공학박사학위논문

# Sharing Logistics Resources in the E-commerce Supply Chain under Uncertainty

## 불확실성을 고려한 이커머스 공급망에서의 물류 자원 공유

2023 년 8 월

서울대학교 대학원

산업공학과

이 준 혁

# Sharing Logistics Resources in the E-commerce Supply Chain under Uncertainty

## 불확실성을 고려한 이커머스 공급망에서의 물류 자원 공유

지도교수 문 일 경

이 논문을 공학박사 학위논문으로 제출함

2023 년 6 월


서울대학교 대학원

산업공학과

이 준 혁


이준혁의 공학박사 학위논문을 인준함

2023 년 7 월

| | | |
|---|---|---|
| 위 원 장 | 이 재 욱 | (인) |
| 부위원장 | 문 일 경 | (인) |
| 위 원 | 박 건 수 | (인) |
| 위 원 | 장 윤 석 | (인) |
| 위 원 | 김 병 수 | (인) |

# Abstract

# Sharing Logistics Resources in the E-commerce Supply Chain under Uncertainty

Junhyeok Lee
Department of Industrial Engineering
The Graduate School
Seoul National University

With the growth of both communications technology and contact-free delivery demand, e-commerce has grown significantly during the past few years. However, fierce competition and the inherent uncertainty in the e-commerce marketplace have made retailers suffer from high operations costs. Because of such circumstances, the concept of the sharing economy has been confirmed as an innovative business model that can answer the need for more flexible logistics. Therefore, we aim to develop decision-making frameworks considering logistics resources sharing under uncertainty. In this thesis, we address three problems in the supply chain management field: (1) the perishable inventory problem, (2) the supply chain network design, and (3) the omnichannel retail operations. In addition, we consider the three different services or strategies to share logistic resources in the abovementioned problems.

First, we address the perishable inventory problem considering transshipment and online-offline channel system. We present a Markov decision process model by

accommodating key attributes of the online-offline channel system. We develop the hybrid deep reinforcement learning algorithm based on the soft actor-critic algorithm to overcome the curse of dimensionality in the large-scale Markov decision process. In addition, we found that transshipment substantially reduces the outdating cost by allowing the offline channel to make good use of the old products that will be discarded in the online channel, which is new to the literature.

Second, we propose the supply chain network design problem considering the on-demand warehousing system. We propose the two-stage stochastic programming model that captures the inherent uncertainties to formulate the presented problem. We solve the proposed model utilizing Sample average approximation combined with the Benders decomposition algorithm. Of particular note, we develop a method to generate effective initial cuts for improving the convergence speed of the Benders decomposition algorithm.

Third, we address the omnichannel retail operations considering the third-party platform channel. We propose the stochastic optimization model considering both the retailer's supply chain and the third-party platform's supply chain for omnichannel retail operations. To tackle the intractability of the stochastic optimization model, we propose a decomposition approach based on the two-phase robust optimization approach. The experimental results suggest that a decomposition approach is scalable to large-scale problems while maintaining its high solution quality.

**Keywords**: Logistics resources sharing, E-commerce, Supply chain management, Reinforcement learning, Stochastic programming, Robust optimization

**Student Number**: 2021-34229

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

## 1.1 Motivations

With the growth of both communications technology and contact-free delivery demand, e-commerce has grown significantly during the past few years [55]. Because of the rapid development of e-commerce marketplaces, the number of *e-commerce retailers* selling products online also has increased.

In the e-commerce marketplaces, retailers can be categorized into two types depending on the size of the business: *large retailers* and *small retailers*. We use the term "large retailers" to refer to retail companies operated with large capital investments. In particular, e-commerce platform companies (e.g., *Amazon*, *Kurly*, and *Coupang*) and the omnichannel company (e.g., *Nike* and *Adidas*) can be categorized as large retailers. These retailers usually have their own e-commerce sales channels (e.g., mobile apps or official online websites) or offline sales channels (e.g., offline stores). In addition, they operate their own logistics system with abundant warehouse space.

On the other hand, the term "small retailers" will be used to indicate e-commerce retailers who run their businesses with low capital investment. Because small retailers have insufficient capital to invest in building their own sales channels, they usu-

ally participate in e-commerce platforms as third-party sellers to sell their products. They operate warehouses with small spaces and also rely on third-party logistics companies to deliver their products to customers. We summarize the differences between large and small retailers in Table 1.1.

Table 1.1: Differences between large and small retailers

| | **Large retailer** | **Small retailer** |
|---|---|---|
| **Size of the business** | Large capital investment | Low capital investment |
| **Sales channels** | Operate their own sales channels (offline stores, apps, online website) | Participate in e-commerce platforms as a third-party seller |
| **Logistics system** | Operate their own logistics system | Rely on third-party delivery service |
| **Warehouse space** | Abundant | Insufficient |
| **Example** | E-commerce platform company, Omnichannel company | Small businesses |

However, both types of retailers have several challenges to operate and run their e-commerce businesses. First, several large retailers in South Korea provide delivery and logistics services for fresh foods because of the ongoing increase in demand. To preserve the freshness of food, they offer a unique delivery service: *dawn deliveries*, which are guaranteed to arrive early morning for orders placed at midnight [69]. However, these large retailers have been suffering from large net losses, and some have stopped providing dawn delivery services [86]. A major factor contributing to the net losses of these companies is the perishability of fresh foods. That is, it is challenging to manage perishable inventories and the substantial outdating cost incurred due to the waste of perishable products.

Small retailers usually operate warehouses with small spaces for dealing with

varying demands [40, 107]. In order to manage the problem of small warehouse capacity, retailers have several older solutions at their disposal [136]. First, retailers can build new warehouses or infrastructure to expand capacity. However, a lot of capital investment is necessary to implement this solution. Another solution is to lease warehouses from traditional warehouse operators or third-party logistics providers. The contract duration for leasing space from traditional providers is usually long and requires a long-term contract. Therefore, this way is not a suitable strategy for e-commerce sellers who need flexible solutions [29].

Because of such circumstances, the concept of the *sharing economy* has been confirmed as an innovative business model that can answer the need for more flexible logistics [28]. From the standpoint of e-commerce retailers, this thesis develops stochastic decision-making frameworks aiming to minimize the expected cost by sharing logistics resources. Among various logistics resources, we have focused on three types of resources: 1) inventory, 2) warehouse space, and 3) sales channel. In Section 1.2, we will briefly explain the concept of logistics resource sharing and related services. In Section 1.3, we introduce three optimization approaches utilized to develop decision-making frameworks considering uncertainty.

## 1.2 Sharing logistics resources

The sharing economy refers to a system in which businesses or customers temporarily share, rent, or borrow resources instead of buying and owning them. This system enables participants to reduce risk, enhance flexibility, and minimize operations costs. The best-known examples of sharing economy models are *Airbnb*, where private individuals can rent their apartments to others, and *Uber*, where an online peer-to-peer

ridesharing service allows people to rent a ride. In addition, logistics practitioners have started to embrace a sharing economy with supply chain collaborations to bring efficiency to fulfillment and delivery services.

As mentioned in Section 1.1, this thesis studies how logistics resources (i.e., inventory, warehouse space, and sales channel) should be shared to minimize operations costs from the perspective of retailers. First, *transshipment* can share inventory between locations of the same echelon to improve the service level [106]. Transshipment can be categorized into two types depending on the timing of implementation: *reactive transshipment* and *proactive transshipment*. The reactive transshipment takes place after observing the demand but before it must be satisfied. On the other hand, proactive transshipment occurs before the demand has been realized [46]. For example, Kranenburg et al. [82] shows that a semi-conductor company, *ASML*, could save up to 50 percent of annual inventory costs by utilizing transshipment.

Second, an *on-demand warehousing platform* has emerged as a new alternative to share warehouse space because of its high flexibility and low risk. In real cases, the platform *FLEXE* provides service for on-demand warehousing in the global market [67]. This platform connects warehouse providers who have excess spaces with e-commerce retailers who require empty space for the short term, as presented in Figure 1.1. There are three advantages of utilizing an on-demand warehousing system, called *ODWS*, as follows: (1) saving setup cost, (2) high flexibility, and (3) high-speed delivery [84]. From the standpoint of the e-commerce retailer, the main advantage of the ODWS is that short-term rent for warehouses is available [135].

Third, retailers can sell their products by utilizing sales channels of e-commerce platform companies, such as *Amazon* and *Coupang*, even when they have their own


Figure 1.1: On-demand warehousing system.

offline and online channels [151]. For example, Coupang launched a service called the *C.AVENUE*, and many omnichannel companies, such as *Nike* and *Adidas*, have participated in this service to sell their products. From the perspective of retailers, there are distinct advantages to adopting the sales channel of the *third-party platform* (3PP) as one of their sales channels. First, the 3PP companies could implement logistics of fulfillment on behalf of the retailer by using their self-supporting logistics service system. Second, the retailer could absorb the additional demand of 3PP. A significant number of customers use 3PP to buy products online. Therefore, in addition to customers who want to buy a retailer's specific product, other users of 3PP could also buy that product while looking around the platform.

## 1.3 Optimization approaches under uncertainty

This thesis utilizes stochastic programming, robust optimization (RO), and reinforcement learning (RL) to deal with problems of making decisions in the presence of different forms of uncertainty, and we briefly introduce the principles of the adopted approaches.

*Stochastic programming* evolves from deterministic linear programming with the adoption of random variables and aims to minimize the expected cost [112]. Depending on the sequences of decisions and information, the research area can be categorized into two types: 1) two-stage stochastic programming (TSSP) and 2) multistage stochastic programming. This thesis concentrates on the first one. TSSP determines the first-stage decisions before the realization of uncertain parameters. Subsequently, the second-stage decisions are determined after the realization of uncertain parameters. However, there is a fundamental difficulty of TSSP in computing the expected value in the objective function. To alleviate this issue, a great deal of research has utilized scenario-based models, in which the problem can be solved as a deterministic model. Because the computational complexity increases depending on the size of scenarios, the approximation method (e.g., Sample average approximation (SAA) [79]) or the decomposition method (e.g., Benders decomposition (BD) [13]) has attracted the interest of many researchers.

*Robust optimization* is one of the approaches that deals with the uncertainty in optimization problems. In contrast to stochastic programming, RO does not need any knowledge about the probability distribution; instead, it assumes that the uncertainty value belongs within a predetermined set, called the *uncertainty set*. RO aims to find the optimal solution under the worst-case scenario, and the obtained solution should be guaranteed to be feasible for any realizations of uncertain parameters in the uncertainty set [10]. In order to make the model tractable, the uncertainty set is generally defined as a convex set (e.g., box [128], ellipsoid [12], and polyhedron[20]). In recent decades, adjustable robust optimization (ARO) has been widely used to deal with multi-stage problems, which commonly assume the

multi-period setting and consider adjustable variables to implement the wait-and-see decision. The wait-and-see decision is less conservative than the here-and-now decision because it can postpone making decisions until some of the uncertain parameters are revealed. However, it is generally intractable to deal with wait-and-see decisions because of the large feasible space. Thus, it is typical to restrict feasible space by optimizing a certain type of parameterized function called the decision rule.

*Reinforcement learning* is one of the tools used to solve the large-scale and complex Markov decision process (MDP). The MDP is a tuple $\langle \mathcal{S}, \mathcal{A}, r, p, \gamma \rangle$, consisting of five components— a set of states, $\mathcal{S}$; a set of actions, $\mathcal{A}$; the reward function, $r$; the state transition probability function, $p$; and the discount factor, $\gamma \in [0, 1)$. The MDP is specialized to solve the sequential decision-making problem, and the sequencing of decisions and information is implemented as follows: decision$\rightarrow$ information $\rightarrow$ decision $\rightarrow$ information $\cdots$. It can be applied to finite horizon problems and also applied to infinite horizon problems. Dynamic programming has been utilized to derive the optimal policy by solving the MDP. However, because of the curse of dimensionality and the curse of modeling, dynamic programming is generally difficult to apply in large-scale and complex MDP [56]. In contrast to dynamic programming, Q-learning could address complex MDP and alleviate the above two issues by approximating the optimal action value [144]. Recently, deep reinforcement learning (DRL), which uses a neural network for a function approximation in RL, has been widely utilized for various domains [23]. The capability of DRL lies in its ability to solve the large-scale MDP, which involves the large dimension of state and action, near optimally.

## 1.4 Contributions

In this thesis, we address three different problems: (1) the perishable inventory problem, (2) the supply chain network design (SCND) problem, and (3) the omnichannel retail operations problem. For each problem, we aim to develop decision-making frameworks considering logistics resources sharing under uncertainty, which can minimize the total expected cost. The proposed problems are closed related in terms of *objective*, *domain*, and *uncertainty* as follows:

- *Objective*: The proposed problems aim to balance the service and inventory levels by considering their trade-offs and minimizing the supply chain cost.
- *Domain*: The proposed problems cope with the application of emerging services, such as ODWS and 3PP, into the e-commerce supply chain.
- *Uncertainty*: The proposed problems address uncertainty in a multi-period situation. Among various approaches dealing with uncertainty in a multi-period planning model, we utilize the appropriate approach for each problem.

The main contributions of this thesis are summarized as follows:

1. For the perishable inventory problem considering proactive transshipment:
   - We propose the MDP model to derive a transshipment policy for perishable products in the online-offline channel system (OOCS). Furthermore, we accommodate key attributes of the OOCS in the mathematical model.
   - To derive a promising policy without assumptions about demand distribution, we customize the soft actor-critic (SAC) algorithm, which is one of the state-of-the-art DRL algorithms, for the proposed problem.

- We observe that the SAC algorithm is unstable during the training process due to large action spaces. To mitigate the issue, we propose a novel hybrid DRL approach by developing two acceleration methods.
- We examine the tendency for transshipment to be effective in high demand variability. In addition, we analyze the impact of a unit transshipment cost parameter shelf life of online and offline channels through a sensitivity analysis.
- By further analyzing the outdating cost regarding each channel, we discovered that the old product discarded in the online channel could be reutilized in the offline channel through transshipment

2. For the SCND problem considering ODWS:
   - We propose the TSSP model for an e-commerce SCND with the ODWS under uncertainties. To estimate the expected function in the proposed model, we employ the SAA method.
   - To alleviate the computational burden in SAA, we utilize the BD algorithm. Furthermore, we develop the acceleration method for improving the convergence of bounds by focusing on the initial iteration in the BD algorithm.
   - We show the potential cost-saving effects of using the ODWS in the supply chain through computational experiments.
3. For the omnichannel retail operations considering the 3PP sales channel:
   - We develop the stochastic optimization model addressing both the retailer' s supply chain and the 3PP supply chain for omnichannel retail operations. Fur-

thermore, we deal with the production capacity of suppliers and transshipment between logistics centers.

- We propose a novel decomposition approach (DECOM) based on the two-phase robust optimization approach (TPA), which is the state-of-the-art method to deal with adjustable binary decisions [93]. By introducing artificial variables, we decompose the total supply chain into two streams, one for the retailer' s supply chain and the other for the 3PP supply chain.

- The experimental results suggest that DECOM is scalable to large-scale problems while maintaining its high solution quality. Finally, even though the production capacity becomes insufficient, the computation time of DECOM does not increase significantly compared to that of the TPA.

## 1.5 Outline of the thesis

The e-commerce supply chain consists of four components, as shown in Figure 1.2: suppliers, distribution centers, offline stores, and demand zones. In the e-commerce supply chain, a retailer replenishes the inventories from multiple suppliers. The inventories are stored at distribution centers to fulfill demand from multiple sales channels. Usually, there are two sales channels in the e-commerce supply chain, one for an offline channel and the other is an online channel. For an offline channel, the inventories are allocated to offline stores from distribution centers. The offline demand is fulfilled by the on-hand inventories held in offline stores. For an online channel, products are delivered directly from distribution centers to customers (i.e., online demand). Considering the distance between locations of suppliers, distribution

centers, offline stores, and demand zones, it is significant to make replenishment, allocation, and fulfillment decisions to minimize the operational cost in the proposed e-commerce supply chain.

This thesis studies three types of e-commerce supply chain problems considering sharing logistics resources and demand uncertainty. However, because of the complicated structure of the e-commerce supply chain, we determine the scope of the study differently for each chapter. In Chapter 2, we propose the perishable inventory management problem considering proactive transshipment and the OOCS. To utilize real-world data without any knowledge about demand distribution, we develop the hybrid DRL approach based on the SAC algorithm. In this study, we only consider the replenishment and transshipment decisions for online and offline channels. In addition, we assume a single supplier, and the locations of four components are not considered. Therefore, this chapter addresses a small part of the e-commerce supply chain, and the scope of the study is depicted as the orange region in Figure 1.2.

In Chapter 3, we present the e-commerce SCND problem considering ODWS by using the TSSP model. We propose the method that was developed by combining SAA and BD algorithms. In particular, we develop a novel acceleration method to improve the convergence speed of the BD algorithm. The scope of the study is shown as the blue region in Figure 1.2. We consider multiple suppliers, and the supply chain for an online channel is considered. However, we do not address the locations of online demand zones and consider the aggregated customer demand to reflect the case of the e-commerce market in South Korea.

In Chapter 4, we address the robust omnichannel retailing problem considering

the 3PP channel. We present a multi-period stochastic optimization model to address uncertainty. Also, we propose the DECOM approach, which could be scalable to various problem instances. This study considers the whole supply chain proposed in Figure 1.2, and the scope of the study is depicted as the green region. This study addresses replenishment, allocation, and fulfillment decisions and also reflects the locations of four components to minimize operational costs. The proposed problem is the most complicated compared to problems in Chapters 2 and 3. However, the problem is the most related to real practice; thus, this study could be instructive to practitioners who are concerned about setting up an effective e-commerce supply chain. In Chapter 5, we provide concluding remarks and possible future research directions of this thesis.

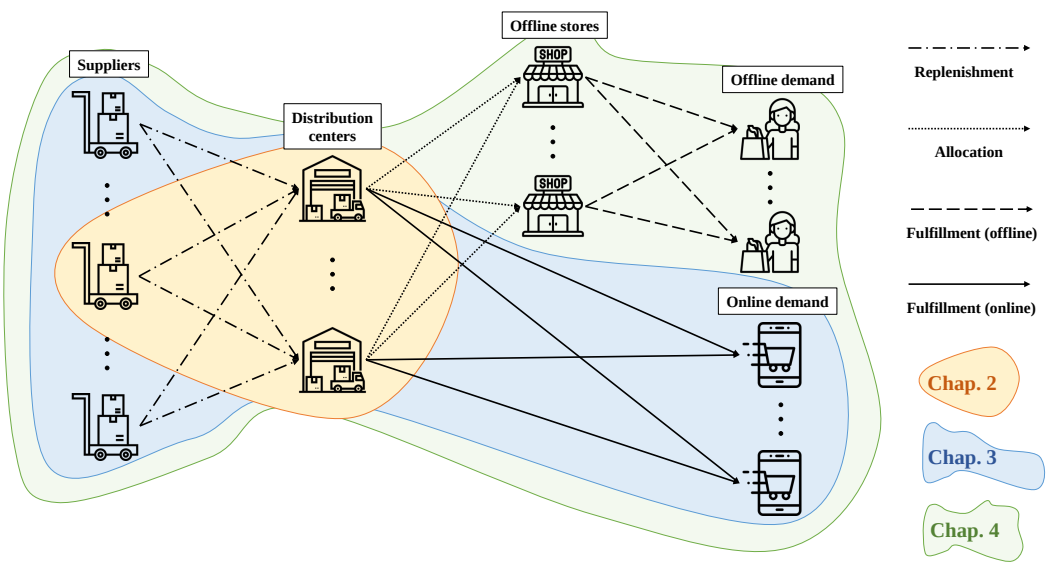



Figure 1.2: E-commerce supply chain and the scope of studies.

# Chapter 2

# A hybrid deep reinforcement learning approach for a proactive transshipment of fresh food in the online-offline channel system

## 2.1 Introduction

In recent years, as customers have become health conscious, the quality of grocery service and the provision of fresh foods has only grown in importance [90]. However, companies providing delivery and logistics services for fresh foods have been suffering from large net losses, and some have stopped providing dawn delivery services [86]. Unlike the other companies, *Oasis Market* is the only company that stays in the black [75]. The majority of companies strive to increase profitability by developing their online platforms, an outlet for selling fresh foods, as well as by improving logistics facilities, such as the cold chain system. Conversely, Oasis Market invests in both online platforms and offline shops simultaneously and has effectively connected the networks. We will use the term *OOCS* to refer to a network of online platforms and offline shops in this thesis. Oasis Market aims to reduce the risk of excess ordering by mutually transshipping leftover products between channels.

Even though the OOCS has achieved success in real practice by Oasis Market, there is room for further study on this system. As far as we know, Oasis Market only

addresses the movement of products from online to offline channels, not from offline to online channels (i.e., *unidirectional transshipment*) and implements transshipment in a simple manner in which the online channel's unsold products on that day are moved to the offline channel. Based on the above considerations, we further studied the OOCS to meet the following two goals. First, in order to minimize the operational costs in the OOCS, it is necessary to develop a method that could derive a near-optimal in policy determining the transshipment quantities. Second, we aim to study the *mutual transshipment* of products between online and offline channels.

The OOCS is related to research areas of lateral transshipment for perishable products. It is necessary to consider the key attributes of the corresponding business model in order to adopt lateral transshipment in OOCS. Before discussing these attributes, we would like to briefly discuss two types of lateral transshipment (i.e., reactive transshipment and proactive transshipment) and clarify the meaning of the term *shelf life*. The reactive transshipment takes place after observing the demand but before it must be satisfied. On the other hand, the proactive transshipment occurs before the demand has been realized [106, 46]. In this chapter, the standard meaning of *shelf life* is used to indicate the length of periods for which fresh food can be sold. For instance, as illustrated in Figure 2.1, fresh food that has been aged more than four days should be considered outdated because the shelf life of the food is three days in the online channel because of delivery time. On the other hand, the offline channel can offer a longer shelf life (i.e., four days) for fresh food since delivery time does not need to be considered.

The three attributes of OOCS with lateral transshipment are as follows:

- *Heterogeneous shelf life:* Throughout the chapter, we use the term *heteroge-*

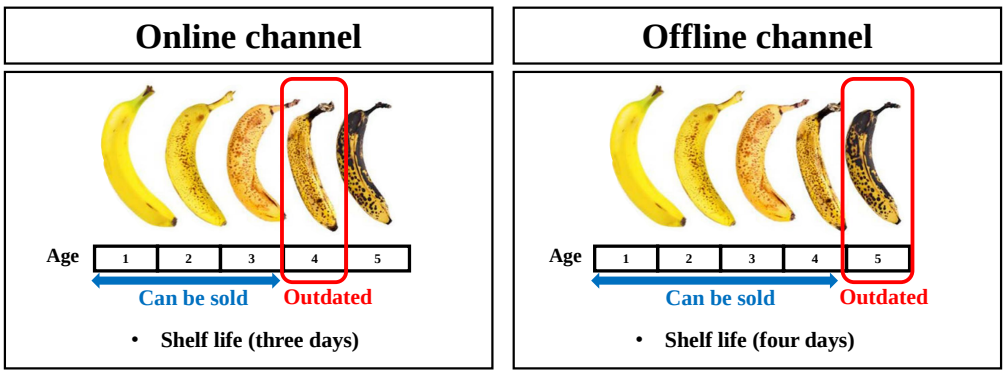



Figure 2.1: Example of heterogeneous shelf life property in OOCS (Photo courtesy of *davidwolfe.com*).

*neous shelf life* to refer to the property that the shelf life of fresh food is different depending on the channel where it is stored. Even though the lifetime of fresh food is homogeneous, the shelf life of fresh food can be heterogeneous depending on the channel [70]. Based on the practice of *Oasis Market*, the shelf life of fresh food in the online channel is shorter than that of fresh food in the offline channel because of the risk that the product can deteriorate during delivery as presented in Figure 2.1. In a practical manner, if fresh food is unsold in the online channel until the end of its shelf life, it is usually transshipped to the offline channel and then resold at a discount.

- *Proactive transshipment:* When the fresh food that customers want to purchase is out of stock in one retail store, they can easily purchase the same kind of fresh food in another store because fresh foods are sold in many stores. Consequently, customers are not apt to wait until the transshipped item arrives, making reactive transshipment inappropriate for fresh food operations.

- *Non-negligible transshipment time:* It is common for the online distribution

center to be located a distance away from the retail store. Typically, an online distribution center is located in a suburban area, whereas an offline retail store is located downtown so that it is more convenient for customers to visit. Moreover, it requires several hours to package and handle inventories that will be transshipped to the other channel. Therefore, the transshipment time should not be negligible when considering inventory movements between online and offline channels.

