parts production of consumer electronics stops. The novelty of this study is that they consider two different options for repairing faulty products, one is repairing and the other is swapping the faulty product with a new one. They found that the second option is favourable when the older generation products depreciate and the new generation products dominate the market.

Another observation is that most studies on spare parts inventory management assume the spare parts demand follows certain probability distributions which are known to manufacturers. This assumption may be invalid in practical operations. For example, when planning for spare parts inventory, manufacturers may not possess enough information on the corresponding demands and they are more likely to estimate the spare parts demands based on the historical data. However, when news product are launching, they may not be able to rely on the historical data because these data is not available. Consequently, the best they can do is to forecast the range of the demand. In this context, how to develop an inventory plan under the demand uncertainty is quite important to the manufacturers but not well studied in the literature. Our research intends to fill these gaps in the studies on the spare parts inventory management of substitute consumer goods.

3.2.2 Adjustable robust optimization

Robust optimization (RO) is a method to solve problems with uncertain parameters residing in a user-specified set, i.e., the uncertainty set. The basic idea of the RO approach is to find an optimal solution which is feasible to all the possible data changes in the uncertainty set. In the literature, the RO models can be classified into two types, one is static robust optimization and the other is ARO. In the static robust optimization approach, all decisions have to be made before the uncertain parameters are realized. However, such a paradigm normally yields overly conservative solutions especially when the problem is multi-period and periodical decisions are taken after uncertainty parameters are realized period by period (Yanıkoğlu et al., 2019). To be specific, the real-life problem is more likely in a dynamic context and involves not only here-and-now but also wait-and-see decisions. The decision makers determine the values of here-and-now decisions before the realizations of uncertainty parameters while the values of wait-and-see decisions are determined after the realization of those parameters.

The concept of ARO is firstly introduced by Ben-Tal et al. (2004). They extend the robust optimization methodology to ARO by introducing the adjustable robust counterpart (ARC). Although the ARC is significantly less conservative than the usual robust counterpart, it is computationally intractable, i.e., NP-hard. To deal with this issue, they propose an approximation of the ARC by restricting the adjustable variables to be affine functions of data. The proposed problem is the so called affinely adjustable robust counterpart (AARC) and the relationship between variables and data is referred to as affine decision rules (ADRs). In general, the ADRs yield an upper bound to the ARC (when it is a minimization problem) because it may be suboptimal or even infeasible for some problems. But for some specific problems, the ADRs are optimal. Bertsimas et al. (2010) prove that for the ARO problems which have convex objective functions with respect to adjustable variables and uncertain parameters, the ADRs are optimal. Besides the ADRs, there are several other forms of decision rules regulating the relationship between adjustable variables and uncertainty parameters. Bertsimas and Goyal (2012) show that the ADRs are optimal for a linear ARO problem with a simplex uncertainty set and only right-hand side uncertainty. Ben-Tal et al. (2020) indicate there exist optimal piece-wise linear decision rules for the linear ARO problems with only right-hand side uncertainty. Ben-Tal and Den Hertog (2014) and Ben-Tal et al. (2015) develop the separable quadratic decision rules for the ARO problems with quadratic objective functions and nonlinear uncertain inequality respectively.

Although the continuous decision rules can be used to solve some linear ARO problems optimally, it generally cannot be applied to the ARO problems with integer adjustable variable or the multi-stage adaptive mixed-integer optimization (AMIO) problems. In the literature, two efforts are made by researchers to solve the AMIO problems. One is to find the special structure of integer problems which can be solved by applying certain decision rules. Bertsimas and Georghiou (2015) derive the piece-wise linear decision rules as optimal decision rules for the AMIO problems involving continuous and binary variables. They present a method to construct such decision rules to solve the multi-stage problems robustly using mixed-integer optimization. Bertsimas and Georghiou (2018) apply the binary decision rules to solve the class of multi-stage adaptive mixed-integer optimization problems. Since the resulting optimization problem grows exponentially in the size of problem data, they develop a systematic methodology to conservatively approximate the associated decision rules so as to balance the optimality and scalability.

The other effort is to design partition-and-bound methods to solve the AMIO problems. Among these methods, the uncertainty set is divided into partition the uncertainty set into subsets with different decisions, such as the studies by Postek and Hertog (2016) and Bertsimas and Dunning (2016). Both studies propose their own partition-and bound method to solve the AMIO problems, but the main difference is the way to obtain the optimal solutions in different partitions. The first is trying to directly optimize the decisions in various partitions while the second is approximating the optimal piece-wise affine policy. They both apply their own methods to solve the same instances of capital budgeting problems and multi-stage lot sizing problems with one product for numerical experiments. The experiments results show that the later one is more efficient and able to obtain better solutions when solving the same problem instances (Bertsimas and Dunning, 2016).

In Bertsimas and Dunning (2016), although the proposed method is able to solve small instances in a reasonable time, it is only efficient in the first few iterations. The reason is the scenario tree will explode as the algorithm iteration number increases. In the scenario tree, each scenario node represents a subset of the uncertainty set and, in that subset, a series of corresponding decisions subjecting to corresponding constraints will be made. In each iteration, each of current subset will be further partitioned into more subsets in the next iteration. Consequently, more and more decision variables and constraints will be added to the model after each iteration, creating more difficulties in solving it. Therefore, the proposed partition-and-bound method will not be efficient in solving the large size problems or even the small size problems with multiple inventory items, because as the iterations increases, the price of improving the solution quality is quite high in terms of computation time. In this chapter, multiple spare parts inventories are considered so that the method in Bertsimas and Dunning (2016) cannot solve this problem efficiently. We will illustrate this point in the first part of numerical experiments in Section 3.5. To apply the partition-and-bound method for solving our problem, we improve it by decreasing the scales of optimization models constructed in each iteration so that they can be solved in a reasonable time even the iteration number is large. In addition, it should be highlighted that the literature on extending and developing this method to solve the AMIO problems is scarce. Our study is trying to fill this gap in the literature.

3.3 Model Description

We focus on managing the spare parts inventory of a product assortment which includes several types of products over multiple time periods. These products are produced by an OEM who is responsible for providing after-sales services which could be either preventive or corrective maintenance. In addition, the products have significant similarities and can substitute each other, but they are differentiated in some characteristics to cater the customers' preferences. Some spare parts are commonly used in these products in the assortment while some others are uniquely used in certain products. Such substitute products are common in durable consumer goods industries such as personal computer and automobile industries. For example, it is not rare to see that one automobile make may contain several different models which have the same engine and transmission system but have different interior decorations and electronics systems.

In our study, each product is assumed to be assembled from several different types of spare parts, and if one type is included, at most one unit will be used. In each period, once a product fails, it is sent to the OEM for repairing within a promised delivery due date. After being inspected, the faulty spare parts will be identified and removed through complete or partial disassembly. Note that, it is possible that several spare parts failures can be caused simultaneously in one product failure. In addition, the failures between different spare parts are assumed to be independent. New spare parts are supplied from the on-site inventory in the first place. To replenish the inventory, orders are placed to external suppliers at the beginning of each period. After a certain period of time, i.e., the lead time, the placed orders will be received. Finally, the defective spare parts are replaced by the new ones and all spare parts are re-assembled into the repaired product. This is the end of the repair process. If the failed products are not repaired within the promised due date, a penalty cost will be incurred.

The inventory decision faced by the OEM is to determine the order cycle and order quantity for each type of spare parts so that the total inventory cost related to the after-sales services of product assortment will be minimized. Such cost includes spare parts procurement cost, inventory holding cost, and delivery delay penalty. Spare parts demand is estimated based on the on-market quantity of each product and the failure rate of each spare part used. For the uncertainty embedded in the spare parts demand, the OEM is assumed to know only the range into which the demand will fall, i.e., the demand uncertainty set, but not the exact probability distributions the demand will follow.

The proposed model for managing spare parts inventory has two parts. The first part of the proposed model is to utilize the MNL model to reflect customer preferences over the products in an assortment so that the sales quantity of each product can be determined in each period. Such sales quantity will be used to project the on-market quantity at first and to estimate the number of product failures in each period finally.

The second part of the model is to decide order quantity for each spare parts and products backorder quantities at each period so that the total cost is minimized. The notation is listed in Table 3.1.

Table 3.1: Notation

Parameters	
t	Planning period $t \in \mathcal{T} = \{1, \ldots, T\}$
c	Spare part $c \in \mathcal{C} = \{1, \ldots, C\}$
n	Product $n \in \mathcal{N} = \{1, \ldots, N\}$
D_{nt}	Demand for repairing failed product n in period t
d_c^t	Demand of spare part c in period t
d_{cn}^t	Demand of spare part c used in product n in period t
L_c	Procurement lead time for spare part c
H_c	Unit inventory holding cost for spare part c
SS_c	Safety stock level for spare part c
B_n	Unit penalty cost for the late repair of each unit of product
P_c	Price of spare part c
Decision variables	
q_{ct}^p	The order quantity of spare part c in period t
l_{ct}	The inventory level of spare part c at the beginning of period t
b_{ct}	The shortage amount of spare part c in period t
θ_{nt}	The delayed amount of repaired product n in period t

3.3.1 Estimate the on-market quantity of each product

The product assortment contains multiple substitute products, offering more options to customers for catering their various preferences over the product family so that the OEM can capture more market share. In the literature, the MNL model is one of the most widely used model to represent customers' decision process of selecting a product from an assortment. In the proposed model, the MNL model is adopted to predict the on-market quantity of each product in the assortment over the planning horizon.

In the MNL model, a representative customer is faced with a product assortment which contains N products ($N \geq 2$). The customer obtains utility U_{nt} if he or she purchases product $n \in \mathcal{N} = \{1, ..., N\}$ during period t and this utility is determined by

$U_{nt} = \alpha_{nt} + \epsilon_{nt}$, where α_{nt} represents a measure of attractiveness of product n to the customers and ϵ_{nt} is a random term which represents the unobserved utility. If the random term ϵ_{nt} is assumed to be independent and identically distributed (i.i.d.) with Gumbel distribution and each customer only purchases one product, then the MNL model gives the choice probability (Train, 2009) that the customer selects product n during period t with probability $p_{nt} = \frac{e^{\alpha_{nt}}}{\sum_{i=1}^{N} e^{\alpha_{it}}}$. If we aggregate the customer preferences, the probability denotes product n's proportion of the total demand in a given period t. In other words, it represents each product's expected share of total sales quantity of the assortment. This proportion is widely referred to as the market share of product n in the relevant studies (Li and Huh, 2011). We will follow this convention in the rest of this chapter. The market share p_{nt} will be used to predict the sales quantity, Q_{nt}, of product n during period t, by multiplying p_{nt} with Q_t which is the estimated sales quantity of all products in the assortment during period t. Such sales quantity of all products at each period can be predicted by utilizing historical data of similar product assortments. Afterwards, the sales quantity of each product in the assortment will be used to track the total on-market quantity of each products in the assortment from the beginning period to period t, which is denoted as M_{nt}. To be specific, the total on-market quantity of product n can be updated by adding new sales quantity of each product in a forthcoming period, i.e., $M_{nt} = M_{n,t-1} + Q_{nt}$. Next, the total on-market quantity of each product is used to determine the product failure quantity. The uncertainty set of demand for each spare part type is estimated based on the failure quantity of the products using it.

3.3.2 Inventory system model

In this subsection, a deterministic inventory model is introduced first and a multi-stage ARO model is then provided to deal with the uncertainties embedded in spare parts demand.

Estimating spare parts demand

To model the spare parts inventory system, we need to consider spare parts demand, inventory policy, spare parts procurement plan, and repair decisions for the failed products simultaneously. Identifying the spare parts demand is a challenging task due to the following reasons. First of all, different types of products usually have different failure probabilities. Second, the number of faulty spare parts to be replaced is highly unpredictable.

To estimate the demand for repairing failed products and the demand for different spare parts in our proposed model, we use the following procedure. In the first step, we calculate the expected failure quantity of product n in a planning period t, D_{nt}, as $D_{nt} = M_{nt} \times \rho_{nt}$, where M_{nt} represents the estimated on-market quantity of product n and ρ_{nt} denotes the failure probability of that product in period t. The estimated on-market quantity M_{nt} is obtained as mentioned in Section 3.3.1 and the failure rate ρ_{nt} can be estimated from historical data. In the second step, the demand quantity for each spare part will be determined. Such quantity is affected by the failure numbers of all products using this spare part during period t, i.e., the total demand of the spare part c is equal to the total quantity of failed products whose failures are induced by the failure of spare part c. In our model, it is assumed that each product in the assortment has a distinct configuration of spare parts. In addition, some spare parts could be commonly used in several products. To be specific, let $\mathcal{N}_c$ represent the product set which contains all products using spare part c and d^t_{cn} denote the demand for repairing faulty spare part c in product n during period t, respectively. Therefore, the demand for spare part c, d^t_c is computed as $d^t_c = \sum_{n \in \mathcal{N}_c} d^t_{cn}$.

Deterministic inventory management model

As mentioned earlier, our model focuses on assisting the OEM managing the spare parts inventory of a product assortment to provide after-sales services. We adopt a periodic review inventory policy with safety stock which indicates the inventory of spare part c is replenished at the beginning of each planning period and the inventory level never drops below the base stock level S_c. The OEM seeks to decide the number of spare parts, q^p_{ct} , purchased at a unit price P_c from external suppliers at the beginning of period t. The order lead time is denoted as L_c. During a certain period, the inventory and new delivered spare parts will be used to meet the spare parts demand for repairing faulty products. However, when there is insufficient inventory, it will lead to unfixed products. The delayed repairs of product n will incur a penalty B_n for each time unit. At the end of each period, when the demand of spare parts c is realized, unused spare parts will be held as inventory at a unit cost of H_c. When the uncertainty embedded in the spare parts demand in not considered, i.e., the demand of each spare part is known, the problem can be formulated as a deterministic mixed integer programming model:

$$\min \sum_{t\in\mathcal{T}\setminus\{T-L_c+1,\cdots,T\}} \sum_{c\in\mathcal{C}} P_c\, q^p_{ct} + \sum_{t\in\mathcal{T}}\sum_{c\in\mathcal{C}} H_c\, l_{ct} + \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} B_n\theta_{nt} \tag{3.1}$$

$$s.t.\quad q^p_{c,t-L_c} + l_{c,t} = \sum_{n\in\mathcal{N}} d^t_{cn} + b_{c,t-1} - b_{ct} + l_{c,t+1}, \quad \forall\, t\in\mathcal{T}\setminus\{1,\cdots,L_c\},\, c\in\mathcal{C}, \tag{3.2}$$

$$\sum_{n\in\mathcal{N}_c} \theta_{nt} \geq b_{ct}, \quad \forall c\in\mathcal{C},\, t\in\mathcal{T}, \tag{3.3}$$

$$l_{ct} \geq SS_c, \quad \forall c\in\mathcal{C},\, t\in\mathcal{T}, \tag{3.4}$$

$$q^p_{ct},\, l_{ct},\, b_{ct}, \theta_{nt} \in \mathbb{Z}^+, \quad \forall c\in\mathcal{C},\, t\in\mathcal{T},\, n\in\mathcal{N}. \tag{3.5}$$

In this model, the objective (3.1) is to minimize the total cost which includes the procurement and inventory costs of spare parts and penalty costs incurred by the delayed product repairs. Constraints (3.2) represent the inventory balance equation for all spare parts. They indicate that during a certain period the amount of spare parts used for repairing should be equal to the inventory level at the beginning of that period plus the amount of received spare parts ordered L_c periods ahead. Constraints (3.3) indicate the amount of delayed product repairs is determined by the maximum shortage of all spare parts used in that product in a certain period. Constrains (3.4) require the inventory level of each spare part not fall below the safety stock level.

Uncertainty in spare parts inventory management

However, the spare parts demand is uncertain and the realized demand may not be exactly the same as what is estimated. In other words, the number of parts that need to be replaced in the products is random, making the spare parts inventory planning more challenging. As mentioned previously, the demand for a spare part is the aggregation of all product failures incurred by the failure of that spare part. Unfortunately, it is hard to estimate the demand in practice, especially when the probability distribution is unknown. Therefore, in our model, the failure quantity of each spare part (spare parts demand) is assumed to be a random variable and the OEM only knows the uncertainty set where it resides based on the estimated product failure quantity, i.e., the spare parts demands fall in an uncertainty set, $\boldsymbol{d} \in \Xi$, where $\boldsymbol{d} = (\boldsymbol{d}_1, \ldots, \boldsymbol{d}_c, \ldots, \boldsymbol{d}_C)^\intercal$ ($\boldsymbol{d}_c = (d_c^1, \ldots, d_c^t, \ldots, d_c^T)$) and Ξ denotes the uncertainty set. To be specific, given the estimated failure number of each product, we assume that the realized spare parts demand takes value from a rectangular uncertainty set $[\underline{d}_c^t, \bar{d}_c^t]$, where $\underline{d}_c^t$ and $\bar{d}_c^t$ are the lower and upper bounds of the set. The lower and upper bounds can be calculated as $\underline{d}_c^t = \underline{\sigma}_c^t \cdot \sum_{n \in \mathcal{N}_c} d_{cn}^t$ and $\bar{d}_c^t = \bar{\sigma}_c^t \cdot \sum_{n \in \mathcal{N}} d_{cn}^t$, respectively, where $\underline{\sigma}_c^t$ and $\bar{\sigma}_c^t$ represents the minimum and maximum possible portion of products failures caused by spare part c.

In this case, the order quantity at the start of the first period, q_{c1}^p, is a non-adaptive here-and-now decision which needs to made before the demand is observed, i.e., it is independent on $\boldsymbol{d}$. The adjustable variables in each period depend on realization of demand in previous periods. To be specific, the adjustable decision variables at period t include the order quantities $q_{ct}^p(d_c^1, \ldots, d_c^{t-1})$ (for $t > 1$), the inventory levels $l_{ct}(d_c^1, , \ldots, d_c^{t-1})$, the shortage amounts of each spare part $b_{ct}(d_c^1 \dots, d_c^t)$, and the product backorder quantity $\theta_{nt}(\boldsymbol{d}_c^1 \dots, \boldsymbol{d}_c^t)$, because these decisions in a period t are made based on the demand realizations preceding time t. For the ease of notation, we will use $q_{ct}^p(\boldsymbol{d}_c)$, $l_{ct}(\boldsymbol{d}_c)$, $b_{ct}(\boldsymbol{d}_c)$, and $\theta_{nt}(\boldsymbol{d})$ to denote $q_{ct}^p(d_c^1, \dots, d_c^{t-1})$(for $t > 1$), $l_{ct}(d_c^1, , \dots, d_c^{t-1})$, $b_{ct}(d_c^1 \dots, d_c^t)$, and $\theta_{nt}(\boldsymbol{d}_c^1 \dots, \boldsymbol{d}_c^t)$ respectively. Based on the deterministic model, we can obtain the multi-stage adaptive robust mixed integer optimization model (AMIO) as

$$\min_{q_{ct}^p(\boldsymbol{d}_c), l_{ct}(\boldsymbol{d}_c), q_{ct}^d(\boldsymbol{d}_c), b_{ct}(\boldsymbol{d}_c), \theta_{nt}(\boldsymbol{d})} \max_{\boldsymbol{d}\in\Xi} z = \sum_{t\in\mathcal{T}\setminus\{T-L_c+1,\cdots,T\}} \sum_{c\in\mathcal{C}} P_c\, q_{ct}^p(\boldsymbol{d}_c) + \sum_{t\in\mathcal{T}} \sum_{c\in\mathcal{C}} H_c\, l_{ct}(\boldsymbol{d}_c) + \sum_{t\in\mathcal{T}} \sum_{n\in\mathcal{N}} B_n \theta_{nt}(\boldsymbol{d}) \tag{3.6}$$

$$s.t. \quad q_{c,t-L_c}^p(\boldsymbol{d}_c) + l_{ct}(\boldsymbol{d}_c) = d_c^t + b_{c,t-1}(\boldsymbol{d}_c) - b_{ct}(\boldsymbol{d}_c) + l_{c,t+1}(\boldsymbol{d}_c), \quad \forall\, t \in \mathcal{T} \setminus \{1, \cdots, L_c\},\, c \in \mathcal{C}, \tag{3.7}$$

$$\sum_{n\in\mathcal{N}_c} \theta_{nt}(\boldsymbol{d}) \geq b_{ct}(\boldsymbol{d}_c), \quad \forall c \in \mathcal{C},\, t \in \mathcal{T}, \tag{3.8}$$

$$l_{ct}(\boldsymbol{d}_c) \geq SS_c, \quad \forall c \in \mathcal{C},\, t \in \mathcal{T}, \tag{3.9}$$

$$q_{ct}^p(\boldsymbol{d}_c),\, l_{ct}(\boldsymbol{d}_c),\, b_{ct}(\boldsymbol{d}_c), \theta_{nt}(\boldsymbol{d}) \in \mathbb{Z}^+, \quad \forall c \in \mathcal{C},\, t \in \mathcal{T}. \tag{3.10}$$

Note that constraints (3.7) can be used to replace $l_{ct}(\boldsymbol{d}_c)$ in the objective 3.6 and constraints (3.9) as shown in Section A1 of Appendix A. In the following contents, we will use q_{ct}^p, l_{ct}, b_{ct}, and θ_{nt} to replace $q_{ct}^p(\boldsymbol{d}_c)$, $l_{ct}(\boldsymbol{d}_c)$, $b_{ct}(\boldsymbol{d}_c)$, and $\theta_{nt}(\boldsymbol{d})$ respectively. Finally, the AMIO problem can be reformulated as

$$\min \quad z \tag{3.11}$$

$$s.t. \quad \sum_{c\in C} \left[\sum_{t=1}^{T-L_c} (P_c + (T-t)H_c) q_{ct}^p + H_c \sum_{t=1}^{T} b_{ct} + TH_c l_{c0} - H_c \sum_{t=1}^{T} (T-t+1) d_c^t \right] + \sum_{n=1}^{N} B_n \sum_{t=1}^{T} \theta_{nt} \leq z, \tag{3.12}$$

$$l_{ct} \geq SS_c, \quad \forall c \in \mathcal{C},\, t \in \mathcal{T}, \tag{3.13}$$

$$\sum_{n \in \mathcal{N}_c} \theta_{nt} \geq b_{ct}, \qquad \forall c \in \mathcal{C},\, t \in \mathcal{T}, \tag{3.14}$$

$$q^p_{ct},\, b_{ct},\, \theta_{nt} \in \mathbb{Z}^+, \qquad \forall c \in \mathcal{C},\, \forall n \in \mathcal{N},\, t \in \mathcal{T}. \tag{3.15}$$

where

$$l_{c,t+1} = \begin{cases} l_{c1} - \sum_{k=1}^{t} d^k_c + b_{ct},\, \forall\, t \in \{1, \cdots, L_c\} \\ l_{c1} + \sum_{k=1}^{t-L_c} q^p_{ck} - \sum_{k=1}^{t} d^k_c + b_{ct},\ \forall\, t \in \mathcal{T} \setminus \{1, \cdots, L_c\} \end{cases}$$

The model above is a mixed integer linear programming model and the decision variables include non-adaptive decision variable q^p_{c1} and the adaptive integer decision variables include q^p_{ct} (for $t > 1$), b_{ct}, and θ_{nt}.

3.4 The Improved Partition-and-bound Method

In the literature, multi-stage AMIO problems can be solved by the partition-and-bound method in Bertsimas and Dunning (2016). Unfortunately, their partition-and-bound method can only deal with small scale problems. The reasons are discussed in Section 3.4.1. To solve larger problems, we need to customize this method and the corresponding procedures are illustrated in Section 3.4.2.

3.4.1 Partition-and-bound method

The partition-and-bound method is proposed by Bertsimas and Dunning (2016) to solve the multi-stage AMIO problem. In this section, we first give a brief introduction to this method and then will discuss its advantages and disadvantages. The readers can refer to Bertsimas and Dunning (2016) for the detailed description of this method.

The partition-and-bound method begins with solving a static policy version of the AMIO problem to get a set of active uncertain parameters. For each constraint i, the active uncertain parameters are defined as the uncertain parameters (i.e., spare parts demand) that make the constraints have zero slack or the lowest slack (note that sometimes the lowest slack is not zero). After the static policy is solved, we use these active uncertain parameters to create a new finitely adaptive version of the studied problem. Solving this new version, in turn, generates a new set of active uncertain parameters which can be used to partition further, improving on the previous solution at each iteration and providing an upper bound. Meanwhile, at the end of each iteration, a “scenario-based" bound is applied to construct a lower bound: The set of active uncertain parameters available at the end of each iteration are used to build a deterministic problem which provides a lower bound.

The general idea of the partition-and-bound method has been described above. Next, we will discuss how to track those parameters across iterations and how the partitions are constructed. For the ease of illustration, a tree $\mathscr{T}$ of uncertain parameters is used to describe the partition construction scheme. For the tree $\mathscr{T}$, we define the following sets. The set of leaves of the tree $\mathscr{T}$ is denoted as $Leaves(\mathscr{T})$. The set of the parent node of $\hat{\boldsymbol{d}}$ in the tree $\mathscr{T}$ is denoted as $Parent(\hat{\boldsymbol{d}})$. The set of children nodes of $\hat{\boldsymbol{d}}$ is represented as $Children(\hat{\boldsymbol{d}})$. Finally, the set of children nodes who have the same parent node of $\hat{\boldsymbol{d}}$ is denoted as $Siblings(\hat{\boldsymbol{d}})$. In the algorithm, each iteration corresponds to each layer of the tree $\mathscr{T}$, and the partition related to the leaf $\hat{\boldsymbol{d}}^i$ is constructed as an intersection of partitions as

$$\begin{aligned}\varXi(\hat{\boldsymbol{d}}^i) =&\{\boldsymbol{d} \mid \| \hat{\boldsymbol{d}}^i - \boldsymbol{d} \|_2 \leq \| \hat{\boldsymbol{d}}^j - \boldsymbol{d} \|_2 \ \forall \hat{\boldsymbol{d}}^j \in Siblings(\hat{\boldsymbol{d}}^i)\} \\ &\cap \{\boldsymbol{d} \mid \| Parent(\hat{\boldsymbol{d}}^i) - \boldsymbol{d} \|_2 \leq \| \hat{\boldsymbol{d}}^j - \boldsymbol{d} \|_2 \ \forall \hat{\boldsymbol{d}}^j \in Siblings(Parent(\hat{\boldsymbol{d}}))\} \\ &\vdots \\ &\cup \varXi, \end{aligned} \tag{3.16}$$

which terminates when the root node is the parent, because it has no siblings. In each iteration, we create subpartitions of $\varXi(\hat{\boldsymbol{d}}^i)$ for the next iteration by simply adding the active uncertain parameters for partition $\varXi(\hat{\boldsymbol{d}}^i)$ in the current iteration as the children of $\hat{\boldsymbol{d}}^i$ in the tree $\mathscr{T}$. Through making decisions in different partitions, the method can reduce the overconservativeness embedded in the static robust optimization methods. Based on the machinery stated above, the steps of the partition-and-bound method for solving the multi-stage AMIO problems are shown in Appendix A2.

As shown in Appendix A2, the partition-and-bound method generates the partitions of uncertainty set based on the active uncertain parameters in the first two constraints which contain uncertain demand. Overall, considering there are C spare part types managed in T periods, we can find that there are $(1 + 2CT)$ types of constraints which will be used for partitioning at each node in each iteration, and for each constraint type, the constraints quantity will be equal to the possible scenarios embedded in the corresponding partition. For example, if the demand for spare part c in a partition has four possible realizations over the planning horizon, then there will be 4 constraints for the first constraint type of this spare part c, and finally lead to $4C$ of first type constraints in total created by this partition. Meanwhile, in each iteration, each node will generate $1 + 2T$ number of children nodes. In this case, the total nodes in the tree is calculated by $\sum_{i=0}^{I}(1+2T)^i = ((1+2T)^I - 1)/(2T)$ (where I is the maximum number of algorithm iterations set by the user).

For the purpose of illustration, we present a small three-stage problem example with two products which commonly use two spare parts in Appendix C. In this example, the lead time is one time period. There are three rounds of demand (one round in each stage) for repairing the faulty products and the uncertainty sets for each spare part demand are $\big((3,6),(7,10),(13,16)\big)$ and $\big((4,7),(6,9),(12,15)\big)$ respectively.

From this example, we can see that for large scale problem, where the uncertainty set is big, the partition model in Step 2 of the method can be hard to solve due to large quantity of constraints. These constraints are induced by the large quantity of possible scenarios for the uncertainty parameters. From this perspective, this method may not work well for the large scale problems. In Section 3.5, this point will be illustrated through a series of numerical analysis.

3.4.2 Reduce the quantity of constraints in the partition-and-bound method

From discussion in last subsection, we can conclude that the partition-and-bound method is highly impacted by the size of uncertainty set. If the uncertain set is big, even the small size problems will be hard to solve. In the next, we will discuss a way to decrease the problem size through diminishing the constraints quantity so that the algorithm can solve the large scale problem.

The key of the partition-and-bound method is to find the active uncertain parameters, i.e., the demand realizations in the constraints having the lowest slack, and then partition the uncertainty set based on them. If we can find the active uncertain parameters without solving the problem and decrease the constraints quantity by eliminating the ones without the active uncertain parameters, the partition-and-bound method will still be efficient. Now we consider a general form of multi-stage AMIO problem with adaptive partitions. Assume the T dimensional uncertain parameters are represented as $\boldsymbol{\xi}$, and decision variables vectors are denoted as $\boldsymbol{x}_j^t$. For the objective, the parameters of decision variables and uncertain parameters are respectively denoted as $\boldsymbol{c}_i^{(1)t}$ and $\boldsymbol{c}_i^{(1)t}$. For each constraint set i, the constant parameters, the parameters of decision variables, and the uncertain parameters are denoted as $b_i(\boldsymbol{\xi})$, $\boldsymbol{a}_i^{(1)t}$, and $\boldsymbol{a}_i^{(2)t}$ individually. In each iteration k of the method, the problems solved for each $Leaves(\mathscr{T}^k)$ with corresponding partitions $\Xi(\hat{\boldsymbol{\xi}}_j)$ are as follows.

$$
\begin{aligned}
z_{alg}(\mathscr{T}^k) = &\min_{\boldsymbol{x}\in\mathbb{Z},z} z \\
\text{s.t.} \quad & \sum_{t=1}^{T} \boldsymbol{c}^{(1)t}(\boldsymbol{\xi}) \cdot \boldsymbol{x}_j^t - \sum_{t=1}^{T} \boldsymbol{c}^{(2)t} \cdot \boldsymbol{\xi}^t \le z,\ \forall \boldsymbol{\xi} \in \Xi(\hat{\boldsymbol{\xi}}_j),\ \forall \hat{\boldsymbol{\xi}}_j \in Leaves(\mathscr{T}^k), \quad (3.17)\\
& \sum_{t=1}^{T} \boldsymbol{a}_i^{(1)t}(\boldsymbol{\xi}) \cdot \boldsymbol{x}_j^t - \sum_{t=1}^{T} \boldsymbol{a}_i^{(2)t} \cdot \boldsymbol{\xi}^t \ge b_i(\boldsymbol{\xi}),
\end{aligned}
$$

$$\forall \boldsymbol{\xi} \in \Xi(\hat{\boldsymbol{\xi}}_j),\, \forall \hat{\boldsymbol{\xi}}_j \in Leaves(\mathscr{T}^k),\, \forall i \in \{1,\dots,m\}, \quad (3.18)$$

$$\boldsymbol{x}_i^t = \boldsymbol{x}_j^t, \qquad \forall \hat{\boldsymbol{\xi}}_i, \hat{\boldsymbol{\xi}}_j \in Leaves(\mathscr{T}^k),\, \forall t : \Xi(\hat{\boldsymbol{\xi}}_i)^{t-1} \cap \Xi(\hat{\boldsymbol{\xi}}_j)^{t-1} \neq \emptyset, \quad (3.19)$$

where parameter vectors $\boldsymbol{c}^{(2)t}$ and $\boldsymbol{a}_i^{(2)t}$ only have non-negative elements. We have the following results.