Although lateral transshipment has been widely studied in the operations management field, a limited number of studies have taken into account perishable products. Depending on the types of lateral transshipment, several researchers focused on reactive transshipment for perishable products [101, 37, 149], and others concentrated on the proactive transshipment [90, 32, 38]. Although previous studies have addressed lateral transshipment with perishable products in various aspects, there are two research gaps in existing studies. The first research gap is that no existing study has simultaneously addressed the three attributes of OOCS. To the best of our knowledge, this is the first study to attempt to analyze heterogeneous shelf lives. A second research gap relates to assumptions about demand distribution Existing studies primarily focus on a specific demand distribution in order to determine the optimal policy except Dehghani et al. [38]. If the gap between real demand and estimated distribution is large, the transshipment policy derived from relying on the estimated demand distribution could show poor performance in real practice. The distinctive differences between our study and Dehghani et al. [38] will be presented in Section 2.2.2.

In order to fill the above two research gaps, our study deals with a lateral trans-

shipment model of fresh foods in the OOCS by accommodating three attributes. Without any assumptions on the demand distribution, we develop the hybrid DRL approach to directly utilize real-world demand data for deriving a promising transshipment policy. We aim to answer the following three research questions through this study:

1. Can the developed hybrid DRL approach increase average profit compared to existing methods in the setting of real-world demand data?

2. What positive effects does transshipment have in terms of profit depending on the variability of demand and value of unit transshipment cost?

3. How does a transshipment policy change the outdating cost of online and offline channels, respectively, compared to a 'no-transshipment' policy?

The main contributions of our research can be summarized as fourfold. First, we propose the MDP model to derive a transshipment policy for fresh foods in the OOCS. Furthermore, we accommodate key attributes of the OOCS in the mathematical model. Second, to derive a promising policy without assumptions about demand distribution, we customize the SAC algorithm, which is one of the state-of-the-art DRL algorithms, for the proposed problem. However, we observe that the SAC algorithm is unstable during the training process due to large action spaces. To mitigate the issue, we propose a novel hybrid DRL approach by developing two acceleration methods. Third, we examine the tendency for transshipment to be effective in high demand variability by conducting computational experiments on various demand data sets. In addition, we analyze the impact of a unit transshipment cost parameter shelf life of online and offline channels through a sensitivity

analysis. Fourth, by further analyzing the outdating cost regarding each channel, we discovered that the old product discarded in the online channel can be reutilized in the offline channel through transshipment.

## 2.2 Literature review

The literature review will focus on two streams of research in operations management: lateral transshipment on perishable inventory management and an RL approach for inventory problems.

### 2.2.1 Overview of perishable inventory management and lateral transshipment

Prior to reviewing the relevant literature, we will provide a brief overview of perishable inventory management and lateral transshipment. It is notoriously difficult to determine an optimal policy for managing perishable inventory because of the short shelf life of the products $M$. The shelf life of the product incurs $M$ dimensional state space, and the positive lead time $L$ makes the problem even more challenging, making the state space dimension $M + L - 1$. Therefore, the inventory model becomes intractable to address with the traditional dynamic programming approach.

It has been shown that several optimal inventory policies can be obtained for an asymptotic case of shelf life based on previous studies [4, 25]. Moreover, approximate ordering policies (e.g., the base stock policy based on stock level levels) have also been developed due to the complexity of the optimal policy, and these approximate policies have been generally evaluated by simulation studies [100, 34, 62]. There have been several methods developed for determining ordering policies that consider in-

formation about the age of units in stock and pipeline inventory [65, 24, 39]. A modified base-stock policy has been proposed by Haijema and Minner [64], in which waste estimation is taken into account in the base-stock policy. Several computational experiments were conducted to compare the performance of existing ordering policies with those that had been developed afresh. The authors conducted computational experiments to compare the performance of existing ordering policies and newly developed ordering policies. In this study, we adopted the best-performing ordering policies presented in Haijema and Minner [64] to evaluate the performance of the developed hybrid DRL approach in our study.

A lateral transshipment can be defined as the movement of inventory between several locations of the same echelon, which can result in a reduction in surplus and shortfall at the different locations by transferring inventory between them [91, 143]. In spite of the complexity of modeling proactive transshipment over reactive transshipment, several studies have focused on developing proactive transshipment models and promising policies [3, 87, 26, 1, 97]. The majority of studies, except Tagaras and Vlachos [131], assumed that transshipment time was negligible in order to make the problem tractable. In addition, Tagaras and Vlachos [131] developed a regular ordering policy and a lateral transshipment policy, as well as conducted extensive simulation studies to verify the advantages of transshipment.

### 2.2.2 Lateral transshipment for perishable products

There have been a number of studies that focus on non-perishable products that involve lateral transshipment. Nevertheless, we found that there were only a few studies that investigated both perishable products and lateral transshipment at the same

time. First, we review the literature concerning reactive transshipment of perishable products in detail. Nakandala et al. [101] developed a periodic review perishable inventory model that considered reactive transshipment in the fresh food supply chain. They embodied five cost components and optimized the trade-off among these components. For the purpose of minimizing the total cost, they developed a simple decision rule that requires only information about the spoilage percentage and the parameters of other cost components. However, logistic practitioners may only use this model if they assume a compound Poisson demand distribution and negligible transshipment times.

It is common for research on lateral transshipment of perishable products to consider the periodic review model. Unlike other research, Dehghani and Abbasi [37] addressed the continuous review perishable inventory model with transshipment. They examined the transshipment of perishable products in a reactive manner and determined stock levels in addition to considering transshipment strategies based on the age threshold of the perishable products. It should be noted, however, that the developed model is only available for Poisson demand distribution. Zhang et al. [149] considered blood platelets inventory which is crucial for blood products and has a perishable nature. For the purposes of issuing and reactive transshipment, they assumed the first-in-first-out (FIFO) rule. A simple transshipment policy, which guarantees near-optimal results, was developed. An application of the policy developed was successfully implemented in a real hospital system, resulting in a substantial reduction in out-of-date platelets.

Additionally, there are very few studies that specifically deal with proactive transshipment of perishable products, which is the most relevant area for our re-

search. Cheong [32] presented a proactive transshipment model incorporating an inventory management system for perishable goods for multiple retail stores. The algorithm determines the optimal order and transshipment quantities on a single-period planning horizon. A perishable product with only a two-period lifecycle was assumed by the authors: old and fresh. Therefore, the proposed model cannot be applied to real-world scenarios such as fresh food supply chains or blood supply chains. Wei et al. [145] emphasized the recycling of perishable products. The authors proposed an approximate dynamic programming model for transshipment policies that included replenishment and recycling decisions. The developed model, however, should have addressed the shelf life of the perishable product, which is the key characteristic of perishable products. Moreover, since the authors dealt with a very short planning horizon (four months), replenishment and transshipment lead times were not taken into account. Li et al. [90] considered offline retailing of perishable goods, which included replenishment decisions and proactive transshipments. Considering that customers usually select the freshest items first, the authors assumed the last-in-first-out (LIFO) rule for issuing. Through rigorous analysis, they showed two roles in transshipment: inventory balancing and inventory separation, which means that new inventory is put in one outlet and the older inventory is put in the other, and this research observed this effect for the first time.

The majority of research in lateral transshipment on perishable inventory management is focused on a specific demand distribution. On the other hand, Dehghani et al. [38] developed a method that does not require an assumption about demand distribution. In addition, the authors assumed that perishable products would arrive immediately, which means the transshipment time is negligible. They utilized

a mixed-integer linear programming (MILP) model, which is equivalent to TSSP model, to decide on proactive transshipment in the blood supply chain. To accommodate a long planning horizon, they developed the rolling horizon algorithm based on the suggested TSSP model, called *RH-TSSP*. However, RH-TSSP required a considerable amount of computational effort to solve one test instance as the MILP model must be solved for every period. For example, if the decision maker wishes to solve a 10,000 period problem, the MILP model must be solved 10,000 times for each instance. Similarly, the RL approach also required several hours to train the RL agent for one instance. In our computational experiments, we found that training both hybrid DRL and RH-TSSP for the same instance took a similar amount of time. However, the trained neural networks of hybrid DRL are capable of reusing other test data sets. In Section 2.5.1 we will compare the performance of the hybrid DRL and the RH-TSSP in detail. In summary, we show several distinctive features of our study compared to existing studies in lateral transshipment for perishable products in Table 2.1.

Table 2.1: Comparison of recent studies related to lateral transshipment on perishable inventory management

| Author | Transshipment type | Replenishment decision | Heterogeneous shelf life | Perishable inventory | Without assumptions on the demand | Solution methodology |
|---|---|---|---|---|---|---|
| Cheong [32] | Proactive | ✓ | | ✓ | | IA[a] |
| Nakandala et al. [101] | Reactive | | | ✓ | | DE[b] |
| Dehghani and Abbasi [37] | Reactive | | | ✓ | | DE[b] |
| Meissner and Senicheva [97] | Proactive | | | | ✓ | Approximate DP[c] |
| Dehghani et al. [38] | Proactive | ✓ | | ✓ | ✓ | RH-TSSP[d] |
| Li et al. [90] | Proactive | ✓ | | ✓ | | DP[c] |
| Wei et al. [145] | Proactive | ✓ | | | | Approximate DP[c] |
| Zhang et al. [149] | Reactive | | | ✓ | | DP[c] |
| **This research** | Proactive | ✓ | ✓ | ✓ | ✓ | RL[e] |

[a] Iterative algorithm; [b] Differential equation; [c] Dynamic programming; [d] Rolling horizon algorithm based on the stochastic programming; [e] Reinforcement learning

### 2.2.3 Reinforcement learning approach for inventory management

Recently, researchers studying inventory management problems have paid considerable attention to the RL approach. Many researchers tried to employ the RL approach to solve intractable inventory problems: perishable inventory management [76, 36], beer game [104], joint replenishment problem [140], multi-product and multi-node inventory management [129], and joint pricing and inventory problem [152]. In addition, Boute et al. [23] suggested a number of research avenues that may help to adopt the DRL approach to practical inventory management problems.

Instead of reviewing all existing studies on the RL approach to inventory management, we present a detailed literature review of four key papers related to our research topic. Gijsbrechts et al. [52] applied the DRL algorithm, namely the asynchronous advantage actor-critic (A3C) [99], to address three classic and intractable inventory problems: lost sales, dual-sourcing, and multi-echelon inventory management. The training of neural networks of DRL for each different inventory problem requires extensive tuning of hyperparameters, which is inevitable when it comes to designing neural networks of DRL. In order to mitigate this burden, the authors proposed a method for automatically tuning several types of hyperparameters. Compared with state-of-the-art heuristics and approximate dynamic programming, the developed A3C algorithm was able to achieve similar performance for each problem. This result implies that DRL could be a promising approach for inventory problems in which effective heuristics are lacking.

Oroojlooyjadid et al. [104] proposed a Deep Q-Network (DQN) algorithm as part of a feedback scheme for the reward function applied to the beer game, a decentralized, multi-agent and cooperative problem. Although there are four players

in the suggested beer game, the authors assumed that only one player could use the DQN, and the others would act irrationally. Without knowledge of the demand probability distribution, real demand data was employed to test the performance of the DQN algorithm. Furthermore, the transfer learning approach was developed to boost the learning speed of DQN in different cost instances. However, the developed DQN algorithm can only be applied in situations where there is only one player using the DQN algorithm. To practically employ the DRL approach to reduce the bullwhip effect, it is more reasonable that all players use DRL than in the above situation. The multi-agent DRL could be a more suitable approach to the beer game problem.

Kara and Dogan [76] and De Moor et al. [36] have both utilized the RL approach to manage perishable inventory. Kara and Dogan [76] employed Q-learning and Sarsa algorithms to solve the problem. They defined two different states: the stock-age-based state and the quantity-based state. Computational experiments showed that Q-learning combined with a stock-age-based state showed the most promising performance. However, the authors should have compared the developed algorithm with state-of-the-art heuristics and evaluated performance systematically using an optimal policy or lower bound; thus, it is difficult to trust the RL algorithm.

In contrast, De Moor et al. [36] evaluated the performance of the developed DQN algorithm in comparison with the optimal policy produced by value iteration (VI). They particularly implemented potential-based reward shaping, which can transfer knowledge of a state-of-the-art inventory policy (teacher policy) into the DQN algorithm, which is called shaped DQN. It was found that the shaped DQN algorithm outperformed the DQN algorithm without reward shaping for the small

size problem (the product has a two-period shelf life). Occasionally, the shaped DQN showed better results than a teacher policy. However, for the practical size problem, the shaped DQN rarely outperformed the teacher policy, which means that the DQN algorithm with potential-based reward shaping is unlikely to surpass the performance of the state-of-the-art inventory policy. As Kara and Dogan [76] and De Moor et al. [36] employed basic reinforcement algorithms, Q-learning and DQN, respectively; thus, there may be room for improving the performance by implementing state-of-the-art RL algorithms, such as SAC and proximal policy optimization (PPO).

In summary, this is the first study that considers the proactive transshipment of perishable products in the OOCS. Specifically, we consider the following three attributes to accommodate key features of the OOCS for fresh foods as mentioned in Section 2.1: heterogeneous shelf life, proactive transshipment, and non-negligible transshipment time. Moreover, we developed the DRL approach based on the SAC algorithm in order to use data directly for decision-making without making any assumptions about demand distribution. In addition, to mitigate the computational burden incurred by large action spaces, we propose the hybrid DRL approach by adopting two novel acceleration methods.

## 2.3 Problem description and mathematical model

### 2.3.1 Lateral transshipment for fresh foods in the online-offline channel system (OOCS)

We consider a periodic review, infinite horizon, perishable inventory problem with stochastic demand and fixed positive lead time. We assume that the unsatisfied demand will become lost sales. There are two heterogeneous outlets, online and

offline channels, indexed by superscript $ON$ and $OF$, owned by the same retailer as in Figure 2.2. Each channel has its own demand for customers, and the demand is random and must be satisfied by the inventories of each channel. The online distribution center satisfies the online demand $D_t^{ON}$, and the offline store satisfies the offline demand $D_t^{OF}$. There are deterministic lead times for online and offline channels ($L^{ON} \geq 0$ and $L^{OF} \geq 0$). We focus on a single type of fresh food, and perishability exists due to the nature of fresh food. From here forward, we will use the term *product* to indicate fresh food.

Furthermore, we accommodate the following assumptions to consider the properties of online and offline channels in real business:

- Shelf life of the product is different depending on the outlet. The shelf life of the product held in the online channel is shorter than in the offline channel.
- Even though online and offline channels sell the same item, the sale price of the offline channel is lower than that of the online channel.
- Demand distributions of online and offline channels are different.
- Because online distribution centers and offline stores are located in different regions, the lead times for replenishment of the online and offline channels are different.

Furthermore, we accommodate the following assumptions to consider the properties of fresh food and the OOCS in real business:

- Shelf life of the product is different depending on the outlet. The shelf life of the product held in the online channel is shorter than in the offline channel.

- Even though online and offline channels sell the same item, the sale price of the offline channel is cheaper than that of the online channel.

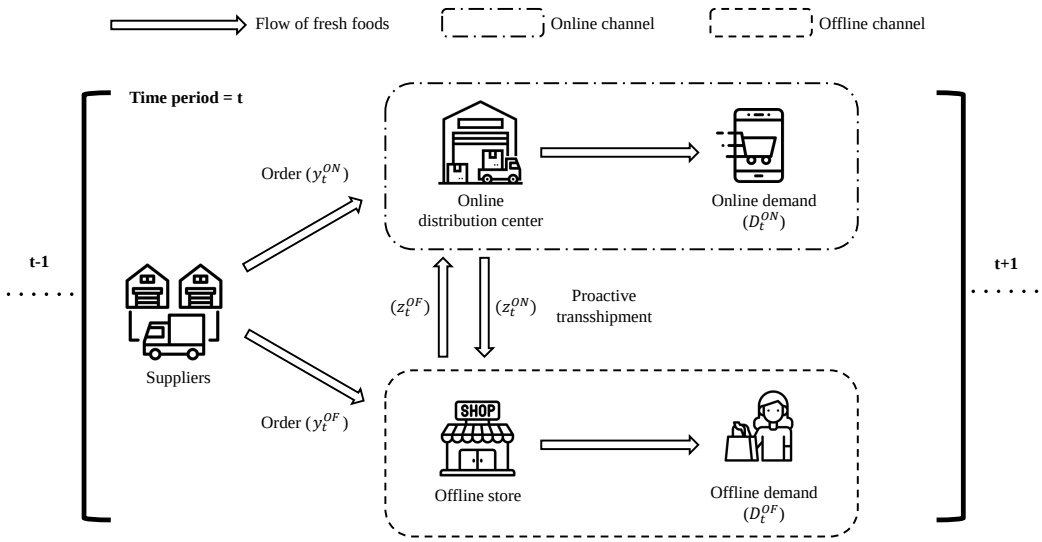



Figure 2.2: Overview of the flow of fresh foods in the OOCS.

The objective of the proposed problem is to maximize the average profit per period of the OOCS. Each channel orders products from suppliers, and the transshipment is mutually implemented between channels. The retailer makes four types of decisions at each period $t$: (1) order quantity of online channel, $y_t^{ON}$, (2) order quantity of offline channel, $y_t^{OF}$, (3) quantity of transshipped units from online channel to offline channel, $z_t^{ON}$, and (4) quantity of transshipped units from offline channel to online channel, $z_t^{OF}$. We consider that the issuing policy is FIFO, and the transshipment policy is first-in-first-transship (FIFT) in both outlets. In particular, the FIFT refers to the strategy of transshipping the older products ahead to the other channel while reserving the younger products [142]. We consider the case that the transshipment time is non-negligible. Specifically, products from the origin

channel are transshipped before demand is realized, and transshipped products arrive at the destination channel in the next period. We assume that realized demand and all variables are integers, to accommodate the situation that most e-commerce companies deal with packaged fresh foods, which counted as units.

To describe the evolution of the proposed system, we use the following notations:

**Indices and sets**

| | |
|---|---|
| $\mathcal{T}$ | set of periods, $t \in \mathcal{T} = \{1, 2, \cdots, T\}$ |
| $\mathcal{M}^{ON}$ | set of the age of product held in the online channel, $m \in \mathcal{M}^{ON} = \{1, 2, \cdots, M^{ON}\}$ |
| $\mathcal{M}^{OF}$ | set of the age of product held in the offline channel, $m \in \mathcal{M}^{OF} = \{1, 2, \cdots, M^{OF}\}$ |

**Parameters**

| | |
|---|---|
| $p^{ON}$ | sale price for a unit of product in online channel |
| $p^{OF}$ | sale price for a unit of product in offline channel ($p^{ON} > p^{OF}$) |
| $M^{ON}$ | shelf life of product held in online channel |
| $M^{OF}$ | shelf life of product held in offline channel ($M^{OF} > M^{ON}$) |
| $y^{ON}_{max}$ | maximum order quantity at each period in online channel |
| $y^{OF}_{max}$ | maximum order quantity at each period in offline channel |
| $z^{ON}_{max}$ | maximum transshipment quantity at each period from online channel to offline channel |
| $z^{OF}_{max}$ | maximum transshipment quantity at each period from offline channel to online channel |
| $L^{ON}$ | lead time of orders in online channel |
| $L^{OF}$ | lead time of orders in offline channel |
| $c_h$ | inventory holding cost for a unit of product |
| $c_p$ | lost sales cost for a unit of product |
| $c_l$ | transshipment cost for a unit of product |
| $c_w$ | outdating cost for a unit of product |
| $c_o$ | ordering cost for a unit of product |

**State variables**

| | |
|---|---|
| $IL_{m,t}^{ON}$ | starting inventory level of age $m$ product at period $t$ in online channel |
| $IL_{m,t}^{OF}$ | starting inventory level of age $m$ product at period $t$ in offline channel |
| $OT_{l,t}^{ON}$ | pipeline inventory that will arrive after $L^{ON} - l$ periods at period $t$ in online channel |
| $OT_{l,t}^{OF}$ | pipeline inventory that will arrive after $L^{OF} - l$ periods at period $t$ in offline channel |
| $LT_{m,t}^{ON}$ | transshipment quantity of age $m$ product at period $t$ from online channel to offline channel |
| $LT_{m,t}^{OF}$ | transshipment quantity of age $m$ product at period $t$ from offline channel to online channel |

**Concatenated vectors of state variables**

| | |
|---|---|
| $IL_t$ | Concatenated vectors of the $\left(IL_{1,t}^{ON}, \cdots, IL_{M^{ON},t}^{ON}, IL_{1,t}^{OF}, \cdots, IL_{M^{OF},t}^{OF}\right)$ |
| $OT_t$ | Concatenated vectors of the $\left(OT_{1,t}^{ON}, \cdots, OT_{L^{ON}-1,t}^{ON}, OT_{1,t}^{OF}, \cdots, OT_{L^{OF}-1,t}^{OF}\right)$ |
| $LT_t$ | Concatenated vectors of the $\left(LT_{1,t}^{ON}, \cdots, LT_{M^{ON}-1,t}^{ON}, LT_{1,t}^{OF}, \cdots, LT_{M^{ON}-2,t}^{OF}\right)$ |

For each period $t$, the following sequence of an event is repeated:

1. At the start of period $t$, the starting inventory level, determined at the end of the previous period $t-1$, is observed. We consider $IL_{1,t}^{ON}$ and $IL_{1,t}^{OF}$ to be the youngest product and $IL_{M^{ON},t}^{ON}$ and $IL_{M^{OF},t}^{OF}$ to be the oldest product that expires at the end of the previous period $t-1$.

2. Four types of decisions, $y_t^{ON}, y_t^{OF}, z_t^{ON}$, and $z_t^{OF}$, are implemented at the start of period $t$. The limit of order quantity exists for each channel, $0 \leq y_t^{ON} \leq y_{max}^{ON}$ and $0 \leq y_t^{OF} \leq y_{max}^{OF}$. Because of the lead time of each channel, the $y_t^{ON}$ will be received at time $t + L^{ON}$, and the $y_t^{OF}$ will be received at time $t + L^{OF}$. For each channel, the limit of transshipment quantity exists and is determined by the current inventory level, $0 \leq z_t^{ON} \leq \min\left\{z_{max}^{ON}, \sum_{m=1}^{M^{ON}-1} IL_{m,t}^{ON}\right\}$ and $0 \leq z_t^{OF} \leq \min\left\{z_{max}^{OF}, \sum_{m=1}^{M^{OF}-1} IL_{m,t}^{OF}\right\}$. The transshipped product is received at each channel at the start of period $t+1$.

3. The random demand, $D_t^{ON}$ and $D_t^{OF}$, is realized; as much as possible, it is satisfied from the inventory level, $IL_{m,t}^{ON}$ and $IL_{m,t}^{OF}$, and the rest is lost. The inventory level and pipeline inventory for the next period $t+1$ are updated.

4. At the end of period $t$, revenue and inventory holding, shortage, outdating, ordering, and the transshipment costs are assessed.

We utilize the state variables $LT_{m,t}^{ON}$ and $LT_{m,t}^{OF}$ to implement the FIFT policy for transshipment. The $LT_{m,t}^{ON}$ and $LT_{m,t}^{OF}$ are decided as follows:

$$LT_{m,t}^{ON} = \min\left\{\left(z_t^{ON} - \sum_{k=m+1}^{M^{ON}-1} IL_{k,t}^{ON}\right)^+, IL_{m,t}^{ON}\right\}, \qquad \forall m \in \left\{1, \cdots, M^{ON}-2\right\}, \tag{2.1}$$

$$LT_{M^{ON}-1,t}^{ON} = \min\left\{IL_{M^{ON}-1,t}^{ON}, z_t^{ON}\right\}, \tag{2.2}$$

$$LT_{m,t}^{OF} = \min\left\{\left(z_t^{OF} - \sum_{k=m+1}^{M^{ON}-2} IL_{k,t}^{OF}\right)^+, IL_{m,t}^{OF}\right\}, \qquad \forall m \in \left\{1, \cdots, M^{ON}-3\right\}, \tag{2.3}$$

$$LT_{M^{ON}-2,t}^{OF} = \min\left\{IL_{M^{ON}-2,t}^{OF}, z_t^{OF}\right\}. \tag{2.4}$$

where $x^+ = \max\{x, 0\}$. The index $m$ of $LT_{m,t}^{ON}$ includes from one to $M^{ON}-1$, and the index $m$ of $LT_{m,t}^{OF}$ includes from one to $M^{ON}-2$, because the shelf life of the product in the online channel is shorter than that in the offline channel, and the transshipment quantity will be received at the next period, $t+1$. Specifically, assume that the age $M^{ON}-1$ product is transshipped from the offline channel to the online channel in period $t$. This type of transshipment will become worthless because the transshipped product's age becomes $M^{ON}$ when the product arrives at the online channel, i.e., the transshipped product cannot be sold to customers.

State variables for inventory level and pipeline inventory in the online channel are updated based on the following equations after $D_t^{ON}$ is realized:

**Online channel**

$$OT_{1,t+1}^{ON} = y_t^{ON}, \tag{2.5}$$

$$OT_{l,t+1}^{ON} = OT_{l-1,t}^{ON}, \quad \forall l \in \left\{2, \cdots, L^{ON} - 1\right\}, \tag{2.6}$$

$$IL_{1,t+1}^{ON} = OT_{L^{ON}-1,t}^{ON}, \tag{2.7}$$

$$IL_{m+1,t+1}^{ON} = \left(\left(IL_{m,t}^{ON} - LT_{m,t}^{ON}\right) - \left(D_t^{ON} - \sum_{k=m+1}^{M^{ON}-1}\left(IL_{k,t}^{ON} - LT_{k,t}^{ON}\right)\right)^+\right)^+ + LT_{m,t}^{OF}, \tag{2.8}$$

$$\forall m \in \left\{1, \cdots, M^{ON} - 2\right\},$$

$$IL_{M^{ON},t+1}^{ON} = \left(\left(IL_{M^{ON}-1,t}^{ON} - LT_{M^{ON}-1,t}^{ON}\right) - D_t^{ON}\right)^+. \tag{2.9}$$

The inventory level for the period $t+1$ is updated by subtracting the current inventory level from the transshipment quantity (i.e., $IL_t^{ON} - LT_t^{ON}$).

State variables for inventory level and pipeline inventory in the offline channel are updated based on the following equations after $D_t^{OF}$ is realized. The pipeline inventory in the offline channel updated as the same transitions in the online channel, Equations (2.5) and (2.6), as follows:

**Offline channel (pipeline inventory)**

$$OT_{1,t+1}^{OF} = y_t^{OF}, \tag{2.10}$$

$$OT_{l,t+1}^{OF} = OT_{l-1,t}^{OF}, \quad \forall l \in \left\{2, \cdots, L^{OF} - 1\right\} \tag{2.11}$$

However, because the shelf life of the online channel is shorter than that of the offline channel, the offline channel's inventory level is updated differently depending on the product's age $m$. If the index $m$ is included in the set $\left\{1, \cdots, M^{ON} - 2\right\}$, the inventory level for the period $t+1$ is also updated by subtracting the current inventory level from the transshipment quantity (i.e., $IL_t^{OF} - LT_t^{OF}$) as follows:

**Offline channel (inventory level)**

$$IL^{OF}_{1,t+1} = OT^{OF}_{L^{OF}-1,t} \tag{2.12}$$

$$IL^{OF}_{m+1,t+1} = \left(\left(IL^{OF}_{m,t} - LT^{OF}_{m,t}\right) - \left(D^{OF}_t - \sum_{k=m+1}^{M^{ON}-2}\left(IL^{OF}_{k,t} - LT^{OF}_{k,t}\right) - \sum_{k=M^{ON}-1}^{M^{OF}-1} IL^{OF}_{k,t}\right)^+\right)^+ + LT^{ON}_{m,t},$$

$$\forall m \in \left\{1, \cdots, M^{ON}-3\right\}, \tag{2.13}$$

$$IL^{OF}_{M^{ON}-1,t+1} = \left(\left(IL^{OF}_{M^{ON}-2,t} - LT^{OF}_{M^{ON}-2,t}\right) - \left(D^{OF}_t - \sum_{k=M^{ON}-1}^{M^{OF}-1} IL^{OF}_{k,t}\right)^+\right)^+ + LT^{ON}_{M^{ON}-2,t}, \tag{2.14}$$

When the index $m$ is included in the set $\left\{M^{ON}-1, \cdots, M^{OF}\right\}$, there is no need to subtract current inventory from the transshipment quantity because the product cannot be transshipped from the offline channel to the online channel. In addition, if the index $m$ is included in the set $\left\{M^{ON}, \cdots, M^{OF}\right\}$, there is no transshipment quantity from online channel to offline channel, thus, $LT^{ON}_{m,t}$ is not considered as follows:

$$IL^{OF}_{M^{ON},t+1} = \left(IL^{OF}_{M^{ON}-1,t} - \left(D^{OF}_t - \sum_{k=M^{ON}}^{M^{OF}-1} IL^{OF}_{k,t}\right)^+\right)^+ + LT^{ON}_{M^{ON}-1,t}, \tag{2.15}$$

$$IL^{OF}_{m+1,t+1} = \left(IL^{OF}_{m,t} - \left(D^{OF}_t - \sum_{k=m+1}^{M^{OF}-1} IL^{OF}_{k,t}\right)^+\right)^+, \quad \forall m \in \left\{M^{ON}, \cdots, M^{OF}-2\right\} \tag{2.16}$$

$$IL^{OF}_{M^{OF},t+1} = \left(IL^{OF}_{M^{OF}-1,t} - D^{OF}_t\right)^+. \tag{2.17}$$

At each period $t$, revenue and costs are defined as follows based on the above state and decision variables:

$$\text{Inventory holding } (HC_t) := \quad c_h \left[\left(\sum_{m=1}^{M^{ON}-2}\left(IL^{OF}_{m,t} - LT^{OF}_{m,t}\right) + \sum_{m=M^{ON}-1}^{M^{OF-1}} IL^{OF}_{m,t} - D^{OF}_t\right)^+\right. \tag{2.18}$$

$$+\left(\sum_{m=1}^{M^{ON}-1}\left(IL_{m,t}^{ON}-LT_{m,t}^{ON}\right)-D_t^{ON}\right)^+\Bigg]$$

$$\text{Shortage } (SC_t) := c_p\left[\left(D_t^{OF}-\sum_{m=1}^{M^{ON}-2}\left(IL_{m,t}^{OF}-LT_{m,t}^{OF}\right)-\sum_{m=M^{ON}-1}^{M^{OF}-1}IL_{m,t}^{OF}\right)^+\right. \tag{2.19}$$

$$\left.+\left(D_t^{ON}-\sum_{m=1}^{M^{ON}-1}\left(IL_{m,t}^{ON}-LT_{m,t}^{ON}\right)\right)^+\right]$$

$$\text{Outdating } (WC_t) := c_w\left(IL_{M^{ON},t+1}^{ON}+IL_{M^{OF},t+1}^{OF}\right) \tag{2.20}$$

$$\text{Ordering } (OC_t) := c_o\left(y_t^{ON}+y_t^{OF}\right) \tag{2.21}$$

$$\text{Transshipment } (TC_t) := c_l\left(z_t^{ON}+z_t^{OF}\right) \tag{2.22}$$

$$\text{Revenue } (RV_t) := p^{ON}\left[\min\left\{\sum_{m=1}^{M^{ON}-1}\left(IL_{m,t}^{ON}-LT_{m,t}^{ON}\right), D_t^{ON}\right\}\right] \tag{2.23}$$

$$+\,p^{OF}\left[\min\left\{\sum_{m=1}^{M^{ON}-2}\left(IL_{m,t}^{OF}-LT_{m,t}^{OF}\right)+\sum_{m=M^{ON}-1}^{M^{OF}-1}IL_{m,t}^{OF}, D_t^{OF}\right\}\right]$$

At last, the profit at period $t$, $PF_t$, is defined as: $PF_t := RV_t - HC_t - SC_t - WC_t - OC_t - TC_t$.

### 2.3.2 Markov decision process for the proposed lateral transshipment problem

The proposed problem can be formalized as the MDP. A MDP is a tuple $\langle \mathcal{S}, \mathcal{A}, r, p, \gamma \rangle$, consisting of five components— a set of states, $\mathcal{S}$; a set of actions, $\mathcal{A}$; the reward function, $r$; the state transition probability function, $p$; and the discount factor, $\gamma \in [0, 1)$. Most existing studies of perishable inventory management defined the state at time $t$ as inventory level and pipeline inventory, $s_t = (IL_t, OT_t)$ [76, 36, 52]. In addition to $IL_t$ and $OT_t$, we include the transshipment information at previous period $t-1$, $LT_{t-1}$, in the state at time $t$, $s_t = (IL_t, OT_t, LT_{t-1})$. The state space

$\mathcal{S}$ is thus $\left(3M^{ON} + M^{OF} + L^{ON} + L^{OF} - 5\right)$ dimensional.

We consider a discrete action space, and the action at time $t$ is defined as $a_t = (y_t^{ON}, y_t^{OF}, z_t^{ON}, z_t^{OF})$. Due to the maximum order and transshipment quantity in each channel, the size of action space $|\mathcal{A}|$ is $\left(y_{max}^{ON} + 1\right) \times \left(y_{max}^{OF} + 1\right) \times \left(z_{max}^{ON} + 1\right) \times \left(z_{max}^{OF} + 1\right)$. The valid actions for transshipment are different depending on the current state $s_t$, specifically the $IL_{m,t}^{ON}$ and $IL_{m,t}^{OF}$ as indicated in Section 2.3.1. The transition probability function is denoted as $p(s_{t+1}|s_t, a_t)$, which indicates the probability that the system is in state $s_{t+1}$ at period $t+1$ when the action $a_t$ is chosen under state $s_t$ at period $t$. The transition probability can be computed if the demand distribution is known. The reward function quantifies how well the immediate action $a_t$ and state $s_t$ are chosen. Because the purpose of the proposed problem is to maximize the average profit per period, the reward function can be defined as follows: $r(s_t, a_t) := PF_t$.

The objective of solving the MDP is to find a policy $\pi : \mathcal{S} \to \mathcal{A}$, mapping each state to an action, that maximizes the expected cumulative reward:

$$\max_{\pi \in \Pi} \mathbb{E}^{\pi} \left[ \sum_{t=0}^{\infty} \gamma^t r(s_t, a_t) \right] \tag{2.24}$$

where $\Pi$ is the set of all policies and $\mathbb{E}^{\pi}$ is the expectation operator when following policy $\pi$. The value function $V^{\pi}(s)$ of a policy $\pi$ is the expected cumulative reward starting from state $s$ under executing $\pi$:

$$V^{\pi}(s) := \mathbb{E}^{\pi} \left[ \sum_{\tau=t}^{\infty} \gamma^{\tau - t} r(s_\tau, a_\tau) | s_t = s \right] \tag{2.25}$$

The optimal policy can be derived from the optimal value function $V^*(s) := \max_{\pi \in \Pi}$

$V^{\pi}(s), \forall s \in \mathcal{S}$, which is the maximum value function overall policies. When the finite state and actions sets are assumed, the optimal value function can be obtained by solving the following Bellman equations recursively (i.e., VI algorithm [14]):

$$V^*(s) := \max_{a \in \mathcal{A}} \left[ r(s,a) + \gamma \sum_{s' \in \mathcal{S}} p(s'|s,a) V^*(s') \right] \tag{2.26}$$

However, when dealing with the large-scale and complex MDP, it becomes challenging to solve with the VI algorithm, due to two issues [56]. First, if several random variables should be manipulated in the MDP, it is generally difficult to compute the transition probability. Furthermore, it is impossible to find transition probability if the probability distribution of random variables is not known. This issue is called the *curse of modeling*. Furthermore, obtaining the accurate transition probability is only possible if the true demand distribution is known. Second, it is challenging to store and handle the value function for all states, $\forall s \in \mathcal{S}$, when the problem involves a large dimension. This issue is called the *curse of dimensionality*.

To solve the proposed problem, we first tried to solve the problem through the VI algorithm with the assumption that the demand distribution is known. However, the dimension of state and action is enormous because our problem deals with two outlets and the property of perishability. Even though we tested on small-size instances, it was impossible to obtain the value function because of the memory issue. Moreover, transition probability cannot be computed directly from real demand data. In order to mitigate the above issues, we adopt the DRL approach to solve the proposed problem. The capability of DRL lies in its ability to solve the complex MDP, which involves the large dimension of state and action, near-

optimally. Furthermore, the Model-free DRL approach does not need to know or learn the transition probability. Instead, the policy is learned through interactions between the environment and agent utilizing the demand data set directly. Among various Model-free DRL algorithms, we employ the SAC algorithm and enhance the performance of the algorithm.

## 2.4 Solution methodology: hybrid deep reinforcement learning (DRL) approach

### 2.4.1 Soft actor-critic algorithm

Model-free DRL algorithms are suffered from poor sample efficiency and sensitivity to hyperparameters. Usually, on-policy algorithms, such as PPO [120] and A3C [99], require new samples at each gradient step. On the other hand, off-policy algorithms can reuse past experience, which increases sample efficiency. Even though deep deterministic policy gradient (DDPG) [92], which is an off-policy algorithm, is proposed to use samples efficiently, this method is too sensitive to hyperparameters in the training process. To mitigate the above issues, Haaranoja et al. [60] introduced the SAC algorithm, which is an off-policy actor-critic DRL algorithm based on the maximum entropy framework. The exploration and robustness are improved by using the maximum entropy framework.

In the maximum entropy MDP problem, the concept of entropy of policy is used as follows:

$$\mathcal{H}\left(\pi\left(\cdot|s\right)\right) := \mathbb{E}_{a\sim\pi(\cdot|s)}\left[-\log\pi\left(a|s\right)\right] \tag{2.27}$$

The goal of the maximum entropy MDP problem is to find a policy that maximizes the maximum entropy objective:

$$\max_{\pi \in \Pi} \sum_{t=0}^{\infty} \mathbb{E}_{(s_t,a_t)\sim\rho_\pi} \left[\gamma^t \left(r(s_t,a_t)\right) + \alpha \mathcal{H}\left(\left(\pi(\cdot|s_t)\right)\right)\right] \tag{2.28}$$

where $\rho_\pi$ is the distribution of trajectories induced by policy $\pi$, and $\alpha$ is the temperature parameter, which is utilized to control the relative importance of the reward and entropy.

From now on, we will introduce methods to derive optimal policy in the maximum entropy MDP in two different settings: (1) tabular setting, and (2) large spaces setting. First, in a tabular setting, the optimal policy can be found by repeating the implementation of soft policy evaluation and soft policy improvement (i.e., soft policy iteration) [60]. For the soft policy evaluation, the following soft Q-function of a policy $\pi$ can be computed by repeatedly applying the modified Bellman backup operator:

$$Q_\pi^{soft}\left(s_t,a_t\right) := r\left(s_t,a_t\right) + \gamma \mathbb{E}_{s_{t+1}\sim p(\cdot|s_t,a_t)} \left[V_\pi^{soft}\left(s_{t+1}\right)\right] \tag{2.29}$$

where

$$V_\pi^{soft}(s_t) := \mathbb{E}_{a_t\sim\pi(\cdot|s_t)} \left[Q_\pi^{soft}\left(s_t,a_t\right) - \alpha \log \pi\left(a_t|s_t\right)\right] \tag{2.30}$$

is called the soft value function. For the soft policy improvement, the policy is updated toward maximizing the rewards, which is the exponential of the new soft Q-function. The set of policies $\Pi$ is considered to constrain policies to a parameterized

family of distributions (e.g., Gaussian). To accommodate the above constraint, $\pi \in \Pi$, we utilize the Kullback-Leibler (KL) divergence for information projection; thus, the policy is updated as follows:

$$\pi_{new} = \arg\min_{\pi \in \Pi} D_{KL} \left( \pi \left( \cdot | s_t \right) \middle\| \frac{\exp\left( \frac{1}{\alpha} Q^{soft}_{\pi_{old}}(s_t, \cdot) \right)}{Z_{\pi_{old}}(s_t)} \right) \tag{2.31}$$

where $Z_{\pi_{old}}(s_t)$ is the partition function used to normalize the distribution but can be ignored because it does not contribute to the gradient descent.

In order to practically employ the soft policy iteration in large spaces (e.g., continuous state), the function approximators, neural networks, are utilized for both the soft Q-function and the policy. A neural network with parameter $\theta$ is used for the soft Q-function, $Q^{soft}_{\theta}(s_t, a_t)$ and a neural network with parameter $\phi$ is used for the policy $\pi_{\phi}(a_t|s_t)$. To mitigate the issue of positive bias, two soft Q-functions and neural networks are utilized, $Q^{soft}_{\theta_i}(s_t, a_t)$, $\forall i \in \{1, 2\}$ [68]. Also, two target soft Q-networks are used to compute the shared target and improve the training stability, $Q^{soft}_{\bar{\theta}_j}(s_t, a_t)$, $\forall j \in \{1, 2\}$. Both target soft Q-networks are updated through the soft update approach, which can be represented as:

$$\bar{\theta}_j \leftarrow \psi \theta_j + (1 - \psi) \bar{\theta}_j \tag{2.32}$$

where $\psi \in [0, 1]$ is the parameter for weight update.

In the large spaces setting, we call soft policy evaluation as *critic* and soft policy improvement as *actor*. We train each soft Q-function parameter $\theta_i$ to minimize the

following critic cost function:

$$J_{Q^{soft}}(\theta_i) = \mathbb{E}_{(s_t,a_t)\sim\mathcal{D}}\left[\left(Q^{soft}_{\theta_i}(s_t,a_t) - \left(r(s_t,a_t) + \mathbb{E}_{s_{t+1}\sim p(\cdot|s_t,a_t)}\left[V_{\bar{\theta}}(s_{t+1})\right]\right)\right)^2\right] \tag{2.33}$$

where:

$$V_{\bar{\theta}}(s_{t+1}) = \mathbb{E}_{a_{t+1}\sim\pi_\phi(\cdot|s_{t+1})}\left[\min_{j\in\{1,2\}} Q^{soft}_{\bar{\theta}_j}(s_{t+1},a_{t+1}) - \alpha\log\pi_\phi(a_{t+1}|s_{t+1})\right] \tag{2.34}$$

The replay buffer storing trajectories of experience is denoted as $\mathcal{D}$. Through sampling experiences from the replay buffer, the soft value function $V_{\bar{\theta}(s_{t+1})}$ is estimated through the Monte-Carlo method.

The policy parameter $\phi$ can be learned by directly minimizing the KL divergence in Equation (2.31) as follows by multiplying $\alpha$ and ignoring the partition function:

$$J_\pi(\phi) = \mathbb{E}_{s_t\sim\mathcal{D}}\left[\mathbb{E}_{a_t\sim\pi_\phi(\cdot|s_t)}\left[\alpha\log\left(\pi_\phi(a_t|s_t)\right) - \min_{j\in\{1,2\}} Q^{soft}_{\bar{\theta}_j}(s_t,a_t)\right]\right] \tag{2.35}$$

Generally, the policy $\pi_\phi(a_t|s_t)$ outputs a mean $\mu_\phi(s_t)$ and a standard deviation $\sigma_\phi(s_t)$ thus the actions are distributed Gaussian distribution, $a_t \sim N(\mu_\phi(s_t), \sigma_\phi(s_t))$. However, because Equation (2.35) cannot be backpropagated in the normal scheme to compute $J_\pi(\phi)$, the *reparameterization trick* is adopted. Given state $s_t$, the squashed Gaussian policy is used; thus, the action is sampled according to:

$$\tilde{a}^\phi_t(s_t,\xi_t) = \tanh\left(\mu_\phi(s_t) + \sigma_\phi(s_t)\odot\xi_t\right), \quad \xi_t \sim N\left(0, I\right) \tag{2.36}$$

where $\xi$ is the noise following the standard normal distribution. By adopting the

reparameterization trick, the policy $\pi_\phi$ is optimized by minimizing the following actor cost function:

$$J_\pi(\phi) = \mathbb{E}_{\substack{s_t \sim \mathcal{D} \\ \xi_t \sim N}} \left[ \alpha \log \left( \pi_\phi(\tilde{a}_t^\phi(s_t, \xi_t) | s_t) \right) - \min_{j \in \{1,2\}} Q_{\bar{\theta}_j}^{soft}(s_t, \tilde{a}_t^\phi(s_t, \xi_t)) \right] \tag{2.37}$$

Instead of deciding the fixed value for temperature parameter $\alpha$, the $\alpha$ can be learned by optimizing the following objective [61]:

$$J(\alpha) = \mathbb{E}_{\substack{s_t \sim \mathcal{D} \\ a_t \sim \pi_\phi(\cdot|s_t)}} \left[ -\alpha \left( \log \pi_\phi(a_t|s_t) + \bar{\mathcal{H}} \right) \right] \tag{2.38}$$

where $\bar{\mathcal{H}}$ is a constant representing the target entropy.

### 2.4.2 SAC for discrete action space and prioritized experience replay

The SAC algorithm was developed to derive a near-optimal policy in a continuous action spaces setting [60, 61]. However, we deal with the discrete action spaces setting because the decision variables of the proposed problem are integers. Therefore, we revised the SAC algorithm to adjust in discrete action spaces setting based on the approach introduced by Christodoulou [33]. In a continuous setting, $\pi_\phi(\cdot|s_t)$ is a probability density function; however, it is now a probability mass function. To revise the SAC algorithm, the following two changes should be considered:

- $Q^{soft} : \mathcal{S} \times \mathcal{A} \to \mathbb{R} \implies Q^{soft} : \mathcal{S} \to \mathbb{R}^{|\mathcal{A}|}$ : Unlike continuous action spaces, which have infinitely many possible actions, there are a limited number of possible actions in discrete actions paces. Therefore, the soft Q-function can be changed as a mapping $Q^{soft} : \mathcal{S} \to \mathbb{R}^{|\mathcal{A}|}$ from a state to a vector containing

the Q-value of each possible action.

- $\pi : \mathcal{S} \to \mathbb{R}^{2|\mathcal{A}|} \implies \pi : \mathcal{S} \to [0,1]^{|\mathcal{A}|}$ : In a continuous setting, $\pi$ outputs the mean and variance of action distribution. On the other hand, the probability for each action can be directly computed because there are finite possible actions in a discrete setting. By applying a softmax function on the output layer in the neural network of $\pi_\phi$, the policy outputs a vector containing the probability of each action.

Due to the above two changes, cost functions of critic $J_{Q^{soft}}(\theta)$, actor $J_\pi(\phi)$, and temperature $J(\alpha)$ should be revised. In terms of critic cost function $J_{Q^{soft}}(\theta)$, the expectation value of $V_{\bar{\theta}}(s_{t+1})$ (Equation (2.34)) can be computed directly because the probability for each possible action can be obtained instead of forming a Monte-Carlo estimate. Through this modification, the variance for the estimate of critic cost function $J_{Q^{soft}}(\theta)$ can be reduced. The soft value function $V_{\bar{\theta}}(s_{t+1})$ can be obtained by applying the following equation:

$$V_{\bar{\theta}}(s_{t+1}) = \pi_\phi(s_{t+1})^\intercal \left( \min_{j \in \{1,2\}} Q^{soft}_{\bar{\theta}_j}(s_{t+1}, a_{t+1}) - \alpha \log \pi_\phi(a_{t+1}|s_{t+1}) \right) \tag{2.39}$$

In a continuous spaces setting, the reparameterization trick is used to optimize actor cost function $J_\pi(\phi)$, so the soft critic cost function is transformed from Equations (2.35) to (2.37). However, in a discrete spaces setting, the expectation can be calculated directly in Equation (2.35) regarding the policy $\pi_\phi(\cdot|s_t)$. Therefore, we do not need the reparameterization trick, and the new actor cost function is defined as:

$$J_\pi(\phi) = \mathbb{E}_{s_t \sim \mathcal{D}} \left[ \pi_\phi(s_t)^\intercal \left( \alpha \log \left( \pi_\phi(s_t) \right) - \min_{j \in \{1,2\}} Q^{soft}_{\bar{\theta}_j}(s_t) \right) \right] \tag{2.40}$$

Similarly, the temperature cost function $J(\alpha)$ is changed from Equation (2.38) to the following equation because the probability for each action can be computed:

$$J(\alpha) = \mathbb{E}_{s_t \sim \mathcal{D}} \left[ \pi_\phi(s_t)^\intercal \left( -\alpha \left( \log \pi_\phi(a_t|s_t) + \bar{\mathcal{H}} \right) \right) \right] \tag{2.41}$$

In summary, we utilize four neural networks for critic (i.e.,$\theta_i$ and $\bar{\theta}_i, i = 1, 2$) and one neural network for actor (i.e, $\phi$). Each neural network has an input layer, at least one hidden layer and an output layer sequentially. Its input is the state vector, and the output is the $|\mathcal{A}|-$dimensional action vector consisting of unnormalized scores, which is called logits. In particular, an actor neural network converts the logits into an action probability distribution using the following softmax function:

$$f(a_i) = \frac{\exp(a_i)}{\sum_{j=1}^{|\mathcal{A}|} \exp(a_j)} \tag{2.42}$$

As mentioned in Section 2.3.2, the valid actions (i.e., available transshipment quantities) in action spaces $\mathcal{A}$ are different depending on the current state $s_t$. To prevent sampling the invalid action in $\mathcal{A}$, we employ the following *action masking* technique [71]:

1. The current inventory level in online and offline channels, $IL_{m,t}^{ON}$ and $IL_{m,t}^{OF}$, are observed. We check the invalid transshipment quantity by comparing the sum of the current inventory level, $\sum_{m=1}^{M^{ON}-1} IL_{m,t}^{ON}$ and $\sum_{m=1}^{M^{OF}-1} IL_{m,t}^{OF}$, and maximum transshipment quantity, $z_{max}^{ON}$ and $z_{max}^{OF}$.