Theorem 3.1. *Given an optimal solution $\boldsymbol{x}_j^*$ and $z^*_{alg}(\mathscr{T}^k)$, for constraints (3.17), the active uncertain parameters are always in the constraints with minimum sum-product of coefficients $\underline{\boldsymbol{c}}^{(2)t}$ and corresponding realized spare parts demands $\underline{\hat{\boldsymbol{\xi}}}_j^t \in \Xi(\hat{\boldsymbol{\xi}}_j)$:*

$$\sum_{t=1}^{T} \boldsymbol{c}^{(1)t}(\boldsymbol{\xi}) \cdot \boldsymbol{x}_j^{*t} - \sum_{t=1}^{T} \underline{\boldsymbol{c}}^{(2)t} \cdot \underline{\hat{\boldsymbol{\xi}}}_j^t \leq z^*_{alg}(\mathscr{T}^k), \qquad \forall \hat{\boldsymbol{\xi}}_j \in Leaves. \quad (3.20)$$

For constraints (3.18), the active uncertain parameters are always in the constraints with maximum sum-product of coefficients $\bar{\boldsymbol{a}}_i^{(2)t}$ and corresponding realized spare parts demands $\bar{\hat{\boldsymbol{\xi}}}_j^t \in \Xi(\hat{\boldsymbol{\xi}}_j)$:

$$\sum_{t=1}^{T} \boldsymbol{a}_i^{(1)t}(\boldsymbol{\xi}) \cdot \boldsymbol{x}_j^{*t} - \sum_{t=1}^{T} \bar{\boldsymbol{a}}_i^{(2)t} \cdot \bar{\hat{\boldsymbol{\xi}}}_j^t \geq b_i(\boldsymbol{\xi}), \qquad \forall \hat{\boldsymbol{\xi}}_j \in Leaves(\mathscr{T}^k), \quad (3.21)$$

Proof. See Appendix A4. □

Remark 3.1. *Theorem 3.1 shows a way to identify the constraints with active uncertain parameters: we only need to find the constraints with minimum sum-product of coefficients $\underline{\boldsymbol{c}}^{(2)t}$ and corresponding realized spare parts demands $\underline{\hat{\boldsymbol{\xi}}}_j^t \in \Xi(\hat{\boldsymbol{\xi}}_j)$. For some problems having uncertainty set with special structure, these active uncertain parameters are easy to find.*

Corollary 3.1. *When the coefficients of realized spare parts demands in (3.20) and (3.21) are all ones and the uncertainty set is a T-dimensional box:*

$$\Xi(\boldsymbol{\xi}_j) = \Xi(\boldsymbol{\xi}_j^1) \times \Xi(\boldsymbol{\xi}_j^2) \times \cdots \times \Xi(\boldsymbol{\xi}_j^T),$$

where

$$\Xi(\boldsymbol{\xi}_j^t) = [\hat{\boldsymbol{\xi}}_{j,\min}^t, \hat{\boldsymbol{\xi}}_{j,\max}], \qquad t = 1, 2, \cdots, T,$$

with $\hat{\boldsymbol{\xi}}_{j,\min}^t < \hat{\boldsymbol{\xi}}_{j,\max}$, *the active uncertain parameters for (3.20) are* $(\hat{\boldsymbol{\xi}}_{j,\min}^1, \hat{\boldsymbol{\xi}}_{j,\min}^2, \cdots, \hat{\boldsymbol{\xi}}_{j,\min}^T)$ *and for (3.21) are* $(\hat{\boldsymbol{\xi}}_{j,\max}^1, \hat{\boldsymbol{\xi}}_{j,\max}^2, \cdots, \hat{\boldsymbol{\xi}}_{j,\max}^T)$.

Corollary 3.1 gives us an easy way to decrease the size of the studied problem: We only need to consider the constraints with active uncertain parameters. Fortunately, the studied problem in this chapter satisfy the conditions of Corollary 1.

Corollary 3.2. *If we apply the results of Corollary 3.1 to our problem (3.11)-(3.15), the final model we are dealing with is*

$$\min\ z \tag{3.22}$$

$$s.t. \sum_{c \in C} \Big[\sum_{t=1}^{T-L_c} (P_c + (T-t)H_c) q_{ct}^p + H_c \sum_{t=1}^{T} b_{ct} + T H_c l_{c0} - H_c \sum_{t=1}^{T} (T-t+1) d_c^t \Big]$$

$$+ \sum_{n=1}^{N} B_n \sum_{t=1}^{T} \theta_{nt} \le z, \quad d_c^t = \underline{\hat{d}_c^t} \tag{3.23}$$

$$l_{ct} \ge SS_c, \quad d_c^t = \bar{\hat{d}}_c^t, \qquad \forall c \in \mathcal{C},\ t \in \mathcal{T}, \tag{3.24}$$

$$\sum_{n \in \mathcal{N}_c} \theta_{nt} \ge b_{ct}, \qquad \forall c \in \mathcal{C},\ t \in \mathcal{T}, \tag{3.25}$$

$$q_{ct}^p,\ b_{ct},\ \theta_t \in \mathbb{Z}^+, \qquad \forall c \in \mathcal{C},\ t \in \mathcal{T}. \tag{3.26}$$

where

$$l_{c,t+1} = \begin{cases} l_{c1} - \sum_{k=1}^{t} d_c^k + b_{ct},\ \forall\, t \in \{1, \cdots, L_c\} \\ l_{c1} + \sum_{k=1}^{t-L_c} q_{ck}^p - \sum_{k=1}^{t} d_c^k + b_{ct},\ \forall\, t \in \mathcal{T} \setminus \{1, \cdots, L_c\}. \end{cases}$$

and $\underline{\hat{d}_c^t}$ *and* $\bar{\hat{d}}_c^t$ *denote the lower bound and upper bound of the demand realization for each spare part respectively.*

From Corollary 3.2, it is clear that all the possible realizations of spare parts demand in constraints (3.12) and (3.13) are replaced by some given realizations as constraints

(3.23) and (3.24) respectively. But one question is if this is valuable, i.e., if the model size will be decreased significantly? We have the following results on how many constraints are kept in our problem after applying Corollary 3.1.

Corollary 3.3. *Let* $\mid \hat{d}_c^t \mid = (\bar{\hat{d}}_c^t - \underline{\hat{d}_c^t} + 1)$*, i.e.,* $\mid \hat{d}_c^t \mid$ *represents the number of possible realizations of spare part* c*'s demand during period* t*. If we apply the results of Corollary 3.1, for constraints (3.12) in the AMIO model, the constraint quantity will be reduced from* $(\Pi_{c=1}^{C} \Pi_{t-1}^{T} \mid \hat{d}_c^t \mid)$ *to 1 in constraints (3.23). Similarly, for constraints (3.13), the constraint quantity will be reduced from* $(\sum_{c=1}^{C} \sum_{t=1}^{T} \Pi_{k=1}^{t} \mid \hat{d}_c^t \mid)$ *to* $(C \times T)$ *in constraints (3.24).*

Proof. See Appendix A5. □

We will further illustrate this result with the example discussed in Appendix A3 as shown in Appendix A6. From the discussion, we demonstrate Corollary 3.2 can significantly decrease the constraints quantity in the studied multistage AMIO model.

3.5 Numerical Experiments

In this section, the numerical experiments of applying the improved partition-and-bound method to the spare parts inventory management problem instances are presented. In the experiments, we randomly generate the instances of the multi-stage AMIO problem to compare the performance of the classical partition-and-bound method proposed by Bertsimas and Dunning (2016) and the improved one proposed in this chapter. These instances have different combinations of product types (N) and spare parts types (C), and the same number of planning periods (T). The parameters in the instances are randomly chosen using the uniform distribution as follows: Market sizes taken by all the products in the assortment are chosen from $M_t \in [40, 100]$ units for all $t = 1, \ldots, T$, respectively. The failure rates of all products in each period are elements of $\rho_{nt} \in [2, 10]$ percents for all $n = 1, \ldots, N$, and $t = 1, \ldots, T$. The minimum portion $\underline{\sigma}_{ct}$ and maximum possible portion $\bar{\sigma}_{ct}$ of the products failures caused by spare part c are respectively picked from $\underline{\sigma}_{ct} \in [60, 80]$ and $\bar{\sigma}_{ct} \in [90, 110]$ percents of the estimated product failure quantity for all $c = 1, \ldots, C$, and $t = 1, \ldots, T$. The unit prices of spare parts ordered from supplier take values from $P_c \in [5, 10]$ dollars and the order lead time is $L_c = 1$ for all $c = 1, \ldots, C$. The holding costs of spare parts are elements of $H_c \in [1, 2]$ for all $c = 1, \ldots, C$ and the backorder costs of products are chosen from $B_n \in [15, 20]$ dollars for all $n = 1, \ldots, N$. The base stock levels of all spare parts are randomly chosen from $SS_c \in [1, 2]$ units and the initial inventory level is 8 units for all $c = 1, \ldots, C$. We also assure that the initial inventory level is a given value which is greater than the maximum possible demand realization for each part. The number of planning periods for each instance is set as $T = 3$. All the numerical experiments are coded in C++ and carried out through the IBM ILOG CPLEX 20.1 optimization package on a PC with an Intel Core i7-10750H 2.60 GHz CPU with 16 GB RAM.

In the first subsection, the advantages brought by the improved method will be illustrated. We solve the small problem instances by the classical partition-and-bound method and the improved one respectively. The complexity of the optimization model and solution time is compared. Afterwards, the large instances are solved by the improved method in this chapter to explore managerial insights on managing the spare parts inventory of the products from an assortment.

3.5.1 Small problem instances

We first illustrate the advantages brought by the improved partition-and-bound method compared to the classical one, i.e., the method in Bertsimas and Dunning (2016). As shown in Tables 3.2 , there are 9 random generated instances, each with different pairs of products type number ($N = 3, 4, 5$) and spare parts type number ($C = 3, 4, 5$). Both methods are evaluated from the perspectives of the variable and constraint quantities in the partition model (A.10) and computation time. Only the results in the first three iterations of the classical and improved methods are listed respectively because both methods terminate within three iterations when the instances are solvable. In the columns of "(Var., Cons.)" and "Time", the corresponding results of the improved method are listed in the first place, followed by those of the classical one. For example, in column 3 and row 2, "$(\mathbf{25}, \mathbf{19})/(25, 1.62 \times 10^6)$" indicates there are 25 variables and 19 constraints in the model of the improved method, while 25 variables and 1.62×10^6 constraints in that of the classical one at the first iteration. In the column 4 and row 4, "$\mathbf{0}/16.70$" shows the first iteration of our improved method takes less than 0.01 seconds (we use 0 to denote any value less than 0.01), while that of the classical one takes 16.70 seconds. In addition, we use "$-$" as placeholder in the results if the instances cannot be solved.

Table 3.2: Numerical results of small instances

Ins.	(N, C)	Iteration 1		Iteration 2		Iteration 3	
		(Var., Cons.)	Time (s)	(Var., Cons.)	Time (s)	(Var., Cons.)	Time (s)
1	$(3,3)$	$(\mathbf{25},\mathbf{19})/(25,1.62\times10^6)$	$\mathbf{0}/16.70$	$(\mathbf{97},\mathbf{103})/(97,611887)$	$\mathbf{0.38}/4.80$	$(\mathbf{385},\mathbf{769})/(385,97912)$	$\mathbf{0.17}/0.97$
2	$(3,4)$	$(\mathbf{30},\mathbf{25})/(-,-)$	$\mathbf{0}/-$	$(\mathbf{117},\mathbf{136})/(-,-)$	$\mathbf{0}/-$	$(\mathbf{465},\mathbf{1016})/(-,-)$	$\mathbf{0}/-$
3	$(3,5)$	$(\mathbf{35},\mathbf{31})/(-,-)$	$\mathbf{0.03}/-$	$(\mathbf{137},\mathbf{169})/(-,-)$	$\mathbf{0}/-$	$(\mathbf{545},\mathbf{1216})/(-,-)$	$\mathbf{0.05}/-$
4	$(4,3)$	$(\mathbf{28},\mathbf{19})/(28,413613)$	$\mathbf{0}/4.90$	$(\mathbf{109},\mathbf{103})/(109,172063)$	$\mathbf{0.31}/1.81$	$(\mathbf{433},\mathbf{744})/(433,29724)$	$\mathbf{0.17}/0.55$
5	$(4,4)$	$(\mathbf{33},\mathbf{25})/(-,-)$	$\mathbf{0}/-$	$(\mathbf{129},\mathbf{136})/(-,-)$	$\mathbf{0}/-$	$(\mathbf{513},\mathbf{1025})/(-,-)$	$\mathbf{0.17}/-$
6	$(4,5)$	$(\mathbf{38},\mathbf{31})/(-,-)$	$\mathbf{0}/-$	$(\mathbf{149},\mathbf{169})/(-,-)$	$\mathbf{0.03}/-$	$(\mathbf{593},\mathbf{1276})/(-,-)$	$\mathbf{0.28}/-$
7	$(5,3)$	$(\mathbf{31},\mathbf{19})/(-,-)$	$\mathbf{0}/-$	$(\mathbf{121},\mathbf{103})/(-,-)$	$\mathbf{0.16}/-$	$(\mathbf{481},\mathbf{766})/(-,-)$	$\mathbf{0.14}/-$
8	$(5,4)$	$(\mathbf{36},\mathbf{25})/(-,-)$	$\mathbf{0.02}/-$	$(\mathbf{141},\mathbf{136})/(-,-)$	$\mathbf{0.97}/-$	$(\mathbf{561},\mathbf{1036})/(-,-)$	$\mathbf{0.08}/-$
9	$(5,5)$	$(\mathbf{41},\mathbf{31})/(-,-)$	$\mathbf{0.02}/-$	$(\mathbf{161},\mathbf{169})/(-,-)$	$\mathbf{0}/-$	$(\mathbf{641},\mathbf{1271})/(-,-)$	$\mathbf{0.03}/-$

From Table 3.2, we can see the improved method dominates the classical one over all tested instances. For Instances 1 and 4, both methods can solve them and find the optimal solutions, but the solution time of the improved one is much less because the quantity of constraints in the partition model is much smaller. The reduced scale of the partition model in the improved method not only leads to the increased solution efficiency, but also makes some unsolvable instances to the classical one solvable. Except for Instances 1 and 4, the classical method cannot solve the rest instances due to the large scale of partition model. On the contrary, the improved method is able to solve all the instances in reasonable time and find optimal solutions in three iterations.

3.5.2 Medium and large problem instances

To further test the performance of the improved partition-and-bound method, we implement it to solve some medium and large problem instances. In total, we evaluated 16 problem instances which include four instance subsets. Each subset has different number of product types for $N \in \{5, 10, 15, 20\}$ but the instances within the same subset have same number of products types. However, the instances in the same subset have various types of spare parts , i.e., there are four instances in each subset and the number of spare parts types in each instance are chosen from $C \in \{5, 10, 15, 20\}$. The improved method stops when one of the following two termination conditions is met: The algorithm iteration number reaches its limits, which is 10; or the percentage gap between the upper bound and lower bound reaches the predetermined value, which is 1%. The corresponding results are illustrated in Table 3.3 and the bound gap of each instance in each iteration of the method is plotted in Figure 3.1.

Table 3.3: Results of the improved partition-and-bound method on the medium and large instances

Set	Ins.	N	C	UB	LB	Gap (%)	Time (s)	Iterations
1	1	5	5	574	572	0.350	6.749	3
	2	5	10	608	608	0.000	138.119	4
	3	5	15	622	622	0.000	240.475	4
	4	5	20	960	960	0.000	5862.420	5
2	5	10	5	716	710	0.845	2.603	3
	6	10	10	805	805	0.000	202.937	4
	7	10	15	1158	1151	0.608	2831.840	5
	8	10	20	1203	1203	0.000	2726.050	5
3	9	15	5	1290	1282	0.624	63.603	4
	10	15	10	1396	1394	0.143	120.806	4
	11	15	15	1568	1568	0.000	3465.600	5
	12	15	20	1710	1705	0.293	61255.400	6
4	13	20	5	1614	1609	0.311	62.944	4
	14	20	10	1637	1631	0.368	115.825	4
	15	20	15	1773	1772	0.056	2658.300	5
	16	20	20	1806	1792	0.781	3042.110	5

As shown in Table 3.3, all 16 instances are solved by the improved method in reasonable time. Even for the most time consuming instance, i.e., Instance 12 with $N = 15$ and

$C = 20$, the computation time is within one day. It is observed that as the size of the problem instance increases, the number of iterations needed to obtain optimal solutions also increases, as well as the computation time. The reason is quite straightforward: The large size problem usually means the corresponding uncertainty set is big, hence more partitions are needed to divide the uncertainty set.

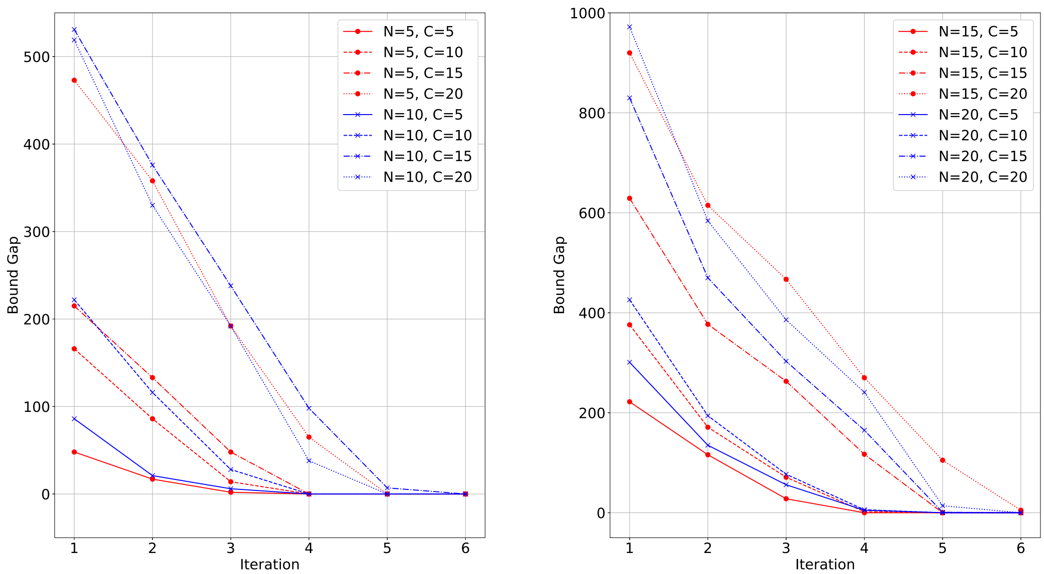

Figure 3.1: Improvements and bound gaps for 16 instances in four subsets with instance sizes $N \in \{5, 10, 15, 20\}$

[†]The four instances in the same subset are in the same colour. They have different types of spare parts $C \in \{5, 10, 15, 20\}$ and corresponding results are represented by the lines with different shapes in the same color from bottom to top respectively.

From Figure 3.1, we can see the convergence trend of the bound gap in each tested instance. The improvement is significant in each iteration and the bound gap decreases sharply especially in the first three or four iterations, based on the size of the problem under study. However, the improvement on the solutions is at the expense of computation time. The computation time explosively grows as the iteration number increases. For example, to solve Instance 11 with $(N = 15, C = 15)$ in Subset 3, the improved method terminates at the fifth iteration and takes 3465.600 seconds. However, when the number of spare parts types increased to $C = 20$ in Instance 12 of the same subset, it terminates at the fifth iteration and solution time raises sharply to 61255.400 seconds. Therefore, in real-life applications, managers have to make a trade-off between solution quality and computation time.

3.5.3 Exploring the factors affecting spare parts inventory management decisions

In this subsection, we will explore the managerial insights on managing the spare parts inventory for a product assortment. In general, there are several factors affecting the spare parts inventory management decisions such as the demand volume, price, backorder cost, etc.

First of all, as mentioned earlier in Section 3.3.2, the sales quantity of each product in the assortment is determined by the attractiveness to customers, i.e., the more attractive the product is, the higher sales quantity will be. The sales quantity affects the spare parts demand in the planning period thus impacting the corresponding inventory replenishment decisions. Note that some common spare parts are simultaneously used by several products in the assortment while some dedicated spare parts are only used by a particular product. Hence, one question is that is there a clear difference in the order quantity between the common spare parts and the dedicated ones of the popular and unpopular products? Secondly, the replenishment decisions of spare parts inventory are also impacted by their prices, inventory and holding costs, and the repair backorder costs of products using them. The repair backorder costs affecting these decisions are not the direct costs incurred by spare parts, but the indirect backorder costs of the products using them. Importantly, in this reassembly repair system, a failed product will not be repaired until all spare parts are available. In other words, if the inventory of a spare part is not ready, then all product repairs using it may under the risk of repair delays with penalty. In this case, another question is if there is any difference in order quantity among the spare parts having different prices and used by the products with different backorder costs?

To address these questions, two tests are conducted. In the first one, we compare the inventory decisions for four types of spare parts in an assortment with three products. These products have various sales quantities: A most popular one (product A), a medium popular one (product B), and a less popular one (product C). Among the four types of spare parts, spare part 1 is a common one used by all products, and the other three are dedicated ones used by each of the three products individually, i.e., spare part 2 is used by product A only, spare part 3 by product B only, and spare part 4 by product C only. To control the impacts of other factors on inventory decisions, spare parts prices, holding costs, and other parameters are set as the same across all spare parts. The products backorder costs are also the same for all products. In the results, the spare parts inventory and product backorder decisions are different for each partition set because decisions are made upon observing different demand realizations in different partitions. Therefore, we take the average value of the same decisions in different partitions and then round the average value to the nearest integer. The corresponding results are illustrated in Table 3.4.

Table 3.4: Inventory decisions for managing inventory of four spare parts in three products with different on-market quantities

q^p_{ct}	$t = 1$	$t = 2$
1	29	41
2	9	18
3	7	15
4	1	7

θ_{nt}	$t = 1$	$t = 2$	$t = 3$
A	22	< 1	< 1
B	0	< 1	< 1
C	0	< 1	< 1

From Table 3.4, there are two order decisions made for each spare part type in three periods. The common spare part 1 is ordered more than the dedicated spare parts 2, 3, and 4 on average. This observation indicates the common spare parts will be ordered more than the dedicated ones because the quantity of products using the former is higher. In addition, among the dedicated spare parts, the order quantity of spare part

2 is the highest, followed by those of spare parts 3 and 4 respectively. This result shows the popularity of products affects the order quantities of spare parts used in them. As for the product backorder quantity, one interesting observation is that, even though all products have the same backorder cost, product A is the only backordered product at the first period due to the insufficient spare parts inventory. This phenomenon may be explained by the following two points. Firstly, the on-market quantity of product A is the highest so the demands for spare parts used in it is not sufficiently covered by the initial inventories in the first period, but this is not the case for products B and C. Secondly, the failures of products B and C caused by the failed common spare part 1 are more likely to be repaired by utilizing the spare part 1 disassembled from the failed product A's with the dedicated spare part 2 failed only, especially when product A is the most popular product which has large on-market quantity. Note that our formulation (constraints (3.25)) allows using a good commonly used spare part in one product to replace a faulty one in another product. Therefore, when the failure quantity is relatively large, product A may have many failures incurred only by the faulty dedicated spare part 2's and these failed products have other spare parts in good condition. After the disassembly, these good spare part 2's can be used to fix other products so that the total backorder quantity will be reduced.

In the second test, we carry out a sensitivity analysis to examine the impacts brought by the prices of spare parts and backorder costs of products on the total costs. We still consider the same instance discussed in the last test. But we will change one parameter of one particular product at one time while fixing other parameters and record the corresponding total inventory costs. The results are summarized as two line charts as shown in Figure 3.2.

In the left line chart of Figure 2, the horizontal axis represents the change of backorder cost of a certain product; the vertical axis represents the total cost, i.e., the optimal objective value. It shows the product backorder cost substantially impacts the total cost.

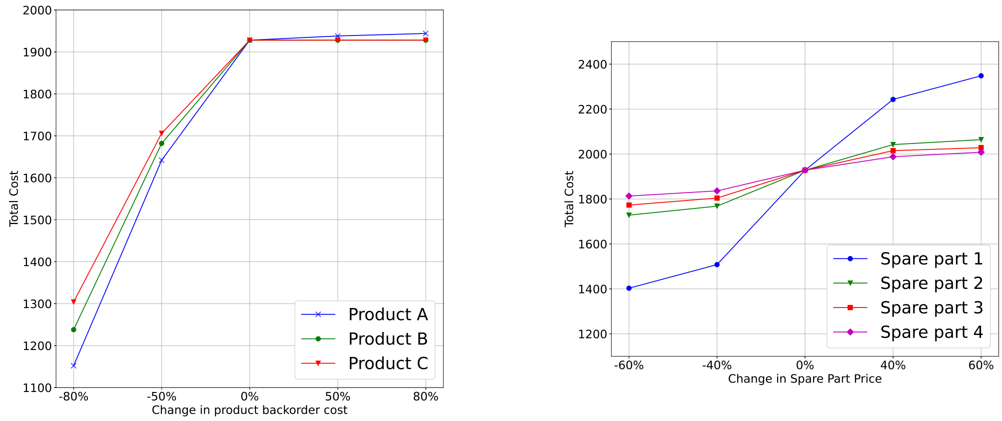

Figure 3.2: Effects of product backorder cost and spare part price on the total cost

The most popular product, product A, has more impacts on total cost than the other two less popular products. In addition, the impacts on total cost are more significant when decreasing the backorder costs of all three products, compared with the increasing corresponding costs. This can be explained by the cost savings through backlogging the repairs of the products with lowered backorder cost as many as possible. However, when raising the backorder cost of the most popular product, the corresponding impacts on total costs are trivial. This is can be explained as follows. When the backorder cost of the popular product is rising, the products with unchanged backorder costs will be backlogged at first so that the total cost will not be raised up. This indicates the demands for the spare parts used by the popular products will be satisfied at the first place. However, the on-hand inventory of dedicated and common spare part for the popular product A may not be sufficient though the common parts disassembled from the other two products are utilized. Therefore, the high demand for repairing the faulty popular products will not be fully met by the on-hand spare parts inventory. As a result, backlogging some demands for repairing the faulty product A is unavoidable. On the other hand, we observe that this phenomenon does not exist when lifting the backorder cost of the other two less popular products. This is because the high failure quantity of the most popular product can contribute some disassembled common spare parts in good condition or simply because the on-hand dedicated spare parts inventory is sufficient. In other words, the unbroken common spare parts in the popular products waiting to be repaired play a role of inventory buffer for replacing those faulty common spare parts in other products. In this case, when the inventory of the dedicated spare parts used in less popular products is sufficient, most of the faulty products with high backorder cost caused by the broken common spare parts can be fixed, leading to the unchanged total cost. A managerial insight from this observation is the managers should be devising the after-sales policies for both popular and less popular products specifically to achieve cost minimization.

As shown in right line chart of Figure 3.2, the total cost is also affected by the prices of both dedicated and common spare parts. The price of dedicated spare parts is less important, because they have less effect on the total cost than common spare part 1. To be specific, spare part 1 is the most used one among all spare parts because it is used by all on-market products. Therefore, lowering the price of spare part 1 can significantly decrease the purchasing cost of spare parts and finally lead to the reduced total cost. Among the dedicated spare parts, the impacts on total cost brought by different prices are also affected by the product popularity. The price of spare part 2, which is used only in the most popular product, is more important than those of spare parts 3 and 4 which are used in the less popular products. This observation indicates that the managers should give decreasing the purchasing prices of both the common spare parts and the dedicated spare parts used in the popular products the highest priority during the procurement negotiations. This conclusion is intuitive qualitatively, but our method can provide managers the quantitative order quantities when the prices of spare parts are changing.

3.6 Conclusions and Future Research Directions

In this chapter, we consider a multi-period spare parts inventory system providing spare parts for several products in an assortment. The OEM follows a repair-replacement policy to fulfill the aftersales services. To handle the uncertainty embedded in the spare parts demand, we formulate a multi-stage adaptive mixed-integer optimization model by assuming the probability distributions of product failures are unknown and the objective is to minimize the total inventory costs including spare parts purchase cost, holding and product backorder cost by determining proper inventory policy. We improve the partition-and-bound method to solve the proposed model and conduct extensive numerical experiments to validate its performance. It is found that the improved method can solve small instances in a fairly short time and dominates the classical one in the medium and large instances. Through sensitivity analysis, we explore the impacts of spare parts purchase cost, product popularity, and product backorder cost on inventory policy and total cost, and provide some managerial insights regarding how to adjust the order quantities for both the dedicated and common spare parts used in the popular and unpopular products and how to adjust the order quantities of those spare parts when the backorder costs of products using them are changing.

There are several directions for future research. Firstly, in this chapter, we assume that the probability distribution is unknown. This assumption can be changed when the OEM has more historical data of products on market. We could use stochastic programming (SP) or distributionally robust optimization (DRO) when more information is available. Secondly, we call for more studies on improving the partition-and-bound method to solve the AMIO. The proposed method in this chapter did improve the classical method significantly because the impact brought by the "curse of dimension" is reduced through refining the constraints in the model. However, it is clear that the computation time of each iteration still increase sharply when the iteration number increases. Thirdly, we note that there are very few studies focusing on the spare parts inventory management on the consumer goods, even though many activities are making this topic more important than before, such as the "right to repair" motion mentioned at the beginning of this chapter. In addition, the repair and remanufacturing operations have already been emphasized by many manufacturers such as Apple, Hyundai, and Microsoft to fulfill their promises to supply chain sustainability (Hanley et al., 2020). Since spare parts inventory management plays an important role in those operations, we call on more researchers to dive into this topic.

Chapter 4

Dynamic assortment planning with uncertainty in customer preferences

4.1 Introduction

In the multi-period selling season, an original equipment manufacturer (OEM) is faced with assortment planning decisions at each period. The customer customer preferences are changing over the season, indicating the dynamic market nature under a multi-period context. Note that the customer preferences of product variants are affected by many factors varying during different periods in the selling seasons, such as seasonality (Caro et al., 2014), promotion activities (Liao et al., 2009), technology changes in products (Tripsas, 2008), new products launching (Li and Calantone, 1998), customer reviews in the online platform (Lim and Lee, 2015), etc. In this case, revising the assortment offered to market at each period based on the estimated customer preferences is beneficial because it can bring more economic benefits to the retailers (Kök et al., 2015). However, it relies on accurate predictions on customer preference trend (Jiang et al., 2019). This is especially suitable to the OEMs who produce and sell product assortments through online retailing platforms. In contrast to brick-and-mortar stores, online retailers have access to tremendous amount of customers' browsing and purchasing data, thus facilitating them to take advantage of data to estimate customer preferences at each period of the selling season. Therefore, online retailers are more capable and likely to periodically revise the assortment to catering the customer preferences in each periods so as to increase their revenue. Such assortment revision can be easily exercised by the online platforms through selecting a product's "in-stock" or "out-of-stock" status information displayed to customers in each period.

In this chapter, we consider an OEM who produces a dynamic assortment of products and sells them through online platform over a selling season with multiple periods. Prior to the selling season, the OEM determines a product line which includes a set of substitute products to release to the market, subject to a cardinality constraint. During each period of the selling season, the OEM picks product variants from the product line to form the assortment that it carries by, based on the estimated preference of customers in each period, and then produces the variants in a build-to-order or assemble-to-order system to fulfill the orders placed by customers. The choice process of customers is modelled through the multinomial logit (MNL) framework and the customers can be offered with dynamic assortments at different periods. In each period, besides the assortment planning decisions, the OEM also decides the quantity of components purchased from suppliers for fulfilling the customers' demand on the offered assortment and repairing the failed products sold in previous periods under warranty contract to maximize the expected revenue. This problem is faced by many manufactures who produce and sell product assortment through online platforms. At the product design phase before the selling season, rather than all the product variants, they have to select a limited number of variants from all potential substitute variants as a product line to produce in the selling season. This is usually due to limited production capacity or budget. In addition, from the marketing perspective, a significant amount of literature points out that too many product variants in an assortment may lead to negative consequences such as information overload, increased cognitive effort, choice uncertainty, choice difficulty, and hence choice avoidance (Sethuraman et al., 2022). Therefore, the manufacturer should limit the product line size offered to the market. In the proposed model, we will limit the size of product line through a cardinality constraint.