2. A large negative number replaces the logit of actions corresponding to the invalid transshipment quantity.

3. The action probability can be obtained by inputting the logit of actions into the softmax function, and the probability of invalid actions will become $\epsilon$, which should be a minimal number.

Due to the property of the off-policy algorithm, the SAC can use the past experiences, $(s_t, a_t, r_t, s_{t+1})$, which are stored in replay buffer $\mathcal{D}$. Experiences can be sampled uniformly from a replay buffer without considering the importance of each experience. Even though this scheme stabilized the training process of DRL, it could impede sampling efficiency because important and unimportant experiences are replayed at the same frequency. Therefore, we employ the *prioritized experience replay* (PER), the method that prioritizes more important experiences by measuring the priority value of each experience using the magnitude of its temporal-difference (TD) error [119]. The TD error of experience $d \in \mathcal{D}$, $|\delta_d|$, is defined using the soft value function, $V_{\bar{\theta}}$, and two soft Q-networks, $Q^{soft}_{\theta_1}$ and $Q^{soft}_{\theta_2}$, as follows:

$$|\delta_d| = \min\left\{\left(Q^{soft}_{\theta_1}(s) - \left(r + \gamma V_{\bar{\theta}}(s')\right)\right)^2, \left(Q^{soft}_{\theta_2}(s) - \left(r + \gamma V_{\bar{\theta}}(s')\right)\right)^2\right\} \tag{2.43}$$

The priority value of each experience, $p_d$, is defined according to:

$$p_d = |\delta_d| + \epsilon_{per} \tag{2.44}$$

where $\epsilon_{per}$ is a small positive constant that prevents the case that the probability of revisiting the experience is zero. Then, the probability of sampling experience $d$ is

defined as:

$$P(d) = \frac{p_d^\eta}{\sum_{k=1}^{N_\mathcal{D}} p_k^\eta} \tag{2.45}$$

where $\eta$ is the prioritization factor that determines how much prioritization is used, and $N_\mathcal{D}$ is the size of replay buffer $\mathcal{D}$.

Usually, the PER can cause inevitable bias because it changes the distribution of expectations in an uncontrolled way. Therefore, we correct this bias by utilizing the following importance sampling weights, $w_d$:

$$w_d = \left( \frac{1}{N_\mathcal{D}} \times \frac{1}{P(d)} \right)^\beta \tag{2.46}$$

where $\beta \in [0, 1]$. In addition, we normalize the above weights by $1/\max_d w_d$, due to stability reasons, and apply the importance of sampling weights to update neural networks. Finally, the SAC algorithm with PER for discrete action setting (SACDPE) is presented in Appendix A.2.

### 2.4.3 Two acceleration methods in the hybrid DRL approach: SQLT policy and reward shaping

Even though the SACDPE could get a promising policy, it suffers from unstable performance because of relatively large action spaces. The output layer of critic $(\theta_i, \bar{\theta}_i,\ i = 1, 2)$ and actor neural networks $(\phi)$ composed of $|\mathcal{A}|$ nodes correspond to the number of available actions. Therefore, the large action spaces lead to neural networks with many parameters to train and many nodes in the output layer, resulting in considerable training time [23]. In addition, the large action space increases

the computational burden for the exploration strategy.

To mitigate the above issues and improve the performance of SACDPE, we developed two methods to accelerate the SACDPE. First, we split decision-making for ordering ($y_t^{ON}$ and $y_t^{OF}$) and transshipment ($z_t^{ON}$ and $z_t^{OF}$) quantity into two stages, as shown in Figure 2.3. In the original SACDPE, these four types of decisions are made simultaneously, which makes the action space extraordinarily large. Instead, we decide the transshipment quantity by relying on the DRL algorithm (SACDPE) in the first stage, and then the order quantity is decided through a specific ordering policy in the second stage. Note that this two-stage approach does not violate the Sequence assumption 2. Even though the decision is made first on $z_t^{ON}$ and $z_t^{OF}$ and then $y_t^{ON}$ and $y_t^{OF}$, four types of decision is made before the random demand ($D_t^{ON}$ and $D_t^{OF}$) is realized (i.e., the start of the period $t$).

Some readers may wonder why the decision on the order quantity is not made before the decision on the transshipment quantity. This method determines the order quantity solely without considering additional transshipment sequentially determined, which may result in inefficiency caused by excessive orders. In other words, the order quantity absorbs the amount that transshipment could supplement. On the other hand, if the decision on the transshipment quantity is made before the order quantity as the proposed manner, the transshipment information can be reflected to decide on the order quantity by utilizing the developed new ordering policy, called *SQLT policy*. By separating the ordering decision from actions in the SACDPE, the size of the action space is reduced from $\left(y_{max}^{ON}+1\right)\times\left(y_{max}^{OF}+1\right)\times\left(z_{max}^{ON}+1\right)\times\left(z_{max}^{OF}+1\right)$ to $\left(z_{max}^{ON}+1\right)\times\left(z_{max}^{OF}+1\right)$. Consequently, the number of nodes at the output layer in the neural network is reduced; thus, the training time

could be reduced significantly.

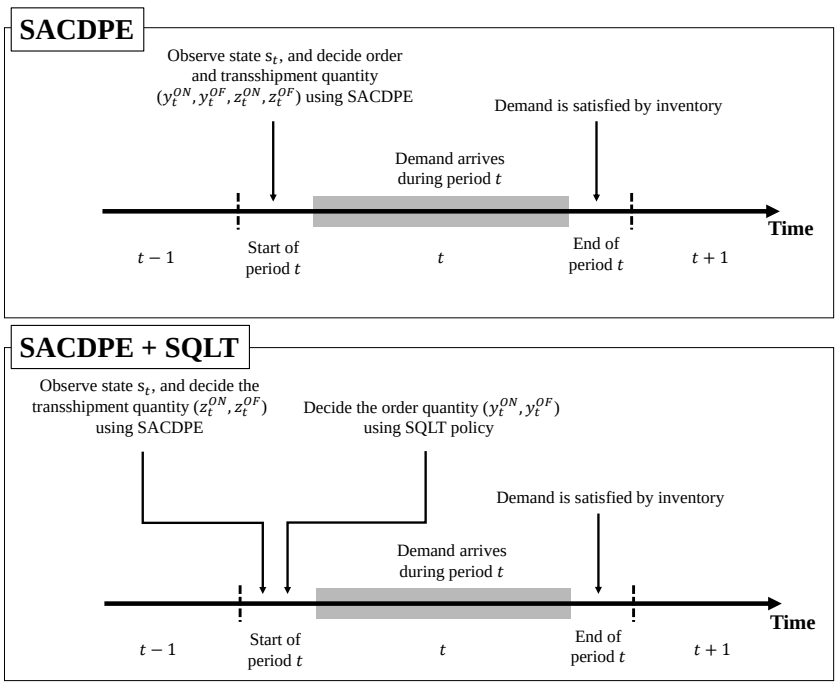



Figure 2.3: Differences between original SACDPE and SACDPE+SQLT.

In the first stage decision, the current state $s_t$ is observed, and the action for transshipment $a_t$ is decided by policy $\pi_\phi(\cdot|s_t)$. In the second stage, any ordering policy, such as the base-stock policy, can be used for the second stage decision. However, we developed the SQLT policy by improving the SQmax policy [63] and reflecting the information about transshipment decisions in the prior stage. The order quantity of policy SQLT is decided as follows:

$$y_t^{ON} = \min\left\{\left(S^{ON} - q_t^{ON}\left(z_t^{ON}, z_t^{OF}\right)\right)^+, Q_{max}^{ON}, y_{max}^{ON}\right\} \tag{2.47}$$

$$y_t^{OF} = \min\left\{\left(S^{OF} - q_t^{OF}\left(z_t^{ON}, z_t^{OF}\right)\right)^+, Q_{max}^{OF}, y_{max}^{OF}\right\} \tag{2.48}$$

where $S^{ON}$ and $S^{OF}$ are parameters for the base-stock level, and $Q_{max}^{ON}$ and $Q_{max}^{OF}$ are parameters for a maximum order quantity in online and offline channels. Because our problem deals with integer values for demand and variables, we find the optimal value of these parameters through a *grid search*. The functions $q_t^{ON}\left(z_t^{ON}, z_t^{OF}\right)$ and $q_t^{OF}\left(z_t^{ON}, z_t^{OF}\right)$ are defined according to:

$$q_t^{ON}\left(z_t^{ON}, z_t^{OF}\right) = \sum_{m=1}^{M^{ON}-1} IL_{m,t}^{ON} + \sum_{l=1}^{L^{ON}-1} OT_{l,t}^{ON} - z_t^{ON} + z_t^{OF} \tag{2.49}$$

$$q_t^{OF}\left(z_t^{ON}, z_t^{OF}\right) = \sum_{m=1}^{M^{OF}-1} IL_{m,t}^{OF} + \sum_{l=1}^{L^{OF}-1} OT_{l,t}^{OF} - z_t^{OF} + z_t^{ON} \tag{2.50}$$

In addition to the above acceleration method, we implement *reward shaping* (RS) to define the more appropriate reward function to maximize the average profit. The RS is a technique to incorporate the exterior knowledge of a teacher heuristic into RL; thus, agents are guided towards more promising policies [36, 153]. In this research, we employ the SQmax policy as a teacher heuristic. The two same environments are declared; one for the RL, $ENV_{RL}$, and the other for a teacher heuristic, the SQmax policy, $ENV_{SQmax}$. At each time step of the training process, the values of realized demand in two environments are equal However, the current state and next state are different (i.e., $(s_t, s_{t+1})$ is obtained from $ENV_{RL}$, and $(\hat{s}_t, \hat{s}_{t+1})$ is obtained from $ENV_{SQmax}$). Even though several methods exist in the RS research area, we could get better solutions by just subtracting the profit of the SQmax policy from the profit of the RL as follows:

$$r(s_t, a_t) = PF_t^{RL} - PF_t^{SQmax} \tag{2.51}$$

where $PF_t^{RL}$ is obtained profit at period $t$ by the RL approach, and $PF_t^{SQmax}$ is obtained profit at period $t$ by the SQmax policy. Intuitively, the value of $PF_t^{SQmax}$ is used as the criteria for assessing the quality of decisions implemented by RL at period $t$. The hybrid DRL approach, SACDPE combining the SQLT policy and RS (SACDPE+SQLT+RS), is presented in Algorithm 1.

**Algorithm 1** SACDPE+SQLT+RS

Initialize $Q^{soft}_{\theta_1} : \mathcal{S} \to \mathbb{R}^{|\mathcal{A}|},\ Q^{soft}_{\theta_2} : \mathcal{S} \to \mathbb{R}^{|\mathcal{A}|},\ \pi_\phi : \mathcal{S} \to [0,1]^{|\mathcal{A}|}$
Initialize $Q^{soft}_{\bar{\theta}_1} : \mathcal{S} \to \mathbb{R}^{|\mathcal{A}|},\ Q^{soft}_{\bar{\theta}_2} : \mathcal{S} \to \mathbb{R}^{|\mathcal{A}|},\ \mathcal{D} \leftarrow \emptyset$
$\bar{\theta}_1 \leftarrow \theta_1,\ \bar{\theta}_2 \leftarrow \theta_2$
Declare the environment for SACDPE ($ENV_{RL}$)
Declare the environment for SQmax policy ($ENV_{SQmax}$)
$e \leftarrow 1$
**for** *each episode* $e = 1, \cdots, E$ **do**
  $t \leftarrow 1$
  **for** *each timestep* $t = 1, \cdots, T$ **do**
    Observe $s_t$ and choose action for transshipment $a_t \sim \pi_\phi(\cdot|s_t)$
    Determine $y^{ON}_t$ and $y^{OF}_t$ by SQLT policy
    Observe $PF^{RL}_t$ and $s_{t+1}$ from $ENV_{RL}$
    Observe state $\hat{s}_t$
    Determine $y^{ON}_t$ and $y^{OF}_t$ by SQmax policy
    Observe $PF^{SQmax}_t$ and $\hat{s}_{t+1}$ from $ENV_{SQmax}$
    $r_t = PF^{RL}_t - PF^{SQmax}_t$
    $\mathcal{D} \leftarrow \mathcal{D} \cup \{(s_t, a_t, r_t, s_{t+1})\}$ with maximal priority $p_t = \max_{i<t} p_i$
    Sample a mini-batch $\mathcal{B}$ from $\mathcal{D}$ according to probability $P(d) = p^\eta_d / \sum_{k=1}^{N_\mathcal{D}} p^\eta_k, \forall d \in \mathcal{D}$
    $\Delta\theta_1, \Delta\theta_2, \Delta\phi, \Delta\alpha = 0$
    **for** $b \in \mathcal{B}$ **do**
      $w_b = \left(\frac{1}{N_\mathcal{D}} \times \frac{1}{P(d)}\right)^\beta / \max_{i\in\mathcal{B}} w_i$
      $|\delta_b| = \min\left\{\left(Q^{soft}_{\theta_1}(s) - (r + \gamma V_{\bar{\theta}}(s'))\right)^2, \left(Q^{soft}_{\theta_2}(s) - (r + \gamma V_{\bar{\theta}}(s'))\right)^2\right\}$
      $p_b \leftarrow |\delta_b| + \epsilon_{per}$
      $\Delta\theta_i \leftarrow \Delta\theta_i + w_b \nabla_{\theta_i} J_{Q^{soft}}(\theta_i)$, for $i \in \{1, 2\}$
      $\Delta\phi \leftarrow \Delta\phi + w_b \nabla_\phi J_\pi(\phi)$
      $\Delta\alpha \leftarrow \Delta\alpha + \nabla_\alpha J(\alpha)$
    **end**
    Update soft Q networks $\theta_i \leftarrow \theta_i - \lambda\Delta\theta_i$, for $i \in \{1, 2\}$
    Update policy network $\phi \leftarrow \phi - \lambda\Delta\phi$
    Adjust temperature $\alpha \leftarrow \alpha - \lambda\Delta\alpha$
    Update target soft Q networks $\bar{\theta}_i \leftarrow \psi\theta_i + (1-\psi)\bar{\theta}_i$, for $i \in \{1, 2\}$
    $t \leftarrow t + 1$
  **end**
  $e \leftarrow e + 1$
**end**
**Return:** $\theta_1,\ \theta_2,\ \pi_\phi$

## 2.5 Computational experiments

Throughout this section, we conduct three types of computational experiments to address research questions 1, 2, and 3. In Section 2.5.1, we evaluate the performance of the developed hybrid DRL approach by comparing it with existing algorithms on the real-world demand data set. In Section 2.5.2, we demonstrate the advantages of transshipment in the OOCS by examining different types of demand and varying the unit transshipment cost parameter. In Section 2.5.3, we examine the outdating costs associated with online and offline channels, respectively. On the basis of the results of the experiment, we suggest several managerial insights in Section 2.5.4. All computational experiments were implemented on a PC with an AMD Ryzen 7 PRO 4750G with a Radeon Graphics 3.60GHz processor and 16GB of RAM with Windows 10 64-bit. In addition, all experiments for the hybrid DRL approach were coded in Python 3.8 and Pytorch 1.12.1.

### 2.5.1 Performance analysis of the developed hybrid DRL approach for real-world data set

Four types of experiments were conducted in this section to validate the hybrid DRL approach, SACDPE+SQLT+RS, with the following purposes:

1. Validating the effects of accelerating approaches by examining learning curves
2. Analyzing the robustness of hybrid DRL's policies on various test instances
3. Analyzing the robustness of hybrid DRL when training multiple times
4. Comparing the hybrid DRL to existing approaches in terms of optimality gap

Except for the first experiment, we reported experimental results by adopting the average profit per period as a performance measure.

By referring the cost parameters in De Moor et al. [36], we set the $c_p = 5$, $c_h = 1$, $c_w = 10$, and $c_o = 3$ for all instances. In this section, we consider the negligible transshipment cost, $c_l = 0$. The sale price in each channel $p^{ON}$ and $p^{OF}$ are set as 10 and 8 by accommodating a property that the sale price online is more expensive than offline. To address the practical size problem, we set the $M^{ON} = 5$, $M^{OF} = 7$, $L^{ON} = 2$, and $L^{OF} = 3$, which determines the dimensions of state. Moreover, we set the $y^{ON}_{max} = 10$, $y^{OF}_{max} = 10$, $z^{ON}_{max} = 5$, and $z^{OF}_{max} = 5$, which determines the size of action spaces. We consider 10,000 periods for the planning horizon to reflect the infinite horizon setting.

In this section, we examine the real-world data set presented in Oroojlooyjadid et al. [104] and Kaggle [74] for demands in online and offline channels, and the demand data set is used directly for training the DRL approach without any assumptions about demand distribution. It is worth to note that we address the different problem with the Oroojlooyjadid et al. [104]. The reasons why we utilize the same data set used in Oroojlooyjadid et al. [104] are presented in Appendix A.3. A total of six instances are considered based on three different types of demand (i.e., Category A, B, and C) for each channel. In each instance, there are two types of demand data sets: training data and test data. As presented in Figure 2.4, the training data consists of 5,000 episodes, and the test data consists of 20 episodes, and each episode contains the demand information within the planning horizon (i.e., 10,000 periods). The training data is utilized for the training process of DRL, and the test data is utilized to assess the performance of the developed algorithms.

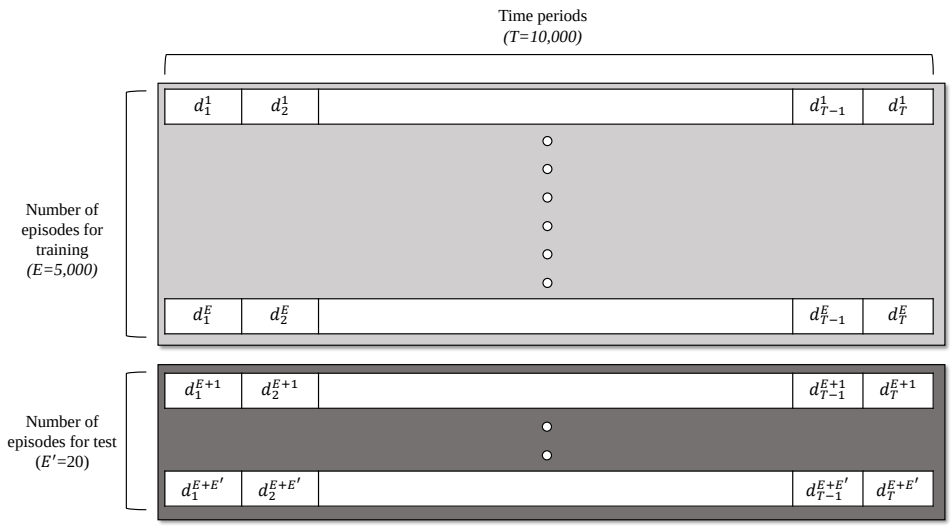



Figure 2.4: Data description.

It has been found that the VI algorithm has not been able to obtain optimal policy as a result of the high complexity of the proposed problem. Therefore, to evaluate the quality of the solution systematically, we used the average profit obtained from the optimal objective value of the integer programming (IP) model under perfect information conditions (i.e., the deterministic demand setting). Due to the fact that the demand for the planning horizon is already known in advance, all costs, with the exception of the ordering cost, are close to zero. Based on Dehghani et al. [38], we developed the IP model using transshipment and replenishment as decision variables. The average profit obtained from perfect information is regarded as the upper bound. We used Python 3.8 and the FICO Xpress Optimizer library to solve the IP model.

In the first experiment, we compared four DRL algorithms: SACDPE, SACDPE+RS, SACDPE+SQLT, and SACDPE+SQLT+RS. The SACDPE did not accommo-

date any proposed acceleration methods, and SACDPE+SQLT and SACDPE+RS adopted SQLT policy and RS, respectively. The SACDPE+RS+SQLT accommodated both acceleration methods. In order to analyze the robustness of DRL algorithms, each algorithm was implemented five times, which means that five actor neural networks $\pi_\phi$ per instance were trained with random weight initialization. We informally conducted hyperparameter tuning instead of conducting the advanced search proposed by previous studies [52]. We determined the values of hyperparameters referring to the setting of the related studies [119, 140, 130]. All experiments were conducted using the same values of hyperparameters as stated in Table A.1. It should be noted that the SAC algorithm, which is the base algorithm of our study, has the advantage of mitigating the brittleness of hyperparameter tuning compared to other RL algorithms [60, 61]. We implemented extensive experiments with varying hyperparameter values, but the performance was not affected significantly. In particular, the SACDPE+SQLT+RS was most robust to different hyperparameter values compared to other algorithms.

Figure 2.5 depicts the learning curves of different DRL algorithms in the training process. The shaded areas around learning curves describe a 95 percent confidence interval for five multiple runs. We trained each DRL algorithm for 5,000 episodes, and one episode consists of 100 time periods. The SACDPE obtained the worst average profit and had the widest confidence interval among every DRL algorithm. The SACDPE+RS shows more stability during training than the SACDPE. Also, the average profit of SACDPE+RS converges to a higher value than SACDPE. The SACDPE+RS requires more than 2,000 episodes for convergence of average profit. On the other hand, SACDPE+SQLT can learn a promising policy within

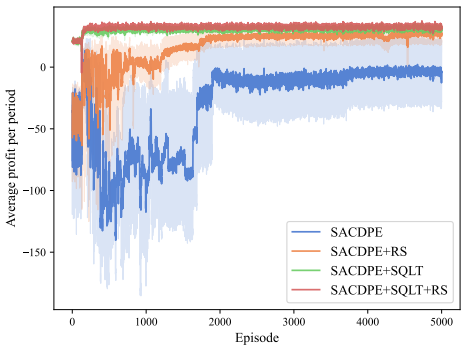
500 episodes. The SACDPE+SQLT+RS also requires about 500 episodes for convergence and obtains the best policy resulting in the highest profit among every DRL algorithm. Because the action space size was reduced by adopting SQLT, the SACDPE+SQLT and SACDPE+SQLT+RS could learn a promising policy within relatively short episodes compared to SACDPE and SACDPE+RS.

In the second experiment, we evaluated the five policies derived from the hybrid DRL algorithm. Each policy was tested for 20 different episodes in test demand data sets. The sample mean and standard deviation of 20 runs were computed for performance measures. As shown in Table 2.2, the sample standard deviation of the five different policies had a relatively low value for 20 episodes in test demand data. These results indicate that policies derived from the hybrid DRL are robust to various test instances.

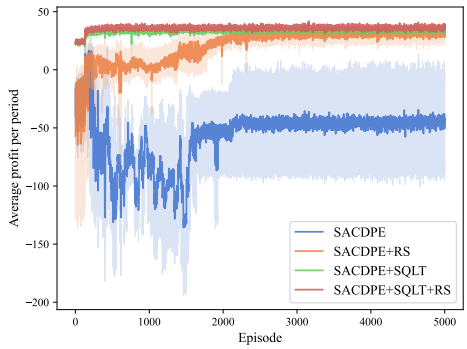
Table 2.2: Comparison between five policies of hybrid DRL on 20 test instances

| Instance | Demand category | | Policy 1 | | Policy 2 | | Policy 3 | | Policy 4 | | Policy 5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Online | Offline | Mean | Std | Mean | Std | Mean | Std | Mean | Std | Mean | Std |
| 1 | A | B | 34.91 | 0.13 | 34.80 | 0.13 | 34.90 | 0.13 | 34.79 | 0.13 | 34.93 | 0.13 |
| 2 | A | C | 37.98 | 0.10 | 38.07 | 0.11 | 37.91 | 0.11 | 38.00 | 0.10 | 38.21 | 0.10 |
| 3 | B | A | 33.76 | 0.14 | 33.88 | 0.14 | 33.70 | 0.13 | 33.74 | 0.14 | 34.08 | 0.13 |
| 4 | B | C | 35.22 | 0.12 | 35.41 | 0.13 | 35.26 | 0.12 | 35.30 | 0.13 | 35.65 | 0.12 |
| 5 | C | A | 38.55 | 0.10 | 38.63 | 0.10 | 38.61 | 0.09 | 38.61 | 0.10 | 38.61 | 0.10 |
| 6 | C | B | 36.90 | 0.12 | 37.02 | 0.11 | 36.94 | 0.12 | 36.92 | 0.12 | 36.81 | 0.12 |

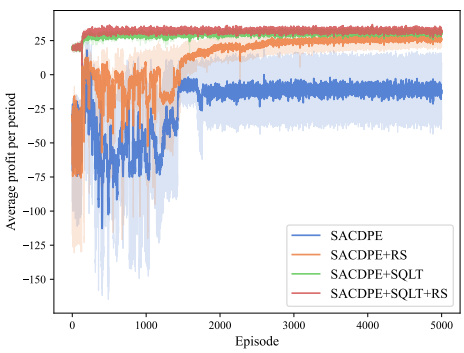
In the third experiment, we analyzed the robustness of hybrid DRL when training multiple times. We compared the performance of hybrid DRL (SACDPE+SQLT+RS) with other DRL algorithms (i.e., SACDPE, SACDPE+RS, and SACDPE+SQLT). We reported the sample mean and standard deviation of five policies for each DRL algorithm. In particular, the result of hybrid DRL could be obtained by computing the sample mean and standard deviation of results in Table 2.2. As presented in



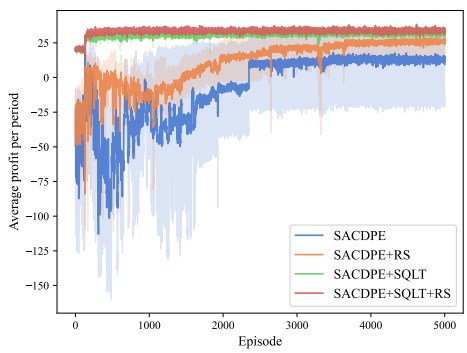

(a) Instance 1



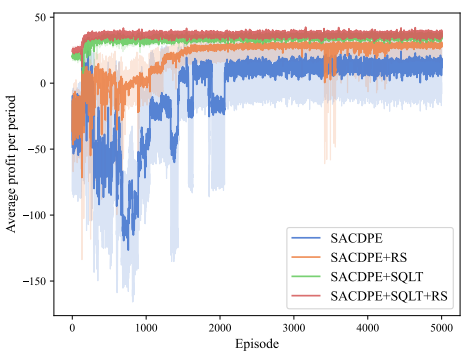

(b) Instance 2



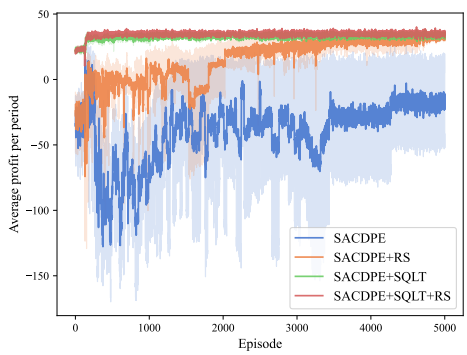

(c) Instance 3




(d) Instance 4




(e) Instance 5




(f) Instance 6

Figure 2.5: Learning curves of different DRL algorithms in training process.