In addition to the product line and assortment planning decisions, we also intend to explore the role of component commonality on the aforementioned decisions in a multi-period context. Based on the orders placed to the product variants in the specific assortment offered at each period, the OEM fulfills the placed orders through an assemble-to-order or build-to-order system. In this system, the OEM decides the stocking decisions for the components used in production at each period. Each product variant is produced based on a bill of materials (BOM) that dictates the components used in its fabrication. Among the components, there are dedicated ones which are product-specific, and common ones which are used to a subset (or all) of the product variants. For example, a motherboard is common to a product line of personal computers (PCs), a CPU is common to a subset of the PCs, and different touch screens are specific to different PC variants respectively. It is assumed that the OEM holds zero or very little inventory of components on the manufacturing site but orders the components from the suppliers when the orders of customers are received and productions begin, and the final production/assembly time is negligible. The reasons for this assumption is that nowadays the OEMs with strong bargaining power are more like to adopt the strategy of vendor managed inventory (VMI) under a lean management philosophy so that they hold very little temporary inventory of components on site for manufacturing. Moreover, the period length in this problem is quite long such that it is not reasonable from economic perspective to hold components inventory over the periods rather than order them directly from the suppliers in each period. Finally, in the studied problem the OEM also has to order more components in each period to repair the returned faulty products under warranty sold in previous periods.

A good example is Dell's production of PCs and retailing through their own online platform. One PC model has abundant specifications, i.e., different combinations of CPU, GPU, memories, drives, screen, keyboard, etc., but not all combinations are available to customers when the new model is released. In addition, at different periods, the available specifications are further reduced. For example, you probably can only find fewer amounts of specifications during a promotion period than a regular period, or after the launch of a new generation of product line than before.To produce the ordered PCs, Dell implements an assemble-to-order system. It also requires suppliers to maintain the ownership of components inventory until they are pulled into the assembly line (Dedrick and Kraemer, 2002).

Even though the OEM with online retailing channel can take advantage of collected data to estimate customer preferences, such estimation is usually coupled with inherent uncertainty of the parameters and/or data. To address such inherent uncertainty, we will formulate the studied problem as a multi-stage stochastic programming model.

The contribution of this study is two-fold. First of all, to our best knowledge, this multi-period dynamic assortment planning problem with a blended setup of uncertain customer preferences and component stocking was unexplored in the literature. This problem models the situation faced by many OEMs who produce and sell product assortments through the online platforms and are able to utilize the historical data to estimate the customer preferences over the selling season. Secondly, a branch-and-price (B&P) algorithm is designed to solve the proposed multi-stage programming model. Through extensive numerical experiments, the complexity of this problem is illustrated and the performance of the proposed algorithm is validated. The advantage of dynamic assortment planning, i.e., dynamically changing the assortment at different periods based on the estimated customer preferences, is also highlighted in the numerical experiments.

The rest of this chapter is organized as follows. In Section 4.2, a brief literature review is presented. In Section 4.3, a deterministic model is formulated for this problem at first, followed by an extension to a multi-stage stochastic programming model when uncertainty in customer preferences is considered. To solve the multi-stage stochastic programming model, a B&P algorithm is proposed in Section 4.4. Section 4.5 presents two sets of numerical experiments. In the first set, the value brought by the dynamic assortment decisions is explored against the static assortment decisions. In addition, we examine the influences of parameters, and product and component structures. In the second set of numerical experiments, the performances of both the B&P algorithm and the CPLEX solver are compared. Finally, we conclude the chapter and discuss future research directions in Section 4.6.

4.2 Literature Review

The study in this chapter is closely related to three streams of literature. The first one is the retail assortment planning literature in which the studies focusing on the assortment planning decisions under static customer preferences with no forecasting uncertainty. A detailed literature review is provided by Kök et al. (2015). Ryzin and Mahajan (1999) study an assortment planning problem for the products in the same category by integrating the newsboy model as inventory model with the MNL model as customer choice model. They show that the optimal assortment policy is to include a certain number of most popular products in the assortment. Mahajan and Van Ryzin (2001) consider dynamic consumer substitution in the assortment planning problem and develop a stochastic gradient algorithm for solving the problem. Talluri and Van Ryzin (2004) find that the revenue ordered assortment is optimal when customer choice is modelled through an MNL model and customer preferences are deterministic and known. Cachon et al. (2005) develop a stylized model for the assortment planning process incorporated with consumer search cost and find that it may be optimal to include an unprofitable product in the assortment when considering consumer search. Honhon et al. (2010) introduce a locational choice model as consumer choice model to a single-period assortment planning problem and consider the stockout-based substitution. They design a dynamic programming algorithm to determine the optimal assortment decisions and inventory levels. Wang (2012a) studies the joint pricing and assortment optimization under the MNL model with cardinality constraints. All the aforementioned literature consider the single-period assortment planning problem with deterministic customer preferences.

The second stream of literature is the dynamic assortment planning problem. The dynamic nature of this problem exists in two dimensions: One is the arriving customers in a single period are heterogeneous, i.e., there are multiple possible realizations (or so-called segments) of customer preferences , and the other is customer preferences are changing over a selling season with multiple periods, resulting in two sub-streams of research. In the first sub-stream, the customer preferences are not deterministic and the goal is to find the optimal assortment that maximizes the expected revenue from customer visits. Rusmevichientong et al. (2014) study a single-period assortment planning problem under the MNL model with random customer preferences and identify two special cases in which the revenue-ordered assortment is optimal under uncertain customer preferences. Méndez-Díaz et al. (2014) propose a branch-and-cut algorithm for the single-period problem with cardinality constraints. Feldman and Topaloglu (2015) derive a tractable upper bound on the expected revenue for the problem. The upper bound can be used to identify the optimality gap of heuristics.

In the second sub-stream, the single-period problem is extended to a problem with a multi-period selling season, thus becoming more complicated. In this problem, the assortment decisions are revised periodically based on the changing customers preferences. There are some research working on dynamic assortment problem with demand learning in which the retailer does not have the information of customer preferences. In this case, the retailer usually learns about the preferences by experimenting with different assortments and observing their sales in different periods. The key trade-off in this problem is between exploration and exploitation products, i.e., there are two kinds of products to be included in the assortment at each period: Exploitation products which are profitable in the current period and exploration products whose demand information can be gathered more to see if they will be profitable in the future periods. Caro and Gallien (2007) first propose the model for this problem by utilizing a stylized multiarmed bandit model. Rusmevichientong et al. (2010) study a dynamic assortment problem with demand learning by incorporating the MNL model and capacity constraint. Ulu et al. (2012) focus on the the dynamic assortment problem with demand learning in which customers' taste is modelled through a locational choice model. Sauré and Zeevi (2013) study a family of stylized dynamic assortment planning problems with demand learning under a limited assortment size and develop a set of policies which can limit the assortment experimentation on exploring the profitable products in the future periods. Aforementioned studies on dynamic assortment planning with demand learning usually relies on the assumption that retailer only has the detailed information on the customers' choice towards a sub set of products variants to put into the assortment and can learn the preferences of the other product variants by putting them in an experimental assortment during the exploration periods. In addition, the customer preferences are updated through observing the sales quantity of product variants in the experimental assortment. On critics on this sub-stream of studies is that the customer preference is assumed to be affected only by the sales of substitute products in the assortments. However, in real-world, there are many factors that affect the customers' preferences such as price changes, customer reviews, seasonality and so on.

The third stream of literature is assortment customization problem which is about revising the product variants in the assortments provided to different types of customers after estimating or observing their preferences (Bernstein et al., 2015). The assortment customization suites well to the online retailers due to their ability of processing and analyzing the information of customers' purchasing history to censor the preferences towards products. Based on the estimated preferences of customers, the online retailers can freely control the product variants exposed to the customers by strategically labelling the product variants displayed on the web page as "in-stock" or "out-of-stock". Current literature on assortment customization is thriving. Bernstein et al. (2015) first introduce the dynamic assortment customization problem with consideration of limited inventories. They consider a stylized model in which the Poisson arrival process is used to model customer arrivals. After observing their types, the retailer will offer a customized assortment to each of arrived customers. The selection of products included in the customized assortment is made based on the preference of arrived customer and the product inventory level upon arrival. They find that the optimal policy to include a product variant in the assortment is a threshold type, i.e., a product variant will be dropped from the assortment offered to an arriving customer only when its stock level is lower than a threshold value. El Housni and Topaloglu (2021) study a joint assortment optimization and customization problem in which the preferences of different customer types are modelled through a MNL model. They propose a two-stage model, i.e., the firm selects a certain amount of product variants from a set of candidate variants to form an assortment carried to the second stage when different types of customers arrive with different arrival probabilities and preferences. The firm will provide each type of customers with an customized assortment through dropping some variants in the pre-determined assortment in the first stage.

It should be noted that the model of Bernstein et al. (2015) assumes the preference of each customer type is unchanged over different periods of the selling season and so are the probabilities of the arriving customer belonging to different types. This assumption may be true when the selling season is short or the attractiveness of products to customer is stable. There are several studies assuming the preferences of customers are changing over the selling season. Caro et al. (2014) study a multi-period assortment planning problem for short-lived products. In the proposed model, the market shares of the product variants in assortment are determined by the MNL model based on customer preferences. The customer preference for a particular product variant in the assortment decays over time because the short-live products usually have a big spike when they are newly offered to market but then their market share would shrink drastically. Rather than releasing all products at once in the beginning of selling season, the firm should revise the product assortment by frequently adding new variants at different periods in the selling season to maximize the total profits. Ghoniem and Maddah (2015) propose a deterministic model for jointly optimizing pricing and assortment decisions over a multi-period selling season under the changing customer preferences to maximize the total retailer profit. In their model, the customer preferences are reflected by the willingness to pay of customers and are changing through the selling periods due to seasonality. Although considering different customer segments (types), they do not customize the assortment for different types of customers. In addition, their assortment planning decision is static, i.e., the variants in the assortment are not changing over the season.

In our proposed model, we further extend the problem raised by El Housni and Topaloglu (2021) to a multi-period context when the distribution of customer preferences is non-static. In addition, we also intend to explore role of the commonality of components for fabricating the products in the assortment on the assortment decisions. A relevant study from this perspective is done by Bernstein et al. (2011). They explore the effect of component commonality on product line decisions in an assemble-to-order system. A single period stylized model in which product line decisions and component procurement decisions are made to maximize the expected is proposed. Their result indicates in some cases introducing the commonality in the manufacturing may decrease the product variety in the product line.

Our study distinguishes the aforementioned studies in the following aspects. First, our problem is under a multi-period context in which customer preferences are changing over the periods. We assume that the OEM can estimate the distributions of customer preferences towards the products in any periods of selling season based on historical data.

Note that there are many statistical methods that can be used to extract customer preferences from historical data on surveys or online reviews such as partial least squares analysis (Nagamachi, 2008), statistical linear regression (You et al., 2006), artificial neural networks (Chen et al., 2006), fuzzy inference systems (Jiang et al., 2019), etc. Unlike the multi-period assortment planning problem in Bernstein et al. (2015), which assumes the preference distributions of customers in different periods are identical and the preferences of each customer type toward each product are unchanged through the periods in selling season, we relax this assumption. This relaxation will make the stylized model in Bernstein et al. (2015) invalid, but can capture the dynamic nature of customer preferences towards the the assortment in a multi-period selling season. Second, we consider the uncertainty embedded in customer preferences estimations in each period and also the component demands for repairing the faulty products sold, which are not discussed in Bernstein et al. (2011). We will develop a multi-stage stochastic programming model to handle such uncertainties in estimating the non-stationary customer preferences over the selling season and finally help OEM jointly decide the product line decisions, assortment decisions, and components procurement decisions under the uncertain customer preferences.

4.3 Problem Formulation

In this section, the formulations of the studied problem will be presented. A deterministic model will be formulated and linearized. Afterwards, we will formulate a multi-stage stochastic programming model to address the inherent uncertainties in estimating customer preferences when making joint product line and assortment planning decisions.

4.3.1 A deterministic mathematical programming model for assortment planning and component management

At the beginning of a selling season with T periods, an OEM is able to produce a set of product variants $\mathcal{N} = \{1, \ldots, N\}$ and let P_{nt} denote the price for each product variant $n \in \mathcal{N}$ at period $t \in \mathcal{T} = \{1, \ldots, T\}$. Meanwhile, the OEM has to select a subset of at most S product variants to form a product line in a multi-period selling season. We denote the product line subset as $\mathcal{S} \subseteq \mathcal{N}$. At a certain period t of the selling season, the OEM will select a subset of the product variants in $\mathcal{S}$ as an assortment offered to the customers based on the estimated customers preferences in that period. The set of product variants included in the assortment at period t is denoted by $\mathcal{S}_t \subseteq \mathcal{S}$.

To model the demand for the product variants in assortment $\mathcal{S}_t$, we consider a consumer choice model. To be specific, consumers' choice to a product variant within the offered assortment is based on the MNL model. In this model, a customer obtains utility U_{it} if he or she purchases product variant $i \in \mathcal{S}_t$, given the displayed assortment $\mathcal{S}_t$ in period t and the utility is determined by $U_{it} = V_{it} + \epsilon_{it}$, in which V_{it} represents the observed mean utility of product variant i to the customer and ϵ_{it} is a random term which represents the unobserved utility. For the ease of exposition, we will denote the vector of mean utilities of the customers for all the products as $\boldsymbol{V}_t = (V_{1t}, \ldots, V_{|\mathcal{S}_t|t})$. If the random term ϵ_{it} is assumed to be independent and identically distributed (i.i.d.) with Gumbel distribution, the MNL model gives the utility maximization choice probability that customer selects product variant i as

$$\gamma_{it}^{\mathcal{S}_t} = \frac{e^{V_{it}}}{\sum_{j \in \mathcal{S}_t} e^{V_{jt}} + e^{V_{0t}}}. \tag{4.1}$$

where V_{0t} represents the customers' utility of no purchase option at period t (Train, 2009).

In the proposed model, a decision variable γ_{nt} is used to denote the estimated probability that a consumer will select product variant $n \in \mathcal{N}$. Let $e_{nt} = e^{V_{nt}}$, $\forall n \in \mathcal{N}$, $\forall t \in \mathcal{T}$, represent the consumers' utility for purchasing product n, then for any $n \in \mathcal{N}$, we have

$$\gamma_{nt} = \frac{x_{nt} \cdot e_{nt}}{\sum_{j \in \mathcal{N}} x_{jt} \cdot e_{jt} + e_{0t}} \tag{4.2}$$

where x_{nt} is binary variable and $x_{nt} = 1$ if and only if the product variant n is included in the assortment $\mathcal{S}_t$ displayed to customers at period t.

Each product $n \in \mathcal{N}$ is a multi-indenture system in which the components are non-identical, i.e., distinct types of components are used in a product. We assume every type of component has at most one unit contained in a unit of product. To be specific, the BOM for all product variants' production consists of a set of components $\mathcal{C} = \{1, 2, \ldots, C\}$ and for a specific product variant n, the corresponding BOM is denote as $\mathcal{C}_n$, which is a subset of $\mathcal{C}$. In addition, from the BOMs of all product variants, we can also obtain the information on all product variants using component c and will denote the set of product variants using component c as $\mathcal{N}_c$.

The notation used in the formulated deterministic mathematical programming model is listed in Table 4.1. In the model, the profits obtained by the OEM are equal to the revenue of selling product variants in the assortments minus the costs of obtaining components used for the fabrication of new products and the repair of faulty products over the selling season.

The OEM first decides the product line before the selling season and then determines the products in the assortment offered to the market based on the estimated customer preferences at each period of the season. The revenue obtained from selling the products offered in the assortment $\mathcal{S}_t$ during period t is calculated as

$$Revenue = \sum_{n \in \mathcal{N}} P_{nt} \cdot d_{nt} \tag{4.3}$$

Table 4.1: Notations used in the deterministic model

Sets and subscripts	
$\mathcal{N}$	Set of products can be selected in the assortment
n	Subscript of product $n \in \mathcal{N}$
$\mathcal{C}$	Set of components used for product fabrication
c	Subscript of component $c \in \mathcal{C}$
$\mathcal{N}_c$	Set of products using component c, $\forall\, c \in \mathcal{C}$
$\mathcal{C}_n$	Set of components used for fabrication of product n, $\forall\, n \in \mathcal{N}$
$\mathcal{T}$	Set of planning periods
t	Subscript of planning period $t \in \mathcal{T}$
$\mathcal{S}$	Set of product variants included in the product line, $\mathcal{S} \subseteq \mathcal{N}$
$\mathcal{S}_t$	Set of product variants included in the assortment at period $t \in \mathcal{T}$, $\mathcal{S}_t \subseteq \mathcal{S}$
Parameters	
P_{nt}	Selling price of product n at period t
K_{ct}	Stocking cost of component c at period t
W_t	Total market demand for product assortment $\mathcal{S}_t$ at period t
r_{cnt}	Average joint failure probability that a faulty product variant n having component c failure at period t
M	A sufficiently large positive integer number
e_{nt}	$e_{nt} = e^{V_{nt}}$ is referred to as the consumers' utility for product n at period t
S	Limit on the number of product variants included in the product line $\mathcal{S}$
T	Number of planning periods
Decision variables	
x_{nt}	Binary variable, $x_{nt} = 1$ if and only if the product variant n is included in the assortment $\mathcal{S}_t$ offered to customers at period t for all $n \in \mathcal{N}$, $t \in \mathcal{T}$
x_{n0}	Binary variable, $x_{n0} = 1$ if and only if the product variant n is included in the assortment $\mathcal{S}$ offered to customers over whole selling season
γ_{nt}	Probability of consumers placing an order on product n in the assortment $\mathcal{S}_t$ at period t
d_{nt}	Fulfilled order amount for product n in period t
d_{ct}	Demand for component c in period t
q_{ct}	The order quantity of component c in period t

in which d_{nt} representing the fulfilled order amounts and satisfies constraints $\gamma_{nt} \cdot W_t - 1 \leq d_{nt} \leq \gamma_{nt} \cdot W_t$ (the definition W_t in Table 4.1), which guarantee the order quantity for product variant n should be the maximum integer no greater than the estimated demand. In addition, γ_{nt} is defined in equation (5.2) and W_t is the total market size faced by the assortment at that period.

In each period, asides from the revenue obtained from selling products in assortment $\mathcal{S}_t$, the OEM has to pay for the costs of procuring the components used for fabricating products in the displayed assortment and repairing the faulty products under warranty sold in previous periods. In the model, we assume that at most one unit of each component type is used in one unit of each product variant. In addition, stocking one unit of component c at period t will induce an aggregate cost K_{ct} which covers inventory costs, handling costs, transportation costs, etc. The total amount of the required component c should be no less than the sum of two demand streams for components. One stream is the total assembly demand for the offered product variants whose fabrication requires component c and the other is the total repairing demand for fixing faulty products having component c failures. To be specific, for the first stream, the assembly demand d^a_{ct} for component c at period t is computed as

$$d^a_{ct} = \sum_{n \in \mathcal{N}_c} d_{nt}. \tag{4.4}$$

The faulty components creates another stream of demand for components. The failure of product is caused by the failures of components used in it. When a sold product fails during warranty periods, it will be returned to the OEM for replacing faulty components. Let r_{cnt} represent the average joint failure probability that a faulty product variant n having faulty component c for all $c \in \mathcal{C}$, $n \in \mathcal{N}$ at period t. In this case, the total failure amount of product variant n caused by the faulty component c at period t is equal to the product of average joint failure rate and the total quantity of that product variant sold in all periods before period t, i.e., $d^r_{cnt} = r_{cnt} \cdot \sum_{k=1}^{t-1} d_{nk}$. In this case, the total repairing demand amount for replacing all faulty component c in the faulty products at period t is

$$d^r_{ct} = \sum_{n\in\mathcal{N}_c} r_{cnt} \cdot \sum_{k=1}^{t-1} d_{nk} \tag{4.5}$$

In real-life, the average joint failure probabilities can be estimated based on the historical data.

To sum up, the OEM seeks to determine the product line decisions at beginning of the selling season, the assortment planning decisions at each period t of the season, i.e., which product variants in the product line should be included in the assortment $\mathcal{S}_t$ based on the estimated customer preferences, and the stocking level decisions of all components at each period, to maximize the total profits obtained in the whole season. The corresponding deterministic mathematical model is formulated as

$$\max \quad \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} P_{nt}\cdot d_{nt} - \sum_{t\in\mathcal{T}}\sum_{c\in\mathcal{C}} K_{ct}q_{ct} \tag{DP1}$$

$$s.t. \quad \gamma_{nt}\cdot W_t - 1 \le d_{nt} \le \gamma_{nt}\cdot W_t \qquad \forall\, n\in\mathcal{N},\, \forall\, t\in\mathcal{T}, \tag{DP2}$$

$$q_{ct} \ge \sum_{n\in\mathcal{N}_c} (d_{nt} + r_{cnt}\cdot \sum_{k=1}^{t-1} d_{nk}), \qquad \forall\, c\in\mathcal{C},\, \forall\, t\in\mathcal{T}, \tag{DP3}$$

$$\gamma_{nt}\Big(\sum_{j\in\mathcal{N}} x_{jt}e_{jt} + e_{0t}\Big) = x_{nt}e_{nt}, \qquad \forall\, n\in\mathcal{N},\, \forall\, t\in\mathcal{T}, \tag{DP4}$$

$$x_{n0} - x_{nt} \ge 0, \qquad \forall\, n\in\mathcal{N},\, \forall\, t\in\mathcal{T}, \tag{DP5}$$

$$\sum_{n\in\mathcal{N}} x_{n0} \le S, \tag{DP6}$$

$$q_{ct} \le M \sum_{n\in\mathcal{N}_c} x_{n0}, \qquad \forall\, c\in\mathcal{C},\, \forall\, t\in\mathcal{T}, \tag{DP7}$$

$$x_{0t},\, x_{nt} \in \{0,1\}, \qquad \forall\, n\in\mathcal{N},\, \forall\, t\in\mathcal{T}, \tag{DP8}$$

$$q_{ct},\, d_{nt} \in \mathbb{Z}^+, \gamma_{nt} \in \mathbb{R}^+, \qquad \forall\, n\in\mathcal{N},\, \forall c\in\mathcal{C},\, \forall\, t\in\mathcal{T}. \tag{DP9}$$

In this model, the objective (DP1) is to maximize the total profits of the OEM when it offers assortment $\mathcal{S}_t$ based on the product line $\mathcal{S}$ over a multi-period selling season.

The profit is represented as the total revenue of selling the product variants in the assortment minus the total stocking costs for all components required for fabrication and repair of those products in the season. Constraints (DP2) with (DP9) all together ensure the fulfilled order quantity should be the maximum integer no greater than the estimated demand for each product by the MNL model in any periods. Constraints (DP3) represent the stocking levels of all components should be no less than the required amounts in each period. These constraints link the demand for product variants to the stocking level decisions of all components. The stocked components will be used for both assembly of new product in that period and the repair of sold products in previous periods. Note that no components inventory will be carried over to the next period. We assume the OEM follows a lean philosophy and keeps low inventory level. So in a period, the stocking components will be used to satisfy the demand in this period and not be kept as inventory for next period. Constraints (DP4) address the choice probabilities of customers over the product variants in the assortment $\mathcal{S}_t$ at period t. Constraints (DP5) ensure the product variants displayed in the assortment $\mathcal{S}_t$ at period t are selected from product line $\mathcal{S}$ which is determined at the beginning of selling season. Constraint (DP6) is the cardinality constraint which limits the maximum number of product variants included in the product line $\mathcal{S}$. Constraints (DP7) restrict the stocking level of a component to be zero if it is not used in fabrication of any variants in the line. Constraints (DP8) and (DP9) regulate the binary variables x_{n0} and x_{nt}, integer variables q_{ct} and d_{nt}, and non-negative real variables γ_{nt} for all $n \in \mathcal{N}$, $c \in \mathcal{C}$, and $t \in \mathcal{T}$, respectively. Note that this model is a mixed integer nonlinear programming due to the nonlinear terms, i.e., $\gamma_{nt} \sum_{j \in \mathcal{N}} x_{jt} e_{jt}$, in constraints (DP4).

4.3.2 Linearization

In this section, the linearization technique is used to adapt the proposed deterministic model to a linear program. Specifically, the nonlinear term $\gamma_{nt} \sum_{j \in \mathcal{N}} x_{jt} e_{jt}$ in constraints

(DP4), for all $n \in \mathcal{N}, t \in \mathcal{T}$, and $j \in \mathcal{N}$. These constraints can be linearized by introducing auxiliary variables $z_{njt} = \gamma_{nt} \cdot x_{jt}$, for all $n \in \mathcal{N}, t \in \mathcal{T}$, and $j \in \mathcal{N}$. In addition, to ensure $z_{njt} = \gamma_{nt} \cdot x_{jt}$ holds when $x_{jt} = 0$ and $x_{jt} = 1$ for all $t \in \mathcal{T}$, and $j \in \mathcal{N}$, we need the following four sets of additional constraints:

$$z_{njt} \leq x_{jt}, \tag{L1}$$

$$z_{njt} \leq \gamma_{nt}, \tag{L2}$$

$$z_{njt} \geq \gamma_{nt} + x_{jt} - 1, \tag{L3}$$

$$z_{njt} \geq 0. \tag{L4}$$

By introducing constraints (L1)-(L4), auxiliary variables z_{njt} will take different values depending on the values of binary variables x_{jt} as follows.

$$z_{njt} = \begin{cases} \gamma_{nt} \cdot x_{jt}, & \text{when } x_{jt} = 1; \\ 0, & \text{when } x_{jt} = 0. \end{cases}$$

In this case, the nonlinear model will be reformulated as a mixed integer linear programming (MILP) model as follows.

$$\max \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} P_{nt} \cdot d_{nt} - \sum_{t \in \mathcal{T}} \sum_{c \in \mathcal{C}} K_c q_{ct} \tag{DPL1}$$

$$s.t. \quad \gamma_{nt} \cdot W_t - 1 \leq d_{nt} \leq \gamma_{nt} \cdot W_t, \qquad \forall n \in \mathcal{N}, \forall t \in \mathcal{T}, \tag{DPL2}$$

$$q_{ct} \geq \sum_{n \in \mathcal{N}_c} (d_{nt} + r_{cnt} \cdot \sum_{k=1}^{t-1} d_{nk}), \qquad \forall c \in \mathcal{C}, \forall t \in \mathcal{T}, \tag{DPL3}$$

$$e_{0t}\gamma_{nt} + \sum_{j \in \mathcal{N}} e_{jt}\, z_{njt} = x_{nt}\, e_{nt}, \qquad \forall n \in \mathcal{N}, \forall t \in \mathcal{T}, \tag{DPL4}$$

$$x_{n0} - x_{nt} \geq 0, \qquad \forall n \in \mathcal{N}, \forall t \in \mathcal{T} \tag{DPL5}$$

$$\textbf{DPL} \qquad \sum_{n \in \mathcal{N}} x_{n0} \leq S, \tag{DPL6}$$

$$q_{ct} \leq M \sum_{n \in \mathcal{N}_c} x_{n0}, \qquad \forall c \in \mathcal{C}, \forall t \in \mathcal{T}, \tag{DPL7}$$

$$z_{njt} \leq x_{jt}, \qquad \forall n \in \mathcal{N}, \forall j \in \mathcal{N}, \forall t \in \mathcal{T}, \tag{DPL8}$$

$$z_{njt} \leq \gamma_{nt}, \qquad \forall n \in \mathcal{N}, \forall j \in \mathcal{N}, \forall t \in \mathcal{T}, \tag{DPL9}$$

$$z_{njt} \geq \gamma_{nt} + x_{jt} - 1, \qquad \forall n \in \mathcal{N}, \forall j \in \mathcal{N}, \forall t \in \mathcal{T}, \tag{DPL10}$$

$$z_{njt} \geq 0, \qquad \forall n \in \mathcal{N}, \forall j \in \mathcal{N}, \forall t \in \mathcal{T}, \tag{DPL11}$$

$$x_{nt} \in \{0, 1\}, \qquad \forall n \in \mathcal{N}, \forall t \in \mathcal{T}, \tag{DPL12}$$

$$q_{ct}, d_{nt} \in \mathbb{Z}^+, \qquad \forall n \in \mathcal{N}, \forall c \in \mathcal{C}, \forall t \in \mathcal{T}. \tag{DPL13}$$

This deterministic model is reasonable when the OEMs prediction on customer preferences is completely correct. However, this can hardly be the case in real-life. In the following, we will incorporate the uncertainty into the proposed model.

4.3.3 The multi-stage stochastic programming model

The deterministic model formulated in the previous subsection does not consider the uncertainty embedded in estimating the choice parameters (customer preferences) of the MNL model. If the estimated mean utility vector $\boldsymbol{V}_t$ in the MNL model is fixed (the OEM is 100% sure about the estimation) and known at each period t, we can easily find the customer choice probability γ_{nt} for each product n. Unfortunately, such estimations are embedded with uncertainty. In this case, the mean utilities that customers attach to the product variants can be treated as random variables with certain probability distributions.

Uncertainty in estimating choice parameters in the MNL model

Assume that there are G_t possible realizations of customer preferences with certain probability distributions at period t, with each realization of customer preferences following the MNL model. To be specific, the mean utility vector $\boldsymbol{V}_t$ is a discrete random vector, which takes G_t different values $\hat{\boldsymbol{V}}_t^1, \ldots, \hat{\boldsymbol{V}}_t^g, \ldots, \hat{\boldsymbol{V}}_t^{G_t}$, where $\hat{\boldsymbol{V}}_t^g = (\hat{V}_{1t}^g, \hat{V}_{2t}^g, \ldots, \hat{V}_{Nt}^g)$ denotes the gth realization of the mean utilities of customers at period t. In addition, ϕ_{gt} is the probability that the customer preferences are revealed as gth realization at period t, where $\sum_{g=1}^{G_t} \phi_{gt} = 1$, $\forall\, t \in \mathcal{T}$. To sum up, this setup corresponds to the situation where the vector of mean utilities has G_t possible realizations $\hat{\boldsymbol{V}}_t^1, \ldots, \hat{\boldsymbol{V}}_t^{G_t}$ and the vector of mean utilities is realized as $\hat{\boldsymbol{V}}_t^g$ with probability ϕ_{gt}. In this setting, the preference weight e_{nt} that customers attach to product variant n at period t is also a random variable. To establish the stochastic programming model, all realizations of the random variable vector $\boldsymbol{e}_t = (e_{1t}, \ldots, e_{nt}, \ldots, e_{Nt})$ are deemed as scenarios. These scenarios all together create the scenario tree of the demand on product variant n and the corresponding realization $\hat{\boldsymbol{e}}_t = \{\hat{e}_{1t}, \ldots, \hat{e}_{nt}, \ldots, \hat{e}_{Nt}\}$ can be determined based on the realizations of the mean utility of customers $\hat{\boldsymbol{V}}_t$. The probability of each node in the scenario tree is equal to the probability of the realization of customer mean utility $\hat{\boldsymbol{V}}_t$ in that node. If the total amount of mean utility vector realizations in each period t is G_t, then the total quantity of nodes in the scenario tree for period T is $\Pi_{t=1}^{T} G_t$. In the scenario tree, the node set will be denoted as $\mathscr{T} = \{0, 1, 2, \ldots, |\mathscr{T}|\}$, where node 0 represents the root node. An illustration of the scenario tree $\mathscr{T}$ is shown in Figure 4.1. In the figure, there are three possible realizations of the mean customer utility vector toward all products during each planning period t, i.e., $\hat{\boldsymbol{V}}_t = (\hat{\boldsymbol{V}}_t^1, \hat{\boldsymbol{V}}_t^2, \hat{\boldsymbol{V}}_t^3)$. Consequently, the scenario tree has a number of 3^t scenarios in each time period t.

A multi-stage stochastic programming model

Considering the aforementioned uncertainty in estimating customer preferences, the deterministic model of the studied problem can be extended to a multi-stage stochastic programming model. Except for the notation provided in Table 4.1, additional notation used in formulation of the multi-stage stochastic programming model are listed in Table 4.2. The model is formulated as follows.