Table 2.3, the hybrid DRL had the smallest value for 'Std' compared to other algorithms. Also, the SACDPE+SQLT had a small value of 'Std' compared to other two algorithms; thus, adopting SQLT as an acceleration method could enhance the robustness of training.

Table 2.3: Comparison between DRL algorithms when training multiple times

| Instance | Demand category | | SACDPE | | SACDPE+RS | | SACDPE+SQLT | | SACDPE+SQLT+RS | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Online | Offline | Mean | Std | Mean | Std | Mean | Std | Mean | Std |
| 1 | A | B | -3.65 | 32.22 | 28.76 | 7.98 | 33.34 | 0.40 | 34.87 | 0.07 |
| 2 | A | C | -47.86 | 57.90 | 32.88 | 5.69 | 36.27 | 0.25 | 38.03 | 0.11 |
| 3 | B | A | -11.13 | 27.11 | 28.11 | 5.52 | 32.51 | 0.13 | 33.83 | 0.16 |
| 4 | B | C | 13.08 | 35.52 | 27.32 | 10.86 | 33.91 | 0.47 | 35.37 | 0.17 |
| 5 | C | A | 13.31 | 25.13 | 30.78 | 10.75 | 37.41 | 0.32 | 38.60 | 0.03 |
| 6 | C | B | -18.98 | 40.88 | 34.38 | 0.79 | 35.44 | 0.41 | 36.92 | 0.07 |

In the fourth experiment, we compared the hybrid DRL to existing approaches in terms of optimality gap. We utilized two existing approaches that do not require prior knowledge of demand distribution. We began by adopting RH-TSSP presented by Dehghani et al. [38]. We revised the TSSP model in Dehghani et al. [38] to be suitable for our proposed problem, and the training data set was used for scenario samples. Like the developed DRL algorithm, RH-TSSP considers replenishment and transshipment decisions simultaneously. Second, we adopted three ordering policies based on an estimate of product waste to the order quantity presented in Haijema and Minner [64]: BSP-EW, BSPlow-EW, and SQmax-EW. Because only replenishment is considered as a decision in these three ordering policies, the order quantity in each channel is determined separately by predefined ordering policies. Among these policies, we reported the SQmax-EW, which showed the best performance.

In Table 2.4, we can see how existing approaches perform in comparison with each other. Even though RH-TSSP considers transshipment as a decision, the per-

formance of RH-TSSP was poorer than the performance of other ordering policies. The solution quality of RH-TSSP cannot be guaranteed because RH-TSSP is also one of the approximation methods for solving the problem with a long planning horizon. SQmax-EW outperformed RH-TSSP in terms of optimality gap. However, we observed that the developed hybrid DRL approach, SACDPE+SQLT+RS, outperformed SQmax-EW and RH-TSSP for every performance measure.

Table 2.4: Comparison between hybrid DRL and existing approaches

| Instance | Demand category | | Perfect | RH-TSSP[b] | | | SQmax-EW[c] | | | SACDPE+SQLT+RS | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Online | Offline | | Mean | Std | Gap[a] | Mean | Std | Gap[a] | Mean | Std | Gap[a] |
| 1 | A | B | 42.05 | 32.77 | 0.06 | 22.06 | 33.67 | 0.11 | 19.92 | 34.87 | 0.07 | **17.07** |
| 2 | A | C | 45.21 | 34.13 | 0.10 | 24.51 | 36.37 | 0.11 | 19.55 | 38.03 | 0.11 | **15.87** |
| 3 | B | A | 41.26 | 31.01 | 0.07 | 24.83 | 32.97 | 0.12 | 20.10 | 33.83 | 0.16 | **18.00** |
| 4 | B | C | 42.46 | 31.63 | 0.06 | 25.50 | 34.41 | 0.11 | 18.95 | 35.37 | 0.17 | **16.70** |
| 5 | C | A | 45.69 | 35.17 | 0.10 | 23.03 | 36.89 | 0.12 | 19.25 | 38.60 | 0.03 | **15.51** |
| 6 | C | B | 43.72 | 34.64 | 0.07 | 20.77 | 35.64 | 0.12 | 18.49 | 36.92 | 0.07 | **15.56** |

[a] Gap: (Perfect−Mean)×100/Perfect
[b] Refer to Dehghani et al. [38]
[c] Refer to Haijema and Minner [64]

In terms of computational efficiency, SQLT reduces the computational burden of the training process. DRL algorithms without SQLT (SACDPE and SACDPE+RS) required about twelve hours to implement 5,000 episodes. In contrast, DRL algorithms that employ SQLT as an acceleration method (SACDPE+SQLT and SACDPE +SQLT+RS) required approximately four hours to implement the same number of episodes. Despite the fact that DRL algorithms require several hours to train the first time, the trained neural networks can be reused and tested on a variety of demand datasets in less than a second. On the other hand, in the case of RH-TSSP, the TSSP model is solved at every period because the algorithm is based on the rolling horizon approach. Thus, RH-TSSP required solving the TSSP model 10,000

times to test on one dataset, and it has low computational efficiency because it takes three and half hours to implement one time.

### 2.5.2 Advantages of transshipment on profit in the OOCS

In this section, we aim to analyze the advantages of transshipment by comparing it with no-transshipment policy. We adopt the SACDPE+SQLT+RS for a transshipment policy and the SQmax-EW for a no-transshipment policy. We set $L^{ON} = 3,\ L^{OF} = 3,\ y^{ON}_{max} = 20$, and $y^{OF}_{max} = 20$, and other parameters are equal to the parameter setting in Section 2.5.1. It should be noted that $L^{ON}$ and $L^{OF}$ were set as the same value because the differences between the lead time of online and offline channels could affect the average profit. The effectiveness of transshipment was evaluated by varying three key factors: (1) demand variability, (2) unit transshipment cost $c_l$, and (3) shelf life of product held in online and offline channels, $M^{ON}$ and $M^{OF}$.

To begin with, we compare the transshipment and no-transshipment policies in terms of average profit for different types of demand. In this experiment, we assume that the transshipment cost is negligible. To demonstrate the effects of transshipment based on demand variability, we generated nine demand data sets as shown in Table 2.5. The determined parameters of the distributions are intended to cover cases of low, medium and high demand variability for each discrete probability distribution. For each demand data set, we trained the SACDPE+SQLT+RS for 2,500 episodes three times (i.e., three actor neural networks). Based on the performance of three neural networks, the neural network that exhibited the most promising performance was selected for analysis.

Table 2.5: Information about distributions utilized to generate demand data sets

| | Distributions | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Discrete uniform | | | Poisson | | | Negative binomial | | |
| Notation | $U\{a_i, b_i\}$ | | | $Pois(\lambda_i)$ | | | $NB(n_i, p_i)$ | | |
| Parameters | $a_i = 0$ | | | $\lambda_i = 5$ | $\lambda_i = 8$ | $\lambda_i = 10$ | $p_i = 0.5$ | | |
| | $b_i = 10$ | $b_i = 16$ | $b_i = 20$ | | | | $n_i = 5$ | $n_i = 8$ | $n_i = 10$ |
| Mean ($\mu$) | 5 | 8 | 10 | 5 | 8 | 10 | 5 | 8 | 10 |
| Variance ($\sigma^2$) | 10 | 24 | 36.67 | 5 | 8 | 10 | 10 | 16 | 20 |
| CV[a] | 0.63 | 0.61 | 0.60 | 0.45 | 0.35 | 0.32 | 0.63 | 0.50 | 0.45 |

[a] CV: coefficient of variation ($\sigma/\mu$)

Table 2.6 shows a comparison of average profits per period with respect to a transshipment policy compared to a no-transshipment policy. For every nine demand data sets, transshipment between online and offline channels resulted in a higher profit than no-transshipment. We use the performance measure 'Gap(diff)' to evaluate the effectiveness of a transshipment policy compared to a no-transshipment policy (i.e., the higher value of Gap(diff) represents that transshipment is more effective than no-transshipment). Under conditions of equal means, the variances of Uniform, Negative binomial, and Poisson distributions are listed in descending order. Also, the Gap(diff) of Uniform, Negative binomial, and Poisson follows the same descending order, which indicates that transshipment is more effective when the variance of demand is greater.

Table 2.6: Average profit of transshipment and no-transshipment policies for different types of demand data sets

| | Average profit per period | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $U\{0, 10\}$ | $U\{0, 16\}$ | $U\{0, 20\}$ | Pois(5) | Pois(8) | Pois(10) | NB(5,0.5) | NB(8,0.5) | NB(10,0.5) |
| No-transshipment | 30.41 | 51.30 | 64.75 | 42.25 | 74.77 | 97.05 | 31.93 | 63.10 | 84.43 |
| Transshipment | 34.31 | 56.89 | 71.62 | 45.70 | 79.31 | 101.99 | 35.39 | 68.11 | 90.64 |
| Gap(diff)[a] | **12.83** | **10.90** | **10.61** | **8.16** | **6.07** | **5.09** | **10.86** | **7.94** | **7.35** |

[a] Gap(diff): (Transshipment − No-transshipment) × 100/No-transshipment

In addition, we examined the correlation between Gap(diff) and demand variability for each distribution. Referring to several existing studies, we utilized the coefficient of variation (CV) to measure demand variability [131]. For every distribution, Figure 2.6 shows a tendency that the higher the value of CV, the more effective the transshipment is.

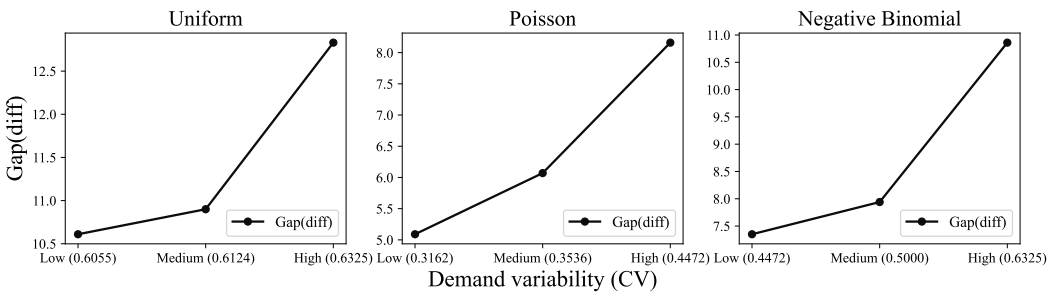



Figure 2.6: Correlation between Gap(diff) and demand variability for three demand distributions.

From the perspective of revenue and inventory holding, lost sales, outdating, and ordering costs, Figure 2.7 illustrates how transshipment improves profitability. The y-axis represents the share of the total improvement according to different components (i.e., improvement percentage). Among all the components of the cost structure, transshipment contributes the most to reducing the outdating cost for every demand distribution. Also, the improvement percentage of inventory holding cost accounts for a relatively large share of total improvement. Due to the fact that the transshipment is implemented before the demand is realized, the holding cost can be saved instead of the transshipment cost, as shown in Equation (2.18).

In the second experiment, a sensitivity analysis on the unit transshipment cost parameter, $c_l$, was implemented. For ease of the expositions, we only consider a demand data set generated from $\mathrm{U}\{0, 20\}$. As with experiments for different demand

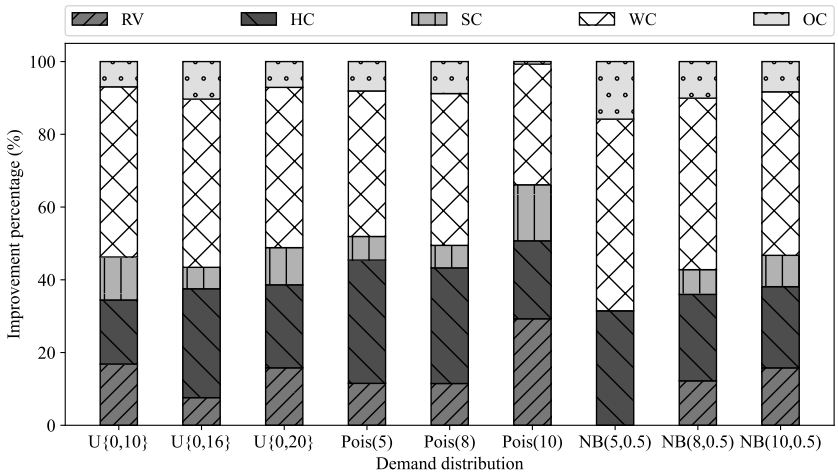



Figure 2.7: Share of total improvement for revenue and each cost component for different demand distributions.

data sets, we also trained the SACDPE+SQLT+RS for 2,500 episodes three times for every value of $c_l$, and we utilized the best one to assess the average profit of the transshipment policy. Table 2.7 shows the average profit of transshipment and no-transshipment policies varying the value of $c_l$. Transshipment can contribute to an increase in the average profit compared to no-transshipment when the $c_l$ is between zero and seven. However, there is no advantage to using transshipment for maximizing profit if the $c_l$ is bigger than seven.

Table 2.7: Average profit of transshipment and no-transshipment policies varying the unit transshipment cost parameter $c_l$

| | Average profit per period | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $c_l = 0$ (0.00)[b] | $c_l = 1$ (0.33)[b] | $c_l = 2$ (0.67)[b] | $c_l = 3$ (1.00)[b] | $c_l = 4$ (1.33)[b] | $c_l = 5$ (1.67)[b] | $c_l = 6$ (2.00)[b] | $c_l = 7$ (2.33)[b] | $c_l = 8$ (2.67)[b] | $c_l = 9$ (3.00)[b] |
| No-transshipment | 64.75 | | | | | | | | | |
| Transshipment | 71.62 | 68.70 | 67.01 | 66.15 | 65.40 | 65.21 | 65.01 | 64.81 | 64.75 | 64.75 |
| Gap(diff)[a] | **10.61** | **6.10** | **3.49** | **2.16** | **1.00** | **0.71** | **0.40** | **0.09** | **0.00** | **0.00** |

[a] Gap(diff): (Transshipment − No-transshipment) × 100/No-transshipment
[b] $(c_l/c_o)$: the ratio of the unit transshipment cost to the unit order cost parameters

Based on the experiment results in Table A.2, we depict Figure 2.8 to show the improvement effect of transshipment on average profit varying the value of $c_l$. The sum of the improvement of revenue and cost will be the same as the improvement of profit. Similar to the results of Figure 2.7, the transshipment improves the outdating cost the most compared to revenue and other cost components. If the transshipment cost is non-negligible ($c_l > 0$), transshipment could not contribute to saving inventory holding cost significantly, unlike the results of the first experiment considering negligible transshipment cost ($c_l = 0$). When the unit transshipment and inventory holding cost parameter is equal ($c_l = 1$), the inventory holding cost is reduced. However, in the case that $c_l$ is bigger than $c_h$, the effect of transshipment to reduce the inventory holding cost was insignificant.

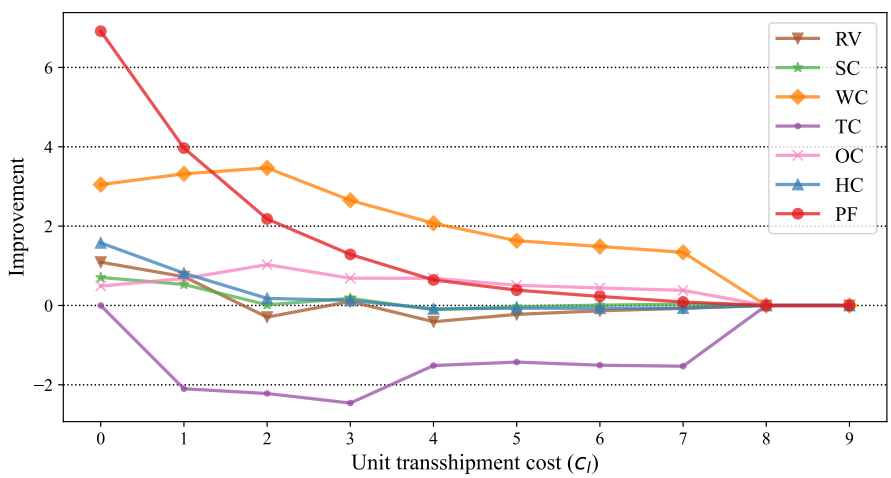



Figure 2.8: Improvement effect of transshipment on average profit varying the unit transshipment cost $c_l$.

In the third experiment, a sensitivity analysis on the shelf life of online and offline channels, $M^{ON}$ and $M^{OF}$, was implemented. The experiment setting is equivalent to the second experiment except that the value of $c_l$ is zero. Table 2.8 shows the

Table 2.8: Average profit of transshipment and no-transshipment policies varying the shelf life of product held in online and offline channels, $M^{ON}$ and $M^{OF}$

| | Short shelf life | | | Long shelf life | | |
|---|---|---|---|---|---|---|
| $M^{ON}$ | 3 | 3 | 3 | 5 | 5 | 5 |
| $M^{OF}$ | 5 | 6 | 7 | 7 | 8 | 9 |
| $M^{OF} - M^{ON}$ | 2 | 3 | 4 | 2 | 3 | 4 |
| No-transshipment | 64.85 | 66.35 | 67.07 | 78.51 | 78.81 | 78.89 |
| Transshipment | 71.54 | 73.73 | 74.35 | 85.61 | 86.10 | 85.96 |
| Gap (diff) [a] | **10.32** | **11.13** | **10.86** | **9.03** | **9.24** | **8.96** |

[a] Gap(diff): (Transshipment − No-transshipment) × 100/No-transshipment

average profit of transshipment and no-transshipment policies by varying the value of $M^{ON}$ and $M^{OF}$. To analyze the effects of shelf life on the profit, we define two settings for the shelf life: 'Short shelf life' ($M^{ON} = 3, M^{OF} = 5, 6, 7$) and 'Long shelf life' ($M^{ON} = 5, M^{OF} = 7, 8, 9$). The transshipment was more effective in the average profit in the setting of short shelf life than the long shelf life. Also, if the difference in the shelf life between channels was slight (i.e., $M^{OF} - M^{ON} = 2$), the positive effect of transshipment was insignificant compared to the case where the difference was more considerable (i.e., $M^{OF} - M^{ON} = 3, 4$). In contrast, in the setting of long shelf life, the variation of Gap (diff) was insignificant even though the value of $M^{OF} - M^{ON}$ was changed. These results could be expected because if the $M^{OF} - M^{ON}$ was equal to two, the transshipped product had a high risk of being outdated as indicated in Table A.3 because of the non-negligible transshipment time.

### 2.5.3 Analysis for saving effect of outdating cost because of transshipment and heterogeneous shelf life

Observing the experiment results in Section 2.5.2, we found that transshipment reduces the outdating cost compared to a no-transshipment policy. To examine the impacts of transshipment on the OOCS, we first analyzed how many products were transshipped from one channel to the other. As a consequence, we identified the outdating cost that can be saved for each channel by using the transshipment on the OOCS. The same experiment setting is applied in Section 2.5.2 for the different types of demand and different values of $c_l$ that will be analyzed.

Figure 2.9 illustrates boxplots of the transshipment quantity according to different types of demand. The planning horizon of 10,000 periods results in 10,000 samples per boxplot. We add a mark for the mean values on boxplots using the white circle. For every type of demand, more products were transshipped from the online channel to the offline channel than transshipped from the offline channel to the online channel. As indicated in Table 2.6, there was no significant difference between the mean values of two types of transshipment quantities for the demand of Poisson distribution, where transshipment is the least effective among three distribution types. In contrast, we can observe that the mean value gap was large for uniform and negative binomial distributions, in which the transshipment is effective due to a high degree of demand variability.

Figure 2.10 presents boxplots of transshipment quantity for varying $c_l$ values generated from the demand data set $\mathrm{U}\{0, 20\}$. As the value of $c_l$ increases, the total transshipment quantity decreases due to a high transshipment cost. If the value of $c_l$ is smaller than eight, more products were transshipped from the online channel to

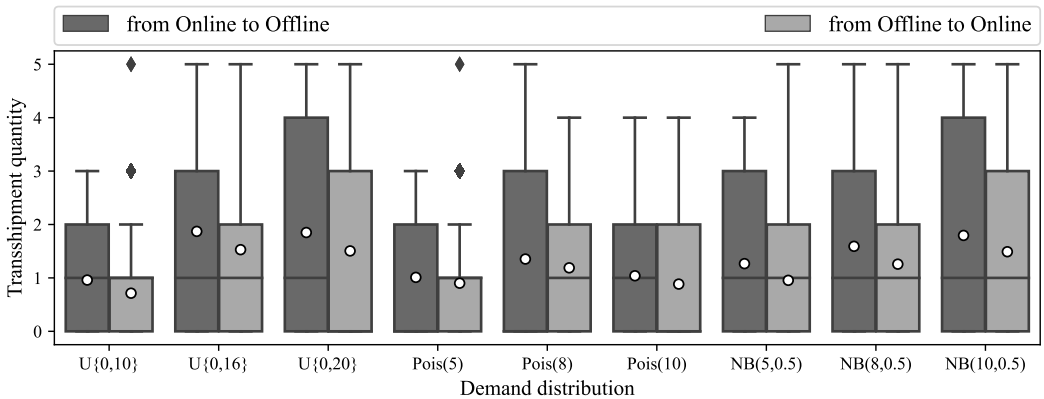



Figure 2.9: Boxplots of transshipment quantity for different types of demand.

the offline channel than transshipped from the offline channel to the online channel. When the value of $c_l$ is larger than eight, the transshipment did not occur in both channels. Consequently, the average profit of transshipment and no-transshipment policies is equal, as shown in Table 2.7.

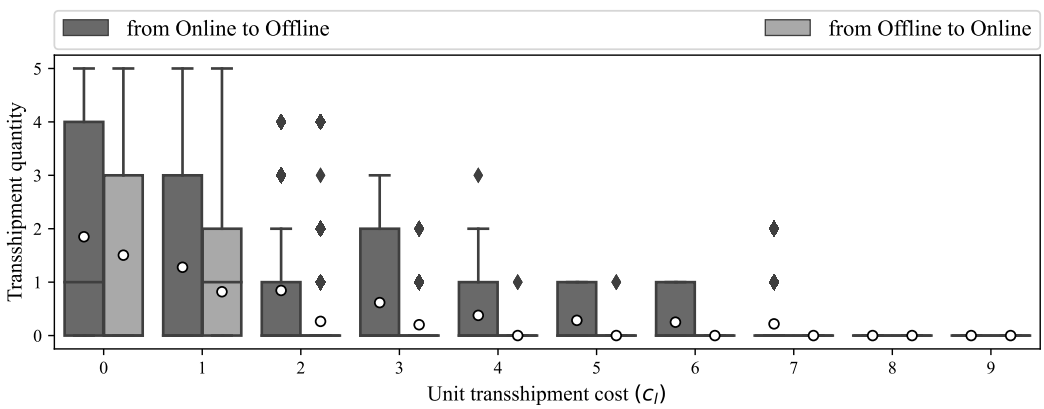



Figure 2.10: Boxplots of transshipment quantity for varying the value of unit transshipment cost parameter $c_l$.

We inspect the outdating cost of each channel for transshipment and no-transshipment policies through Tables 2.9 and 2.10. For every experiment setting, we could observe

that the outdating cost increased in the offline channel when utilizing transshipment compared to no-transshipment. However, the outdating cost decreased substantially in the online channel compared to the offline channel; thus, the transshipment can save the outdating cost from the standpoint of the total system. It has been found that this tendency is the result of two different factors:

- The shelf life of the online channel is shorter than that of the offline channel (heterogeneous shelf life property).
- It is evident from Figures 2.9 and 2.10 that a greater number of products are transshipped from an online channel to an offline channel, rather than from an offline channel to an online channel.

We observe that products that must be discarded in the online channel were transshipped to the offline channel. Most of them were used to satisfy demand in the offline channel, and few products were abandoned in the offline channel. As can be observed from Figure 2.11, in the case of this experiment setting, $M^{ON} = 3$ and $M^{OF} = 5$, it is necessary to discard the product in the online channel once the product reaches the age of three. However, this product can be used in the offline channel because the product will be disposed of when the age is five in the offline channel. Therefore, when utilizing transshipment in the OOCS, we found that the offline channel, which has a longer shelf life, plays the role of making good use of the old product that will be discarded in the online channel if not transshipped to the offline channel.

Table 2.9: Analysis about the outdating cost of each channel in different types of demand data sets

| | Channel | U$\{0,10\}$ | U$\{0,16\}$ | U$\{0,20\}$ | Pois(5) | Pois(8) | Pois(10) | NB(5,0.5) | NB(8,0.5) | NB(10,0.5) |
|---|---|---|---|---|---|---|---|---|---|---|
| $WC_{NT}$[a] | ON | 5.85 | 8.76 | 11.54 | 2.44 | 2.55 | 1.95 | 4.58 | 4.43 | 5.02 |
| | OF | 1.09 | 1.71 | 2.34 | 0.24 | 0.03 | 0.01 | 0.85 | 0.59 | 0.34 |
| | ON+OF | 6.95 | 10.47 | 13.88 | 2.68 | 2.57 | 1.95 | 5.43 | 5.02 | 5.35 |
| $WC_T$[b] | ON | 3.50 | 5.00 | 7.49 | 0.89 | 0.55 | 0.37 | 1.97 | 1.42 | 1.73 |
| | OF | 1.78 | 2.90 | 3.33 | 0.46 | 0.08 | 0.03 | 1.66 | 1.19 | 0.77 |
| | ON+OF | 5.27 | 7.90 | 10.83 | 1.35 | 0.63 | 0.40 | 3.63 | 2.61 | 2.50 |
| Diff[c] | ON | **2.36** | **3.76** | **4.05** | **1.55** | **2.00** | **1.58** | **2.61** | **3.01** | **3.29** |
| | OF | **-0.68** | **-1.19** | **-1.00** | **-0.22** | **-0.06** | **-0.02** | **-0.81** | **-0.60** | **-0.43** |
| | ON+OF | 1.67 | 2.57 | 3.05 | 1.33 | 1.94 | 1.55 | 1.80 | 2.41 | 2.85 |
| Saving(%)[d] | ON+OF | **24.07** | **24.57** | **21.96** | **49.70** | **75.48** | **79.55** | **33.20** | **47.96** | **53.32** |

[a] $WC_{NT}$ : Outdating cost of no-transshipment policy (SQmaxEW)
[b] $WC_T$ : Outdating cost of transshipment policy (SACDPE+SQLT+RS)
[c] Diff: Effects of transshipment on outdating cost ($WC_{NT} - WC_T$)
[d] Saving(%) : $(WC_{NT} - WC_T) \times 100/WC_{NT}$

Table 2.10: Analysis about the outdating cost of each channel varying the unit transshipment cost parameter $c_l$

| | Channel | $c_l = 0$ | $c_l = 1$ | $c_l = 2$ | $c_l = 3$ | $c_l = 4$ | $c_l = 5$ | $c_l = 6$ | $c_l = 7$ | $c_l = 8$ | $c_l = 9$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $WC_{NT}$[a] | ON | 11.54 | | | | | | | | | |
| | OF | 2.34 | | | | | | | | | |
| | ON+OF | 13.88 | | | | | | | | | |
| $WC_T$[b] | ON | 7.49 | 7.32 | 7.30 | 8.33 | 8.96 | 9.56 | 9.77 | 9.99 | 11.54 | 11.54 |
| | OF | 3.33 | 3.24 | 3.11 | 2.90 | 2.85 | 2.69 | 2.61 | 2.54 | 2.34 | 2.34 |
| | ON+OF | 10.83 | 10.56 | 10.41 | 11.22 | 11.80 | 12.24 | 12.39 | 12.54 | 13.88 | 13.88 |
| Diff[c] | ON | **4.05** | **4.22** | **4.24** | **3.21** | **2.58** | **1.98** | **1.77** | **1.55** | **0.00** | **0.00** |
| | OF | **-1.00** | **-0.90** | **-0.77** | **-0.56** | **-0.51** | **-0.35** | **-0.28** | **-0.21** | **0.00** | **0.00** |
| | ON+OF | 3.05 | 3.32 | 3.47 | 2.65 | 2.07 | 1.63 | 1.49 | 1.34 | 0.00 | 0.00 |
| Saving(%)[d] | ON+OF | **21.96** | **23.92** | **24.99** | **19.11** | **14.94** | **11.76** | **10.72** | **9.66** | **0.00** | **0.00** |

[a] $WC_{NT}$ : Outdating cost of no-transshipment policy (SQmaxEW)
[b] $WC_T$ : Outdating cost of transshipment policy (SACDPE+SQLT+RS)
[c] Diff: Effects of transshipment on outdating cost ($WC_{NT} - WC_T$)
[d] Saving(%) : $(WC_{NT} - WC_T) \times 100/WC_{NT}$

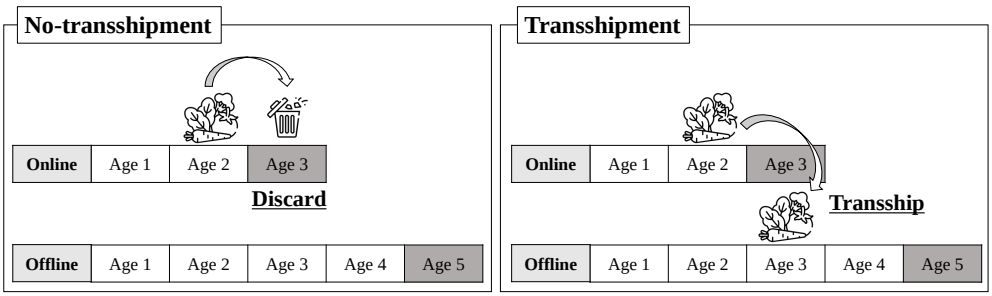



Figure 2.11: Example of saving outdating cost in the OOCS through transshipment.

### 2.5.4 Managerial insights

According to the results of the experiment, we suggest the following managerial insights that are relevant to logistics practitioners who are concerned about setting up an effective transshipment policy within the OOCS:

- As a result of the rapid development of computational technology in recent years, multiple e-commerce companies have been able to secure lots of data about the historical demand for fresh foods. For companies that have an abundance of demand data, it is recommended that they utilize the data directly with the developed DRL approach to obtain a practical transshipment policy since the neural network of the DRL can be trained more accurately as more data is gathered. On the other hand, when the company does not have enough data to train a neural network of DRL, it is necessary to generate artificial data utilizing estimated demand distribution in order to train the neural network. Alternatively, it is also possible to obtain a transshipment policy using traditional methods, such as VI and heuristics, although they are difficult to apply in practice.

- In a condition that the value of unit transshipment cost $c_l$ is relatively small compared to the value of unit outdating cost $c_w$, a transshipment policy is effective in increasing the average profit from the perspective of the total system by saving the outdating cost. In comparison with a no-transshipment policy, the transshipment could not contribute to ramping up the average profit if the value of $c_l$ is not much smaller than the value of $c_w$ (i.e., $c_l/c_w \approx 1$). Thus, we recommend that logistic managers estimate the accurate value of $c_l$ and $c_w$ before deciding whether to implement the transshipment policy within the OOCS.

- Although the hybrid DRL approach developed has proved to be effective in maximizing profit, training the DRL once takes several hours. A further reason for not relying on DRL's transshipment policy is that it does not follow a simple rule and is difficult to interpret. Therefore, several logistics practitioners have not been able to rely on DRL's transshipment policy in their business practices. Hence, if logistics managers do not have time to train the DRL from scratch and require an interpretable policy, a simple decision rule for transshipment could be considered. According to Figures 2.9 and 2.10, in the OOCS with a heterogeneous shelf life, we can observe that more products are transshipped from the online channel (short shelf life) to the offline channel (long shelf life) than from the offline channel to the online channel. As a result of this trend, if logistics practitioners do not have enough time or desire a more interpretable policy, developing a simple policy that only covers transshipments from online to offline could be acceptable in business practice.

## 2.6 Summary

We developed the lateral transshipment model for fresh food by accommodating the key attributes of the OOCS: heterogeneous shelf life, proactive transshipment, and non-negligible transshipment time. In the field of lateral transshipment research, the majority of studies focus on a specific distribution of demand to determine the policy of transshipment. Conversely, we seek to directly derive a transshipment policy based on demand data by developing the DRL approach based on the SAC algorithm, which does not need any assumptions about demand distribution.

Unfortunately, the action space of the proposed model is extraordinarily large because four types of decisions must be made simultaneously. In our experience, the DRL approach suffers from unstable performance during training, which is due to the difficult task of computing large action spaces in the DRL approach, and therefore requires considerable computation time. As a way to mitigate these issues, we developed a hybrid DRL approach that combines two novel acceleration methods: SQLT and RS, to create a hybrid DRL approach. First, we split the decision-making process into two stages. Transshipment decisions are handled by the DRL approach, while replenishment decisions are handled by the SQLT approach. Second, to enhance the performance of DRL, we implement the RS by adopting the SQmax policy as a teacher heuristic into DRL. By conducting computational experiments, we observed that adopting two acceleration methods enabled the training process to be stabilized and the average profit to be maximized.

We analyzed the impacts of transshipment in the OOCS by differing types of demand and varying the unit transshipment cost parameter and shelf life of online and offline channels. In line with our expectations, transshipment was more effective

when demand variability was high. Transshipment could lead to an increase in average profit as a result of a substantial reduction in outdating cost, as compared to revenue and other components of the cost. Transshipment resulted in a slight increase in the outdating cost in the offline channel, compared to the case where there was no-transshipment. However, the outdating cost in the online channel was reduced substantially by implementing transshipment. Also, we found that more fresh foods are transshipped from online to offline channels than from offline to online channels. These findings suggest that the offline channel could be utilized to resell old products planned to be discarded in the online channel. Finally, we presented several managerial insights instructive to logistics practitioners who require a transshipment policy with the OOCS.

This study could serve as a starting point for future research related to the DRL approach to lateral transshipment of perishable products in the future. Even though this study has focused on the proactive transshipment, we expect that the proposed DRL approach could be applied to the reactive transshipment by adding the observed demand to the state in the MDP. Moreover, by differing the types of demand and cost parameters, analyzing the effectiveness between proactive and reactive transshipment strategies for each case could provide better guidance and managerial insights for real-business operators.

# Chapter 3

# E-commerce supply chain network design using on-demand warehousing system under uncertainty

## 3.1 Introduction

In recent cases, because it cannot be sure how long the pandemic-driven consumer spending will last, many small-medium sized e-commerce companies prefer to utilize the ODWS [88]. From the standpoint of the e-commerce retailer, the main advantage of the ODWS is that a short-term rent for warehouses is available [135]. Throughout this chapter, we will use the terminology *commitment* to indicate the short-term rent contract for warehouses in the ODWS.

Because of the distinctive advantages of the ODWS, several recent studies have focused on solving the supply chain problem with the ODWS to derive a cost-saving strategy based on optimization-based methods [138, 125, 133, 29, 137]. Even though previous studies have dealt with the ODWS in various aspects, this study seeks to fill two research gaps in the ODWS research area. The first research gap is that previous studies did not address the main characteristic of the ODWS, the short-term rent contract (i.e., commitment), except Unnu and Pazour [137]. Although Unnu and Pazour [137] addressed the property of commitment, they did not deal with the decisions for the commitment period for using the ODWS because the

available commitment period was a given parameter. The second research gap is scarcity studies that consider the inherent uncertainties systematically involved in making decisions that might occur in the supply chain with the ODWS. While several studies considered uncertainties of demand [138, 125], as far as we know, there was no research that dealt with the properties of commitment and inherent uncertainties simultaneously.

To fill these research gaps, this study aims to deal with the SCND problem considering the characteristics of the ODWS and the decisions for the commitment period. Furthermore, because demand and supply have inherent uncertainties, our research addresses the SCND problem with the ODWS under uncertain environments. To the best of our knowledge, this study is the first attempt to solve the problem considering the properties of commitment and uncertainties simultaneously in the ODWS research area. Of special note, we define the supply uncertainty form as *yield uncertainty*, which means the amount actually supplied is random and different from the amount ordered.

This study extends the conference paper Lee et al. [85] by considering decisions for supplier selections and inherent uncertainties of demand and supply. Motivated by the above research gaps in existing ODWS literature, this study defines the following four research questions to address:

1. How would it be best to consider the uncertainties for the SCND with the ODWS and devise the solution approach for reducing computational efforts?

2. How does the ODWS affect the supply chain network and the total cost of the resulting supply chain?

3. What impact does the total cost and utilization of warehouses have when the commitment and stockout costs vary?

4. What impact does the lead time have in the supply chain with the ODWS?

The main contributions of this chapter are threefold. First, we propose the TSSP for an e-commerce SCND with the ODWS under uncertainties. To estimate the expected function in the proposed model, we employ the SAA method. Second, to alleviate the computational burden in SAA, we utilize the multi-cut version of BD algorithm. Furthermore, we develop the acceleration method for improving the convergence of bounds by focusing on the initial iteration in the BD algorithm. Third, we show the potential cost-saving effects of using the ODWS in the supply chain through computational experiments.

## 3.2 Literature review

Our study is directly related to three streams of literature in operations management. First, we review the literature on the dynamic facility location model (DFLM), which is the general supply chain model of our study. Second, we investigate relevant literature on scenario-based stochastic programs for the SCND within a methodological context for our research problem. Third, we review literature that considers the properties of ODWS in supply chain problems. In addition, we present distinctive features of our study compared to relevant studies on the ODWS.

### 3.2.1 Dynamic facility location and supply chain network design under uncertainty

The facility location (FL) model is roughly categorized using six classifications, and detailed taxonomy is presented in Klose and Drexl [80]. Our SCND model is developed based on the multi-stage, capacitated, multiple-sourcing, multi-item, and dynamic FL model. Among several categories, the dynamic property is the most essential for accommodating the features of ODWS. The DFLM considers the multi-period problem, and the input parameters (e.g., cost, capacity, and demand) differ depending on the time period. Due to this property, facilities can be opened or closed in every period throughout a given planning horizon [80, 95].

Instead of reviewing all the works related to the DFLM, we present three papers covering the capacity adjustment through the lens of opening or closing a facility, which is related to one of the properties of the ODWS. Melo et al. [98] proposed the DFLM that considered the gradual relocation of facilities over the planning horizon. In this model, the capacity could be transferred from existing facilities to new facilities. To accommodate fluctuations of demands, two extended mathematical models were suggested for dealing with scenarios of capacity expansion and reduction. In addition, because the above two scenarios considered capacity transfer size as continuous, the authors presented the modular case model that permits discrete amounts. However, they did not consider any commitment properties for opening or closing facilities.

Several related works considered different time resolutions for strategic and tactical periods over a planning horizon [132, 8, 7, 44]. In this literature, the decision to open or close facilities could be allowed only in strategic periods. Badri et al. [7]

developed a MILP model for capacity expansion in four echelons of the multiple commodity supply chain. The budget constraint for the expansion of the supply chain was determined according to cumulative net profits and funds supplied by external sources. Two types of warehouses, private and public, were considered, and public warehouses could be used at any time if contracted to be utilized. However, they also did not accommodate commitment constraints for using public warehouses. Fattahi et al. [44] proposed a multi-stage, multi-item, and DFLM, which considered price dependent demand. The authors also considered private and public warehouses, and decisions for product shipments were made in tactical periods. While public warehouses could be opened or closed at any time period, private warehouses could not be closed if opened once.

The FL problem is applied to various domains. Especially, SCND has been considered as an appropriate application area for the FL problem [49, 80]. In general, large investments are required to make strategic decisions for determining locations and the number of facilities in SCND. However, if these strategic decisions are made in a deterministic environment, a huge amount of costs can be incurred due to the fluctuations of demands and supplies. Therefore, in both practice and academia, the necessity of considering uncertainty in SCND has obtained substantial attention [57]. To cope with uncertainty in SCND, our study proposes a mathematical framework based on scenario-based stochastic programs. Owing to the nature of scenario-based stochastic programs, the problem size increases depending on the number of scenarios. The emphasis in our review of the literature is on how existing studies address the scenario-generation issue and solution approach for the proposed stochastic programming model.

Through reviewing several previous studies, we could observe that the SAA and scenario tree construction are broadly used for scenario generation. First, several studies adopting the SAA will be introduced. Santoso et al. [118] dealt with the large-scale problem for the global SCND. They used the SAA method and single-cut BD algorithm. In the single-cut BD algorithm, only a single optimality cut is applied at each iteration. Schütz et al. [121] considered the SCND problem for the Norwegian meat industry. They used the SAA method and dual decomposition algorithm to solve the problem. Fazeli et al. [45] proposed the two-stage stochastic mixed-integer nonlinear programming (MINLP) to design an electric vehicle charging station network. They compared the single-cut and multi-cut BD algorithms and showed that multi-cut BD outperformed single-cut BD. Different from the single-cut BD, several optimality cuts are generated at each iteration in the multi-cut BD. Nur et al. [103] addressed a biofuel SCND incorporating biomass quality properties. They proposed a parallelized decomposition algorithm that combined the SAA and an enhanced progressive hedging algorithm to solve real-life problem instances in a reasonable time. Azaron et al. [5] developed a multi-objective TSSP for taking into account the decision about production, inventory, and shipping among the entities of the supply chain network. The $\epsilon-$constraint method and SAA were utilized to solve the proposed multi-objective TSSP.

To generate efficient scenarios, several studies utilized scenario tree construction. Khatami et al. [77] addressed closed-loop supply chains and used the single-cut BD algorithm for the solution approach. They generated scenarios based on the demand distribution function using Chloesky's factorization method. Fattahi and Govindan [43] introduced the SCND problem for an integrated forward/reverse logistics setup

over a planning horizon. The Latin Hypercube Sampling method generated a fan of scenarios for demand and potential return uncertainty. Zahiri et al. [148] presented the multi-stage stochastic programming approach with a combined scenario tree for an integrated supply chain planning for blood products. The meta-heuristic algorithm was used to alleviate the high complexity of the model. Azizi et al. [6] addressed the SCND problem with multi-period reverse logistics with lot-sizing. Scenarios were generated with the moment matching technique, and the number of scenarios was reduced using forward selection. Ghorashi Khalilabadi et al. [51] developed the multi-stage stochastic integer programming model for prior planning for disruptions in the supply chain. A scenario tree was constructed, and a progressive hedging algorithm was used to alleviate the computational burden.

### 3.2.2 Supply chain problems in the ODWS and distinctive features of this study

The last few years have seen a huge growth in the problem of utilizing different types of warehouses to mitigate capacity and demand shortage issues. In particular, the two warehouse system that utilizes rented warehouse has become a central issue for reducing product shortage or expiration [134, 73, 59]. In addition, recent developments in third-party logistics and online platforms have led to many researchers proposing novel problems [124, 110, 116]. Although on-demand warehousing is a very popular trend in real business, it is underexplored, and only a few researchers dealt with the problems regarding the supply chain using the ODWS.

There are two significant characteristics of the ODWS compared to other warehouse systems: *capacity granularity* and *commitment granularity* [108]. *Capacity*

*granularity* means the minimum capacity that can be acquired by a chosen distribution alternative (e.g., warehouses). In terms of the ODWS, the minimum capacity requirement is very small. *Commitment granularity* means the minimum commitment periods (in time units) a user of the system must maintain their decision. As mentioned in Section 3.1, the minimum commitment periods of the ODWS are usually very short (e.g., monthly or weekly commitments) compared to leasing warehouses. Throughout this study, we will use the term *duration constraint* to refer to the constraint that the firm must utilize the ODWS at least the minimum of specified commitment periods. On the other hand, the firm can commit for a period of use that is longer than the minimum commitment periods and shorter than the maximum commitment periods allowed by the ODWS. The cost structures for using the ODWS and other facilities are usually different, depending on the commitment periods. The term *period decision* will be used to indicate this decision for commitment periods.

We reviewed related studies that accommodated the properties of the ODWS. Thanh et al. [132] proposed a MILP model based on DFLM to design a production-distribution system in a deterministic demand setting. In their model, two types of warehouses, public and private warehouses, were considered. Even though the authors did not directly refer to the ODWS, the concept of public warehouses was similar to the ODWS. Public warehouses could be opened and closed multiple times, but their status only can be changed after at least two periods. This property was similar to the duration constraint in commitment granularity. Van der Heide et al. [138] analyzed the benefits of utilizing dynamic shipments in shared warehouse and transportation networks motivated by the ODWS. They defined the model as a se-

quential decision making problem and accommodated the demand uncertainty. They applied a mathematical framework to compute optimal ordering and transportation decisions using the MDP and VI method. Even though capacity and commitment granularity were not considered, several numerical experiments provided managerial insights into improving demand fulfillment and transport efficiency through dynamic shipments and a high degree of consolidation.

Tian and Zhang [133] dealt with the problem of renting warehouses and allocating products among the warehouses in the e-commerce supply network with the ODWS. The authors suggested the MINLP model and converted the proposed MINLP to MILP form. However, when demand uncertainty is considered, it is impossible to convert the MINLP to MILP, as stated in Tian and Zhang [133]. Moreover, the commitment granularity of the ODWS was not accommodated because the problem was defined as the single-period setting. Shi et al. [125] suggested a periodic review warehouse model that considers the ODWS and third-party retailers. They showed the optimality of base stock policy and monotonicity of optimal space allocation decisions in the suggested model. To address a multiple items situation that incurs the curse of dimensionality, the heuristic based on approximate dynamic programming was developed. However, the commitment granularity was also not considered. Ceschia et al. [29] proposed the supply matching problem from the perspective of platform providers in the ODWS. In contrast to related studies in ODWS, the objective was to maximize the number of transactions between customers and warehouse space suppliers. In addition, they developed a list-based heuristic to reduce the time for solving the problem. However, because customer requests and supplier availability were given, it is necessary to consider the dynamic situation for enhancing the

applicability in the ODWS.

Unnu and Pazour [137] proposed the MILP based on the DFLM that determines location-allocation decisions of three distribution warehouses types—self-distribution, third-party logistics company(3PL)/lease, and the ODWS. In the proposed MILP, the duration constraint in commitment granularity was considered, and the stochastic parameter was replaced with the expected value of demand. By using the obtained solution of this model, the authors evaluated distribution network design with and without the ODWS by adding the randomness of demand in the simulation. Although they tried to accommodate the demand uncertainty, it is difficult to confirm that the stochastic nature is properly considered. If a shortage of demand can occur, the quality of the solution from the MILP model replacing the stochastic parameter with the expected value could be poorer than the solution obtained by the stochastic approach (e.g., SAA+BD). We will show this stochastic solution gap in Section 3.5.3.

We show several distinctive features of our study in Table 3.1. As far as we know, this is the first study to consider the period decision for commitment in the ODWS. In addition, we consider a realistic situation in which the longer the commitment period, the greater the discount is that's applied. The novelties of our study can be summed up from three perspectives, as follows:

- *Modeling:* We develop a mathematical model based on the multi-stage, capacitated, multiple-sourcing, multi-item, and dynamic FL model. In the presented model, we accommodate the period decision in commitment granularity for the first time. In addition, we consider the aggregated customer demand to reflect the case of the e-commerce market supply chain in South Korea.

- *Uncertainty:* We propose the TSSP model that makes the decision considering the uncertainty of demand. Also, because supplier selections are included as decisions in our model, supply uncertainty (i.e., yield uncertainty) is also considered. We utilize the SAA method to estimate the expected function accurately with the reasonable size of scenarios.