Table 4.2: Additional notation used in the multi-stage stochastic programming model

Sets and subscripts	
$\mathcal{N}$	Set of products can be selected in the assortment
n	Subscript of product $n \in \mathcal{N}$
$\mathcal{C}$	Set of components used for product fabrication
c	Subscript of component $c \in \mathcal{C}$
$\mathcal{N}_c$	Set of products using component c, $\forall\, c \in \mathcal{C}$
$\mathcal{C}_n$	Set of components used for fabrication of product n, $\forall\, n \in \mathcal{N}$
$\mathscr{T}$	Set of all nodes in the scenario tree
$\mathscr{T}_0$	Set of scenario tree nodes except the root node, i.e., $\mathscr{T}_0 = \mathscr{T} \setminus \{0\}$
m	Subscript of node in the scenario tree, $\forall m \in \mathscr{T}$
$\mathscr{A}(m)$	Set of all predecessors of node m, $\forall\, m \in \mathscr{T}$
$\mathscr{S}(m)$	Set of all successors of node m, $\forall\, m \in \mathscr{T}$
Parameters	
P_{nm}	Selling price of product n at node m, $\forall\, n \in \mathcal{N}$, $\forall\, m \in \mathscr{T}$
ϕ_m	Probability of node m, $\forall\, m \in \mathscr{T}$
K_{cm}	Stocking cost of component c at node m, $\forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}$
W_m	Total market size at node m, $\forall\, m \in \mathscr{T}$
M	A sufficiently large positive integer number
e_{nm}	$e_{nm} = e^{V_{nm}}$ is referred to as the preference weight that customers attach to product n at node m, $\forall\, n \in \mathcal{N}$, $\forall\, m \in \mathscr{T}$
e_{0m}	Preference weight that customers attach to no-purchase option at node m, $\forall\, m \in \mathscr{T}$
r_{cnm}	Average joint failure probability that a faulty product variant n having faulty component c at node m, $\forall\, c \in \mathcal{C}$, $\forall\, n \in \mathcal{N}_c$, $\forall\, m \in \mathscr{T}$
Decision variables	
x_{nm}	Binary variable, $x_{nm} = 1$ if and only if the product variant n is included in the assortment $\mathcal{S}_m$ offered to the market at node m, $\forall\, m \in \mathscr{T}$
x_{n0}	Binary variable, $x_{n0} = 1$ if and only if the product variant n is included in the assortment $\mathcal{S}$ offered to the market over the selling season
γ_{nm}	Probability of consumers placing an order on product n in the assortment $\mathcal{S}$ at node m, $\forall\, m \in \mathscr{T}$
d_{nt}	Fulfilled order quantity for product n at node m, $\forall\, m \in \mathscr{T}$
q_{cm}	The order quantity of component c at node m, $\forall\, m \in \mathscr{T}$
z_{njm}	$z_{njm} = \gamma_{nm} \cdot x_{jm}$, auxiliary variables used for linearization, $\forall\, n \in \mathcal{N}$, $\forall\, j \in \mathcal{N}$, $\forall\, m \in \mathscr{T}$

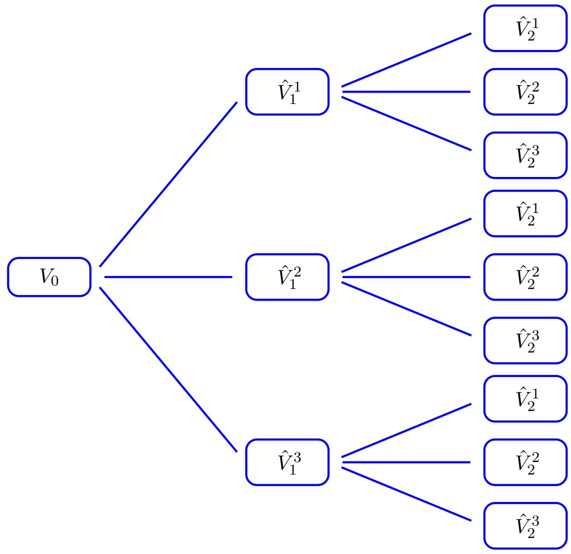

Figure 4.1: An illustration of the scenarios tree

$$\max \quad \sum_{m\in\mathscr{T}} \phi_m \Bigg(\sum_{n\in\mathcal{N}} P_{nm} \cdot d_{nm} - \sum_{c\in\mathcal{C}} K_{cm} q_{cm} \Bigg) \tag{SPR1}$$

$$s.t. \quad \gamma_{nm} \cdot W_m - 1 \leq d_{nm} \leq \gamma_{nm} \cdot W_m, \qquad \forall\, n \in \mathcal{N}, \forall\, m \in \mathscr{T}, \tag{SPR2}$$

$$q_{cm} \geq \sum_{n\in\mathcal{N}_c} \Bigg(d_{nm} + r_{cnm} \sum_{m'\in\mathscr{A}(m)} d_{nm'} \Bigg), \qquad \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}, \tag{SPR3}$$

$$e_{0m}\gamma_{nm} + \sum_{j\in\mathcal{N}} e_{jm}\, z_{njm} = x_{nm}\, e_{nm}, \qquad \forall\, n \in \mathcal{N}, \forall\, m \in \mathscr{T}, \tag{SPR4}$$

$$x_{n0} - x_{nm} \geq 0, \qquad \forall\, n \in \mathcal{N}, \forall\, m \in \mathscr{T}, \tag{SPR5}$$

$$\textbf{SPR} \quad \sum_{n\in\mathcal{N}} x_{n0} \leq S, \tag{SPR6}$$

$$q_{cm} \leq M \sum_{n\in\mathcal{N}_c} x_{n0}, \qquad \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}, \tag{SPR7}$$

$$z_{njm} \leq x_{jm}, \qquad \forall\, n \in \mathcal{N}, \forall\, j \in \mathcal{N}, \forall\, m \in \mathscr{T}, \tag{SPR8}$$

$$z_{njm} \leq \gamma_{nm}, \qquad \forall\, n \in \mathcal{N}, \forall\, j \in \mathcal{N}, \forall\, m \in \mathscr{T}, \tag{SPR9}$$

$$z_{njm} \geq \gamma_{nm} + x_{jm} - 1, \qquad \forall\, n \in \mathcal{N}, \forall\, j \in \mathcal{N}, \forall\, m \in \mathscr{T}, \tag{SPR10}$$

$$z_{njm} \geq 0, \qquad \forall\, n \in \mathcal{N}, \forall\, j \in \mathcal{N}, \forall\, m \in \mathscr{T}, \tag{SPR11}$$

$$x_{n0},\, x_{nm} \in \{0, 1\}, \qquad \forall\, n \in \mathcal{N},\, \forall\, m \in \mathscr{T}, \quad \text{(SPR12)}$$

$$q_{cm},\, d_{nm} \in \mathbb{Z}^+,\, \gamma_{nm},\, z_{njm} \in \mathbb{R}^+, \qquad \forall\, n \in \mathcal{N},\, \forall\, c \in \mathcal{C},\, \forall\, m \in \mathscr{T}. \quad \text{(SPR13)}$$

where $\mathscr{A}(m)$ represents the set of all the predecessors of node $m \in \mathscr{T}$.

4.4 The Branch-and-Price Algorithm

The multi-stage stochastic programming model is a large-scale mixed integer programming model, making it difficult to solve by commercial solvers such as CPLEX. However, it is notable that the original model has a coefficient matrix with special structure, i.e., block-angular structure. This allows us to use the Dantzig-Wolfe decomposition method to reduce solution difficulty. In this section, we design a B&P algorithm for solving the proposed multi-stage stochastic programming model. The B&P algorithm is based on the branch-and-bound approach, and applies the Dantzig-Wolfe decomposition to each node of the branch-and-bound tree. The basic idea of Dantzig-Wolfe decomposition is "divide and conquer", i.e., to transform a large-scale original problem into many smaller sub-problems by utilizing the special structure of the original problem to reduce solution time.

4.4.1 Outline of the B&P algorithm

In the model, constraints (SPR2), (SPR4), (SPR8), (SPR9), and (SPR10) are specific to each node $m \in \mathscr{T}$ of the scenario tree and constraints (SPR6) is specific to the root node. In contrast, constraints (SPR3), (SPR5), and (SPR7) are coupling constraints in which variables from different scenario tree nodes are involved. According to the Dantzig-Wolfe decomposition approach, they are the constraints staying in the master problem while the other constraints are included in the pricing problem for each node m of the scenario tree.

For each node $m \in \mathscr{T}$, define $\mathcal{X}_m = \{(\boldsymbol{x}_m, \boldsymbol{d}_m, \boldsymbol{q}_m) | \forall m \in \mathscr{T}\}$, where $\boldsymbol{x}_m = \{x_{nm} | \forall n \in \mathcal{N}\}$ with $x_{nm} \in \{0,1\}$, $\boldsymbol{d}_m = \{d_{nm} | \forall n \in \mathcal{N}\}$, with $d_{nm} \in \mathbb{Z}^+$ and $\boldsymbol{q}_m = \{q_{cm} | \forall c \in \mathcal{C}\}$ with $q_{cm} \in \mathbb{Z}^+$, All $\boldsymbol{x}_m$, $\boldsymbol{d}_m$ and $\boldsymbol{q}_m$ satisfy the corresponding constraints (SPR2), (SPR4), (SPR8), (SPR9), and (SPR10). Since the variables are either binary or bounded integers (it is easy to find the upper bounds for the integer variables), the set $\mathcal{X}_m$ has finite point and can be written as $\mathcal{X}_m = \{(\boldsymbol{x}_m^k, \boldsymbol{q}_m^k, \boldsymbol{d}_m^k) | k = 1, \ldots, K_m\}$. Based on the Minkowski's Representation Theorem (Wolsey and Nemhauser, 1999), any points $(\boldsymbol{x}_m, \boldsymbol{q}_m, \boldsymbol{d}_m)$ in $\mathcal{X}_m$ can be represented as $\boldsymbol{x}_m = \sum_{k=1}^{K_m} \lambda_m^k \boldsymbol{x}_m^k$, $\boldsymbol{q}_m = \sum_{k=1}^{K_m} \lambda_m^k \boldsymbol{q}_m^k$, and $\boldsymbol{d}_m = \sum_{k=1}^{K_m} \lambda_m^k \boldsymbol{d}_m^k$, respectively, where $\sum_{k=1}^{K_m} \lambda_m^k = 1$ and $\lambda_m^k \in \{0,1\}$, $\forall k = 1, \cdots, K_m$. Finally, the master problem (MP) can be formulated as

$$
\begin{aligned}
& \max \sum_{m \in \mathscr{T}} \phi_m \sum_{k=1}^{K_m} \Big(\sum_{n \in \mathcal{N}} P_{nm} \cdot d_{nm}^k - \sum_{c \in \mathcal{C}} K_{cm} q_{cm}^k \Big) \lambda_m^k && \text{(MP1)} \\
& s.t. \sum_{k=1}^{K_m} q_{cm}^k \lambda_m^k \geq \sum_{n \in \mathcal{N}_c} \Big(\sum_{k=1}^{K_m} d_{nm}^k \lambda_m^k \\
& \qquad + r_{cnm} \sum_{m' \in \mathscr{A}(m)} \sum_{k=1}^{K_{m'}} d_{nm'}^k \lambda_{m'}^k \Big), && \forall c \in \mathcal{C}, \forall m \in \mathscr{T}, \quad \text{(MP2)} \\
& \sum_{k=1}^{K_0} x_{n0}^k \lambda_0^k - \sum_{k=1}^{K_m} x_{nm}^k \lambda_m^k \geq 0, && \forall n \in \mathcal{N}, \forall m \in \mathscr{T}, \quad \text{(MP3)} \\
\textbf{MP} & \sum_{k=1}^{K_m} q_{cm} \leq M \sum_{n \in \mathcal{N}_c} \sum_{k=1}^{K_0} x_{n0}^k, && \forall c \in \mathcal{C}, \forall m \in \mathscr{T}, \quad \text{(MP4)} \\
& \sum_{k=1}^{K_m} \lambda_m^k = 1, && \forall m \in \mathscr{T}, \quad \text{(MP5)} \\
& \lambda_m^k \in \{0,1\}, && \forall k = 1, \ldots, K_m, \forall m \in \mathscr{T}. \quad \text{(MP6)}
\end{aligned}
$$

By replacing $\lambda_m^k \in \{0,1\}$ by $\lambda_m^k \geq 0$, we can get a linear relaxation of MP, which can be optimally solved to obtain the lower bound of the branch-and-bound tree node in the algorithm. However, the cardinality of $\mathcal{X}_m$ might be huge for any nodes $m \in \mathscr{T}$, such that finding all the points (a.k.a. columns) for $\mathcal{X}_m$ is time-consuming and computationally challenging. Therefore, the column generation method is used to solve the linear relaxation of MP. In this case, instead of the entire set, a subset of $\mathcal{X}_m$ will be used to construct a restricted version of the master problem. Furthermore, using a subset $\mathcal{K}_m$ of $\mathcal{X}_m$ instead of the entire set, we can obtain a restricted version of the master problem (RMP) as follows.

$$\max \sum_{m\in\mathscr{T}} \phi_m \sum_{k\in\mathcal{K}_m} \Big(\sum_{n\in\mathcal{N}} P_{nm} \cdot d^k_{nm} - \sum_{c\in\mathcal{C}} K_{cm} q^k_{cm} \Big) \lambda^k_m \tag{RMP1}$$

$$s.t. \sum_{n\in\mathcal{N}_c} \Big(\sum_{k\in\mathcal{K}_m} d^k_{nm}\lambda^k_m + r_{cnm} \sum_{m'\in\mathscr{A}(m)} \sum_{k\in\mathcal{K}_{m'}} d^k_{nm'}\lambda^k_{m'} \Big)$$

$$- \sum_{k\in\mathcal{K}_m} q^k_{cm}\lambda^k_m \leq 0, \qquad \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}, \tag{RMP2}$$

$$\sum_{k\in\mathcal{K}_m} x^k_{nm}\lambda^k_m - \sum_{k\in\mathcal{K}_0} x^k_{n0}\lambda^k_0 \leq 0, \qquad \forall\, n \in \mathcal{N}, \forall\, m \in \mathscr{T}, \tag{RMP3}$$

$$\sum_{k\in\mathcal{K}_m} q_{cm} - M \sum_{n\in\mathcal{N}_c} \sum_{k\in\mathcal{K}_0} x^k_{n0} \leq 0, \qquad \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}, \tag{RMP4}$$

$$\sum_{k\in\mathcal{K}_m} \lambda^k_m = 1, \qquad \forall\, m \in \mathscr{T}, \tag{RMP5}$$

$$\lambda^k_m \geq 0, \qquad \forall\, k \in \mathcal{K}_m, \forall\, m \in \mathscr{T}. \tag{RMP6}$$

Let $\pi^{(1)}_{mc}$, $\pi^{(3)}_{mc}$ be the dual variables associated with constraints (RMP2) and (RMP4), $\forall\, m \in \mathscr{T}$ and $\forall\, c \in \mathcal{C}$, and $\pi^{(2)}_{mn}$ be the dual variables associated with constraints (RMP3), $\forall\, m \in \mathscr{T}$ and $\forall\, n \in \mathcal{N}$ respectively. The dual variables associated with constraints (RMP5) for each $m \in \mathscr{T}$ is denoted as μ_m. Afterwards, we can construct the corresponding pricing problem ($\boldsymbol{SP}(m)$) for each node $m \in \mathscr{T} \setminus \{0\}$ in the scenario tree as follows.

$$\max \sum_{n\in\mathcal{N}} \Big(\phi_m P_{nm} - \sum_{c\in\mathcal{C}_n} (\pi^{(1)}_{mc} + \sum_{m'\in\mathscr{S}(m)} r_{cnm'}\pi^{(1)}_{m'c}) \Big) d_{nm}$$

$$+\sum_{c\in\mathcal{C}}\Big(-\phi_m K_{cm}+\pi^{(1)}_{mc}-\pi^{(3)}_{mc}\Big)q_{cm}$$

$$-\sum_{n\in\mathcal{N}}\pi^{(2)}_{mn}x_{nm}-\mu_m \tag{SP1}$$

$$s.t. \gamma_{nm}\cdot W_m-1\le d_{nm}\le\gamma_{nm}\cdot W_m, \qquad \forall n\in\mathcal{N},\forall m\in\mathscr{T}, \tag{SP2}$$

$$e_{0m}\gamma_{nm}+\sum_{j\in\mathcal{N}}e_{jm}\,z_{njm}=x_n\,e_{nm}, \qquad \forall n\in\mathcal{N},\forall m\in\mathscr{T}, \tag{SP3}$$

$$z_{njm}\le x_{jm}, \qquad \forall n\in\mathcal{N},\forall j\in\mathcal{N},\forall m\in\mathscr{T}, \tag{SP4}$$

$$z_{njm}\le\gamma_{nm}, \qquad \forall n\in\mathcal{N},\forall j\in\mathcal{N},\forall m\in\mathscr{T}, \tag{SP5}$$

$$z_{njm}\ge\gamma_{nm}+x_{jm}-1, \qquad \forall n\in\mathcal{N},\forall j\in\mathcal{N},\forall m\in\mathscr{T}, \tag{SP6}$$

$$z_{njm}\ge 0, \qquad \forall n\in\mathcal{N},\forall j\in\mathcal{N},\forall m\in\mathscr{T}, \tag{SP7}$$

$$x_{nm}\in\{0,1\}, \qquad \forall n\in\mathcal{N},\forall m\in\mathscr{T}, \tag{SP8}$$

$$q_{cm},d_{nm}\in\mathbb{Z}^+, \qquad \forall n\in\mathcal{N},\forall c\in\mathcal{C},\forall m\in\mathscr{T}. \tag{SP9}$$

Exceptionally, for the root node of the scenario tree, i.e., node 0, the cardinality constraints (SPR6) should be added to the corresponding pricing problem and the objective should be changed as follows.

$$\max \quad \phi_0\Big[\sum_{n\in\mathcal{N}}\Big(P_{n0}-\sum_{c\in\mathcal{C}_n}(\pi^{(1)}_{0,c}+\sum_{m'\in\mathscr{S}(0)}r_{cnm'}\pi^{(1)}_{m',c})\Big)d_{n0}$$

$$+\sum_{c\in\mathcal{C}}\Big(-K_{c0}+\pi_{0,c}^{(1)}-\pi_{0,c}^{(3)}\Big)q_{c0}\Big]+\sum_{n\in\mathcal{N}}(-\pi_{0,n}^{(2)}+\sum_{m'\in\mathscr{S}(0)}\pi_{m',n}^{(2)})x_{n0}-\mu_0 \quad \text{(SP10)}$$

In the pricing problem $SP(m)$, the constraints (SP8) and (SP9) which indicate the integrity of decision variables x_{nm}, d_{nm}, and q_{cm} are kept so that the integer columns will be generated when solving the $SP(m)$.

Algorithm 1 B&P algorithm for solving multi-stage stochastic programming model

Input: Parameters $P_{nm}, K_{cm}, W_m, r_{cnm}, e_{nm}$, $\forall\, c\in\mathcal{C}$, $\forall\, n\in\mathcal{N}, \forall\, m\in\mathscr{T}$, scenario Tree $\mathscr{T}$ with probability ϕ_m, $\forall\, m\in\mathscr{T}$, $UB=0$, and $LB=-\infty$.
Output: z_{alg}, the objective value of multistage stochastic programming model.
Step 1. Reformulate the studied problem in the form of MP. Initialize the feasible solution to MP and the sub-problem set in the branch-and-bound tree with MP.
Step 2. Choose and remove a sub-problem from the sub-problem set and initialize the column pool with the a feasible solution.
Step 3. Conduct Dantzig-Wolf decomposition to the selected sub-problem to obtain the RMP and solve it. If RMP is feasible, obtain the corresponding dual solution $\{(\pi_{m,c}^{(1)}, \pi_{m,n}^{(2)}, \pi_{m,c}^{(3)}, \mu_m) \mid \forall\, n\in\mathcal{N}, \forall\, c\in\mathcal{C}, \forall\, m\in\mathscr{T}\}$ and update LB when the solutions are integers and the corresponding objective value is higher than LB; Otherwise, go to Step 2.
Step 4. For each tree node $m\in\mathscr{T}$, formulate and solve the corresponding $SP(m)$ to generate columns, whose objective is to maximize the reduced cost of any columns associated with variables for this specific node m. If there is positive reduced cost, add the corresponding generated columns to column pool and go to Step 3; Otherwise, go to Step 5.
Step 5. Let the objective value of RMP be the upper bound on this sub-problem, i.e., $UB=obj(RMP)$. If the UB is higher than the LB, move to Step 6; Otherwise, move to Step 7.
Step 6. Find an x_{nm} variable which is required to be an integer, then do the branching by defining two sub-problems adding the constraints $x_{nm}=0$ or $x_{nm}=1$, respectively. Add the new generated sub-problems to the sub-problem set.
Step 7. If there are no more active nodes in the branch-and-bound tree, then we declare the incumbent solution as the optimal one and stopc. Else, we select an active node (sub-problem) in the branch-and-bound tree and go to Step 2.

The detailed steps for the B&P algorithm is shown in Algorithm 1. The basic procedures of B&P algorithm are as follows. Firstly, the algorithm starts with the column generation process, or the so-called pricing procedure. The RMP is solved to obtain the dual solutions $\{(\pi^{(1)}_{m,c}, \pi^{(2)}_{m,n}, \pi^{(3)}_{m,c}, \mu_m) \mid \forall\, n \in \mathcal{N}, \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}\}$. Afterwards, the dual solution $\{(\pi^{(1)}_{m,c}, \pi^{(2)}_{m,n}, \pi^{(3)}_{m,c}, \mu_m) \mid \forall\, n \in \mathcal{N}, \forall\, c \in \mathcal{C}, \}$ will be used to construct the corresponding $SP(m)$ whose objective is to maximize the reduced cost of any columns associated with variables for this specific node m. The optimal solutions to all $SP(m)$'s are referred to as the generated columns and only the columns with positive reduced costs will be added to the RMP. This procedure will be repeated till there is no columns with positive reduced costs generated by solving $SP(m)$ for all $m \in \mathscr{T}$, i.e., the current solution to RMP is optimal and the corresponding objective value of RMP is the upper bound of this sub-problem in the branch-and-bound tree. Finally, the integrity of the optimal RMP solution will be checked and the lower bound of the original problem will be updated if the solutions are integers. After the column generation approach, the algorithm will continue with branching procedure in which some decision variables are fixed at integer values to create sub-problems in the branch-and-bound tree. Some details of the B&P algorithm will be discussed in the following subsections.

4.4.2 Initialization

The initialization of the B&P algorithm for this problem is trivial. At the beginning of the algorithm, the global lower bound can be set to negative infinity. For each sub-problem in the branch-and-bound tree, we need to formulate the corresponding RMP based on the column pools $\mathcal{K}_m, \forall\, m \in \mathscr{T}$, which need to be initialized with initial columns at the start of the column generation approach. A straightforward way to initialize the column pools $\mathcal{K}_m, \forall\, m \in \mathscr{T}$ is to construct a feasible solution $(\boldsymbol{x}^1, \boldsymbol{q}^1, \boldsymbol{d}^1) = \{(\boldsymbol{x}^1_m, \boldsymbol{q}^1_m, \boldsymbol{d}^1_m) | \forall\, m \in \mathscr{T}\}$ to the sub-problem by finding a product line whose cardinality is less than the required quantity, then including each variants from the product line in the assortment at each specific node, and finally determining the demand for each product and components. To be specific, at the beginning of the column generation procedure for a sub-problem, we can simply assign 1 to all unfixed x_{n0} until the total number of $x_{n0}, \forall n \in \mathcal{N}$ taking value of 1 is larger than S and then also assign 1 to all $x_{nm}, \forall n \in \mathcal{N}, \forall m \in \mathscr{T}_0$, if $x_{n0} = 1$. Note that S is the maximum number of product variants included in the product line. The rest variables $x_{nm}, \forall n \in \mathcal{N}, \forall m \in \mathscr{T}$ which are not assigned with value 1 will be assigned with 0. Finally, based on the initialized product line decision $x_{n0}, \forall n \in \mathcal{N}$ at the root node 0, and assortment decision $x_{nm}, \forall n \in \mathcal{N}$ at each node $m \in \mathscr{T}_0$, we can easily determine the values of d_{nm} and $q_{cm}, \forall c \in \mathcal{C}, \forall n \in \mathcal{N}, m \in \mathscr{T}$ based on constraints (SPR2)-(SPR4). Clearly, this solution $(\boldsymbol{x}^1, \boldsymbol{q}^1, \boldsymbol{d}^1)$ complies with the coupling constraints in the RMP and the branching schemes so that $(\boldsymbol{x}_m^1, \boldsymbol{q}_m^1, \boldsymbol{d}_m^1)$ is an initial column in the column pool $\mathcal{K}_m$ for any node $m \in \mathscr{T}$. If there exists initial column $(\boldsymbol{x}_m^1, \boldsymbol{q}_m^1, \boldsymbol{d}_m^1)$ not feasible to the pricing problem $SP(m)$ for any node $m \in \mathscr{T}$, then this sub-problem is infeasible and we should prune it from the branch-and-bound tree.

4.4.3 Upper bounds on sup-problems

As illustrated in Algorithm 1, each sub-problem in the branch-and-bound tree will be solved by column generation procedure, i.e., the RMP solution is optimal only when there exists no pricing problems with positive optimal objective values. In this case, the optimal objective value of the RMP provides an upper bound to the sub-problems in the branch-and-bound tree (because RMP is constructed to obtain the optimal value of the linear programming relaxation of the MP). However, in practice, some optimal objective values of the pricing problems remain positive because of computation precision in the solver, thus only checking the positiveness of reduced costs of pricing problems is inefficient. In addition, instead of solving the linear programming to optimality, it may be more efficient to prematurely end the column generation process and work with bounds on the final linear programming value of the MP (Barnhart et al., 1998). In this case, a new upper bound needs to be constructed to improve the efficiency of column generation process. In the literature, there are some simple and relatively easy to compute bounds on the final linear programming value based on the objective value of RMP and the reduced costs of pricing problems such as Lasdon (2002), Farley (1990), and Vanderbeck and Wolsey (1996). In our proposed algorithm, the upper bound is constructed as follows.

Theorem 4.1. *(Farley, 1990) Let Z^{MP-LP} represent the optimal objective value of the linear relaxation of MP. In each iteration of the column generation, denote the optimal objective value of RMP as Z^{RMP} and the reduced cost of $SP(m)$ as $\xi_m, \forall\, m \in \mathscr{T}$, then*

$$Z^{RMP} + \sum_{\xi_m > 0} \xi_m \geq Z^{MP-LP}. \tag{4.6}$$

This result is well known so we omit the corresponding proof. The theorem provides a way to construct the upper bond as $Z^{RMP} + \sum_{\xi_m>0} \xi_m$ during the column generation when the reduced costs of some pricing problems are still positive. In addition, the column generation process can be prematurely ended when this upper bound is good enough.

4.4.4 Feasible solutions and lower bounds

In the B&P algorithm, the integrity of all variables in the final solution to the MP linear relaxation of each sub-problem will be checked when the column generation procedure is done. If the final solution happens to be integral and the corresponding objective value of the linear relaxation of MP is higher than the current lower bound, then both lower bound and current best solution will be updated. Nevertheless, feasible solutions can be captured in each iteration of column generation by checking the integrity of x variables in optimal RMP solutions. If all x variables in an optimal RMP solution are integers, we can simply determine the values of γ and z variables based on constraints (SPR4) and (SPR8) - (SPR11) and then determine the values of d and q based on constraints (SPR2) and (SPR3). This is because if both the product line and assortment decisions (x) are given, then product order probability (γ) can be determined through applying the MNL model at each node. Finally, product order quantities (d) and component procurement quantities (q) will be determined by utilizing product order probabilities, market sizes, and BOM information at each node. In this way, a feasible solution to current sub-problem is generated based on the RMP optimal solution and corresponding objective value can be calculated. If the feasible solution is better than the current best solution and the lower bound of the problem, then both will be updated.

4.4.5 Branching rules

After updating current best solution and the lower bound, the branching procedure in the B&P algorithm will start. In the literature, two branching rules are usually used. One rule is that branching on the λ variables in the MP creates two sub-problems along two branches where a variable is set to either 0 or 1, respectively. However, this rule is not efficient because there are too many variables to be branched on and most of them will not take value 0 in the optimal solution. Fortunately, there is a simple remedy to this issue, which leads to the second branching rule. In this rule, we first branch on the original variables x_{nm} and then branch on d_{nm} and q_{cm} variables, $\forall\, n \in \mathcal{N}$, $\forall\, c \in \mathcal{C}$, $\forall\, m \in \mathscr{T}$, instead of branching on the λ variables in the MP.

In our proposed B&P algorithm, this branching procedure can be further simplified. Note that for the studied problem, when x_{nm}'s are fixed, the optimal integer solution to the problem can be easily obtained by calculating the product choice probability of each product variant through the MNL model at each node and further calculating the product variant order quantities and component procurement quantities according to constraints (SPR2) and (SPR3), and also by the maximization nature of the problem. Therefore, it is sufficient to branch only on the unfixed x_{nm} variables (which are binary) to generate sub-problems so that the size of branch-and-bound tree is greatly reduced.

To implement this branching rule in the proposed algorithm with Dantzig-Wolfe decomposition framework, a constraint $\sum_{k\in\mathcal{K}_m} x^k_{nm}\lambda^k_m = 0$ or $\sum_{k\in\mathcal{K}_m} x^k_{nm}\lambda^k_m = 1$ can be simply added to the RMP. However, the RMP model for each sub-problem in the branch-and-bound tree will change, thus making the algorithm implementation more complicated. A quick remedy to this procedure is to move the added constraint to the pricing problem $SP(m)$ as $x_{nm} = 0$ or $x_{nm} = 1, \forall n \in \mathcal{N}, \forall m \in \mathscr{T}$. Meanwhile, the initial columns generated at the beginning of the column generation procedure for this sub-problem need to satisfy the constraints $x_{nm} = 0$ or $x_{nm} = 1, \forall n \in \mathcal{N}, \forall m \in \mathscr{T}$, respectively. In this case, all columns in the column pool for the column generation procedure of this sub-problem are ensured to satisfy the constraints indicated by the branching rules while the RMP is kept unchanged during the branching procedure. One merit of this approach is that the dual solution of the RMP and the objective functions of the pricing problems remain the same and the coding work is simple and concise.

Another implementation detail on the branching procedure in the B&P algorithm is to take advantage of the principle that if one product variant does not appear in the product line, then it will not be selected in any assortments offered in the selling season. To be specific, once x_{nm} is fixed to 0 at the root node (node 0) in the scenario tree $\mathscr{T}$, it is reasonable to fix all $x_{nm'}, \forall m' \in \mathscr{S}(0)$ to 0 (note that $\mathscr{S}(0)$ represents the children nodes set of root node 0 in the scenario tree $\mathscr{T}$). Even though the B&P algorithm functions without this procedure, the procedure makes it convenient when constructing the feasible solutions to the sub-problems and has potentials to shrink the size of branch-and-bound tree by decreasing the number of x variables to be branched.

4.5 Numerical Experiments

In this section, various numerical experiments are performed. Firstly, a set of problem instances are solved to show the value of dynamic assortment decisions. Afterwards, the structure of optimal solution will be explored through solving the second instance sets in Section 4.5.2. Finally, the impacts of component commonality and unit cost on the total expected profit will be demonstrated in the experiments in Section 4.5.3. Finally, 24 problem instances are randomly generated to test the performance of the proposed B&P algorithm compared to the CPLEX solver. All numerical experiments are coded in C++ and carried out through the IBM ILOG CPLEX 20.1 optimization package on a PC with an Intel Core i7-10750H 2.60 GHz CPU and 16 GB RAM.

4.5.1 Value of dynamic assortment decisions

We will show the value of dynamic assortment decisions. As discussed in Section 4.1, the OEM produces and sells the dynamic product assortment through an online platform and is able to collect and take advantage of the historical data to estimate the customer preferences. This feature allows the OEM to adjust the assortment decisions based on the estimated customer preferences in each period of the selling season. In this case, an interesting question is that how much more profits the OEM obtains by adopting dynamic assortments. In the following, we denote the value of dynamic assortment as

$$\tau = 100 \cdot \frac{z_{dyn} - z_{stat}}{z_{stat}}, \tag{4.7}$$

where z_{dyn} denotes the total expected profits obtained by offering the dynamic assortment which allows changing product variants in the assortment at each period over the selling season, while z_{stat} represents the total expected profits obtained by offering the static assortment, i.e., the product line decisions determined at the beginning of the selling season is treated as the assortment decisions offered over all periods in the season.

There are 8 instances generated in this experiment, i.e., instances DA1-8 (DA represents “dynamic assortment”.) and they are generated based on the following rules. Each instance has five product variants and five components. The product prices and component costs in all instances are randomly generated. We implement different scenario tree structures in these instances as follows. The number of periods included in the selling season is set as $T = 2$, 3, 4, or 5 in different instances, leading to the scenario trees with 2, 3, 4, or 5 layers respectively. In each non-leaf layer of the tree, one parent node has two branches (or children nodes in the next layer), resulting to 7, 15, 31, or 63 nodes in the whole tree respectively. The corresponding results are shown in Table 4.3 and the higher profit obtained by either of two strategy is highlighted in boldface.

Table 4.3: The value of dynamic assortment decisions

Inst. No.	Inst. size				z_{stat}	z_{dyn}	τ
	$\lvert\mathcal{N}\rvert$	$\lvert\mathcal{C}\rvert$	T	$\lvert\mathscr{T}\rvert$			
DA1	5	5	2	7	29633.4	**29739.9**	0.36%
DA2	5	5	2	7	29005.4	**29261.0**	0.88%
DA3	5	5	3	15	29723.8	**30336.8**	2.06%
DA4	5	5	3	15	24047.1	**24945.5**	3.83%
DA5	5	5	4	31	21292.6	**21964.6**	3.16%
DA6	5	5	4	31	29678.9	**31233.5**	5.24%
DA7	5	5	5	63	37995.6	**39483.6**	3.91%
DA8	5	5	5	63	21155.8	**23000.8**	8.72%

From Table 4.3, we can see that dynamic assortment generally performs better than the static one, and the value of dynamic assortment is increasing as the number of periods in the selling season rises. Moreover, it reveals the importance of capturing and utilizing the historical data: If the OEM possesses historical data, they will obtain higher expected profits by adjusting the assortment decisions based on the estimated customer preferences in each period of the season. More importantly, this can be easily achieved in an e-commerce environment.

4.5.2 Impact of product unit profit

Numerical experiments are carried out to explore the structure of optimal solution to the problem in this subsection. In the literature, the optimal assortment planning decisions for a single period problem with deterministic customer preferences are usually in the form of revenue-ordered assortment when there is no product cost considered (Talluri and Van Ryzin, 2004). In addition, Rusmevichientong et al. (2014) find the revenue-ordered assortments perform well even when the uncertainty in customer preferences is considered in the single period assortment planning problem. In the revenue-ordered assortment, a product will not be included in the assortment when its unit revenue is lower than those of products in the assortment. This may not be true when the uncertainties in customer preferences are considered under a multi-period context. In our problem, each product has its own cost which is equal to the sum of costs of all components used in it. In this case, the revenue-ordered assortment with no consideration of cost in the literature will be equivalent to the profit-ordered assortment in our studied problem. In this series of experiments, 10 instances are generated to test if the profit-ordered assortment is optimal to our problem and they are titled as PA1-10 (PA represents "product line and assortment"). In each instance, there are 5 products and 5 components. The maximum number of products deployed in the product line is set as $S = 4$. The number of periods is set as $T = 3$ and the corresponding tree structure can be found in Figure B2.1 at Section B1 of Appendix B. Specifically, the number of nodes in the scenario tree for each instance is set as $|\mathscr{T}| = 15$. The corresponding product line and assortment decisions in each instance are shown in Table 4.4. In this table, product indexes are in the order of unit product profits from high to low, i.e., product 1 has the highest unit profit, product 2 has the second highest unit profit and so on. The unit profit for each product variant is fixed through the selling season. In the "Product line" column, we count the number of scenario tree nodes when a particular product variant is included in the dynamic assortment. Correspondingly, the column of "No. of assort. in" counts the number of nodes in which a particular product variant is included in the dynamic assortment. The last column, "No. of assort. out" indicates the number of assortments which do not have a particular product variant. Note that the maximum number of assortments in which a variant included or not included is 14, because there are 14 nodes in the scenario tree (excluding the root node), and each node will have a specific assortment determined based on the corresponding customer preferences.

Table 4.4: Product line and assortment decision structures

Inst. No.	Product line					No. of assort. in					No. of assort. out				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
PA1	1	1	1	0	0	14	14	10	0	0	0	0	4	14	14
PA2	1	1	1	1	0	14	14	14	11	0	0	0	0	3	14
PA3	1	1	1	0	0	14	14	11	0	0	0	0	3	14	14
PA4	1	1	0	1	0	14	13	0	10	0	0	1	14	4	14
PA5	1	1	1	1	0	14	14	12	10	0	0	0	2	4	14
PA6	1	1	1	1	0	14	14	13	11	0	0	0	1	3	14
PA7	1	1	1	0	0	14	10	12	0	0	0	4	2	14	14
PA8	1	1	1	0	0	14	11	12	0	0	0	3	2	14	14
PA9	1	1	0	0	0	14	13	0	0	0	0	1	14	14	14
PA10	1	1	1	1	0	14	14	9	5	0	0	0	5	9	14

From the table, it is clear that the profit-ordered assortments may not always be optimal in both product line decisions at the beginning of the selling season and assortment decisions in each period of the season. For example, the product line decisions in the instance PA4 are not in the form of profit-ordered assortment, i.e., product 3 does not show in the optimal solution while product 4 does. Additionally, in the instances PA7 azd PA8, product 3 appears in more assortments than product 2 does, indicating the optimal assortments for some scenario nodes are not profit-ordered.

Another interesting observation is that in the instances when profit-ordered assortments are optimal, the quantity of product variants included in the optimal product line decisions is not same as the quantity of variants in the assortment at each node. For example, product variant 3 in instance PA1 appears in the product line decision but is not included in the assortment of 4 scenario nodes. This is caused by the flexibility in changing the assortment decisions based on the estimated customer preferences.

The aforementioned observations illustrate the complexity in the structure of the optimal solution to the studied problem and show that the heuristic of profit-ordered assortments may not work well. Even for some instances with optimal solutions in the structure of profit-ordered assortment, it is hard to construct the optimal assortment for each node in the scenario tree because the amounts of variants included in the optimal assortments in various nodes are different. In this case, the experiments results demonstrate that the proposed B&P algorithm is necessary for solving the problem in this chapter.

4.5.3 Impacts of component commonality and component unit cost

In this subsection, a series of numerical experiments will be conducted to explore the impacts of component commonality and component unit cost on the product line and assortment planning decisions and the total expected profits.

Component commonality

Firstly, we will test if the component commonality will impact the product line and assortment decisions by using one common component which is shared by all product variants. There are two settings in this experiments. It is assumed that each product are assembled from two different components with the same unit cost. In the first setting, a system referred to as commonality system (denoted as $\mathscr{C}$) is considered. In this system, each product uses one unit of dedicated component which is uniquely used by one product variant, and one unit of common component which is shared by all product variants. In the second setting, a dedicated system (denoted as $\mathscr{D}$) in which each product variant is built from two dedicated components is considered. In other words, the common component in the $\mathscr{C}$ system is replaced with a dedicated component with the same cost for each product variant, resulting in a $\mathscr{D}$ system. We will compare the decisions and total expected profits in these two systems to find if there are any changes.

Ten problem instances are generated to test the role of component commonality in impacting the product line and assortment decisions. We use CD1-10 to denote those instances (CD represents "common and dedicated"). In each instance, the tree structure is shown in Figure B2.1 at Section B1 of Appendix B. In addition, all product variants have the same selling price. The number of components used in each product variant is fixed and the failure rates of components are same. In this way, the unit profit and failure probability do not vary over different product variants so that their impacts on both product line and assortment decisions will be controlled. Table 4.5 illustrates the comparison between the dedicated system and commonality system of the tested instances. The second and third columns report the corresponding total expected profits of commonality and dedicated systems respectively. The column, "If decisions change?", represents if the product line and assortment decisions in both system are identical or not. For the same instance, the higher expected profits will be highlighted in boldface for the ease of comparison.

From the table, the last column indicates there are no changes in product line and assortment decisions when replacing common components by dedicated components. However, such replacement will result in the dedicated systems with lower total expected profits. The reason lies upon the pooling effect of the usage of common component, i.e., that usage will decrease the component quantity required in the system such that the uncertainties embedded in the dedicated component demands in the $\mathscr{D}$ system will be pooled into the uncertainty of one common component demand in the $\mathscr{C}$. For example, consider a commonality system with five product variants, each of them using one dedicated component and one common component. The total number of components required is six. If the common one is replaced by a dedicated one for each variant, then four more dedicated components will be handled in the $\mathscr{D}$ system. In this case, more

Table 4.5: The comparison between $\mathscr{C}$ and $\mathscr{D}$ systems

Inst. No.	$\mathscr{C}$ system expected profits	$\mathscr{D}$ system expected profits	If decisions change?
CD1	**75925.4**	75908.4	NO
CD2	**92465.6**	92443.8	NO
CD3	**106259.0**	106236.0	NO
CD4	**96733.9**	96718.8	NO
CD5	**93959.1**	93924.7	NO
CD6	**101413.0**	101385.0	NO
CD7	**86625.3**	86600.3	NO
CD8	**98952.5**	98937.5	NO
CD9	**99441.4**	99419.2	NO
CD10	**91046.3**	91032.5	NO

components are exposed to demand uncertainties, leading to more units to be ordered for manufacturing and repair. This results indicate the commonality of components will impact the total expected profits and introducing the common components to the system may potentially increase the expected total profits.

Component unit cost

In this series of numerical experiments, we want to examine if the changes in the costs of common and dedicated components will affect the expected total profits earned by the OEM. Instances CD1-10 in Section 4.5.3 will still be used but the unit costs (c_c) of common components in the $\mathscr{C}$ system and those (c_d) of corresponding dedicated components in the $\mathscr{D}$ system will be changed. Let Δc_c and Δc_d denote the percentage changes in common component unit cost in $\mathscr{C}$ system and corresponding component unit cost in $\mathscr{D}$ system respectively. For each instance, Δc_c and Δc_d are selected from $\{-40\%, -20\%, 20\%, 40\%\}$. The results are summarized in Table 4.6.

Table 4.6: Impacts of price changes in common and dedicated components

Inst.	Percentage change in $\mathscr{C}$ system expected profits				Percentage change in $\mathscr{D}$ system expected profits			
No.	$\Delta c_c = -40\%$	$\Delta c_c = -20\%$	$\Delta c_c = 20\%$	$\Delta c_c = 40\%$	$\Delta c_d = -40\%$	$\Delta c_d = -20\%$	$\Delta c_d = 20\%$	$\Delta c_d = 40\%$
CD1	5.519%	2.759%	-2.759%	-5.519%	5.529%	2.764%	-2.764%	-5.529%
CD2	5.533%	2.767%	-2.767%	-5.533%	5.544%	2.772%	-2.772%	-5.544%
CD3	5.446%	2.723%	-2.724%	-5.446%	5.456%	2.728%	-2.729%	-5.457%
CD4	5.548%	2.774%	-2.774%	-5.548%	5.554%	2.777%	-2.777%	-5.555%
CD5	5.466%	2.733%	-2.733%	-5.466%	5.483%	2.741%	-2.742%	-5.483%
CD6	5.537%	2.769%	-2.769%	-5.538%	5.549%	3.775%	-2.775%	5.550%
CD7	5.449%	2.725%	-2.725%	-5.449%	5.463%	2.731%	-2.731%	-5.463%
CD8	5.435%	2.718%	-2.718%	-5.435%	5.442%	2.721%	-2.721%	-5.442%
CD9	5.453%	2.727%	-2.753%	-5.453%	5.464%	2.732%	-2.732%	-5.463%
CD10	5.465%	2.733%	-2.733%	-5.465%	5.472%	2.736%	-2.736%	-5.472%
AVG.	5.485%	2.743%	-2.745%	-5.485%	5.496%	2.748%	-2.748%	-5.496%

From the table, it is clear that there is no significant differences between the impacts of the common component cost on the total expected profits in $\mathscr{C}$ system and those of the dedicated component cost on the total expected profits in $\mathscr{D}$ system. The average changes in both systems are very closed to each other, 2.745% for commonality system and 2.748% for dedicated system. In this case, one managerial insight may be it is better to lower the costs of both common and dedicated components and there is no significant priorities for lowering the cost of a particular component type.

4.5.4 Branch-and-price algorithm performance

In this subsection, a total of 24 instances are randomly generated to compare the efficiencies and performances of the proposed B&P algorithm and the CPLEX MIP solver. The instances are titled as BP1-24 (BP represents "branch-and-price") respectively. When generating the instances with random data, we limit the total cost of all components used in a product to be lower than the unit price of that product to guarantee a positive unit profit. The number of variables, constraints, and nonzero elements for all tested instances are shown in Table B2.1 enclosed in Section B2 of Appendix B. As shown in the table, each instance has $|\mathcal{N}| \cdot |\mathscr{T}|$ binary variables x_{nm}. In addition, there are $|\mathcal{N}| \cdot |\mathscr{T}|$ integer variables d_{nm} and $|\mathcal{C}| \cdot |\mathscr{T}|$ integer variables q_{cm} so that the total amount of integer variables in in each instance is $(|\mathcal{N}| + |\mathcal{C}|) \cdot |\mathscr{T}|$. The scenario tree structures used in the instances are shown in Section B1 of Appendix B.

For each instance, the time limits for running both B&P algorithm and CPLEX MIP solver are set as 7200 seconds and the optimality gaps are set as 0.5 %. The performances of both solution methods are compared based on the solution quality and time as shown in Table 4.7. Within the same instance, if one method performs better than the other, the corresponding results will be highlighted. For example, in Instance 1, both methods can obtain the same best objective value, i.e., $z_{MIP} = z_{BP} = 30336.8$, so both z_{MIP} and z_{BP} are highlighted in boldface. However, the B&P algorithm can solve the instance in 11 seconds while the solution time of CPLEX solver is 361.77 seconds. To highlight this difference, we put the solution time of B&P algorithm in boldface to indicate it performs better than CPLEX solver in computation time.

From Table 4.7, it can be concluded that the B&P algorithm performs better than the CPLEX solver in general. From the perspective of solution quality, the B&P algorithm obtains better solutions in 16 out of 24 instances. In the other 8 instances, the solutions obtained by the B&P algorithm are same as those obtained by the CPLEX solver. From the perspective of computation time, the B&P algorithm dominates the CPLEX solver: The CPLEX solver can only solve 7 instances in two hours while the B&P algorithm can solve all instances in minutes and for these 7 instances, the computation time of the B&P algorithm is shorter than that of the CPLEX.

4.6 Conclusions and Future Research Directions

In this chapter, we consider a dynamic assortment planning problem for an OEM who launches and sells the product variants through an online platform under uncertain customer preferences over a multi-period selling season. The OEM first determines the

Table 4.7: The efficiency of the B&P algorithm

Inst.	Inst. size			CPLEX			B&P Algorithm		
No.	$\|\mathcal{N}\|$	$\|\mathcal{C}\|$	$\|\mathscr{T}\|$	z_{MIP}	Time (s)	Gap	z_{BP}	Time (s)	Gap
BP1	5	5	15	**30336.80**	361.77	$< 0.5\%$	**30336.80**	**10.21**	$< 0.5\%$
BP2	5	5	15	59931.10	46.44	$< 0.5\%$	**59933.40**	**7.60**	$< 0.5\%$
BP3	5	10	15	**46712.60**	40.27	$< 0.5\%$	**46712.60**	**14.02**	$< 0.5\%$
BP4	5	10	15	**50947.30**	181.53	$< 0.5\%$	**50947.30**	**22.14**	$< 0.5\%$
BP5	5	15	15	**30983.20**	> 7200	2.33%	**36983.20**	**27.24**	$< 0.5\%$
BP6	5	15	15	**59390.10**	165.41	$< 0.5\%$	**59390.10**	**28.20**	$< 0.5\%$
BP7	10	5	15	**49379.70**	> 7200	12.97%	**49379.70**	**61.58**	$< 0.5\%$
BP8	10	5	15	**42259.80**	> 7200	16.41%	**42259.80**	**29.72**	$< 0.5\%$
BP 9	10	10	15	58382.50	> 7200	2.28%	**59756.00**	**98.22**	$< 0.5\%$
BP10	10	10	15	63386.00	> 7200	3.03%	**63793.00**	**100.88**	$< 0.5\%$
BP11	10	15	15	66950.70	> 7200	3.87%	**67502.60**	**170.77**	$< 0.5\%$
BP12	10	15	15	79031.10	> 7200	4.20%	**79257.90**	**60.55**	$< 0.5\%$
BP13	5	5	40	**21889.50**	356.29	$< 0.5\%$	**21889.50**	**58.50**	$< 0.5\%$
BP14	5	5	40	**28761.10**	314.26	$< 0.5\%$	**28761.10**	**82.30**	$< 0.5\%$
BP15	5	10	40	54971.50	> 7200	7.75%	**55043.30**	**62.26**	$< 0.5\%$
BP16	5	10	40	**35149.10**	> 7200	2.14%	**35149.10**	**34.68**	$< 0.5\%$
BP17	5	15	40	46169.20	> 7200	3.14%	**46173.80**	**137.27**	$< 0.5\%$
BP18	5	15	40	**36871.40**	> 7200	5.48%	**36871.40**	**92.73**	$< 0.5\%$
BP19	10	5	40	**61868.90**	> 7200	25.59%	**61868.90**	**146.05**	$< 0.5\%$
BP20	10	5	40	**41104.70**	> 7200	18.68%	**41104.70**	**86.47**	$< 0.5\%$
BP21	10	10	40	**40693.90**	> 7200	22.72%	**40693.90**	**244.33**	$< 0.5\%$
BP22	10	10	40	69192.65	> 7200	27.70%	**69266.10**	**188.37**	$< 0.5\%$
BP23	10	15	40	**52761.60**	> 7200	26.71%	**52761.60**	**217.59**	$< 0.5\%$
BP24	10	15	40	**66507.00**	> 7200	27.21%	**66507.00**	**333.21**	$< 0.5\%$

variants in the product line at the beginning of the selling season. During each period of the season, the OEM can adjust the assortment produced based on the predetermined product line decisions and the estimated customer preferences. Such adjustment can be easily exercised by the online platform through changing a product's "in-stock" or "out-of-stock" status displayed to customers in each period. In addition, we consider the stocking decisions of components used for both manufacturing the new products and repairing the sold products under warranty. This problem models the situation faced by many OEMs who sell product assortments through the online platforms and are able to utilize the historical data to estimate the customer preferences over the selling season. To our best knowledge, this multi-period dynamic assortment planning problem with a blended setup of uncertain customer preferences and multiple component was unexplored in the literature.

To handle the uncertainty embedded in the estimation of customer preferences over the selling season, a multi-stage stochastic programming model is proposed for this problem. The proposed model is becoming hard to solve when its size increases. Therefore, we propose a B&P algorithm based on the block-angular structure of the stochastic programming model. In the numerical experiments, we first study the advantage brought by the dynamic assortment compared to the static assortment. Afterwards, the structure of the optimal solution of this problem is investigated. Through some problem instances, we find the well-known revenue-ordered assortment which is optimal to the static assortment planning problem cannot guarantee the optimality to the studied problem. This observation also highlights the necessity of using B&P algorithm as a solution method. In addition, the impacts of component commonality and unit cost on the decisions and total expected profits are also explored. Finally, through the extensive numerical experiments, the performance of the B&P algorithm is confirmed through comparing with CPLEX solver. The B&P algorithm can provide a high quality solution and the corresponding computation time is reasonable.

The future research on this topic can be extended to incorporating more conditions or constraints in the real-life scenario, such as spare parts or components inventory control, product recycling and remanufacturing, and so on. As for handling uncertainty in customer preferences, more techniques such as robust optimization can be applied when less information on preferences is available to the customer preferences.

Chapter 5

Integration of assortment planning and spare parts procurement and remanufacturing under warranty service: A product lifecycle perspective

5.1 Introduction

In this chapter, we study an assortment planning problem from the perspective of the product lifecycle. In this problem, the product lifecycle consists of two phases: the selling season and the end-of-life (EOL) phase. In the selling season, the original equipment manufacturer (OEM) selects the products included in the assortment based on the customer preferences and produce them in an assemble-to-order system. After the selling season, the production stops and the supply of certain parts is permanently discontinued by suppliers. This is often caused by the fact that those parts cannot be utilized in the new generation of products. This is popular in the industries of electronic equipment such as PCs and cell phones. In these products, some key components, like the CPU, GPU, etc., are developed in a fast manner, which is normally one year. However, the demand for those components still exists during the product EOL phase because the OEM has to provide warranty services to customers when the components fail in the products sold.

To ensure sufficient component inventory meeting the repair demand during the product EOL phase, the two typical countermeasures adopted by the OEM is resorting to alternative suppliers who are able to provide extra production service of components or the last-time buy (LTB) operation which requires the OEM place a final order of components to the suppliers at the beginning of the EOL phase to guarantee the repair demand in the rest periods of that phase. In practice, these two approaches are not economical-friendly to a lot of stakeholders, e.g., the OEM, the society, etc. On one hand, the extra production of components during product EOL phase for a supplier can take more than doubled costs than regular production during the selling season due to loss of scale economies (Inderfurth and Kleber, 2013). One the other hand, the LTB strategy incurs extremely high inventory levels being held over the product EOL phase, generating a high level of holding costs and putting OEM under the jeopardy of high component obsolescence risks. In recent years, new solutions such as remanufacturing components from used products (hereafter referred to as the installed base) which are returned by customers are implemented in practice as feasible and relatively low-cost alternatives. In the remanufacutring approach, the returned products are usually collected by OEMs through various trade-in or buyback programs when the customers switch to next-generation products. The OEM disassembles the returned products to obtain the recoverable components which will be turned into ready-for-use components through remanufacturing process as another source of component inventory during the product EOL phase (Shi, 2019). In this case, one important decision faced by the OEM is to determine the quantity of the returned products to be disassembled and remanufactured. Indeed, if too many returns are disassembled, the inventory holding costs of the remanufactured components will be high, even though the repair demand is satisfied. On the contrary, if the amount of returns to be disassembled is low, the component inventory may not be sufficient to cover the repair demand and finally results in the penalty on backordered repairs.

However, when making the remanufacturing decisions, the OEM is usually faced up with the uncertainties embedded in both the return rates of products and the failures rate of components. The former type of uncertainty is brought by the unknown attitudes of customers towards the trade-in or buyback programs. The latter type is due to the insufficient historical data on the failures of the components, especially for the newer implemented ones. Under these two types of uncertainties, the remanufacturing decisions are more difficult to be made by the OEM.

The study in this chapter is closely related to two strands of literature. The first strand is the literature on the retail assortment planning and the second is the literature on the spare parts inventory management during product EOL phase. The retail assortment planning literature focus on the assortment planning decisions based on the customer preferences to maximize the gross sales or profits subjected to various constraints (Ryzin and Mahajan, 1999; Ryzin, 2001). Among the studies in this field, the customer preferences are captured by various choice models, such as multinomial logit (MNL) model (Talluri and Van Ryzin, 2004; Wang, 2012b; Rusmevichientong et al., 2014), locational choice model (Honhon et al., 2010), etc. A recent literature review on this filed can be found at Kök et al. (2015). From the perspective of retailers, considering only the impact of the customer preferences on the assortment decisions is reasonable because they are not involving any after-sales services over the product lifecycle. However, this may not be true from the perspective of OEMs, because the assortment decisions made by them determines the products produced and offered to the market and the OEM are responsible to provide the after-sales services of the products sold. In this case, other factors such as the costs of spare parts inventory and supply, and products repairing costs should be considered when making the product assortment decisions to maximize the total profits over the product lifecycle.

The study in the chapter also related to the other strand of the literature on the spare parts inventory management during the product EOL phase. The literature from this strand mainly focus on securing the supplies for the spare parts used for the after-sales services through various strategies particularly the warranty services during the product EOL phase. This problem is valuable because the suppliers will not continue to provide the spare parts to the OEM due to no regular demand for the products during their EOL phase (Shi, 2019). However, during this phase, there is repair demand for the spare parts from providing warranty services to the sold products by replacing the failed components. To fulfill such demand, there are many strategies for resupplying the spare parts discussed in the literature, such as LTB (Behfard et al., 2018; Hur et al., 2018), remanufacutring recoverable parts collected by disassembling the returned products (Inderfurth and Mukherjee, 2008; Behfard et al., 2015; Pourakbar et al., 2014) and so on. A comment to this literature strand is that the product assortments handled is usually given in advance and no assortment decisions are considered in the studied problem.

In this chapter, we integrate the warranty service operations into the strategic assortment planning for the OEMs. The contributions of this study are as follows. Firstly, different from the existing literature on the assortment planning, we consider the lifecycle costs of the products when making strategic product assortment planning decisions. To be specific, the expected costs related to the warranty services for assortment products during the EOL phase are included in the decision-making. Secondly, both components LTB and remanufacturing strategies are adopted as the supply sources of the components inventory during the EOL phase. To the best of our knowledge, this setting is novel in both the literature of assortment planning and those of spare parts inventory management for the products in the EOL phase. Finally, through the numerical experiments, we explore the advantages of joint optimization on the assortment planning decisions and the spare parts procurement and remanufacturing decisions compared to the separate optimizations on those decisions. Afterwards, we discuss the impacts of the uncertainty levels of those two uncertainties on the the expected total profits obtained by the OEM.

5.2 Problem Formulation

In this section, the formulations of the studied problem will be presented. A deterministic model will be formulated at first. Afterwards, we will formulate a multistage stochastic programming model to address the inherent uncertainties in both the return rates of installed bases and the failure rates of components.

5.2.1 A deterministic mathematical programming model

In the studied problem, the product lifecycle have T periods and can be split into two phases. Each period can be a season or a year.

The selling season

The lifecycle starts with a selling season, which is the the first phase. In this chapter, it is assumed that the selling season only contains one period. At the beginning of the selling season, the OEM is able to produce a set of product variants $\mathcal{N} = \{1, \ldots, N\}$ and has to select a subset of at most S product variants to form an assortment $\mathcal{A}$ to be produced during the season, based on the estimated customer preferences. The selling price of each product variant $n \in \mathcal{N}$ is denoted as P_n.

To model the demand for the product variants in the assortment $\mathcal{A}$, we consider a consumer choice model. To be specific, customers' choice to a product variant within the offered assortment is based on the multinomial logit (MNL) model. In this model, a customer obtains utility U_{it} if he or she purchases product variant $i \in \mathcal{A}$, given the offered assortment $\mathcal{A}$ and the utility is determined by $U_i = V_i + \epsilon_i$, in which V_i represents the mean utility of product variant i to the customer and ϵ_i is a random term which represents the unobserved utility. For the ease of notation, we will denote the vector of mean utilities of the customers towards all the product variants as $\boldsymbol{V} = (V_1, \ldots, V_N)$. If the random term ϵ_i is assumed to be independent and identically distributed (i.i.d.) with Gumbel distribution, the MNL model gives the utility maximization choice probability (Train, 2009) that customer selects product variant i from assortment $\mathcal{A}$ as

$$\gamma_i^{\mathcal{A}} = \frac{e^{V_i}}{\sum_{j\in\mathcal{A}} e^{V_j} + e^{V_0}}, \forall\, i \in \mathcal{A}, \tag{5.1}$$

where V_0 represents the customers' utility of no purchase option.

In the proposed model, a decision variable γ_i is used to denote the estimated probability that a consumer will select product variant $n \in \mathcal{N}$. Let $e_n = e^{V_n}, \forall\, n \in \mathcal{N}$ represent the consumers' utility for purchasing product n, we have

$$\gamma_n = \frac{x_n \cdot e_n}{\sum_{j\in\mathcal{N}} x_j \cdot e_j + e_0}, \forall\, n \in \mathcal{N} \tag{5.2}$$

where x_n is binary variable and $x_n = 1$ if and only if the product variant n is included in the assortment $\mathcal{A}$ offered to customers.

Each product variant $n \in \mathcal{N}$ is a multi-indenture system in which the components are non-identical and has a bill-of-material (BOM) which indicates the components used in that product. The set of all the components used for the assembly of all products variants is denoted as $\mathcal{C} = \{1, \ldots, C\}$. In addition, from the BOMs of product variants, we can also obtain the information on all the variants using component c and will denote the set of those variants as $\mathcal{N}_c$.

The OEM first decides the assortment offered to the market based on the estimated customer preferences at the beginning of the selling season. The expected revenue obtained from selling products offered in the assortment $\mathcal{A}$ during the selling season $(t = 1)$

is calculated as

$$Revenue = \sum_{n \in \mathcal{N}} P_n \cdot w_{n1} \tag{5.3}$$

where w_{n1} represents the fulfilled order amount (installed base) during the selling season and satisfies constraints $\gamma_{nt} \cdot M - 1 \leq w_{n1} \leq \gamma_{nt} \cdot M$. In this constraint, γ_{nt} is defined in equation (5.2) and M is the total market size faced by the assortment $\mathcal{A}$. It guarantees the ordered quantity for product variant n should be the maximum integer lower than estimated demand.

During the selling season, asides from the revenue obtained from selling the product variants in the assortment $\mathcal{A}$, the OEM has to pay for the costs of procuring and holding the components used for both fabricating those product variants and repairing the faulty sold products under warranty. In the model, we assume that at most one unit of each component type is used in one unit of each product variant. In addition, the unit purchasing cost of component c is denoted as K_c. At the beginning of the selling season, the OEM has to order components from the suppliers. The order quantity for component c is denoted as q_{c1}. For a component c, the total required amount d_{c1} includes two demand streams for components. One stream is the total assembly demand d^a_{c1} for the offered product variants whose fabrication requires component c. The other is the total repairing demand d^a_{c1} for fixing faulty products having component c failures under warranty, i.e., $d_{c1} = d^a_{c1} + d^r_{c1}$. To be specific, for the first stream, the assembly demand d^a_{c1} for component c at the selling season is computed as

$$d^a_{c1} = \sum_{n \in \mathcal{N}_c} w_{n1}. \tag{5.4}$$

For the second stream, when the failure of product is caused by the failures of components used in it. When a sold product fails during warranty periods, it will be sent to the OEM for replacing faulty components. In this chapter, it is assumed each product failure only has one faulty component. Let ρ_{cnt} represent the average joint failure probability that a faulty product variant n having component c failure at period t. In this case, the failure amount of product n caused by the failure of component c at period t is equal to the product of ρ_{cnt}, the average joint failure rate, and w_{nt}, the installed base of that product at period t. Therefore, for the selling season, i.e., when $t = 1$, the total repair demand of component c is calculated as

$$d_{c1}^{r} = \sum_{n \in \mathcal{N}_c} \rho_{cn1} \cdot w_{n1}. \tag{5.5}$$

The average failure rate of each component can be easily estimated based on the historical data.

Product EOL phase

The selling season is followed by the product EOL phase which contains $(T-1)$ periods (i.e., $t = \{2, \ldots, T\}$). During this phase, the production stops and there is no market demand for the product variants in the assortment. In this case, there will be no assembly demand and only repair demand, i.e., $d_{ct} = d_{ct}^{r}$, $\forall\, t \in \mathcal{T} \setminus \{1\}$. Moreover, the supply of components is permanently discontinued in this phase. The OEM has to implement two countermeasures to support the repair operations for faulty products under warranty during the phase. The first one is final order policy, i.e., the OEM places the final order for the components at the beginning of this phase (when $t = 2$) to replenish the spare parts inventory. Let us denote the final order quantity for component c as q_{c2}. In addition, it is assumed that the OEM starts the buyback or trade-in programs at the beginning of the product EOL phase, i.e., $t = 2$, to collect the returned products and finally remanufactures those products to the ready-for-use spare parts to replenish the inventory. Assume that, at period $t \in \mathcal{T} \setminus \{1\}$, there is r_{nt} fraction of w_{nt}, the product n installed base, is collected by the OEM as the returned products. All the collected returns can be either disassembled into components for remanufacturing or disposed, because during the product EOL phase, there will be no future demand for the end product such that there is no need for refurbishing the returned products for resale. The components disassembled from the returned products can be remanufactured with unit cost C_r and the OEM has to determine the quantity of returned products to be remanufactured. However, such return collections and the remanufacturing process may not be completed immediately. It is more realistic that there is considerable time lag between when the returned products are collected and when the spare parts are available to the OEM after the remanufacturing (Inderfurth and Kleber, 2013). In this study, we restrict that the time lag to be one planning period, i.e., it takes one period to recollect the returns and remanufacture the spare parts. As a result, the remanufacturing quantity of the returned products is denoted as $f_{nt} \in [0, r_{nt}\, w_{nt}]$, $\forall\, t \in \mathcal{T} \setminus \{1\}$, and the corresponding spare parts obtained through the remanufacturing at period $t+1$ is equal to $\sum_{n\in\mathcal{N}_c} f_{nt}, \forall\, t \in \mathcal{T} \setminus \{1\}$. It is also assumed that the returned products not being disassembled for remanufacturing will be disposed and not be carried to next period. In this case, the total repair demand for component c are generated by the remaining $(1 - r_{nt})w_{nt}$ units of the installed base at period t and can be modelled as follows.

$$d^r_{ct} = \sum_{n\in\mathcal{N}_c} \rho_{cn1}(1 - r_{nt})w_{nt}, \;\forall\, t \in \{2, \ldots, T\}. \tag{5.6}$$

It should be noted that the product installed bases are changing over the periods. Specifically, a portion of the installed base will not generate warranty claims and component repair demand in the future because they are not in use or their warranty coverage is terminated. In this case, at period t, we introduce the "remaining ratio" to represent the percentage of the installed base which will exist in the future and denote it as $\delta_{nt} \in (0, 1)$. Therefore, the relationship between the installed bases of product n in two consecutive periods t and $t+1$ is captured as $w_{n(t+1)} = \delta_{nt}(1-r_{nt})\, w_{nt}$, $\forall\, t \in \{2, \ldots, T\}$. To ensure the installed base in each period to be a positive integer, such a relationship will be represented by $\delta_{nt}(1 - r_{nt}) - 1 \leq w_{n(t+1)} \leq \delta_{nt}(1 - r_{nt})\, w_{nt}$, $\forall\, t \in \{2, \ldots, T\}$. In addition, the product returns appear in the period 2 when the OEM starts the buyback or trade-in programs, and the installed bases in period 1 do not have the percentage loss brought by the product returns. Therefore, the relationship between the installed bases is represented as $\delta_{n1}w_{n1} - 1 \leq w_{n2} \leq \delta_{n1}w_{n1}$. For the ease of exposition, such a relationship will also be represented by equation (5.6) with r_{n1} set as 0 when $t = 1$.