- *Computational time:* Through our use of a commercial solver, the scenario-based model can be solved with a large number of scenarios. However, because the problem size increases depending on the number of scenarios, the solver cannot solve the practical large-scale problem in reasonable times. To alleviate the computational burden, we propose a methodology combined with SAA and a multi-cut version of BD.

Table 3.1: Comparison of recent studies related to dynamic facility location and on-demand warehousing

| Author | On-demand warehousing | Multi-item | Multi-period | Capacity granularity | Commitment granularity | | Uncertainty (factors) | Solution methodology |
|---|---|---|---|---|---|---|---|---|
| | | | | | duration constraint | period decision | | |
| Melo et al. [98] | | ✓ | ✓ | ✓ | | | | Solver (Cplex) |
| Thanh et al. [132] | | ✓ | ✓ | ✓ | ✓ | | | Solver (Xpress) |
| Badri et al. [7] | | ✓ | ✓ | ✓ | | | | LR[a] |
| Fattahi et al. [44] | | ✓ | ✓ | ✓ | | | | Solver (Cplex) |
| Van der Heide et al. [138] | ✓ | | ✓ | | | | ✓(demand) | MDP[b], VI[c] |
| Shi et al. [125] | ✓ | | ✓ | | | | ✓(demand) | ADP[d] |
| Tian and Zhang [133] | ✓ | ✓ | | ✓ | | | | Solver (Cplex) |
| Ceschia et al. [29] | ✓ | | ✓ | | | | | Heuristics |
| Unnu and Pazour [137] | ✓ | | ✓ | ✓ | ✓ | | | Solver (Cplex), SIM[e] |
| **This research** | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓(demand, supply) | TSSP, SAA + BD |

[a] Lagrangian relaxation; [b] Markov decision process; [c] Value iteration; [d] Approximate dynamic programming; [e] Simulation

## 3.3 Problem description and mathematical model

This section presents a problem and mathematical formulation for the supply chain considering the ODWS. The detailed problem description for the SCND utilizing an ODWS is presented in Section 3.3.1. Section 3.3.2 presents the TSSP to represent the problem under uncertainty. In Section 3.3.3, we represent a compact formulation and explain the well-defined property briefly.

### 3.3.1 The supply chain with the ODWS

We describe the supply chain network for e-commerce retailers using the ODWS. We use the case of the e-commerce market in South Korea for the supply chain network description. From here forward, we will use the term *retailer* to indicate the e-commerce retailer and the term *provider* to indicate the warehouse operator who has excess capacity. We deal with the multi-items and multi-period problem, and the decision-maker corresponds to a retailer. An overview of the supply chain with an ODWS is shown in Figure 3.1.

We define the two types of decisions determined based on the before and after the realization of uncertainties. Before the realization of uncertainties, the decisions for choice of suppliers and warehouses are made because they are in the *strategic* levels of decision [127]. First, we will illustrate the decisions for the selection of suppliers, $y_j$. Among many suppliers, $j \in \mathcal{J}$, the retailer tries to cooperate with suppliers who provide a better quality of items or who provide a number of items with low variability. Also, the locations of the suppliers are significant in order to minimize the transportation costs from suppliers to warehouses, $c_j^r, c_j^k$, and $c_j^e$. The different value of investment cost, $F_j$, is charged to engage cooperation according to

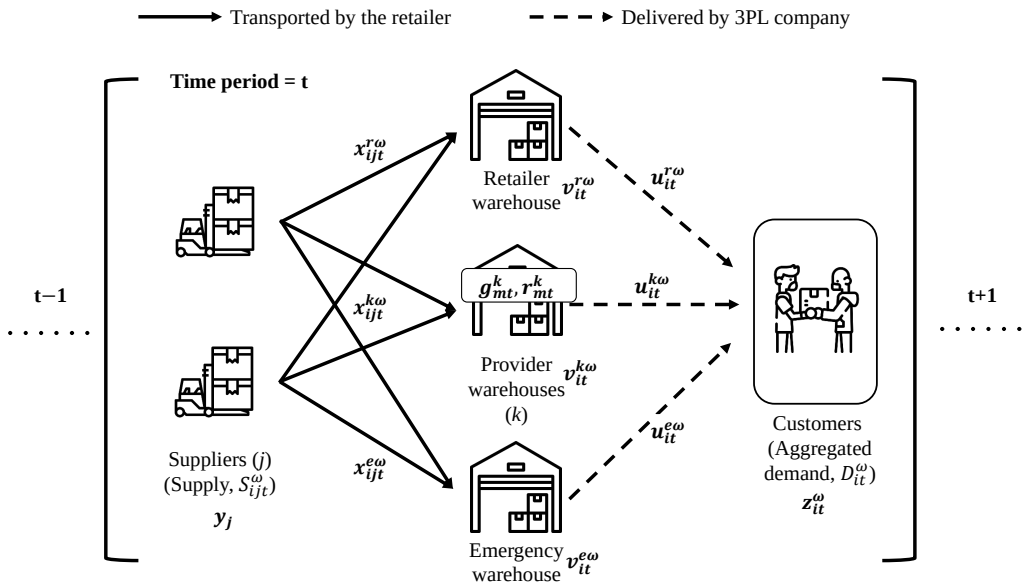



Figure 3.1: Overview of the supply chain with an ODWS.

suppliers.

We assume that the retailer can utilize three types of warehouses: (1) the retailer's own warehouse (*retailer warehouse*), (2) the warehouse of providers connected by the ODWS platform (*provider warehouse*), and (3) the warehouse that charges higher unit holding and transportation costs than other types of warehouses (*emergency warehouse*). We assume that there is one retailer warehouse, one emergency warehouse, and several provider warehouses, $k \in \mathcal{K}$. Note that the problem can easily be extended to multiple retailer and emergency warehouses by increasing the set size for warehouses. We propose the mathematical model and solution methodology considering multiple retailer and emergency warehouses, but every computational experiment is conducted in the setting of one retailer and one emergency warehouse.

The transportation capacity from suppliers to warehouses, as well as the storage

capacity, is assigned for every warehouse, $C^r$, $C^k$, and $C^e$. Every warehouse has the same role with distribution centers as follows:

1. Shipments from the suppliers will be assembled, and vehicle loads will be de-aggregated.

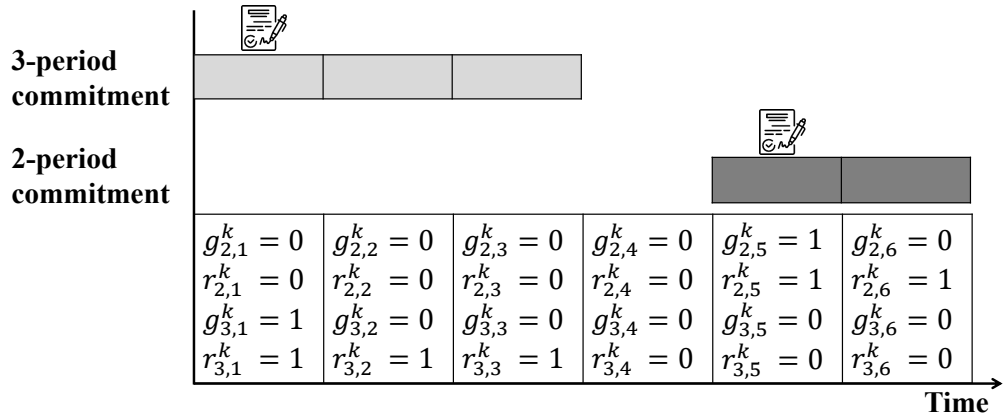
2. If the capacity of the warehouses is not full, every item can be held in warehouses for the short or long term.

3. Items will be assorted according to customers' demands and will be processed or packaged for bringing to customers.

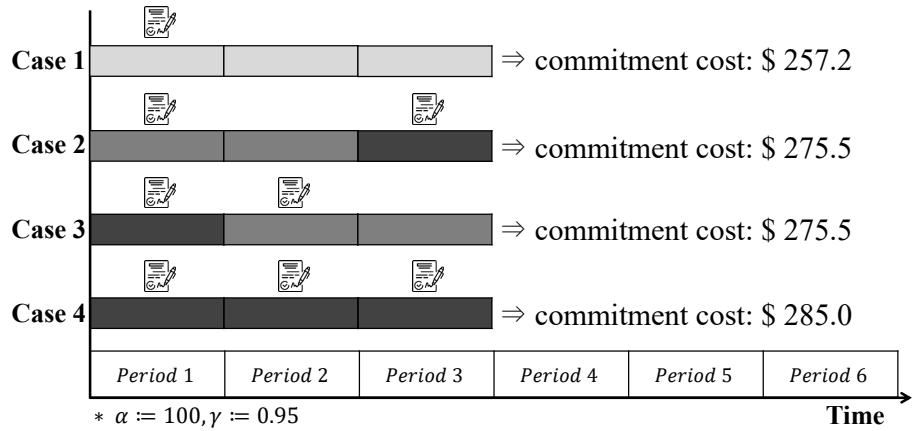
In the case of the provider warehouse, the above roles can only be applied when the retailer has committed to using the provider warehouse for a designated period. We introduce the detailed procedure for the commitment decisions, $g^k_{mt}$ and $r^k_{mt}$, for the provider warehouse using the simple example that is depicted in Figure 3.2a. For a brief explanation, we consider two types of commitments (2-period and 3-period) over a six-period planning time horizon with a provider warehouse, $k$. In period one, the retailer made the 3-period commitment; thus, the provider warehouse, $k$, can be used from period one to period three. However, because the commitment for using the warehouse in period four has not been made, the retailer cannot utilize the provider warehouse, $k$, at this period. On the other hand, the warehouse is available for use from period five to period six because the retailer made the 2-period commitment in period five.

We take into account the realistic situation in which the retailer takes a greater discount when a longer commitment period is made. Therefore, the cost function for committing warehouses for the $m$-period is defined as $m\alpha\gamma^m$, where $\gamma$ is the

discount factor, and $\alpha$ is the commitment cost to utilize a provider warehouse for a period. In Figure 3.2b, we describe the effects of commitment periods on the cost. When retailers plan to utilize the provider warehouse from periods one to three, there are three ways to make the commitment in these periods. First, the retailer can use the warehouse from period one to three by making the 3-period commitment in period one (Case 1). Furthermore, the retailer can use 2-period and 1-period commitments (Cases 2 and 3) or make the 1-period commitment for each period to utilize the warehouse for three periods (Case 4). Because of the cost function for committing warehouses $m\alpha\gamma^m$, Case 1 is the cheapest way to utilize the warehouse (i.e., commitment cost: \$257.2). However, committing for a long period to use warehouses could incur unnecessary costs due to the long-term use of warehouses, although there is small customer demand.




(a) Commitment decisions using two types of commitments




(b) Commitment cost using four cases

Figure 3.2: Simple example of commitment decisions for provider warehouse $k$.

After the realization of uncertainties, operational decisions are made. We describe the decision procedure following the flow of items from suppliers to customers according to the process from left to right in Figure 3.1. In the beginning, the transportation decisions from suppliers to the arriving warehouses, $x_{ijt}^{r\omega}$, $x_{ijt}^{k\omega}$, and $x_{ijt}^{e\omega}$, are made for the ordered items. The lead time between suppliers and warehouses exists, $L_s$. After items have arrived at the designated warehouses, items are processed for sending to customers. Inventory holding decisions, $v_{it}^{r\omega}$, $v_{it}^{k\omega}$, and $v_{it}^{e\omega}$, and delivery decisions, $u_{it}^{r\omega}$, $u_{it}^{k\omega}$, and $u_{it}^{e\omega}$, will be made at warehouses. Depending on the type of warehouses, different inventory holding costs, $h_i^r$, $h_i^k$, and $h_i^e$, will be incurred. In particular, because most retailers commonly use the services of a logistics company for last-mile deliveries in the case of the South Korean e-commerce market, we consider the aggregated customer demand for the proposed model and assume that items will be delivered from warehouses to customers by the 3PL company. Furthermore, the delivery cost per parcel of items, $b_i$, is identical without taking into account the weights of items and locations of destinations. There exists lead time between warehouses and aggregated customer demands, $L_d$. Finally, in order to address the stock-out issue, we assume that unsatisfied demand will become lost sales, $z_{it}^{\omega}$. This assumption is reasonable because customers are more likely to switch to another website to search for substitute items rather than wait for insufficient items to be stocked. Additionally, the corresponding penalty cost, $\beta_i$, for lost sales will be incurred.

We consider two additional assumptions. First, we exclude perishable items in the proposed problem. In order to deal with perishable products, it is necessary to install the cold storage system that is available to maintain the specific tempera-

ture and humidity conditions that do not alter the products' original characteristics. However, it is difficult to use this system in the ODWS because various users store heterogeneous products in the same space. Second, lateral transshipment between warehouses is not considered. The lateral transshipment could increase the complexity of the problem because the number of decision variables related to lateral transshipment could increase exponentially depending on the number of warehouses. Furthermore, because of the property that ensures that the provider warehouse can be opened or closed at each period, the connections for lateral transshipment can be negated.

### 3.3.2 The two-stage stochastic programming model

This section presents the SCND model, which is developed based on the multi-stage, capacitated, multiple-sourcing, multi-item, and dynamic FL model. In order to model the problem under uncertainties, we extend the deterministic model as the TSSP. As mentioned in Section 3.3.1, the operational decisions have been considered after the prior decisions. By considering these characteristics of decision-making, we employ the TSSP to represent the situation of the SCND with an ODWS. We assume that demands, $D$, and supplies, $S$, are random parameters with full knowledge of probability distributions, defined as *stochastic parameters*. Therefore, we use $\zeta = (D, S)$, which stands for the stochastic parameters vector with finite and discrete support, which can be represented as a finite number of realizations (scenarios). Let $\Omega$ be a set of scenarios, and each scenario is denoted as $\omega$. Then, $\zeta^{\omega}$, $\forall \omega \in \Omega$, is a particular realization of stochastic parameters. The sample space of stochastic parameters is represented as set $\left\{\zeta^1, \cdots, \zeta^{|\Omega|}\right\}$ with the following proba-

bilities, $p_1, \cdots, p_{|\Omega|}$.

In the proposed TSSP, decisions for supplier selection and commitments for the provider warehouses are made in the first-stage problem. The first-stage decisions are the *here-and-now* decisions that are determined before the realization of stochastic parameters. Subsequently, in the second-stage, operational decisions such as transportation, inventory holding, and lost sales are made after realizations of stochastic parameters. The following notations are utilized in the proposed mathematical formulation.

**Indices and sets**

| | |
|---|---|
| $\mathcal{T}$ | set of periods, $t \in \mathcal{T} = \{1, 2, \cdots, T\}$ |
| $\mathcal{I}$ | set of items, $i \in \mathcal{I} = \{1, 2, \cdots, I\}$ |
| $\mathcal{J}$ | set of suppliers, $j \in \mathcal{J} = \{1, 2, \cdots, J\}$ |
| $\mathcal{K}$ | set of provider warehouses, $k \in \mathcal{K} = \{1, 2, \cdots, K\}$ |
| $\mathcal{R}$ | set of retailer warehouses, $r \in \mathcal{R} = \{1, 2, \cdots, R\}$ |
| $\mathcal{E}$ | set of emergency warehouses, $e \in \mathcal{E} = \{1, 2, \cdots, E\}$ |
| $\mathcal{M}$ | set of available commitment periods, $m \in \mathcal{M} = \{1, 2, \cdots, M\}$ |
| $\Omega$ | set of scenarios, $\omega \in \Omega$ |

**Parameters**

| | |
|---|---|
| $D_{it}^{\omega}$ | aggregated demand of item $i$ at period $t$ under scenario $\omega$ |
| $S_{ijt}^{\omega}$ | supply of item $i$ from supplier $j$ at period $t$ under scenario $\omega$ |
| $C^r$ | capacity of the retailer warehouse $r$ |
| $C^k$ | capacity of the provider warehouse $k$ |
| $C^e$ | capacity of the emergency warehouse $e$ |
| $L_s$ | lead time between suppliers and warehouses |
| $L_d$ | lead time between warehouses and customers |
| $F_j$ | investment cost to select supplier $j$ |

| | |
|---|---|
| $h_i^r$ | inventory holding cost of the retailer warehouse $r$ for a unit of item $i$ per period |
| $h_i^k$ | inventory holding cost of provider warehouse $k$ for a unit of item $i$ per period |
| $h_i^e$ | inventory holding cost of the emergency warehouse $e$ for a unit of item $i$ per period |
| $\alpha$ | commitment cost to utilize provider warehouse for a period |
| $\beta_i$ | lost sales cost for a unit of item $i$ |
| $b_i$ | cost of delivery for a unit of item $i$ from warehouses to customers |
| $c_j^r$ | transportation cost for a unit of item from supplier $j$ to the retailer warehouse $r$ |
| $c_j^k$ | transportation cost for a unit of item from supplier $j$ to provider warehouse $k$ |
| $c_j^e$ | transportation cost for a unit of item from supplier $j$ to the emergency warehouse $e$ |
| $\gamma$ | discount factor of commitment cost |
| $p_\omega$ | probability that scenario $\omega$ occurred |

**Decision variables**

| | |
|---|---|
| $g_{mt}^k$ | 1 if an $m$ period commitment is made at period $t$ for provider warehouse $k$, 0 otherwise |
| $r_{mt}^k$ | 1 if provider warehouse $k$ can be utilized because of the $m$ period commitment at period $t$, 0 otherwise |
| $y_j$ | 1 if supplier $j$ is selected, 0 otherwise |
| $v_{it}^{r\omega}$ | number of item $i$ held in inventory at the retailer warehouse $r$ from period $t$ to $t+1$ under scenario $\omega$ |
| $v_{it}^{k\omega}$ | number of item $i$ held in inventory at provider warehouse $k$ from period $t$ to $t+1$ under scenario $\omega$ |
| $v_{it}^{e\omega}$ | number of item $i$ held in inventory at the emergency warehouse $e$ from period $t$ to $t+1$ under scenario $\omega$ |
| $x_{ijt}^{r\omega}$ | number of item $i$ transported from supplier $j$ to the retailer warehouse $r$ at period $t$ under scenario $\omega$ |
| $x_{ijt}^{k\omega}$ | number of item $i$ transported from supplier $j$ to provider warehouse $k$ at period $t$ under scenario $\omega$ |
| $x_{ijt}^{e\omega}$ | number of item $i$ transported from supplier $j$ to the emergency warehouse $e$ at period $t$ under scenario $\omega$ |

| | |
|---|---|
| $u_{it}^{r\omega}$ | number of item $i$ delivered to satisfy aggregated demand from the retailer warehouse $r$ at period $t$ under scenario $\omega$ |
| $u_{it}^{k\omega}$ | number of item $i$ delivered to satisfy aggregated demand from provider warehouse $k$ at period $t$ under scenario $\omega$ |
| $u_{it}^{e\omega}$ | number of item $i$ delivered to satisfy aggregated demand from the emergency warehouse $e$ at period $t$ under scenario $\omega$ |
| $z_{it}^{\omega}$ | lost sales of item $i$ at period $t$ under scenario $\omega$ |

By considering the above problem descriptions and notations, the extensive form of the TSSP is formulated as follows:

**First-stage problem**

$$\min \quad \sum_{j\in\mathcal{J}} F_j y_j + \sum_{k\in\mathcal{K}} \sum_{m\in\mathcal{M}} \sum_{t\in\mathcal{T}} m\alpha\gamma^m g_{mt}^k + \mathbb{E}_\zeta \left[Q(y,r,\zeta^\omega)\right] \tag{3.1}$$

$$\text{s.t.} \quad \sum_{\tau=t}^{\min\{t+m-1,|\mathcal{T}|\}} g_{m\tau}^k \leq 1, \qquad \forall k\in\mathcal{K}, m\in\mathcal{M}, t\in\mathcal{T}, \tag{3.2}$$

$$\sum_{\tau=\max\{t-m+1,1\}}^{t} g_{m\tau}^k = r_{mt}^k, \qquad \forall k\in\mathcal{K}, m\in\mathcal{M}, t\in\mathcal{T}, \tag{3.3}$$

$$\sum_{m\in\mathcal{M}} r_{mt}^k \leq 1, \qquad \forall k\in\mathcal{K}, t\in\mathcal{T}, \tag{3.4}$$

$$\sum_{m\in\mathcal{M}} g_{mt}^k \leq 1, \qquad \forall k\in\mathcal{K}, t\in\mathcal{T}, \tag{3.5}$$

$$r_{mt}^k, g_{mt}^k \in \{0,1\}, \qquad \forall k\in\mathcal{K}, m\in\mathcal{M}, t\in\mathcal{T}, \tag{3.6}$$

$$y_j \in \{0,1\}, \qquad \forall j\in\mathcal{J}. \tag{3.7}$$

where $Q(y,r,\zeta^\omega)$ is the value function for the optimal objective value of the second-stage problem with a given scenario $\omega$. By applying the scenario-based approach, the expected second-stage cost can be denoted with $\sum_{\omega\in\Omega} p_\omega Q(y,r,\zeta^\omega)$. The objective function of the first-stage problem (3.1) minimizes the total cost incurred in the

supply chain. Constraint (3.2) ensures that other commitments for provider warehouses cannot be made until the ongoing commitment expires. Constraints (3.3) and (3.4) ensure that every provider warehouse can be utilized only in the case when commitments are made for the designated period. Constraint (3.5) guarantees that just one type of commitment can be made among available commitment periods for each provider warehouse at each period. Constraints (3.6) and (3.7) enforce that first-stage decision variables are binary variables. Given the values of $y_j$ and $r^k_{mt}$ and a scenario $\omega$, the second-stage problem that determines the recourse function $Q(y, r, \zeta^\omega)$ is as follows:

**Second-stage problem**

$$Q(y, r, \zeta^\omega) =$$

$$\min \quad \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}}\left(\sum_{r\in\mathcal{R}} h^r_i v^{r\omega}_{it} + \sum_{e\in\mathcal{E}} h^e_i v^{e\omega}_{it} + \sum_{k\in\mathcal{K}} h^k_i v^{k\omega}_{it} + b_i\left(\sum_{r\in\mathcal{R}} u^{r\omega}_{it} + \sum_{e\in\mathcal{E}} u^{e\omega}_{it} + \sum_{k\in\mathcal{K}} u^{k\omega}_{it}\right)\right.$$
$$\left. +\beta_i z^\omega_{it} + \sum_{j\in\mathcal{J}}\left(\sum_{r\in\mathcal{R}} c^r_j x^{r\omega}_{ijt} + \sum_{e\in\mathcal{E}} c^e_j x^{e\omega}_{ijt} + \sum_{k\in\mathcal{K}} c^k_j x^{k\omega}_{ijt}\right)\right) \tag{3.8}$$

$$\text{s.t.} \quad \sum_{r\in\mathcal{R}} x^{r\omega}_{ijt} + \sum_{k\in\mathcal{K}} x^{k\omega}_{ijt} + \sum_{e\in\mathcal{E}} x^{e\omega}_{ijt} \le S^\omega_{ijt} y_j, \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}, t\in\mathcal{T}, \tag{3.9}$$

$$u^{r\omega}_{it} + v^{r\omega}_{it} = v^{r\omega}_{it-1} + \sum_{j\in\mathcal{J}} x^{r\omega}_{ijt-L_s}, \qquad \forall i\in\mathcal{I}, r\in\mathcal{R}, t\in\mathcal{T}, \tag{3.10}$$

$$u^{k\omega}_{it} + v^{k\omega}_{it} = v^{k\omega}_{it-1} + \sum_{j\in\mathcal{J}} x^{k\omega}_{ijt-L_s}, \qquad \forall i\in\mathcal{I}, k\in\mathcal{K}, t\in\mathcal{T}, \tag{3.11}$$

$$u^{e\omega}_{it} + v^{e\omega}_{it} = v^{e\omega}_{it-1} + \sum_{j\in\mathcal{J}} x^{e\omega}_{ijt-L_s}, \qquad \forall i\in\mathcal{I}, e\in\mathcal{E}, t\in\mathcal{T}, \tag{3.12}$$

$$\begin{aligned}&\textstyle\sum_{r\in\mathcal{R}} u^{r\omega}_{it-L_d} + \sum_{k\in\mathcal{K}} u^{k\omega}_{it-L_d} \\ &+ \textstyle\sum_{e\in\mathcal{E}} u^{e\omega}_{it-L_d} + z^\omega_{it} \ge D^\omega_{it},\end{aligned} \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}, \tag{3.13}$$

$$\sum_{i\in\mathcal{I}} v^{r\omega}_{it} \le C^r, \qquad \forall r\in\mathcal{R}, t\in\mathcal{T}, \tag{3.14}$$

$$\sum_{i\in\mathcal{I}} v_{it}^{e\omega} \leq C^e, \qquad \forall e \in \mathcal{E}, t \in \mathcal{T}, \tag{3.15}$$

$$\sum_{i\in\mathcal{I}} v_{it}^{k\omega} \leq C^k \sum_{m\in\mathcal{M}} r_{mt}^k, \qquad \forall k \in \mathcal{K}, t \in \mathcal{T}, \tag{3.16}$$

$$\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} x_{ijt}^{r\omega} \leq C^r, \qquad \forall r \in \mathcal{R}, t \in \mathcal{T}, \tag{3.17}$$

$$\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} x_{ijt}^{e\omega} \leq C^e, \qquad \forall e \in \mathcal{E}, t \in \mathcal{T}, \tag{3.18}$$

$$\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} x_{ijt}^{k\omega} \leq C^k \sum_{m\in\mathcal{M}} r_{mt}^k, \qquad \forall k \in \mathcal{K}, t \in \mathcal{T}, \tag{3.19}$$

$$x_{ijt}^{r\omega}, x_{ijt}^{k\omega}, x_{ijt}^{e\omega} \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, r \in \mathcal{R}, k \in \mathcal{K}, e \in \mathcal{E}, t \in \mathcal{T}, \tag{3.20}$$

$$u_{it}^{r\omega}, u_{it}^{k\omega}, u_{it}^{e\omega}, v_{it}^{r\omega}, v_{it}^{k\omega}, v_{it}^{e\omega} \geq 0, \qquad \forall i \in \mathcal{I}, r \in \mathcal{R}, k \in \mathcal{K}, e \in \mathcal{E}, t \in \mathcal{T}, \tag{3.21}$$

$$z_{it}^{\omega} \geq 0, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}. \tag{3.22}$$

In the second-stage problem, every constraint is defined within the entire time horizon, $t \in \mathcal{T}$. For a realization of $\omega$, the objective function of the second-stage problem (3.8) minimizes the costs for the inventory holding, delivery, stockout, and transportation within the entire time horizon. Constraint (3.9) requires that the total number of items transported from the supplier, $j$, to every warehouse should be less than the given supplies. Constraints (3.10), (3.11), and (3.12) are the balance equations representing the flow of items from retailer, provider, and emergency warehouses to customers, respectively. The inventories stored in warehouses during the previous period, $t-1$, are transferred to the current period, $t$. Moreover, these constraints ensure the lead time between suppliers and warehouses, $L_s$. Constraint (3.13) ensures that the demand is satisfied by delivered items from each warehouse and that the lead time between warehouses and customers, $L_d$, exists. Furthermore, this constraint enforces that unsatisfied demand is lost. Constraints (3.14), (3.15), and (3.16) express the storage capacity for the retailer, emergency, and

provider warehouses, respectively. Constraints (3.17), (3.18), and (3.19) represent the transportation capacity between suppliers and the retailer, emergency, and provider warehouses, respectively. Finally, Constraints (3.20), (3.21), and (3.22) ensure that decision variables for the second-stage problems are non-negative real variables.