During the product EOL phase, the repair demand of component c is satisfied by the component inventory hold by the OEM. Let l_t, $\forall\, t \in \{1, \ldots T\}$ denote the number of component c avaialable in stock in period t. Both overage and underage costs for component c are incurred with unit holding and backlogging costs of H_c and B_c respectively. Note the assumption that each faulty product only has one component failure. Hence, the backlogging cost of a certain component can be estimated based on the cost related to the unsatisfied warranty claims of the products using that component.

A deterministic model

In this part, the deterministic model of the studied problem will be formulated. The deterministic model incorporates the product assortment planning decisions and component procurement decisions during the product selling season, the final component order decisions and remanufacturing decisions during the product EOF phase, and the component inventory related decisions during the whole planning horizon. With loss of generality, it is assumed that the product selling season consists one planning period while the EOL phase consists $T - 1$ periods. This simplification can be further relaxed to a multi-period selling season with minor changes in the proposed model.

All the notation used in the deterministic model (DM) is presented in Table 5.1 and the corresponding model is formulated as follows.

$$\begin{aligned}
\max \quad & \sum_{n \in \mathcal{N}} P_n \cdot w_{n1} - \Big(\sum_{t=1}^{2} \sum_{c \in \mathcal{C}} K_c q_{ct} \\
& + \sum_{t \in \mathcal{T}} \sum_{c \in \mathcal{C}} (H_c l_{ct} + B_c b_{ct}) + \sum_{t \in \mathcal{T} \setminus \{1\}} \sum_{n \in \mathcal{N}} C_r f_{nt} \Big) \qquad \text{(DM1)}
\end{aligned}$$

Table 5.1: Notation used in the deterministic model

Sets and subscripts	
$\mathcal{N}$	Set of products can be selected in the assortment
n	Subscript of product $n \in \mathcal{N}$
$\mathcal{C}$	Set of components used for product fabrication
c	Subscript of component $c \in \mathcal{C}$
$\mathcal{N}_c$	Set of products using component c, $\forall\, c \in \mathcal{C}$
$\mathcal{C}_n$	Set of components used for fabrication of product n, $\forall\, n \in \mathcal{N}$
$\mathcal{T}$	Set of planning periods
t	Subscript of planning period $t \in \mathcal{T}$
$\mathcal{A}$	Set of product variants included in the assortment, $\mathcal{A} \subseteq \mathcal{N}$
Parameters	
M	Total market size for product assortment $\mathcal{A}$
P_n	Unit selling price of product n during the selling season
C_r	Unit remanufacturing cost of returned product during the EOL phase terminal period T
r_{nt}	Return rates of product n at period t, $\forall\, n \in \mathcal{N}$, $t \in \mathcal{T} \setminus \{1\}$
δ_{nt}	Percentage of the installed base of product n will exist in the future at period t
K_c	Unit purchasing cost of component c at period $t \in \{1, 2\}$
H_c	Unit inventory holding cost per period for spare part c
B_c	Backlogging cost of component c
ρ_{cnt}	Average joint failure probability that a faulty product variant n having component c failure at period t
e_n	$e_n = e^{V_n}$ is the consumers' utility for product n during the selling season
S	Limit on the number of product variants included in the assortment $\mathcal{A}$
T	Number of planning periods
Decision variables	
x_n	Binary variable, $x_n = 1$ if and only if the product variant n is included in the assortment $\mathcal{A}$ offered to customers over the selling season
γ_n	Probability of consumers placing an order on product n in the selling season
d_n	Fulfilled order amount for product n in the selling season
w_{nt}	Installed base of product n at period t
f_{nt}	Remanufacturing quantity of returned product n at period t
d_{ct}	Demand for component c in period t
q_{ct}	The order quantity of component c in period $t \in \{1, 2\}$
l_{ct}	The number of component c available in stock in period t
b_{ct}	The shortage amount of component c in period t

$$
\begin{aligned}
\text{s.t.} \quad & \gamma_n\Big(\sum_{j\in\mathcal{N}} x_j e_j + e_0\Big) = x_n e_n, && \forall\, n\in\mathcal{N}, && \text{(DM2)}\\
& \sum_{n\in\mathcal{N}} x_n \le S, && && \text{(DM3)}\\
& \gamma_n\cdot M - 1 \le w_{n1} \le \gamma_n\cdot M, && \forall\, n\in\mathcal{N}, && \text{(DM4)}\\
& d_{c1} \ge \sum_{n\in\mathcal{N}_c} (1+\rho_{cn1})\cdot w_{n1}, && \forall\, c\in\mathcal{C}, && \text{(DM5)}\\
\textbf{DM} \quad & d_{ct} \ge \sum_{n\in\mathcal{N}_c} \rho_{cnt}\cdot(1-r_{nt})\cdot w_{nt}, && \forall\, c\in\mathcal{C},\, \forall\, t\in\mathcal{T}\setminus\{1\}, && \text{(DM6)}\\
& \delta_{n(t-1)}(1-r_{n(t-1)})w_{n(t-1)} - 1 \le w_{nt}, && && \\
& w_{nt} \le \delta_{n(t-1)}(1-r_{n(t-1)})w_{n(t-1)}, && \forall\, n\in\mathcal{N},\, \forall\, t\in\mathcal{T}\setminus\{1\}, && \text{(DM7)}\\
& f_{nt} \le r_{nt}\cdot w_{nt}, && \forall\, n\in\mathcal{N}, \forall\, t\in\mathcal{T}\setminus\{1\}, && \text{(DM8)}\\
& q_{c1} + l_{c0} = d_{c1} + l_{c1}, && \forall\, c\in\mathcal{C}, && \text{(DM9)}\\
& q_{c2} + l_{c1} = d_{c2} + l_{c2} - b_{c2}, && \forall\, c\in\mathcal{C}, && \text{(DM10)}\\
& \sum_{n\in\mathcal{N}_c} f_{n(t-1)} + l_{c(t-1)} = d_{ct} + b_{c(t-1)} - b_{ct} + l_{ct}, && && \\
& && \forall\, c\in\mathcal{C},\, \forall\, t\in\mathcal{T}\setminus\{1,2\}, && \text{(DM11)}\\
& x_n \in \{0,1\}, && \forall\, n\in\mathcal{N}, && \text{(DM12)}\\
& q_{ct} \in \mathbb{Z}^+, && \forall\, c\in\mathcal{C},\, \forall\, t\in\{1,2\}, && \text{(DM13)}\\
& w_{nt}, f_{nt} \in \mathbb{Z}^+, \gamma_n \in \mathbb{R}^+, && \forall\, n\in\mathcal{N},\, \forall\, t\in\mathcal{T}, && \text{(DM14)}\\
& d_{ct}, l_{ct}, b_{ct}, && \forall\, c\in\mathcal{C},\, \forall\, t\in\mathcal{T}. && \text{(DM15)}
\end{aligned}
$$

In the model, the objective (DM1) represents the total lifecycle profits of offering product assortment $\mathcal{A}$ to the market. It includes the revenue brought by the product variants in the assortment during the selling season, the procurement cost of components prior to the EOL phase, the product remanufacturing cost during the EOL phase, and the inventory holding and backlogging costs of the components over the whole planning horizon. In constraints (DM2), the choice probability of customers to each product is determined based on the MNL model and is ensured to be a non-negative real number in constraints (DM14). Constraints (DM3) restrict the maximum number of the product variants in the assortment. Constraints (DM4) with (DM14) together ensure the installed base of each product variant should be the maximum integer lower than or equal to the estimated demand for that variant given by the MNL model in the selling season. The demands for components during the selling season and the EOL phase are given in constraints (DM5) and (DM6) respectively. Specifically, the component demand involves both the needs for assembly and repairing faulty products under warranty in the selling season as shown in constraints (DM5). After the selling season, there is no assembly, so the component demand during the EOL phase only involves the needs for repairing as shown in constraints (DM6). Constraints (DM7) dictate the changing patterns of the installed base for each product variant during the planning horizon. Constraints (DM8) and (DM14) restrict the quantity of product returns used for remanufacturing to be a non-negative integer not higher than the total collected returns in a certain period during the EOL phase. Constraints (DM9)-(DM11) are the component inventory balance constraints. Constraints (DM12) guarantee the assortment decision for all product variants be binary variables. The decision variables regarding to the product variant installed base, the quantity of the remanufactured product returns, and the order quantity, the demand level, the inventory level, and the shortage amount of components are defined as non-negative integers in constraints (DM14) and (DM15).

5.2.2 A multistage stochastic programming model

In the studied problem, when strategically planning an assortment of the product variants for the new generation of products, the OEM needs to estimate ρ_{cnt}, the average joint failure probability that a faulty product variant n having the component c failure, and r_{nt}, the return rates of product n's installed base at each period t of the planning horizon. Such estimations normally rely on the historical data collected from the old product generations. In the literature, the mainstream methods used for such estimations include regression methods, machine learning methods, and so on (Limbourg, 2008). However, such estimations for the new product generation are usually coupled with errors, especially when new types of components are adopted to upgrade the products to a new generation, and the customers' willingness to participate in the buyback or trade-in programs is altering.

Uncertainties in component failures and product returns

In this chapter, we will study a problem where a new type of component c' is used in the new product generation. To consider the aforementioned uncertainties embedded in the estimations, we treat $\rho_{c'nt}$ and r_{nt}, $n \in \mathcal{N}$, $t \in \mathcal{T}$, as random variables. In other words, the average joint failure probability regarding component c' which is a new implemented component is estimated with uncertainty. On contrary, those of component $c \in \mathcal{C} \setminus \{c'\}$ are estimated with no uncertainty because their failure probabilities are stable when they are already used in the older product generations. Let $\mathbf{P}_t$ be a random vector which has $G_t^{\mathbf{P}}$ different realizations, $\hat{\mathbf{P}}_t^1, \ldots, \hat{\mathbf{P}}_t^{g_1}, \ldots, \hat{\mathbf{P}}_t^{G_t^{\mathbf{P}}}$, where $\hat{\mathbf{P}}_t^{g_1} = (\hat{\rho}_{c'1t}^{g_1}, \ldots, \hat{\rho}_{c'nt}^{g_1}, \ldots, \hat{\rho}_{c'Nt}^{g_1})$ denotes the g_1th realization of the average joint failure probability regarding component c' at period t. In the same fashion, we can define a random vector $\mathbf{R}_t$ for the return rates of the product installed bases. It has $G_t^{\mathbf{R}}$ different realizations with $\hat{\mathbf{R}}_t^{g_2} = (\hat{r}_{1t}, \ldots, \hat{r}_{nt}, \ldots, \hat{r}_{Nt})$, $\forall\, g_2 \in \{1, \ldots, G_t^{\mathbf{R}}\}$. To sum up, this setup corresponds to the situation where the vectors of the average joint failure probability of component c' and the return rates of the product installed bases all together have $G_t = G_t^{\mathbf{P}} \cdot G_t^{\mathbf{R}}$ possible realizations at each period, i.e., $(\hat{\mathbf{P}}_t^1, \hat{\mathbf{R}}_t^1), \ldots, (\hat{\mathbf{P}}_t^{g_1}, \hat{\mathbf{R}}_t^{g_2}), \ldots, (\hat{\mathbf{P}}_t^{G_t^{\mathbf{P}}}, \hat{\mathbf{R}}_t^{G_t^{\mathbf{R}}})$ and each realization can be deemed as a scenario. In addition, let ϕ_{gt} represent the probability that the joint random vector is revealed as $(\hat{\mathbf{P}}_t^{g_1}, \hat{\mathbf{R}}_t^{g_2})$ at period t, such that $\sum_{g=1}^{G_t} \phi_{gt} = 1$, $\forall\, t \in \mathcal{T}$. Finally, all those scenarios will create the scenario tree $\mathscr{T}$ and the total amount of scenario nodes in the tree is calculated as $\Pi_{t=1}^{T} G_t$. For the ease of notation, we will index the scenario node of the tree as $m \in \mathscr{T}$ and denote the corresponding probability of that node as ϕ_m. An illustration of the scenario tree with two planning periods is presented in Figure 1. As shown in the figure, in the selling season (note that we index the stages as 0, 1, $\cdots$, T, so the selling season is indexed as stage 1), there is no returns collected so that only the average joint failure probability has two realizations, i.e., $\mathbf{P}_1 = (\hat{\mathbf{P}}_1^1, \hat{\mathbf{P}}_1^2)$. In the product EOL phase (the second period/layer), there are four possible realizations of the joint random vector with two realizations for the average joint failure probability and two realizations for the return rate, i.e., $(\mathbf{P}_2, \mathbf{R}_2) = \{(\hat{\mathbf{P}}_2^1, \hat{\mathbf{R}}_2^1), (\hat{\mathbf{P}}_2^2, \hat{\mathbf{R}}_2^1), (\hat{\mathbf{P}}_2^1, \hat{\mathbf{R}}_2^2), (\hat{\mathbf{P}}_2^2, \hat{\mathbf{R}}_2^2)\}$. In this case, the number of scenario nodes in the tth tree layer is equal to $2 \cdot 4^{(t-1)}$, $\forall\, t \in \mathcal{T}$.

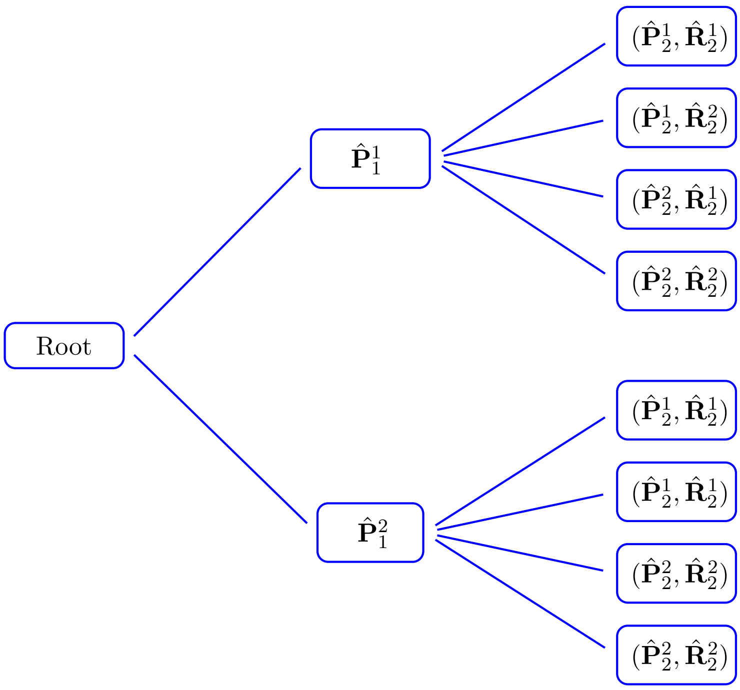

Figure 5.1: An illustration of the scenarios tree

A multistage stochastic programming model

To handle the inherent uncertainties in both the component failure probabilities and the product return rates in this problem, the deterministic model will be extended to a multistage stochastic programming model as follows. Additional notation not listed in Table 5.1 for the stochastic programming model is shown in Table 5.2. The model is

Table 5.2: Additional notation used in the multistage stochastic programming model

Sets and subscripts	
$\mathscr{T}$	Set of all nodes in the scenario tree
$\mathscr{T}_t$	Set of scenario nodes in the first t layers of the scenario tree
$\mathscr{T}_t'$	Set of scenario nodes excluding the first t layers of the scenario tree, i.e., $\mathscr{T}_t' = \mathscr{T} \setminus \mathscr{T}_t$
m	Subscript of node in the scenario tree, $\forall m \in \mathscr{T}$
$a(m)$	Direct predecessor of node m, $\forall m \in \mathscr{T}$
$\mathscr{S}(m)$	Set of direct successors of node m, $\forall m \in \mathscr{T}$
Parameters	
ϕ_m	Probability of node m, $\forall m \in \mathscr{T}$
ρ_{cnm}	Average joint failure probability that a faulty product variant n having faulty component c at node m, $\forall c \in \mathcal{C}$, $\forall n \in \mathcal{N}_c$, $\forall m \in \mathscr{T}$
r_{nm}	Return rates of product n at node m, $\forall n \in \mathcal{N}$, $\forall m \in \mathscr{T}$
δ_{nm}	Percentage of the installed base of product n will exist in the future at node m, $\forall n \in \mathcal{N}$, $\forall m \in \mathscr{T}$
Decision variables	
d_{cm}	Demand for component c at node m, $\forall c \in \mathcal{C}$, $\forall m \in \mathscr{T}$
w_{nm}	Installed base of product n at node m, $\forall n \in \mathcal{N}$, $\forall m \in \mathscr{T}$
q_{cm}	The order quantity of component c at node m, $\forall m \in \mathscr{T}$
f_{nm}	Remanufacturing quantity of the returned product n at node m, $\forall n \in \mathcal{N}$, $\forall m \in \mathscr{T}$
l_{cm}	The number of component c available in stock at node m, $\forall c \in \mathcal{C}$, $\forall m \in \mathscr{T}$
b_{cm}	The shortage amount of component c at node m, $\forall c \in \mathcal{C}$, $\forall m \in \mathscr{T}$
z_{njm}	$z_{njm} = \gamma_{nm} \cdot x_{jm}$, auxiliary variables used for linearization, $\forall n \in \mathcal{N}$, $\forall j \in \mathcal{N}$, $\forall m \in \mathscr{T}$

formulated as follows.

$$\max \sum_{m\in\mathscr{T}_1} \phi_m \sum_{n\in\mathcal{N}} P_n w_{nm} - \sum_{m\in\mathscr{T}_2} \phi_m \sum_{c\in\mathcal{C}} K_c q_{cm} -$$
$$\sum_{m\in\mathscr{T}} \phi_m \sum_{c\in\mathcal{C}} (H_c l_{cm} + B_c b_{cm}) - \sum_{m\in\mathscr{T}_1'} \phi_m \sum_{n\in\mathcal{N}} C_r f_{nm} \tag{SP1}$$

$$
\begin{aligned}
\textbf{SP}\quad & s.t.\ \gamma_n\Big(\sum_{j\in\mathcal{N}} x_j e_j + e_0\Big) = x_n e_n, && \forall\, n \in \mathcal{N}, && \text{(SP2)}\\
& \sum_{n\in\mathcal{N}} x_n \le S, && && \text{(SP3)}\\
& \gamma_n \cdot M - 1 \le w_{nm} \le \gamma_n \cdot M, && \forall\, n \in \mathcal{N}, \forall\, m \in \mathscr{T}_1, && \text{(SP4)}\\
& d_{cm} \ge \sum_{n\in\mathcal{N}_c} (1+\rho_{cnm}) \cdot w_{nm}, && \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}_1, && \text{(SP5)}\\
& d_{cm} \ge \sum_{n\in\mathcal{N}_c} \rho_{cnm} \cdot (1 - r_{nm}) \cdot w_{nm}, && \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}_1', && \text{(SP6)}\\
& \delta_{n\,a(m)}(1 - r_{n\,a(m)}) w_{n\,a(m)} - 1 \le w_{nm}, && && \\
& w_{nm} \le \delta_{n\,a(m)}(1 - r_{n\,a(m)}) w_{n\,a(m)}, && \forall\, n \in \mathcal{N}, \forall\, m \in \mathscr{T}_1', && \text{(SP7)}\\
& f_{nm} \le r_{nm} \cdot w_{nm}, && \forall\, n \in \mathcal{N}, \forall\, m \in \mathscr{T}_1', && \text{(SP8)}\\
& q_{cm} + l_{c0} = d_{cm} + l_{cm}, && \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}_1, && \text{(SP9)}\\
& q_{cm} + l_{c\,a(m)} = d_{cm} + l_{cm} - b_{cm}, && \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}_2 \setminus \mathscr{T}_1, && \text{(SP10)}\\
& \sum_{n\in\mathcal{N}_c} f_{n\,a(m)} + l_{c\,a(m)} = d_{cm} + b_{c\,a(m)} - b_{cm} + l_{cm}, && \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}_2', && \text{(SP11)}\\
& x_n \in \{0,1\}, && \forall\, n \in \mathcal{N}, && \text{(SP12)}\\
& q_{cm} \in \mathbb{Z}^+, && \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}_2, && \text{(SP13)}\\
& w_{nm}, f_{nm} \in \mathbb{Z}^+, \gamma_n \in \mathbb{R}^+, && \forall\, n \in \mathcal{N}, \forall\, m \in \mathscr{T}, && \text{(SP14)}\\
& d_{cm}, l_{cm}, b_{cm} \in \mathbb{Z}^+, && \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}. && \text{(SP15)}
\end{aligned}
$$

Linearization

The proposed multistage stochastic programming model is a mixed integer nonlinear programming model due to the nonlinear term $\gamma_n \sum_{j\in\mathcal{N}} x_j e_j$ in constraints (SP2), for all $n \in \mathcal{N}$. In this section, the reformulation-linearization technique (RLT) (Sherali and Adams, 2013) will be used to reformulate the model as a mixed integer linear programming model through introducing auxiliary variables $z_{nj} = \gamma_n \cdot x_j$, $\forall n, j \in \mathcal{N}$. Furthermore, we need the four more sets of constraints to guarantee $z_{nj} = \gamma_n \cdot x_j$ holds when $x_j = 0$ and $x_j = 1$, $\forall\, j \in \mathcal{N}$. Those additional constraints are

$$z_{nj} \leq x_j, \tag{L1}$$

$$z_{nj} \leq \gamma_n, \tag{L2}$$

$$z_{nj} \geq \gamma_n + x_j - 1, \tag{L3}$$

$$z_{nj} \geq 0. \tag{L4}$$

Through introducing above constraints, we ensure that auxiliary variables z_{nj}, $\forall\, n,\, j \in \mathcal{N}$, will take different values based on the values of binary variables x_j as follows.

$$z_{nj} = \begin{cases} \gamma_n \cdot x_j, & \text{when } x_j = 1; \\ 0, & \text{when } x_j = 0. \end{cases}$$

As a result, the mixed integer nonlinear model will be reformulated as a mixed integer linear model (labelled as SPL model) as follows.

$$\max \sum_{m \in \mathscr{T}_1} \phi_m \sum_{n \in \mathcal{N}} P_n w_{nm} - \sum_{m \in \mathscr{T}_2} \phi_m \sum_{c \in \mathcal{C}} K_c q_{cm} -$$

$$\sum_{m \in \mathscr{T}} \phi_m \sum_{c \in \mathcal{C}} \left(H_c l_{cm} + B_c b_{cm}\right) - \sum_{m \in \mathscr{T}_1'} \phi_m \sum_{n \in \mathcal{N}} C_r f_{nm} \tag{SPL1}$$

$$s.t. \quad e_0 \gamma_n + \sum_{j \in \mathcal{N}} e_j z_{nj} = x_n e_n, \qquad \forall\, n \in \mathcal{N}, \tag{SPL2}$$

$$\sum_{n \in \mathcal{N}} x_n \leq S, \tag{SPL3}$$

$$\gamma_n \cdot M - 1 \leq w_{nm} \leq \gamma_n \cdot M, \qquad \forall\, n \in \mathcal{N}, \forall\, m \in \mathscr{T}_1, \tag{SPL4}$$

$$d_{cm} \geq \sum_{n \in \mathcal{N}_c} \left(1 + \rho_{cnm}\right) \cdot w_{nm}, \qquad \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}_1, \tag{SPL5}$$

$$d_{cm} \geq \sum_{n \in \mathcal{N}_c} \rho_{cnm} \cdot (1 - r_{nm}) \cdot w_{nm}, \qquad \forall\, c \in \mathcal{C}, \forall\, m \in \mathscr{T}_1', \tag{SPL6}$$

$$\delta_{n\,a(m)}(1 - r_{n\,a(m)}) \cdot w_{n\,a(m)} - 1 \leq w_{nm},$$

$$w_{nm} \leq \delta_{n\,a(m)}(1 - r_{n\,a(m)}) \cdot w_{n\,a(m)}, \qquad \forall\, n \in \mathcal{N},\, \forall\, m \in \mathscr{T}_1', \tag{SPL7}$$

$$f_{nm} \leq r_{nm} \cdot w_{nm}, \qquad \forall\, n \in \mathcal{N}, \forall\, m \in \mathscr{T}_1', \tag{SPL8}$$

$$q_{cm} + l_{c0} = d_{cm} + l_{cm}, \qquad \forall\, c \in \mathcal{C},\, \forall\, m \in \mathscr{T}_1, \tag{SPL9}$$

$$q_{cm} + l_{c\,a(m)} = d_{cm} + l_{cm} - b_{cm}, \qquad \forall\, c \in \mathcal{C},\, \forall\, m \in \mathscr{T}_2 \setminus \mathscr{T}_1, \tag{SPL10}$$

$$\sum_{n \in \mathcal{N}_c} f_{n\,a(m)} + l_{c\,a(m)} = d_{cm} + b_{c\,a(m)} - b_{cm} + l_{cm}, \qquad \forall\, c \in \mathcal{C},\, \forall\, m \in \mathscr{T}_2', \tag{SPL11}$$

$$z_{nj} \leq x_j, \qquad \forall\, n \in \mathcal{N},\, \forall\, j \in \mathcal{N}, \tag{SPL12}$$

$$z_{nj} \leq \gamma_n, \qquad \forall\, n \in \mathcal{N},\, \forall\, j \in \mathcal{N}, \tag{SPL13}$$

$$z_{nj} \geq \gamma_n + x_j - 1, \qquad \forall\, n \in \mathcal{N},\, \forall\, j \in \mathcal{N}, \tag{SPL14}$$

$$z_{nj} \geq 0, \qquad \forall\, n \in \mathcal{N},\, \forall\, j \in \mathcal{N}, \tag{SPL15}$$

$$x_n \in \{0, 1\}, \qquad \forall\, n \in \mathcal{N}, \tag{SPL16}$$

$$q_{cm} \in \mathbb{Z}^+, \qquad \forall\, c \in \mathcal{C},\, \forall\, m \in \mathscr{T}_2, \tag{SPL17}$$

$$w_{nm},\, f_{nm} \in \mathbb{Z}^+, \gamma_n,\, z_{nj} \in \mathbb{R}^+, \qquad \forall\, n,\, j \in \mathcal{N},\, \forall\, m \in \mathscr{T}, \tag{SPL18}$$

$$d_{cm},\, l_{cm},\, b_{cm} \in \mathbb{Z}^+, \qquad \forall\, c \in \mathcal{C},\, \forall\, m \in \mathscr{T}. \tag{SPL19}$$

5.3 Numerical Experiments

In this section, two sets of numerical experiments are performed. Firstly, problem instances are generated to show the value of the joint optimization of the assortment decisions in the selling season and the spare parts procurement and remanufacturing decisions in the EOL phase. Afterwards, the impacts on total expected profits brought by the uncertainties embedded in the product return rates and component failure rates are explored. All numerical experiments are coded in C++ and carried out through the IBM ILOG CPLEX 20.1 optimization package on a PC with an Intel Core i7-10750H 2.60 GHz CPU and 16 GB RAM.

5.3.1 Joint optimization v.s. separate optimization

To highlight the benefits brought by incorporating the assortment planning decisions with the considerations of the spare parts procurement and remanufacturing under warranty service, we generate a set of instances to compare the total expected profits obtained through both the joint optimizations and the separate optimizations. To be specific, for one problem instance, the expected profits obtained through solving the SPL model (i.e., joint optimization model) will be compared with those of determining the assortment first at the selling season and then deciding the spare parts procurement and remanufacturing decisions at the EOL product phase (i.e., separate optimization). In the following, we denote the value of joint optimization as

$$\tau = 100 \cdot \frac{z_J - z_S}{z_S}, \tag{5.7}$$

where z_J denotes the total expected profits obtained by the joint optimization while z_S represents the total expected revenue obtained by the separate optimization. In this experiment, we generate 12 problem instances in total and the instances are labelled from JS1 to JS12 (JS represents "joint v.s. separate"). To be specific, the number of product variants ranges from 5 to 7 and so does the number of components. The product prices, component purchasing costs, remanufacturing costs, holding costs, and backlogging costs in all instances are randomly generated. We implement different scenario tree structures in these instances as follows. The number of periods included in the product lifecycle is set as $T = 3$ or 4 in different instances, leading to the scenario trees with 3 or 4 layers respectively. The whole lifecycle consists a one-period selling season and a 2-period or 4-period product EOL phase. The number of possible realizations of both product return rate and component failure rate is set as 2. In this case, in the selling season layer of the tree, one parent node has two branches (or children nodes in the next layer) while four branches for the parent nodes in the non-leaf EOL layer of the tree. This results to 43 or 171 nodes in the whole tree respectively. The corresponding results are shown in Table 5.3 and the higher profit obtained by either of two optimization strategies is highlighted in boldface.

Table 5.3: The value of dynamic assortment decisions

Instance No.	Inst. size				z_S	z_J	τ
	$\lvert\mathcal{N}\rvert$	$\lvert\mathcal{C}\rvert$	T	$\lvert\mathscr{T}\rvert$			
JS1	5	5	3	43	7611.67	**7702.89**	1.20%
JS2	5	5	3	43	12024.8	**12095.4**	0.59%
JS3	6	6	3	43	9679.99	**9847.64**	1.73%
JS4	6	6	3	43	3885.13	**3952.58**	1.73%
JS5	7	7	3	43	1139.47	**1264.33**	10.96%
JS6	7	7	3	43	902.379	**1123.62**	24.51%
JS7	5	5	4	171	8141.54	**8283.72**	1.75%
JS8	5	5	4	171	5721.77	**10010.4**	1.82%
JS9	6	6	4	171	9831.78	**6527.4**	14.08%
JS10	6	6	4	171	3640.73	**4139.57**	13.70%
JS11	7	7	4	171	4753.42	**5640.13**	18.65%
JS12	7	7	4	171	2330.7	**2758.52**	18.36%

From the results in Table 5.3, we can see that the benefits brought by joint optimization is significantly increased as the quantities of both the product variants and the components in the instance rise. This indicates that when the OEM has many products variants which can be put into the assortment or many components used in those variants, they should determine the assortment decisions from a lifecycle perspective, i.e., they need to consider incorporating the spare parts procurement and remanufacturing with assortment planning. Meanwhile, the benefits of joint optimization is more significant when more periods are included in the planning horizon. Specifically, the values of joint optimization in the instances JS7-JS12 are higher than those in the instances JS1-JS6, which have one less period. This observation is due to that, as the number of periods in the product EOL phase rises, the costs for providing the warranty service to a product will increase, weakening the role of the customer preferences on the assortment planning decisions.

5.3.2 Impact of uncertainty levels

In this chapter, we consider two types of uncertainties, one is embedded in the product return rates and the other is in the component failure rates. To investigate the impacts of those uncertainties on the total expected profits gained by the OEM through selling and maintaining the assortment over the whole product lifecycle, we set up numerical experiments. Specifically, we aim to explore the impacts of the uncertainty levels of both uncertainties on the decisions made by the OEM.

In the experiments, we randomly generate 30 problem instances. These instances consist of 10 sets of problem instances and each set has 3 instances with different scales of uncertainty. Those instance sets are labelled from IU1 to IU10 (IU represents "impact of uncertainty"). In the same instance set, the three instances have various uncertainty levels for both the product return rates and the component failure rates. Specifically, the parameters of the product return rates and the component failure rates in the three instances are generated from three different uniform distributions: Uniform[5,30], Uniform[10, 25], and Uniform[15, 20], respectively. The corresponding results are shown in Table 5.4.

Table 5.4: The expected total profits gained under different uncertainty levels in the product return rates and component failure rates

Instance set	The expected total profits		
	$[5, 30]$	$[10, 25]$	$[15, 20]$
IU1	10522.90	12165.30	11928.50
IU2	10502.20	10490.90	10548.60
IU3	9311.65	9607.29	9768.81
IU4	7654.94	8611.81	8833.56
IU5	9451.09	9891.33	9749.79
IU6	8214.87	8500.24	8346.56
IU7	15795.30	15942.50	16146.30
IU8	14733.80	14760.00	15092.10
IU9	12769.50	12510.80	12648.30
IU10	10652.40	10720.60	11055.90
Average	10960.87	11320.08	11411.84

From the table, we can see that generally the OEM could obtain higher expected total profits when the size of the uncertainty set decreases. The average expected total profits of the instances increase from 10960.87 to 11320.08 when the uncertainty set is changed from $Uniform[5, 30]$ to $Uniform[10, 25]$ and finally reaches the greatest value 11411.84 when the uncertainty set becomes $Uniform[15, 20]$. This observation indicates the uncertainty levels significantly impacts the expected total profits gained by the OEM. A managerial insight can be derived accordingly, i.e., the more the errors the OEM can diminish when estimating both the product return rates and component failure rates, the higher the expected total profits could be obtained.

5.4 Conclusions and Future Research Directions

In this chapter, we studied an assortment planning problem from the perspective of the product lifecycle. Specifically, we assume that the product lifecycle consists of two phases: one is the selling season and the other is the EOL phase. In the selling season, the OEM selects the products put in the assortment based on the customer preferences and produce them in an assemble-to-order system. After the selling season, the production stops and the supplies of certain parts are permanently discontinued by suppliers. We first developed a deterministic model for this problem and then extended it in to a multistage stochastic programming model when the uncertainties embedded in the estimations of both the products return rates and components failure rates. The aim of this study is to explore the impacts on product assortment decisions brought by implementing the LTB and remanufacturing strategies to supply the spare parts through the warranty periods. Through the numerical experiments, we first demonstrated the advantages of joint optimization on the assortment planning decisions and the spare parts procurement and remanufacturing decisions compared to the separate optimizations on those decisions. Afterwards, we explored the impacts of the uncertainty levels of those two uncertainties on the the expected total profits obtained by the OEM. We found that the OEM can obtain higher expected total profits on average when the levels of the uncertainty set decrease. This observation indicates the OEM can optimize their decision-making by improving their estimation tools used for predicting both the product return rates and component failure rates.

The future research directions include the following two aspects. Firstly, more research can focus on developing solution methods for the proposed multistage stochastic programming model in this chapter. This model is a mixed integer programming and the commercial software (e.g., CPLEX and Gurobi) may not be able to solve the large scale instances in the reasonable time. Therefore, new solution methods for solving the large scale instances are valuable in practice. Secondly, as indicated in the numerical experiments, the joint optimization on the product assortment decisions and spare parts procurement and remanufacturing decisions can increase the expected total profits but subjects to the estimation accuracy of both the products return rates and components failure rates. In this sense, developing reliable and accurate estimation techniques by utilizing the data-mining and machining learning methods is another future direction under this topic.

Chapter 6

Conclusions

In this chapter, the major contributions of this thesis are summarized. In this thesis, we build up mathematical models based on adaptive robust optimization (ARO) and stochastic programming (SP) for the spare parts inventory management under various uncertainties in a multi-period planning horizon, especially for the substitute consumer durable products offered in an assortment. In other words, our main focus is to incorporate the spare parts inventory management with assortment planning under uncertainties. The uncertainties considered in this thesis are normally embedded in the prediction and estimation process for many parameters needed in the assortment planning and spare parts inventory management. To be specific, the uncertainties considered in this thesis include the products and spare parts failure rates, the customer preferences, the used products return rates, etc.

In Chapter 2, we present a comprehensive literature review on spare parts inventory management and 142 papers are surveyed and classified. Two different typologies are used for the literature classification. One typology classifies the literature based on the systematic characteristics of spare parts inventory systems while the other typology is built based on the research methodologies and topics. This review presents a big picture on the spare parts supply chains to discuss the studies on spare parts inventory management. This big picture links the important aspects relevant to managing spare parts inventory system, such as product and spare part types, after-sales services, maintenance operations, inventory management strategies and policies, supply sources, demand patterns and so on. Furthermore, we classify the research methodology from the perspective of supply chain analytics and current studies using descriptive, predictive and prescriptive analytics are identified. Several research directions regarding reverse logistics, spare parts demand pattern modelling, and big data analytics, are also highlighted for future research in this field.

In Chapter 3, we consider a multi-period spare parts inventory system providing spare parts for several products in an assortment and formulate a multi-stage adaptive mixed-integer robust optimization model to minimize the total inventory costs when the uncertainty embedded in the spare parts demand is involved. We aim to develop the spare parts inventory policies for the original equipment manufacturer (OEM). We improve the partition-and-bound method proposed by Bertsimas and Dunning (2016) to solve the proposed model and conduct extensive numerical experiments to validate its performance. Through the sensitivity analysis, we explore the impacts of spare parts purchase cost, product popularity, and product backorder cost on the inventory policy and total cost, and provide some managerial insights regarding how to adjust the order quantities for both the dedicated and common spare parts used in the popular and unpopular products and how to determine the order quantities of those spare parts when the backorder costs of products using them are changing.

In Chapter 4, we consider a dynamic assortment planning problem for an OEM who launches and sells the product variants through an online platform under uncertain customer preferences over a multi-period selling season. We aim to develop the dynamic products assortment decisions and spare parts inventory policies when there are uncertainties in the customer preferences to the products in the assortment. A multi-stage stochastic programming model is proposed for this problem. We design a branch-and-price (B&P) algorithm based on the block-angular structure of the stochastic programming model. In the numerical experiments, we evaluate the advantage brought by the dynamic assortment compared to the static assortment and the structure of the optimal solutions. In addition, the impacts of component commonality and unit cost on the decisions and total expected profits are explored. Finally, the performance of the B&P algorithm is confirmed in the experiments through comparing it with CPLEX solver.

In Chapter 5, we study an assortment planning problem from the perspective of the product lifecycle and intend to simultaneously determine the products assortment decisions, spare parts inventory policies, and returned product remanufacturing decisions when there are uncertainties in the failure rates of both the products and the spare parts, and the return rates of the used products. We first develop a deterministic model for this problem and then extend it in to a multistage stochastic programming model when the uncertainties embedded in the estimations of both the products return rates and components failure rates are considered. The aim of this study is to explore the impacts on product assortment decisions brought by implementing the last-time buy (LTB) and remanufacturing strategies to supply the spare parts over the warranty periods. In the numerical experiments, we first evaluate the advantages of joint optimization of the assortment planning decisions and the spare parts procurement and remanufacturing decisions. Afterwards, we explore the impacts of the uncertainty levels of those two uncertainties on the the expected total profits obtained by the OEM. Finally, we observe that the OEM can optimize their decision making by improving their estimation tools used for predicting both the product return rates and component failure rates.

Appendix A

Chapter 3 Supplements

A1 Reformulation to the AMIO Model

To replace $l_{ct}(\boldsymbol{d}_c)$ in the objective (3.6) and constraints (3.9) with constraints (3.7), the AMIO model (3.6)-(3.10) can be rewritten as:

$$\begin{aligned}\min_{q^p_{ct}(\boldsymbol{d}_c), b_{ct}(\boldsymbol{d}_c), \theta_{nt}(\boldsymbol{d})} \max_{\boldsymbol{d}\in\Xi} z = & \sum_{t\in\mathcal{T}\setminus\{T-L_c+1,\cdots,T\}} \sum_{c\in\mathcal{C}} P_c\, q^p_{ct}(\boldsymbol{d}_c) \\ & + \sum_{t\in\mathcal{T}} \sum_{c\in\mathcal{C}} H_c\, l_{ct}(\boldsymbol{d}_c) + \sum_{t\in\mathcal{T}} \sum_{n\in\mathcal{N}} B_n \theta_{nt}(\boldsymbol{d})\end{aligned} \tag{A.1}$$

$$s.t. \sum_{n\in\mathcal{N}_c} \theta_{nt}(\boldsymbol{d}) \geq b_{ct}(\boldsymbol{d}_c), \qquad \forall c\in\mathcal{C},\, t\in\mathcal{T}, \tag{A.2}$$

$$l_{ct}(\boldsymbol{d}_c) \geq SS_c, \qquad \forall c\in\mathcal{C},\, t\in\mathcal{T}, \tag{A.3}$$

$$q^p_{ct}(\boldsymbol{d}_c),\, b_{ct}(\boldsymbol{d}_c), \theta_{nt}(\boldsymbol{d}) \in \mathbb{Z}^+, \qquad \forall c\in\mathcal{C},\, t\in\mathcal{T}, \tag{A.4}$$

where,

$$l_{c,t+1}(\boldsymbol{d_c}) = l_{c,t}(\boldsymbol{d_c}) - \left(d_c^{t+1} + b_{c,t}(\boldsymbol{d}_c) - b_{c,t+1}(\boldsymbol{d}_c)\right), \quad \forall c\in\mathcal{C},\, t\in\{1,\cdots,L_c\}, \tag{A.5}$$

$$l_{c,t+1}(\boldsymbol{d}_c) = l_{ct}(\boldsymbol{d}_c) + q^p_{c,t-L_c}(\boldsymbol{d}_c) - \left(d_c^t + b_{c,t-1}(\boldsymbol{d}_c) - b_{ct}(\boldsymbol{d}_c)\right), \quad \forall c\in\mathcal{C},\, t\in\mathcal{T}\setminus\{1,\cdots,L_c\},$$

$$d_c^t = \sum_{n\in\mathcal{N}} d_{cn}^t, \qquad \forall c\in\mathcal{C},\, t\in\mathcal{T},\, \forall\, d_{cn}^t \in \boldsymbol{d}_{cn}, \tag{A.6}$$

and $\boldsymbol{d}_{cn}$ represents the demand for spare part c used in product n. In addition, the inventory level at a certain period t can be represented as follows:

$$l_{c,t+1}(\boldsymbol{d}_c) = \begin{cases} l_{c1}(\boldsymbol{d}_c) - \sum_{k=1}^{t} d_c^k + b_{ct}(\boldsymbol{d}_c), \, \forall\, t \in \{1, \cdots, L_c\}, \\ l_{c,1+L_c}(\boldsymbol{d}_c) + \sum_{k=1}^{t-L_c} q_{ck}^p(\boldsymbol{d}_c) - \sum_{k=L_c+1}^{t} d_c^k - b_{c,L_c} + b_{ct}(\boldsymbol{d}_c), \\ \qquad\qquad\qquad\qquad \forall\, t \in \mathcal{T} \setminus \{1, \cdots, L_c\}. \end{cases} \tag{A.7}$$

In the following contents, we will use q_{ct}^p, l_{ct}, b_{ct}, and θ_{nt} to replace $q_{ct}^p(\boldsymbol{d}_c)$, $l_{ct}(\boldsymbol{d}_c)$, $b_{ct}(\boldsymbol{d}_c)$, and $\theta_{nt}(\boldsymbol{d})$ respectively. Note that we can find the summation of inventory levels in k consecutive planning periods in the above equation, i.e., $\sum_{t=2}^{k+1} l_{ct}$ (note initial inventory level is l_{c1}), as

$$\sum_{t=2}^{k+1} l_{ct} = \begin{cases} k l_{c1} - \sum_{t=1}^{k} \sum_{n=1}^{t} d_c^n + \sum_{t=1}^{k} b_{ct}, \;\; \forall\, k \in \{1, \cdots, L_c\}, \\ k l_{c1} + \sum_{t=1}^{k-L_c} \sum_{n=1}^{t} q_{cn}^p + \sum_{t=1}^{k} b_{ct} - \sum_{t=1}^{k} \sum_{n=1}^{t} d_c^n, \, \forall\, k \in \mathcal{T} \setminus \{1, \cdots, L_c\}. \end{cases} \tag{A.8}$$

Using equation (A.8) to replace the term $\sum_{t \in \mathcal{T}} l_{ct}(\boldsymbol{d_c})$, i.e., $\sum_{t=2}^{T+1} l_{ct}$ in the objective (3.6), we can obtain a new form of objective function without inventory level as:

$$\begin{aligned} z &= \sum_{c \in \mathcal{C}} \sum_{t=1}^{T-L_c} P_c\, q_{ct}^p + \sum_{c \in \mathcal{C}} H_c (T l_{c1} + \sum_{t=1}^{T-L_c} \sum_{n=1}^{t} q_{cn}^p + \sum_{t=1}^{T} b_{ct} - \sum_{t=1}^{T} \sum_{n=1}^{t} d_c^n) + \sum_{n=1}^{N} B_n \sum_{t=1}^{T} \theta_{nt} \\ &= \sum_{c \in C} \Big[\sum_{t=1}^{T-L_c} (P_c q_{ct}^p + H_c \sum_{n=1}^{t} q_{ct}^p) + H_c \sum_{t=1}^{T} b_{ct} + T H_c l_{c0} - H_c \sum_{t=1}^{T} \sum_{n=1}^{t} d_c^n \Big] + \sum_{n=1}^{N} B_n \sum_{t=1}^{T} \theta_{nt} \\ &= \sum_{c \in C} \Big[\sum_{t=1}^{T-L_c} (P_c + (T-t) H_c) q_{ct}^p + H_c \sum_{t=1}^{T} b_{ct} + T H_c l_{c0} - H_c \sum_{t=1}^{T} (T - t + 1) d_c^t \Big] \\ &\quad + \sum_{n=1}^{N} B_n \sum_{t=1}^{T} \theta_{nt} \end{aligned} \tag{A.9}$$

A2 Partition-and-Bound Method for Multistage AMIO

Algorithm 2 Partition-and-bound for multistage AMIO

Results: z_{alg}, the objective value of multi-stage adaptive robust optimization problem obtained using adaptive partition algorithm.
Step 1. Initialization. Input the parameters and initialize number of iteration $k \leftarrow 1$ and the root node (any $\boldsymbol{d} \in \Xi$).
Step 2. Solve the partition model of the problem, with one partition ($\Xi(\hat{\boldsymbol{d}}^j)$) for every $\hat{\boldsymbol{d}}^j \in Leaves(\mathscr{T}^k)$, where $Leaves(\mathscr{T}^k)$ represents the set of leaves of the scenario tree. The problem is

$$z_{alg}(\mathscr{T}^k) = \min z$$

$$s.t. \sum_{c \in C} \left[\sum_{t=1}^{T-L_c} \left(P_c + (T-t)H_c\right) q^p_{ct} + H_c \sum_{t=1}^{T} b_{ct} + TH_c l_{c0} - H_c \sum_{t=1}^{T} (T-t+1) d^t_c \right]$$
$$+ \sum_{n=1}^{N} B_n \sum_{t=1}^{T} \theta_{nt} \leq z, \forall c \in \mathcal{C},\, t \in \mathcal{T},\, d^t_c \in \Xi(\hat{\boldsymbol{d}}^j),\, \hat{\boldsymbol{d}}^j \in Levaes(\mathscr{T}^k), \tag{A.10}$$
$$l_{ct} \geq SS_c, \qquad \forall c \in \mathcal{C},\, t \in \mathcal{T},$$
$$\sum_{n \in \mathcal{N}_c} \theta_{nt} \geq b_{ct}, \qquad \forall c \in \mathcal{C},\, t \in \mathcal{T},$$
$$\boldsymbol{q}^t_i = \boldsymbol{q}^t_j, \qquad \forall\ \hat{\boldsymbol{d}}^i, \hat{\boldsymbol{d}}^j \in Leaves(\mathscr{T}^k),$$
$$\forall\ t : \Xi(\hat{\boldsymbol{d}}^i)^{t-1} \cap \Xi(\hat{\boldsymbol{d}}^j)^{t-1} \neq \emptyset \tag{A.11}$$
$$q^p_{ct},\, b_{ct},\, \theta_{nt} \in \mathbb{Z}^+, \qquad \forall c \in \mathcal{C}, \forall n \in \mathcal{N},\, t \in \mathcal{T}.$$

where $\boldsymbol{q}^t_i$ is equivalent to $(q^p_{1t}, \ldots, q^p_{Ct})^T$, Equation (A.11) represents nonanticipativity constraints (Bertsimas and Dunning, 2016), and

$$l_{c,t+1} = \begin{cases} l_{c1} - \sum_{k=1}^{t} d^k_c + b_{ct}, \forall\, t \in \{1, \cdots, L_c\} \\ l_{c1} + \sum_{k=1}^{t-L_c} q^p_{ck} - \sum_{k=1}^{t} d^k_c + b_{ct},\ \forall\, t \in \mathcal{T} \setminus \{1, \cdots, L_c\}. \end{cases}$$

Step 3. Grow the tree. Initialize $\mathscr{T}^{k+1} \leftarrow \mathscr{T}^k$ and do the following. For each leaf node in $\mathscr{T}^{k+1}$, add children to that leaf for each $\hat{\boldsymbol{d}}$ in the set of active uncertain parameters $\hat{\boldsymbol{d}}^j$ (obtained by finding the constraints with zero slack) for the solution to (A.10).
Step 4. Determine partition $\Xi(\hat{\boldsymbol{d}}^j)$ for each child node $\hat{\boldsymbol{d}}^j \in \mathscr{T}^{k+1}$ by the scheme:

$$\Xi(\hat{\boldsymbol{d}}^i) = \{\boldsymbol{d} |\ \| \hat{\boldsymbol{d}}^i{}_{t_{i,j}} - \boldsymbol{d}_{t_{i,j}} \|_2 \leq \| \hat{\boldsymbol{d}}^j{}_{t_{i,j}} - \boldsymbol{d}_{t_{i,j}} \|_2, \forall \hat{\boldsymbol{d}}^j \in Siblings(\hat{\boldsymbol{d}}^i)\} \tag{A.12}$$
$$\cap \{\boldsymbol{d} |\ \| Parent(\hat{\boldsymbol{d}}^i)_{t'_{i,j}} - \boldsymbol{d}_{t'_{i,j}} \|_2 \leq \| \hat{\boldsymbol{d}}^j{}_{t'_{i,j}} - \boldsymbol{d}_{t'_{i,j}} \|_2, \forall \hat{\boldsymbol{d}}^j \in Siblings(Parent(\hat{\boldsymbol{d}}^j))\}$$
$$\cap \cdots \cap \Xi,$$

where, $t_{i,j}$ is the minimum t such that $\hat{\boldsymbol{d}}^i_t \neq \hat{\boldsymbol{d}}^j_t \in Siblings(\hat{\boldsymbol{d}}^i)$, i.e., the first time stage for which the demand of $\hat{\boldsymbol{d}}^i$ differs from that of $\hat{\boldsymbol{d}}^j$.

Algorithm 3 Partition-and-bound for multi-stage AMIO (continued)

Step 5. Calculate a lower bound $z_{low}(\mathscr{T}^{k+1})$ of the fully adaptive solution by solving following program:

$$
\begin{aligned}
& z_{low}(\mathscr{T}^{k+1}) = \min\ z \\
s.t. & \sum_{c \in C} \left[\sum_{t=1}^{T-L_c} \left(P_c + (T-t)H_c\right) q^p_{ct} + H_c \sum_{t=1}^{T} b_{ct} + T H_c l_{c0} - H_c \sum_{t=1}^{T} (T-t+1) d^t_c \right] \\
& + \sum_{n=1}^{N} B_n \sum_{t=1}^{T} \theta_{nt} \leq z, \qquad \forall c \in \mathcal{C},\, t \in \mathcal{T},\, d^t_c \in \hat{\boldsymbol{d}}^j,\, \hat{\boldsymbol{d}}^j \in Levaes(\mathscr{T}^k), \qquad (A.13) \\
& l_{ct} \geq SS_c, \qquad \forall c \in \mathcal{C},\, t \in \mathcal{T}, \\
& \sum_{n \in \mathcal{N}_c} \theta_{nt} \geq b_{ct}, \qquad \forall c \in \mathcal{C},\, t \in \mathcal{T}, \\
& \boldsymbol{q}^t_i = \boldsymbol{q}^t_j, \qquad \forall\ \hat{\boldsymbol{d}}^i, \hat{\boldsymbol{d}}^j \in Leaves(\mathscr{T}^k), \\
& \qquad\qquad \forall\ t : \Xi(\hat{\boldsymbol{d}}^i)^{t-1} \cap \Xi(\hat{\boldsymbol{d}}^j)^{t-1} \neq \emptyset \\
& q^p_{ct},\, b_{ct},\, \theta_t \in \mathbb{Z}^+, \qquad \forall c \in \mathcal{C},\, t \in \mathcal{T}.
\end{aligned}
$$

where

$$
l_{c,t+1} = \begin{cases} l_{c1} - \sum_{k=1}^{t} d^k_c + b_{ct},\ \forall\, t \in \{1, \cdots, L_c\} \\ l_{c1} + \sum_{k=1}^{t-L_c} q^p_{ck} - \sum_{k=1}^{t} d^k_c + b_{ct},\ \forall\, t \in \mathcal{T} \setminus \{1, \cdots, L_c\}. \end{cases}
$$

Terminate if the bound gap

$$
\frac{z_{alg}(\mathscr{T}^k) - z_{low}(\mathscr{T}^{k+1})}{|z_{low}(\mathscr{T}^{k+1})|} \qquad (A.14)
$$

is less than ϵ_{gap}. Otherwise, set $k \leftarrow k+1$ and go to Step 2.

A3 A Three-stage Example with Two Products and Two Spare Parts

The corresponding three-stage AMIO model in this example is as follows.

$$
\begin{aligned}
& \min \quad z \\
s.t.\ & (P_1 + 2H_1) q^p_{11} + (P_1 + H_1) q^p_{12} + H_1(b_{11} + b_{12} + b_{13}) + 3H_1 l_{11} - H_1(3d^1_1 + 2d^2_1 + d^3_1) \\
& + (P_2 + 2H_2) q^p_{21} + (P_2 + H_2) q^p_{22} + H_2(b_{21} + b_{22} + b_{23}) + 3H_2 l_{21} - H_2(3d^1_2 + 2d^2_2 \\
& + d^3_2) + B_1(\theta_{11} + \theta_{12} + \theta_{13}) + B_2(\theta_{21} + \theta_{22} + \theta_{23}) \leq z,\ \forall\, d^t_c \in \hat{\boldsymbol{d}}^j,\ \hat{\boldsymbol{d}}^j \in Leaves(\mathscr{T}^k),
\end{aligned}
\qquad (A.15)
$$

$$l_{11} + b_{11} - d_1^1 \geq SS_1,\, l_{21} + b_{21} - d_2^1 \geq SS_2, \quad \forall\, d_c^t \in \hat{\boldsymbol{d}}^j,\, \hat{\boldsymbol{d}}^j \in Leaves(\mathscr{T}^k), \tag{A.16}$$

$$l_{11} + q_{11}^p + b_{12} - (d_1^1 + d_1^2) \geq SS_1, \quad l_{21} + q_{21}^p + b_{22} - (d_2^1 + d_2^2) \geq SS_2,$$
$$\forall\, d_c^t \in \hat{\boldsymbol{d}}^j,\, \hat{\boldsymbol{d}}^j \in Leaves(\mathscr{T}^k), \tag{A.17}$$

$$l_{11} + q_{11}^p + q_{12}^p + b_{13} - (d_1^1 + d_1^2 + d_1^3) \geq SS_1,$$
$$l_{21} + q_{21}^p + q_{22}^p + b_{23} - (d_2^1 + d_2^2 + d_2^3) \geq SS_2,\; \forall\, d_c^t \in \hat{\boldsymbol{d}}^j,\, \hat{\boldsymbol{d}}^j \in Leaves(\mathscr{T}^k), \tag{A.18}$$

$$\theta_{11} + \theta_{21} \geq b_{11},\, \theta_{11} + \theta_{21} \geq b_{21},$$
$$\theta_{12} + \theta_{22} \geq b_{12},\, \theta_{12} + \theta_{22} \geq b_{22},$$
$$\theta_{13} + \theta_{23} \geq b_{13},\, \theta_{13} + \theta_{23} \geq b_{23},$$
$$q_{11}^p,\, q_{21}^p,\, q_{12}^p,\, q_{22}^p,\, q_{13}^p,\, q_{23}^p,\, b_{11},\, b_{21},\, b_{12},\, b_{22},\, b_{13},\, b_{23} \geq 0.$$

For this small instance, we will solve it and then find the active uncertain parameters in constraints (A.15), (A.16), (A.17), and (A.18) respectively. Even the instance has only two products and two spare parts, the quantity of constraints is quite big. For example, constraints (A.15) include $(4 \times 4 \times 4) \times (4 \times 4 \times 4) = 64^2$ constraints. To be specific, for each of the two spare parts in each period, the corresponding demand has four possible realizations and each spare part has three rounds of demands, thus there are $4 \times 4 \times 4 = 64$ possible combinations for each one and 64^2 constraints in total. Similarly, constraints (A.16) have $4 + 4 = 8$ constraints, constraints (A.17) have $4 \times 4 + 4 \times 4 = 32$ and constraints (A.18) have $4 \times 4 \times 4 + 4 \times 4 \times 4 = 128$ constraints. Clearly, the constraint quantity increases as the iteration number increases, making the problem in later iteration harder to solve.

A4 Proof of Theorem 3.1

For constraints (3.17), we prove it by contradiction. Given an optimal solution $\boldsymbol{x}_j^*$ and z^*, assume there exists an non-minimum sum-product of coefficients $\boldsymbol{c'}^{(2)t}$ and corresponding realized spare parts demands $\hat{\boldsymbol{\xi}'}_j^t \in \Xi(\hat{\boldsymbol{\xi}}_j)$, i.e., $\sum_{t=1}^{T} \boldsymbol{c'}^{(2)t} \cdot \hat{\boldsymbol{\xi}'}_j^t - \delta = \sum_{t=1}^{T} \underline{\boldsymbol{c}}^{(2)t} \cdot \hat{\underline{\boldsymbol{\xi}}}_j^t$, such that the active uncertain parameters are in the constraints with it:

$$\sum_{t=1}^{T} \boldsymbol{c}^{(1)t}(\boldsymbol{\xi}) \cdot \boldsymbol{x}_j^{*t} - \sum_{t=1}^{T} \boldsymbol{c'}^{(2)t} \cdot \hat{\boldsymbol{\xi}'}_j^t \leq z^*, \qquad \forall \hat{\boldsymbol{\xi}}_j \in Leaves. \tag{A.19}$$

If we replace $\sum_{t=1}^{T} \boldsymbol{c'}^{(2)t} \cdot \hat{\boldsymbol{\xi}'}_j^t$ in constraints (A.19) with $\sum_{t=1}^{T} \underline{\boldsymbol{c}}^{(2)t} \cdot \underline{\hat{\boldsymbol{\xi}}}_j^t + \delta$:

$$\begin{aligned} &\sum_{t=1}^{T} \boldsymbol{c}^{(1)t}(\boldsymbol{\xi}) \cdot \boldsymbol{x}_j^{*t} - \sum_{t=1}^{T} \underline{\boldsymbol{c}}^{(2)t} \cdot \underline{\hat{\boldsymbol{\xi}}}_j^t - \delta \leq z^*, \qquad \forall \hat{\boldsymbol{\xi}}_j \in Leaves. \\ \Leftrightarrow &\sum_{t=1}^{T} \boldsymbol{c}^{(1)t}(\boldsymbol{\xi}) \cdot \boldsymbol{x}_j^{*t} - \sum_{t=1}^{T} \underline{\boldsymbol{c}}^{(2)t} \cdot \underline{\hat{\boldsymbol{\xi}}}_j^t \leq z^* + \delta, \qquad \forall \hat{\boldsymbol{\xi}}_j \in Leaves. \end{aligned} \tag{A.20}$$

In addition, for the constraints with minimum sum-product of coefficients $\underline{\boldsymbol{c}}^{(2)t}$ and corresponding realized spare parts demands $\underline{\hat{\boldsymbol{\xi}}}_j^t \in \varXi(\hat{\boldsymbol{\xi}}_j)$, they still hold when optimal solution is achieved:

$$\sum_{t=1}^{T} \boldsymbol{c}^{(1)t}(\boldsymbol{\xi}) \cdot \boldsymbol{x}_j^{*t} - \sum_{t=1}^{T} \underline{\boldsymbol{c}}^{(2)t} \cdot \underline{\hat{\boldsymbol{\xi}}}_j^t \leq z^*, \qquad \forall \hat{\boldsymbol{\xi}}_j \in Leaves. \tag{A.21}$$

Comparing (A.20) and (A.21), we can conclude that (A.21) have lower slack than (A.20), i.e., the active uncertain parameters are $\underline{\hat{\boldsymbol{\xi}}}_j^t$, which contradicts our assumption. For (3.18), we can prove it in the same way as that of (3.17).

A5 Proof of Corollary 3.3

Constraints (3.12) lead to $\Pi_{c=1}^{C}\Pi_{t-1}^{T} \mid \hat{d}_c^t \mid$ constraints, since for the demand of spare part c in a period t, there are $\mid \hat{d}_c^t \mid$ possible realization such that there are $\Pi_{c=1}^{C}\Pi_{t=1}^{T} \mid \hat{d}_c^t \mid$ combinations of realizations of all spare parts in every period. In contrast, constraints (3.23) contain single constraint which has the combination of all spare parts demand realization as $\underline{\hat{\boldsymbol{d}}} = (\underline{\hat{\boldsymbol{d}}}_1, \ldots, \underline{\hat{\boldsymbol{d}}}_c, \ldots, \underline{\hat{\boldsymbol{d}}}_C)^T$, where $\underline{\hat{\boldsymbol{d}}}_c = (\underline{\hat{d}}_c^1, \underline{\hat{d}}_c^2, \ldots, \underline{\hat{d}}_c^T)$

Similarly, for constraints (3.13), given period t and spare part c, there are $(\Pi_{k=1}^{t} \mid \hat{d}_c^k \mid)$ number of constraints induced. The total quantity of constraints included is

$(\sum_{c=1}^{C}\sum_{t=1}^{T}\Pi_{k=1}^{t} \mid \hat{d}_c^t \mid)$. For constraints (3.24), given a period t and a spare part c, only one constraint which has the combination of the spare parts demand realization as $\underline{\hat{\boldsymbol{d}}}_c^t = (\underline{\hat{d}}_c^1, \underline{\hat{d}}_c^2, \ldots, \underline{\hat{d}}_c^t)$, thus leading to $(C \times T)$ constraints.

A6 Illustration of Corollary 3.3 on the Example in A3

At the first iteration of partition-and-bound method, we take constraints (A.15) for instance and write down all the constraints:

$$
\begin{aligned}
&(P_1 + 2H_1)\, q_{11}^p + (P_1 + H_1)\, q_{12}^p + H_1\,(b_{11} + b_{12} + b_{13}) + 3H_1\, l_{11} - H_1(3 + 7 + 13)\\
&+ (P_2 + 2H_2)\, q_{21}^p + (P_2 + H_2)\, q_{22}^p + H_2\,(b_{21} + b_{22} + b_{23}) + 3H_2\, l_{21} - H_2\,(4 + 6 + 12)\\
&+ B_1(\theta_{11} + \theta_{12} + \theta_{13}) + B_2(\theta_{21} + \theta_{22} + \theta_{23}) \le z,\\
&(P_1 + 2H_1)\, q_{11}^p + (P_1 + H_1)\, q_{12}^p + H_1\,(b_{11} + b_{12} + b_{13}) + 3H_1\, l_{11} - H_1(4 + 7 + 13)\\
&+ (P_2 + 2H_2)\, q_{21}^p + (P_2 + H_2)\, q_{22}^p + H_2\,(b_{21} + b_{22} + b_{23}) + 3H_2\, l_{21} - H_2\,(4 + 6 + 12)\\
&+ B_1(\theta_{11} + \theta_{12} + \theta_{13}) + B_2(\theta_{21} + \theta_{22} + \theta_{23}) \le z,\\
&\ldots\\
&(P_1 + 2H_1)\, q_{11}^p + (P_1 + H_1)\, q_{12}^p + H_1\,(b_{11} + b_{12} + b_{13}) + 3H_1\, l_{11} - H_1(6 + 10 + 16)\\
&+ (P_2 + 2H_2)\, q_{21}^p + (P_2 + H_2)\, q_{22}^p + H_2\,(b_{21} + b_{22} + b_{23}) + 3H_2\, l_{21} - H_2\,(7 + 9 + 14)\\
&+ B_1(\theta_{11} + \theta_{12} + \theta_{13}) + B_2(\theta_{21} + \theta_{22} + \theta_{23}) \le z,\\
&(P_1 + 2H_1)\, q_{11}^p + (P_1 + H_1)\, q_{12}^p + H_1\,(b_{11} + b_{12} + b_{13}) + 3H_1\, l_{11} - H_1(6 + 10 + 16)\\
&+ (P_2 + 2H_2)\, q_{21}^p + (P_2 + H_2)\, q_{22}^p + H_2\,(b_{21} + b_{22} + b_{23}) + 3H_2\, l_{21} - H_2\,(7 + 9 + 15)\\
&+ B_1(\theta_{11} + \theta_{12} + \theta_{13}) + B_2(\theta_{21} + \theta_{22} + \theta_{23}) \le z,
\end{aligned}
$$

The only difference among the above constraints is that they have various uncertainty parameters, i.e., the demand realization for spare parts varies. If we apply the results of Corollary 3.3, it is clear that, given a optimal solution $\boldsymbol{q}^p$ and z^*, the constraints with the lowest demand $(3, 7, 13)$ for the first spare part, and $(4, 6, 12)$ for the second spare part, are the constraints with active uncertain parameters. In this sense, we can see that the rest constraints do not affect the partition result, thus can be treated as redundant constraints and eliminated from the model. Same logic can be applied to constraints (A.16), (A.17), and (A.18). Finally we can find that the constraints with the highest demand $(6, 10, 16)$ and $(7, 9, 15)$ for two spare parts respectively are the constraints with active uncertain parameters. With this treatment, we only need to reserve the constraints with active uncertain parameters in the model, leading to only one constraint in (A.15), two in (A.16), and two in (A.17) respectively.

Appendix B

Chapter 4 Supplements

B1 Scenario Tree Structures

The scenario tree structure used in the Instances 1 to 12 is illustrated in Figure 2 while that in Instances 13 to 24 is shown in Figure 3. The probability of each node in the scenario tree is randomly generated.

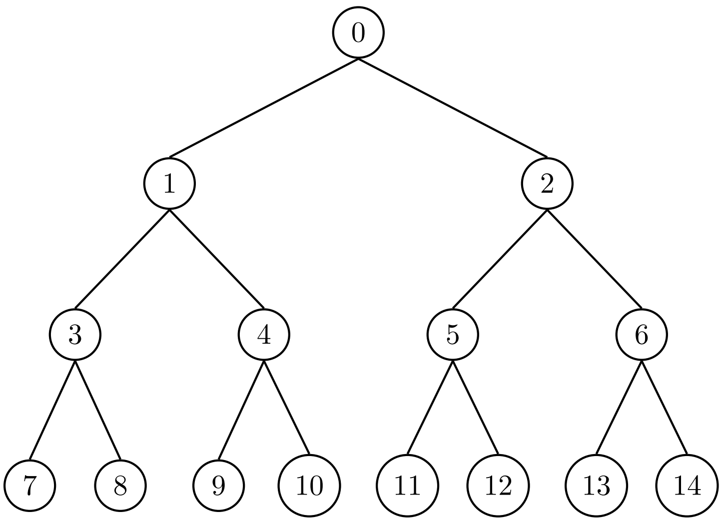

Figure B2.1: The scenario tree used in Instances 1-12

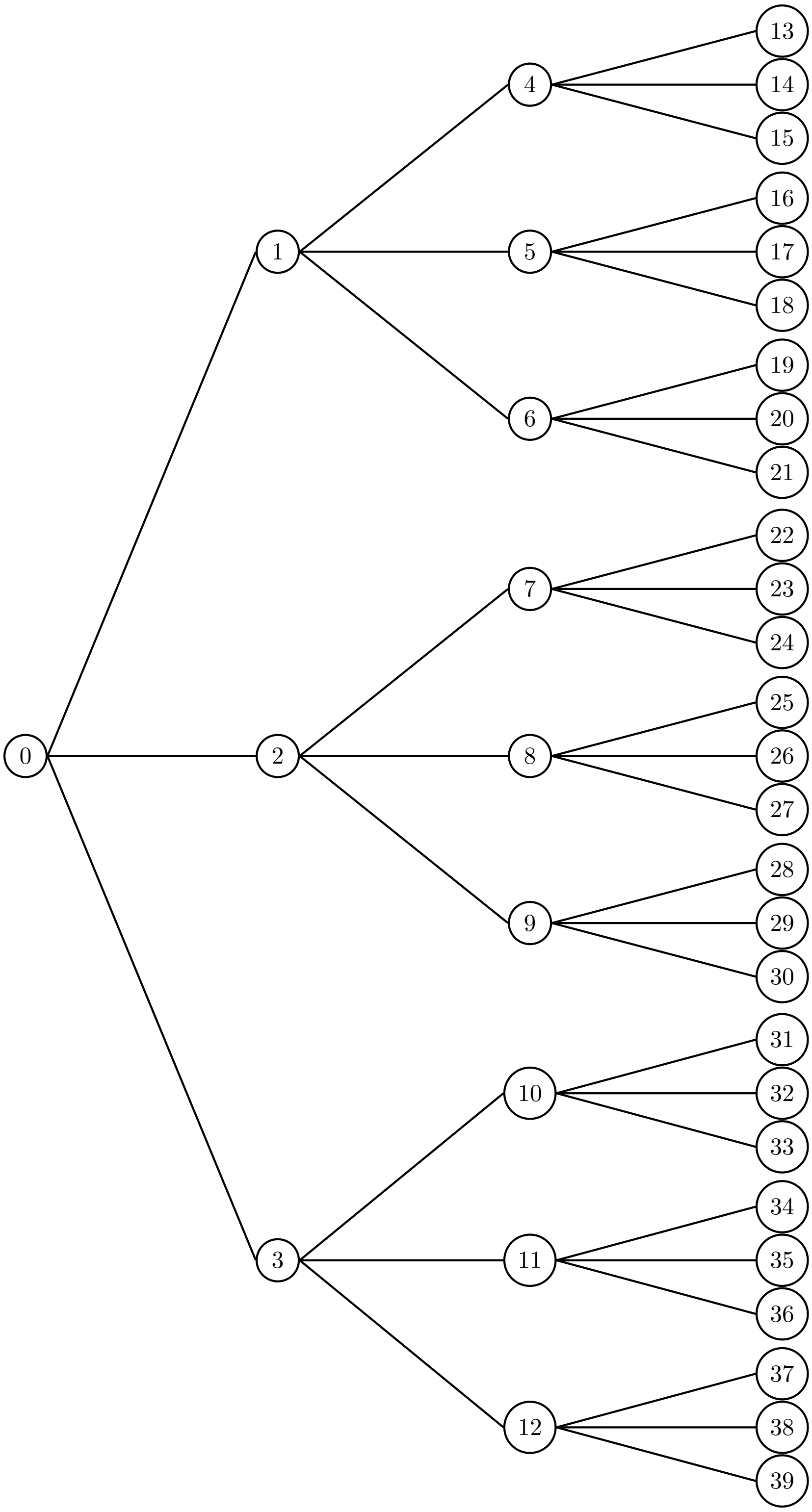

Figure B2.2: The scenario tree used in Instances 13-24

B2 Problem Instance Scales

Table B2.1: The scales of problem instances

Inst. No.	Inst. size			No. of variables			No. of constraints	No. of nonzeros
	$\lvert\mathcal{N}\rvert$	$\lvert\mathcal{C}\rvert$	$\lvert\mathcal{T}\rvert$	Binary	Integer	Total		
BP1	5	5	15	75	150	675	1576	5345
BP2	5	5	15	75	150	675	1576	5345
BP3	5	10	15	75	225	750	1726	7095
BP4	5	10	15	75	225	750	1726	7095
BP5	5	15	15	75	300	825	1876	8845
BP6	5	15	15	75	300	825	1876	8845
BP7	10	5	15	150	225	2025	5251	16540
BP8	10	5	15	150	225	2025	5251	16540
BP9	10	10	15	150	300	2100	5401	19890
BP10	10	10	15	150	300	2100	5401	19890
BP11	10	15	15	150	375	2175	5551	23240
BP12	10	15	15	150	375	2175	5551	23240
BP13	5	5	40	200	400	1800	4201	14545
BP14	5	5	40	200	400	1800	4201	14545
BP15	5	10	40	200	600	2000	4601	19495
BP16	5	10	40	200	600	2000	4601	19495
BP17	5	15	40	200	800	2200	5001	24445
BP18	5	15	40	200	800	2200	5001	24445
BP19	10	5	40	400	600	5400	14001	44690
BP20	10	5	40	400	600	5400	14001	44690
BP21	10	10	40	400	800	5600	14401	54190
BP22	10	10	40	400	800	5600	14401	54190
BP23	10	15	40	400	1000	5800	14801	63690
BP24	10	15	40	400	1000	5800	14801	63690

Bibliography

Abdallah, T., Diabat, A., and Simchi-Levi, D. (2012). Sustainable supply chain design: a closed-loop formulation and sensitivity analysis. *Production Planning & Control*, 23(2-3):120–133.

Agrawal, S., Singh, R. K., and Murtaza, Q. (2015). A literature review and perspectives in reverse logistics. *Resources, Conservation and Recycling*, 97:76–92.

Aguilera-Caracuel, J. and Guerrero-Villegas, J. (2018). How corporate social responsibility helps mnes to improve their reputation. the moderating effects of geographical diversification and operating in developing regions. *Corporate Social Responsibility and Environmental Management*, 25(4):355–372.

Ahiska, S. S., Gocer, F., and King, R. E. (2017). Heuristic inventory policies for a hybrid manufacturing/remanufacturing system with product substitution. *Computers & Industrial Engineering*, 114:206–222.

Apple Inc. (2020). iPhone 12 product environmental reports. Apple Inc.

Arts, J. (2014). Spare parts planning and control for maintenance operations. In *Second International Conference on Railway Technology: Research, Development and Maintenance (Railways 2014), April 8-11, 2014, Ajaccio, France*, pages 301–1. Civil-Comp Press.

Arts, J. (2017). A multi-item approach to repairable stocking and expediting in a fluctuating demand environment. *European Journal of Operational Research*, 256(1):102–115.

Asiedu, Y. and Gu, P. (1998). Product life cycle cost analysis: State of the art review. *International Journal of Production Research*, 36(4):883–908.

Axsäter, S. (2015). *Inventory Control*, volume 225. Springer.

Bacchetti, A., Plebani, F., Saccani, N., and Syntetos, A. (2013). Empirically-driven hierarchical classification of stock keeping units. *International Journal of Production Economics*, 143(2):263–274.

Barnhart, C., Johnson, E. L., Nemhauser, G. L., Savelsbergh, M. W., and Vance, P. H. (1998). Branch-and-price: Column generation for solving huge integer programs. *Operations research*, 46(3):316–329.

Basten, R. J. and van Houtum, G.-J. (2014). System-oriented inventory models for spare parts. *Surveys in Operations Research and Management Science*, 19(1):34–55.

Behfard, S., Al Hanbali, A., van der Heijden, M. C., and Zijm, W. H. (2018). Last time buy and repair decisions for fast moving parts. *International Journal of Production Economics*, 197:158–173.

Behfard, S., van der Heijden, M. C., Al Hanbali, A., and Zijm, W. H. (2015). Last time buy and repair decisions for spare parts. *European Journal of Operational Research*, 244(2):498–510.

Ben-Tal, A. and Den Hertog, D. (2014). Hidden conic quadratic representation of some nonconvex quadratic optimization problems. *Mathematical Programming*, 143(1):1–29.

Ben-Tal, A., Den Hertog, D., and Vial, J.-P. (2015). Deriving robust counterparts of nonlinear uncertain inequalities. *Mathematical Programming*, 149(1):265–299.

Ben-Tal, A., El Housni, O., and Goyal, V. (2020). A tractable approach for designing piecewise affine policies in two-stage adjustable robust optimization. *Mathematical Programming*, 182(1):57–102.

Ben-Tal, A., Goryashko, A., Guslitzer, E., and Nemirovski, A. (2004). Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376.

Bernstein, F., Kök, A. G., and Xie, L. (2011). The role of component commonality in product assortment decisions. *Manufacturing & Service Operations Management*, 13(2):261–270.

Bernstein, F., Kök, A. G., and Xie, L. (2015). Dynamic assortment customization with limited inventories. *Manufacturing & Service Operations Management*, 17(4):538–553.

Bertsimas, D. and Dunning, I. (2016). Multistage robust mixed-integer optimization with adaptive partitions. *Operations Research*, 64(4):980–998.

Bertsimas, D. and Georghiou, A. (2015). Design of near optimal decision rules in multistage adaptive mixed-integer optimization. *Operations Research*, 63(3):610–627.

Bertsimas, D. and Georghiou, A. (2018). Binary decision rules for multistage adaptive mixed-integer optimization. *Mathematical Programming*, 167(2):395–433.

Bertsimas, D. and Goyal, V. (2012). On the power and limitations of affine policies in two-stage adaptive optimization. *Mathematical Programming*, 134(2):491–531.

Bertsimas, D., Iancu, D. A., and Parrilo, P. A. (2010). Optimality of affine policies in multistage robust optimization. *Mathematics of Operations Research*, 35(2):363–394.

Bijvank, M., Koole, G., and Vis, I. F. (2010). Optimising a general repair kit problem with a service constraint. *European Journal of Operational Research*, 204(1):76–85.

Boylan, J. E. and Syntetos, A. A. (2010). Spare parts management: a review of forecasting research and extensions. *IMA Journal of Management Mathematics*, 21(3):227–237.

Cachon, G. P., Terwiesch, C., and Xu, Y. (2005). Retail assortment planning in the presence of consumer search. *Manufacturing & Service Operations Management*, 7(4):330–346.

Çapar, İ. (2013). Joint shipment consolidation and inventory decisions in a two-stage distribution system. *Computers & Industrial Engineering*, 66(4):1025–1035.

Caro, F. and Gallien, J. (2007). Dynamic assortment with demand learning for seasonal consumer goods. *Management science*, 53(2):276–292.

Caro, F., Martínez-de Albéniz, V., and Rusmevichientong, P. (2014). The assortment packing problem: Multiperiod assortment planning for short-lived products. *Management Science*, 60(11):2701–2721.

Carson, Y. and Maria, A. (1997). Simulation optimization: Methods and applications. In *Proceedings of the 29th Conference on Winter Simulation*, pages 118–126.

Chen, C.-H., Khoo, L. P., and Yan, W. (2006). An investigation into affective design using sorting technique and kohonen self-organising map. *Advances in Engineering Software*, 37(5):334–349.

Chen, H., Chiang, R. H., and Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, pages 1165–1188.

Cho, D. I. and Parlar, M. (1991). A survey of maintenance models for multi-unit systems. *European Journal of Operational Research*, 51(1):1–23.

Dedrick, J. and Kraemer, L. (2002). Dell computer: Organization of a global production network. *Center for Research on Information Technology and Organizations*, pages 1–17.

Diabat, A., Abdallah, T., and Henschel, A. (2015). A closed-loop location-inventory problem with spare parts consideration. *Computers & Operations Research*, 54:245–256.

Dreyfuss, M., Giat, Y., and Stulman, A. (2018). An analytical approach to determine the window fill rate in a repair shop with cannibalization. *Computers & Operations Research*, 98:13–23.

Du, C., Cooper, W. L., and Wang, Z. (2016). Optimal pricing for a multinomial logit choice model with network effects. *Operations Research*, 64(2):441–455.

Duran, O. and Afonso, P. S. L. P. (2019). An activity based costing decision model for life cycle economic assessment in spare parts logistic management. *International Journal of Production Economics*, 222:107499.

El Housni, O. and Topaloglu, H. (2021). Joint assortment optimization and customization under a mixture of multinomial logit models: On the value of personalized assortments. *Available at SSRN 3830082*.

Ellram, L. M. (1995). Total cost of ownership: An analysis approach for purchasing. *International Journal of Physical Distribution & Logistics Management*, 25(8):4–23.

Farley, A. A. (1990). A note on bounding a class of linear programming problems, including cutting stock problems. *Operations Research*, 38(5):922–923.

Feldman, J. and Topaloglu, H. (2015). Bounding optimal expected revenues for assortment optimization under mixtures of multinomial logits. *Production and Operations Management*, 24(10):1598–1620.

Frenk, J., Javadi, S., Pourakbar, M., and Sezer, S. O. (2019a). An exact static solution approach for the service parts end-of-life inventory problem. *European Journal of Operational Research*, 272(2):496–504.

Frenk, J., Javadi, S., and Sezer, S. O. (2019b). An optimal stopping approach for the end-of-life inventory problem. *Mathematical Methods of Operations Research*, 90(3):329–363.

Gallego, G. and Wang, R. (2014). Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities. *Operations Research*, 62(2):450–461.

Ghoniem, A. and Maddah, B. (2015). Integrated retail decisions with multiple selling periods and customer segments: optimization and insights. *Omega*, 55:38–52.

Guajardo, M. and Rönnqvist, M. (2015). Cost allocation in inventory pools of spare parts with service-differentiated demand classes. *International Journal of Production Research*, 53(1):220–237.

Guide, V. D. R. and Srivastava, R. (1997). Repairable inventory theory: Models and applications. *European Journal of Operational Research*, 102(1):1–20.

Hanley, D., Kelloway, C., and Vaheesan, S. (2020). Fixing America: Breaking manufacturers' aftermarket monopoly and restoring consumers' right to repair. Open Markets Institute.

Honhon, D., Gaur, V., and Seshadri, S. (2010). Assortment planning and inventory decisions under stockout-based substitution. *Operations Research*, 58(5):1364–1379.

Hu, Q., Boylan, J. E., Chen, H., and Labib, A. (2018). OR in spare parts management: A review. *European Journal of Operational Research*, 266(2):395–414.

Hu, Q., Chakhar, S., Siraj, S., and Labib, A. (2017). Spare parts classification in industrial manufacturing using the dominance-based rough set approach. *European Journal of Operational Research*, 262(3):1136–1163.

Huiskonen, J. (2001). Maintenance spare parts logistics: Special characteristics and strategic choices. *International Journal of Production Economics*, 71(1-3):125–133.

Hur, M., Keskin, B. B., and Schmidt, C. P. (2018). End-of-life inventory control of aircraft spare parts under performance based logistics. *International Journal of Production Economics*, 204:186–203.

Inderfurth, K. and Kleber, R. (2013). An advanced heuristic for multiple-option spare parts procurement after end-of-production. *Production and Operations Management*, 22(1):54–70.

Inderfurth, K. and Mukherjee, K. (2008). Decision support for spare parts acquisition in post product life cycle. *Central European Journal of Operations Research*, 16(1):17–42.

Jiang, H., Kwong, C., Kremer, G. O., and Park, W.-Y. (2019). Dynamic modelling of customer preferences for product design using denfis and opinion mining. *Advanced Engineering Informatics*, 42:100969.

Jin, T., Taboada, H., Espiritu, J., and Liao, H. (2017). Allocation of reliability–redundancy and spares inventory under poisson fleet expansion. *IISE Transactions*, 49(7):737–751.

Jin, T. and Tian, Y. (2012). Optimizing reliability and service parts logistics for a time-varying installed base. *European Journal of Operational Research*, 218(1):152–162.

Karsten, F. and Basten, R. J. (2014). Pooling of spare parts between multiple users: How to share the benefits? *European Journal of Operational Research*, 233(1):94–104.

Kennedy, W., Patterson, J. W., and Fredendall, L. D. (2002). An overview of recent literature on spare parts inventories. *International Journal of production economics*, 76(2):201–215.

Kök, A. G., Fisher, M. L., and Vaidyanathan, R. (2015). Assortment planning: Review of literature and industry practice. *Retail Supply Chain Management*, pages 175–236.

Kumar, S. and Putnam, V. (2008). Cradle to cradle: Reverse logistics strategies and opportunities across three industry sectors. *International Journal of Production Economics*, 115(2):305–315.

Lasdon, L. S. (2002). *Optimization theory for large systems*. Courier Corporation.

Li, H. and Huh, W. T. (2011). Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications. *Manufacturing & Service Operations Management*, 13(4):549–563.

Li, R. and Ryan, J. K. (2011). A bayesian inventory model using real-time condition monitoring information. *Production and Operations Management*, 20(5):754–771.

Li, T. and Calantone, R. J. (1998). The impact of market knowledge competence on new product advantage: conceptualization and empirical examination. *Journal of Marketing*, 62(4):13–29.

Li, Y., Zhao, Q., and He, S. (2018). Determination of the warranty service method switch time of old generation products after new generation products launched. *International Journal of Quality & Reliability Management*, 35(1):291–311.

Liao, S.-L., Shen, Y.-C., and Chu, C.-H. (2009). The effects of sales promotion strategy, product appeal and consumer traits on reminder impulse buying behaviour. *International Journal of Consumer Studies*, 33(3):274–284.

Lim, S. and Lee, B. (2015). Loyalty programs and dynamic consumer preference in online markets. *Decision Support Systems*, 78:104–112.

Limbourg, P. (2008). Dependability modelling under uncertainty. *Springer*, 117:118.

Liu, X. and Tang, L. C. (2016). Reliability analysis and spares provisioning for repairable systems with dependent failure processes and a time-varying installed base. *IIE Transactions*, 48(1):43–56.

López-Pérez, M. E., Melero, I., and Javier Sese, F. (2017). Management for sustainable development and its impact on firm value in the sme context: Does size matter? *Business Strategy and the Environment*, 26(7):985–999.

Mahajan, S. and Van Ryzin, G. (2001). Stocking retail assortments under dynamic consumer substitution. *Operations Research*, 49(3):334–351.

Martinez, V., Radnor, Z., Cavalieri, S., Gaiardelli, P., and Ierace, S. (2007). Aligning strategic profiles with operational metrics in after-sales service. *International Journal of Productivity and Performance Management*, 56(5/6):436–455.

Méndez-Díaz, I., Miranda-Bront, J. J., Vulcano, G., and Zabala, P. (2014). A branch-and-cut algorithm for the latent-class logit assortment problem. *Discrete Applied Mathematics*, 164:246–263.

Mirzahosseinian, H. and Piplani, R. (2011). A study of repairable parts inventory system operating under performance-based contract. *European Journal of Operational Research*, 214(2):256–261.

Mirzahosseinian, H., Piplani, R., and Jin, T. (2016). The impact of fleet size on performance-based incentive management. *Journal of the Operational Research Society*, 67(2):165–175.

Mo, D. Y., Tseng, M. M., and Wang, Y. (2016). Mass customizing spare parts support services based on response time with inventory pooling strategies. *IEEE Transactions on Engineering Management*, 63(3):305–315.

Moharana, U. and Sarmah, S. P. (2016). Determination of optimal order-up to level quantities for dependent spare parts using data mining. *Computers & Industrial Engineering*, 95:27–40.

Nagamachi, M. (2008). Perspectives and the new trend of kansei/affective engineering. *The TQM Journal*.

Nguyen, T. K., Yeung, T. G., and Castanier, B. (2013). Optimal maintenance and replacement decisions under technological change with consideration of spare parts inventories. *International Journal of Production Economics*, 143(2):472–477.

Oliva, R. and Kallenberg, R. (2003). Managing the transition from products to services. *International Journal of Service Industry Management*, 14(2):160–172.

Öner, K. B., Kiesmüller, G. P., and van Houtum, G.-J. (2010). Optimization of component reliability in the design phase of capital goods. *European Journal of Operational Research*, 205(3):615–624.

Öner, K. B., Scheller-Wolf, A., and Van Houtum, G.-J. (2013). Redundancy optimization for critical components in high-availability technical systems. *Operations Research*, 61(1):244–264.

Panagiotidou, S. (2014). Joint optimization of spare parts ordering and maintenance policies for multiple identical items subject to silent failures. *European Journal of Operational Research*, 235(1):300–314.

Paterson, C., Kiesmüller, G., Teunter, R., and Glazebrook, K. (2011). Inventory models with lateral transshipments: A review. *European Journal of Operational Research*, 210(2):125–136.

Pfohl, H.-C. and Ester, B. (1999). Benchmarking for spare parts logistics. *Benchmarking: An International Journal*, 6(1):22–45.

Pokharel, S. and Mutha, A. (2009). Perspectives in reverse logistics: A review. *Resources, Conservation and Recycling*, 53(4):175–182.

Postek, K. and Hertog, D. d. (2016). Multistage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set. *INFORMS Journal on Computing*, 28(3):553–574.

Pourakbar, M., Frenk, J., and Dekker, R. (2012). End-of-life inventory decisions for consumer electronics service parts. *Production and Operations Management*, 21(5):889–906.

Pourakbar, M., van der Laan, E., and Dekker, R. (2014). End-of-life inventory problem with phaseout returns. *Production and Operations Management*, 23(9):1561–1576.

Rezapour, S., Allen, J. K., and Mistree, F. (2016). Reliable flow in forward and after-sales supply chains considering propagated uncertainty. *Transportation Research Part E: Logistics and Transportation Review*, 93:409–436.

Rosa-Quino, P. (2020). Fix, or Toss? The 'Right to Repair' Movement Gains Ground.

Rusmevichientong, P., Shen, Z.-J. M., and Shmoys, D. B. (2010). Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations Research*, 58(6):1666–1680.

Rusmevichientong, P., Shmoys, D., Tong, C., and Topaloglu, H. (2014). Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management*, 23(11):2023–2039.

Ryzin, G. v. (2001). Analyzing inventory cost and service in supply chain.

Ryzin, G. v. and Mahajan, S. (1999). On the relationship between inventory costs and variety benefits in retail assortments. *Management Science*, 45(11):1496–1509.

Sahba, P., Balcioglu, B., and Banjevic, D. (2018). Multilevel rationing policy for spare parts when demand is state dependent. *OR Spectrum*, 40(3):751–780.

Sahyouni, K., Savaskan, R. C., and Daskin, M. S. (2010). The effect of lifetime buys on warranty repair operations. *Journal of the Operational Research Society*, 61(5):790–803.

Sauré, D. and Zeevi, A. (2013). Optimal dynamic assortment planning with demand learning. *Manufacturing & Service Operations Management*, 15(3):387–404.

Sethuraman, R., Gázquez-Abad, J. C., and Martínez-López, F. J. (2022). The effect of retail assortment size on perceptions, choice, and sales: Review and research directions. *Journal of Retailing*.

Shafiee, M. and Chukova, S. (2013). Maintenance models in warranty: A literature review. *European Journal of Operational Research*, 229(3):561–572.

Sherali, H. D. and Adams, W. P. (2013). *A reformulation-linearization technique for solving discrete and continuous nonconvex problems*, volume 31. Springer Science & Business Media.

Shi, Z. (2019). Optimal remanufacturing and acquisition decisions in warranty service considering part obsolescence. *Computers & Industrial Engineering*, 135:766–779.

Simonson, I. (1999). The effect of product assortment on buyer preferences. *Journal of Retailing*, 75(3):347–370.

Solomon, R., Sandborn, P. A., and Pecht, M. G. (2000). Electronic part life cycle concepts and obsolescence forecasting. *IEEE Transactions on Components and Packaging Technologies*, 23(4):707–717.

Song, Y. and Yang, J.-h. (2015). Simulation of dynamics behaviors for shipping equipment support with system dynamics analysis approach. *Journal of Industrial Engineering and Management*, 8(3):636–657.

Souza, G. C. (2014). Supply chain analytics. *Business Horizons*, 57(5):595–605.

Talluri, K. and Van Ryzin, G. (2004). Revenue management under a general discrete choice model of consumer behavior. *Management Science*, 50(1):15–33.

Tiacci, L. and Saetta, S. (2011). Reducing the mean supply delay of spare parts using lateral transshipments policies. *International Journal of Production Economics*, 133(1):182–191.

Togwe, T., Eveleigh, T. J., and Tanju, B. (2019). An additive manufacturing spare parts inventory model for an aviation use case. *Engineering Management Journal*, 31(1):69–80.

Topan, E. and Bayindir, Z. P. (2012). Multi-item two-echelon spare parts inventory control problem with batch ordering in the central warehouse under compound poisson demand. *Journal of the Operational Research Society*, 63(8):1143–1152.

Topan, E., Bayındır, Z. P., and Tan, T. (2017). Heuristics for multi-item two-echelon spare parts inventory control subject to aggregate and individual service measures. *European Journal of Operational Research*, 256(1):126–138.

Train, K. E. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

Tripsas, M. (2008). Customer preference discontinuities: A trigger for radical technological change. *Managerial and Decision Economics*, 29(2-3):79–97.

Turan, H. H., Sleptchenko, A., Pokharel, S., and ElMekkawy, T. Y. (2018). A clustering-based repair shop design for repairable spare part supply systems. *Computers & Industrial Engineering*, 125:232–244.

Ukko, J., Saunila, M., Rantala, T., and Havukainen, J. (2019). Sustainable development: Implications and definition for open sustainability. *Sustainable Development*, 27(3):321–336.

Ulu, C., Honhon, D., and Alptekinoğlu, A. (2012). Learning consumer tastes through dynamic assortments. *Operations Research*, 60(4):833–849.

Van der Auweraer, S. and Boute, R. (2019). Forecasting spare part demand using service maintenance information. *International Journal of Production Economics*, 213:138–149.

Van Horenbeek, A., Buré, J., Cattrysse, D., Pintelon, L., and Vansteenwegen, P. (2013). Joint maintenance and inventory optimization systems: A review. *International Journal of Production Economics*, 143(2):499–508.

van Jaarsveld, W., Dollevoet, T., and Dekker, R. (2015). Improving spare parts inventory control at a repair shop. *Omega*, 57:217–229.

van Wijk, A., Adan, I. J., and Van Houtum, G. (2019). Optimal lateral transshipment policies for a two location inventory problem with multiple demand classes. *European Journal of Operational Research*, 272(2):481–495.

van Wingerden, E., Basten, R. J. I., Dekker, R., and Rustenburg, W. (2014). More grip on inventory control through improved forecasting: A comparative study at three companies. *International Journal of Production Economics*, 157:220–237.

Vanderbeck, F. and Wolsey, L. A. (1996). An exact algorithm for ip column generation. *Operations Research Letters*, 19(4):151–159.

Wang, G., Gunasekaran, A., and Ngai, Eric W.T.and Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, 176:98–110.

Wang, R. (2012a). Capacitated assortment and price optimization under the multinomial logit model. *Operations Research Letters*, 40(6):492–497.

Wang, W. (2012b). A stochastic model for joint spare parts inventory and planned maintenance optimisation. *European Journal of Operational Research*, 216(1):127–139.

Wang, W. and Yue, S. (2015). An inventory pooling model for spare units of critical systems that serve multi-companies. *Transportation Research Part E: Logistics and Transportation Review*, 76:34–44.

Wolsey, L. A. and Nemhauser, G. L. (1999). *Integer and combinatorial optimization*, volume 55. John Wiley & Sons.

Wong, H., van Houtum, G.-J., Cattrysse, D., and Van Oudheusden, D. (2006). Multi-item spare parts systems with lateral transshipments and waiting time constraints. *European Journal of Operational Research*, 171(3):1071–1093.

Wong, H., Van Oudheusden, D., and Cattrysse, D. (2007). Cost allocation in spare parts inventory pooling. *Transportation Research Part E: Logistics and Transportation Review*, 43(4):370–386.

Yanıkoğlu, İ., Gorissen, B. L., and den Hertog, D. (2019). A survey of adjustable robust optimization. *European Journal of Operational Research*, 277(3):799–813.

You, H., Ryu, T., Oh, K., Yun, M.-H., and Kim, K.-J. (2006). Development of customer satisfaction models for automotive interior materials. *International journal of Industrial Ergonomics*, 36(4):323–330.

Zhao, Q., Chang, R., Ma, J., and Wu, C. (2019). System dynamics simulation-based model for coordination of a three-level spare parts supply chain. *International Transactions in Operational Research*, 26(6):2152–2178.

Zheng, M. and Wu, K. (2017). Smart spare parts management systems in semiconductor manufacturing. *Industrial Management & Data Systems*, 117(4):754–763.

Zhu, S., Dekker, R., Van Jaarsveld, W., Renjie, R. W., and Koning, A. J. (2017). An improved method for forecasting spare parts demand using extreme value theory. *European Journal of Operational Research*, 261(1):169–181.

Zipkin, P. H. (2000). *Foundations of Inventory Management*. Boston: McGraw-Hill.