Because of the decision variables for lost sales, $z_{it}^{\omega}$, the second-stage problem remains feasible under any first-stage feasible solution, $y_j, g_{mt}^k$, and $r_{mt}^k, \forall j, m$, and $t$. In this case, we say that the stochastic programming (3.1)−(3.22) has the property called *relatively complete recourse* [22]. This is a key property for implementing the SAA and BD algorithms, and we will explain this in detail in Section 3.4.

We now define six cost components as follows:

$$\text{Delivery cost} := \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}}\sum_{\omega\in\Omega} p_\omega b_i \left(\sum_{r\in\mathcal{R}} u_{it}^{r\omega} + \sum_{e\in\mathcal{E}} u_{it}^{ew} + \sum_{k\in\mathcal{K}} u_{it}^{k\omega}\right) \tag{3.23}$$

$$\text{Commitment cost} := \sum_{k\in\mathcal{K}}\sum_{m\in\mathcal{M}}\sum_{t\in\mathcal{T}} m\alpha\gamma^m g_{mt}^k \tag{3.24}$$

$$\text{Stockout cost} := \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}}\sum_{\omega\in\Omega} p_\omega \beta_i z_{it}^{\omega} \tag{3.25}$$

$$\text{Supplier investment cost} := \sum_{j\in\mathcal{J}} F_j y_j \tag{3.26}$$

$$\text{Transportation cost} := \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\sum_{\omega\in\Omega} p_\omega \left(\sum_{r\in\mathcal{R}} c_j^r x_{ijt}^{r\omega} + \sum_{e\in\mathcal{E}} c_j^e x_{ijt}^{e\omega} + \sum_{k\in\mathcal{K}} c_j^k x_{ijt}^{k\omega}\right) \tag{3.27}$$

$$\text{Inventory holding cost} := \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}}\sum_{\omega\in\Omega} p_\omega \left(\sum_{r\in\mathcal{R}} h_i^r v_{it}^{r\omega} + \sum_{e\in\mathcal{E}} h_i^e v_{it}^{e\omega} + \sum_{k\in\mathcal{K}} h_i^k v_{it}^{k\omega}\right) \tag{3.28}$$

### 3.3.3 Compact formulation

For ease of the expositions, we represent the extensive form, (3.1)−(3.22), by the compact form using the concatenated vectors of decision variables, which are defined

as follows:

**Concatenated vectors of decision variables**

| | |
|---|---|
| $\mathbf{g}$ | Concatenated vector of the $g_{mt}^{k}$, $\forall k \in \mathcal{K}, m \in \mathcal{M}$, and $t \in \mathcal{T}$ |
| $\mathbf{r}$ | Concatenated vector of the $r_{mt}^{k}$, $\forall k \in \mathcal{K}, m \in \mathcal{M}$, and $t \in \mathcal{T}$ |
| $\mathbf{y}$ | Concatenated vector of the $r_{j}^{y}$, $\forall j \in \mathcal{J}$ |
| $\mathbf{u}_{\omega}$ | Concatenated vector of the $\left(u_{it}^{r\omega}, u_{it}^{k\omega}, u_{it}^{e\omega}\right)$, $\forall i \in \mathcal{I}, r \in \mathcal{R}, k \in \mathcal{K}, e \in \mathcal{E}$, and $t \in \mathcal{T}$ under scenario $\omega$ |
| $\mathbf{v}_{\omega}$ | Concatenated vector of the $\left(v_{it}^{r\omega}, v_{it}^{k\omega}, v_{it}^{e\omega}\right)$, $\forall i \in \mathcal{I}, r \in \mathcal{R}, k \in \mathcal{K}, e \in \mathcal{E}$, and $t \in \mathcal{T}$ under scenario $\omega$ |
| $\mathbf{v}_{\omega}^{r}$ | Concatenated vector of the $v_{it}^{r\omega}$, $\forall i \in \mathcal{I}, r \in \mathcal{R}$, and $t \in \mathcal{T}$ under scenario $\omega$ |
| $\mathbf{v}_{\omega}^{k}$ | Concatenated vector of the $v_{it}^{k\omega}$, $\forall i \in \mathcal{I}, k \in \mathcal{K}$, and $t \in \mathcal{T}$ under scenario $\omega$ |
| $\mathbf{v}_{\omega}^{e}$ | Concatenated vector of the $v_{it}^{e\omega}$, $\forall i \in \mathcal{I}, e \in \mathcal{E}$, and $t \in \mathcal{T}$ under scenario $\omega$ |
| $\mathbf{x}_{\omega}$ | Concatenated vector of the $\left(x_{ijt}^{r\omega}, x_{ijt}^{k\omega}, x_{ijt}^{e\omega}\right)$, $\forall i \in \mathcal{I}, j \in \mathcal{J}, r \in \mathcal{R}, k \in \mathcal{K}$, $e \in \mathcal{E}$, and $t \in \mathcal{T}$ under scenario $\omega$ |
| $\mathbf{x}_{\omega}^{r}$ | Concatenated vector of the $x_{ijt}^{r\omega}$, $\forall i \in \mathcal{I}, j \in \mathcal{J}, r \in \mathcal{R}$, and $t \in \mathcal{T}$ under scenario $\omega$ |
| $\mathbf{x}_{\omega}^{k}$ | Concatenated vector of the $x_{ijt}^{k\omega}$, $\forall i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K}$, and $t \in \mathcal{T}$ under scenario $\omega$ |
| $\mathbf{x}_{\omega}^{e}$ | Concatenated vector of the $x_{ijt}^{e\omega}$, $\forall i \in \mathcal{I}, j \in \mathcal{J}, e \in \mathcal{E}$, and $t \in \mathcal{T}$ under scenario $\omega$ |
| $\mathbf{z}_{\omega}$ | Concatenated vector of the $z_{it}^{\omega}$, $\forall i \in \mathcal{I}$, and $t \in \mathcal{T}$ under scenario $\omega$ |

With the above vectors of decision variables, the extensive form can be simplified as follows:

**Compact formulation**

$$\min \quad \mathbf{f}^{\intercal}\mathbf{y} + \mathbf{e}^{\intercal}\mathbf{g} + \mathbb{E}_{\zeta}\left[Q(\mathbf{y}, \mathbf{r}, \zeta^{\omega})\right] \tag{3.29}$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{g} \leq \mathbf{1}, \tag{3.30}$$

$$\mathbf{B}\mathbf{g} = \mathbf{r}, \tag{3.31}$$

$$\mathbf{W}\mathbf{r} \leq \mathbf{1}, \tag{3.32}$$

$$\mathbf{y} \in \{0,1\}^{|\mathcal{J}|}, \tag{3.33}$$

$$\mathbf{g}, \mathbf{r} \in \{0,1\}^{|\mathcal{K}||\mathcal{M}||\mathcal{T}|}. \tag{3.34}$$

where $\quad Q(\mathbf{y}, \mathbf{r}, \zeta^{\omega}) =$

$$\min \quad \mathbf{h}^{\intercal}\mathbf{v}_{\omega} + \mathbf{b}^{\intercal}\mathbf{u}_{\omega} + \boldsymbol{\beta}^{\intercal}\mathbf{z}_{\omega} + \mathbf{c}^{\intercal}\mathbf{x}_{\omega} \tag{3.35}$$

$$\text{s.t.} \quad \mathbf{P}\mathbf{x}_{\omega} \leq \mathbf{S}_{\omega}\mathbf{y}, \tag{3.36}$$

$$\mathbf{U}\mathbf{u}_{\omega} + \mathbf{V}\mathbf{v}_{\omega} - \mathbf{T}\mathbf{x}_{\omega} = \mathbf{0}, \tag{3.37}$$

$$\mathbf{K}\mathbf{u}_{\omega} + \mathbf{J}\mathbf{z}_{\omega} \geq \mathbf{D}_{\omega}, \tag{3.38}$$

$$\mathbf{M}\mathbf{v}^{r}_{\omega} \leq \mathbf{C}^{r}, \tag{3.39}$$

$$\mathbf{G}\mathbf{v}^{e}_{\omega} \leq \mathbf{C}^{e}, \tag{3.40}$$

$$\mathbf{H}\mathbf{v}^{k}_{\omega} \leq \mathbf{C}^{k}\mathbf{r}, \tag{3.41}$$

$$\mathbf{E}\mathbf{x}^{r}_{\omega} \leq \mathbf{C}^{r}, \tag{3.42}$$

$$\mathbf{R}\mathbf{x}^{e}_{\omega} \leq \mathbf{C}^{e}, \tag{3.43}$$

$$\mathbf{L}\mathbf{x}^{k}_{\omega} \leq \mathbf{C}^{k}\mathbf{r}, \tag{3.44}$$

$$\mathbf{v}_{\omega}, \mathbf{u}_{\omega}, \mathbf{z}_{\omega}, \mathbf{x}_{\omega} \geq \mathbf{0}. \tag{3.45}$$

In the case in which the objective function of the stochastic programming model is well-defined, the model possesses the optimal solution [22]. As mentioned earlier, because of the relatively complete recourse property, the feasibility of the proposed

model (3.29)−(3.34) will always be guaranteed for all $\mathbf{y} \in \mathcal{Y}, \mathbf{g} \in \mathcal{G}, \mathbf{r} \in \mathcal{V}$ and $\omega \in \Omega$, where $Y, G$, and $R$ are feasible sets of the corresponding decision variables. The objective function is to minimize the sum of the cost for the first-stage problem, $\mathbf{f}^\intercal \mathbf{y} + \mathbf{e}^\intercal \mathbf{g}$, and the expected cost for the second-stage problem, $\mathbb{E}_\zeta \left[Q(\mathbf{y}, \mathbf{r}, \zeta^\omega)\right]$ . Therefore, the lost sales cost term, $\boldsymbol{\beta}^\intercal \mathbf{z}_\omega$, guarantees that $Q(\mathbf{y}, \mathbf{r}, \zeta^\omega) \leq \infty$ for all $\mathbf{y}, \mathbf{r}$, and $\omega$. Moreover, because we assume that all cost parameters are non-negative, it is obvious that $Q(\mathbf{y}, \mathbf{r}, \zeta^\omega) \geq -\infty$ for all $\mathbf{y}, \mathbf{r}$, and $\omega$. Thus, $Q(\mathbf{y}, \mathbf{r}, \zeta^\omega)$ is finite for all $\mathbf{y}, \mathbf{r}$ and every realization of $\omega$, and it can be assumed that the expected value, $\mathbb{E}_\zeta \left[Q(\mathbf{y}, \mathbf{r}, \zeta^\omega)\right]$, is well defined. Finally, the objective function of the first-stage variables is well-defined in the proposed model (3.29)−(3.45), and the optimal solutions exist because the set $\mathcal{Y}, \mathcal{G}$, and $\mathcal{V}$ is nonempty and finite.

## 3.4 Solution methodology: Sample average approximation combined with the Benders decomposition algorithm

In this section, we develop the solution methodology, specifically the SAA and BD algorithms, for solving the proposed TSSP. There are several advantages of using the SAA and BD algorithms compared to other methods [141, 115]. First, the SAA approach is quite general, so that can be combined with various algorithms that are specialized in solving the deterministic optimization problem. Also, the SAA approach has valuable convergence properties. The BD algorithm can efficiently solve complicated problems due to several variables, which makes the problem easier to handle when temporarily fixed. The BD algorithms converge to the optimal of the MILP rather than to a relaxation of the problem. Section 3.4.1 presents the

concept and procedure of the SAA. Section 3.4.2 examines the BD algorithm, and Section 3.4.3 illustrates the acceleration method for the BD algorithm.

### 3.4.1 Sample average approximation

The fundamental difficulty of solving the *true problem* (3.29)−(3.45) is computing the expected value function, $\mathbb{E}_{\zeta}\left[Q(\mathbf{y}, \mathbf{r}, \zeta^{\omega})\right]$. Let $\zeta^1, \cdots, \zeta^N$ be an independently and identically distributed (i.i.d) random sample of $N$ realizations (scenarios) of the stochastic parameter vector $\zeta$. By solving the following *SAA problem* with a larger $N$, the objective function of the SAA problem converges to the true objective function with a probability of one [79].

$$\min_{\mathbf{y}\in\mathcal{Y},\mathbf{g}\in\mathcal{G},\mathbf{r}\in\mathcal{V}} \left\{ \mathbf{f}^{\intercal}\mathbf{y} + \mathbf{e}^{\intercal}\mathbf{g} + \frac{1}{N}\sum_{n=1}^{N} Q(\mathbf{y}, \mathbf{r}, \zeta^{n}) \right\} \tag{3.46}$$

Let $\hat{\psi}_N$ denote the optimal value of the SAA problem (3.46), and $\hat{\psi}_N$ is random because the value will be different depending on the corresponding random sample.

However, the computational complexity for solving the SAA problem (3.46) often increases exponentially with the size of $N$. In order to overcome these challenges, we utilize the SAA algorithm, which estimates the objective value of the true problem and requires less computational effort than solving the SAA problem with a large-sized $N$.

In the SAA algorithm, we employ the number $M$ of replications, generating and solving the SAA problem with the same size $N$. It is more efficient to utilize several SAA problems with a smaller-sized $N$ than it is to solve one SAA problem with a large-sized $N$. Based on the number $M$ of SAA replications, the solution quality

of each replication is measured with an optimality gap. In this chapter, the *SAA gap* stands for an optimality gap used for stopping criteria in the SAA algorithm. When the SAA gap can not satisfy the predefined threshold $\epsilon_{SAA}$, we increase the sample size $N$ for every SAA replication to obtain solutions with better quality. The procedure for the SAA algorithm is described as follows:

**SAA algorithm**

1. Generate i.i.d. samples with size $N$ scenarios for each replication of $m$ (i.e., $\left(\zeta_m^1, \cdots, \zeta_m^n\right)$, $\forall m \in \{1, \cdots, M\}$), and solve the corresponding SAA problem. Let $\hat{\psi}_N^m$ and $\hat{\mathbf{y}}_N^m, \hat{\mathbf{g}}_N^m$, and $\hat{\mathbf{r}}_N^m$ be the optimal objective value and the optimal solution of the $m$th SAA replication, respectively.

2. Compute the following equation to obtain the statistical lower bound for $\psi^*$, where $\psi^*$ is the optimal objective value for the true problem.

$$\bar{\psi}_{MN} := \frac{1}{M} \sum_{m=1}^{M} \hat{\psi}_N^m \tag{3.47}$$

   It is well known that the expected value of the $\hat{\psi}_N$ is less than or equal to the $\psi^*$ [102, 41]. Because $\bar{\psi}_{MN}$ is the unbiased estimator for the $\mathbb{E}[\hat{\psi}_N]$, it is clear that $\bar{\psi}_{MN}$ provides the statistical lower bound for $\psi^*$, $\mathbb{E}[\bar{\psi}_{MN}] \leq \psi^*$. Let $\sigma^2_{\bar{\psi}_{MN}}$ be an estimate of the variance of $\bar{\psi}_{MN}$. It can be obtained by computing the following equation, which is derived from the *Central Limit Theorem*.

$$\sigma^2_{\bar{\psi}_{MN}} := \frac{1}{M(M-1)} \sum_{m=1}^{M} \left(\hat{\psi}_N^m - \bar{\psi}_{MN}\right)^2 \tag{3.48}$$

3. Select a feasible first-stage solution, $\hat{\mathbf{y}} \in \mathcal{Y}, \hat{\mathbf{g}} \in \mathcal{G}$, and $\hat{\mathbf{r}} \in \mathcal{V}$. This feasible

first-stage solution was determined from the obtained solution by solving the SAA problem for each replication, $\hat{\mathbf{y}}_N^m, \hat{\mathbf{g}}_N^m$, and $\hat{\mathbf{r}}_N^m$. With a newly-generated sample of $N'$ scenarios, $\left(\zeta^1, \cdots, \zeta^{N'}\right)$, the optimal value of the true problem is estimated from the following equation.

$$\bar{f}_{N'}(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}}) := \mathbf{f}^\intercal \widehat{\mathbf{y}} + \mathbf{e}^\intercal \widehat{\mathbf{g}} + \frac{1}{N'} \sum_{n=1}^{N'} Q(\widehat{\mathbf{y}}, \widehat{\mathbf{r}}, \zeta^n) \tag{3.49}$$

Note that the size of $N'$ is much larger than the sample size of $N$ used to obtain the estimate for the lower bound ($N \ll N'$). Among obtained solutions $\hat{\mathbf{y}}_N^m, \hat{\mathbf{g}}_N^m$, and $\hat{\mathbf{r}}_N^m, \forall m$, a solution that has the smallest value, $\bar{f}_{N'}$, is commonly chosen for $\hat{\mathbf{y}}, \hat{\mathbf{g}}$, and $\hat{\mathbf{r}}$ to estimate the upper bound. Let $f(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}})$ be the optimal objective value of the true problem with the solution $\hat{\mathbf{y}}, \hat{\mathbf{g}}$, and $\hat{\mathbf{r}}$. The inequality $\psi^* \leq f(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}})$ holds because $\hat{\mathbf{y}}, \hat{\mathbf{g}}$, and $\hat{\mathbf{r}}$ are the feasible solutions of the true problem. Then, because $\bar{f}_{N'}(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}})$ is the unbiased estimator of $f(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}})$, $\bar{f}_{N'}(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}})$ provides an upper bound for $\psi^*$. Similar to the way of deriving $\sigma^2_{\bar{\psi}_{MN}}$, the estimate of the variance of $\bar{f}_{N'}(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}})$ can be obtained by the following equation.

$$\sigma^2_{N'}(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}}) := \frac{1}{N'(N'-1)} \sum_{n=1}^{N'} \left(\mathbf{f}^\intercal \widehat{\mathbf{y}} + \mathbf{e}^\intercal \widehat{\mathbf{g}} + Q(\hat{\mathbf{y}}, \hat{\mathbf{r}}, \zeta^n) - \bar{f}_{N'}(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}})\right)^2 \tag{3.50}$$

4. Obtain the SAA gap of the feasible solution $\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}}$ and its variance by calculating the following equations:

$$Gap_{MNN'}(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}}) := \bar{f}_{N'}(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}}) - \bar{\psi}_{MN} \tag{3.51}$$

The relative SAA gap is computed by the following equation:

$$Gap^{rel}_{MNN'} := \frac{(\bar{f}_{N'}(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}}) - \bar{\psi}_{MN})}{\bar{\psi}_{MN}} \times 100(\%) \tag{3.52}$$

The estimate of the variance of $Gap_{MNN'}$ can be calculated as follows:

$$\sigma^2_{Gap_{MNN'}} := \sigma^2_{N'}(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}}) + \sigma^2_{\hat{\psi}_{MN}} \tag{3.53}$$

### 3.4.2 Benders decomposition algorithm

By applying the SAA algorithm, we can obtain a stochastic solution. However, for the large problem, a lot of computational effort is required to solve the SAA problem (3.46) even with the moderate size of $N$ scenarios. Therefore, we alleviate the computational burden by utilizing a special property of the TSSP. It is well known that the TSSP has the block structure of the extensive form. When taking the dual of the extensive form, a *dual block-angular structure* appears, and the BD algorithm is a suitable approach to exploit this structure [13, 22]. As mentioned in Section 3.4.1, because the SAA problem (3.46) is itself the TSSP, we use the BD algorithm to solve the SAA problem.

Without loss of generality, we explain the BD algorithm with the model (3.29)−(3.45), which will be referred to as the *original problem*. We present the multi-cut version of the BD algorithm, which generates several optimality cuts in one iteration. The $1/N$ and $\{\zeta^1, \cdots, \zeta^N\}$ replace $p_\omega$ and $\Omega$, respectively, when applying the BD algorithm for the SAA problem. In order to devise the BD algorithm, the proposed stochastic mathematical model is decomposed into one master problem (MP) and

several subproblems (SUB($\omega$), $\forall\omega \in \Omega$). MP and the corresponding SUB($\omega$), $\forall\omega \in \Omega$, in the ($itr$+1)th iteration are presented as follows:

**MP**

$$\min \quad \mathbf{f}^\intercal\mathbf{y} + \mathbf{e}^\intercal\mathbf{g} + \sum_{\omega\in\Omega} p_\omega\theta_\omega \tag{3.54}$$

$$\text{s.t.} \quad \mathbf{Ag} \leq \mathbf{1}, \tag{3.55}$$

$$\mathbf{Bg} = \mathbf{r}, \tag{3.56}$$

$$\mathbf{Wr} \leq \mathbf{1}, \tag{3.57}$$

$$\theta_\omega \geq (\mathbf{a}_\omega^{itr})^\intercal\mathbf{y} + (\mathbf{c}_\omega^{itr})^\intercal\mathbf{r} + d_\omega^{itr}, \qquad \forall itr \in \mathfrak{I}, \omega \in \Omega, \tag{3.58}$$

$$\mathbf{y} \in \{0,1\}^{|\mathcal{J}|}, \tag{3.59}$$

$$\mathbf{g}, \mathbf{r} \in \{0,1\}^{|\mathcal{K}||\mathcal{M}||\mathcal{T}|}. \tag{3.60}$$

where $\mathfrak{I} := \{1, \cdots, itr\}$ and $\theta_\omega$, $\forall\omega \in \Omega$, are free variables. Constraint (3.58) is called as *optimality cuts* at iteration $itr$, and coefficients $(\mathbf{a}_\omega^{itr})^\intercal$, $(\mathbf{c}_\omega^{itr})^\intercal$, and $d_\omega^{itr}$ will be explained in the latter part of this section. After solving the MP with current optimality cuts, obtained optimal solutions are denoted as $\bar{\mathbf{y}}, \bar{\mathbf{g}}, \bar{\mathbf{r}}$, and $\bar{\theta}_\omega$, $\forall\omega \in \Omega$. Because MP is the relaxed problem to the model (3.29)−(3.45), the optimal objective value of MP provides the lower bound, $Z_{lb}$, for the original problem.

Based on the obtained solution from MP, we solve the SUB($\omega$) for each $\omega \in \Omega$. $SUB(\omega)$ is presented as follows:

**SUB($\omega$)**

$$\min \quad \mathbf{h}^\intercal\mathbf{v}_\omega + \mathbf{b}^\intercal\mathbf{u}_\omega + \boldsymbol{\beta}^\intercal\mathbf{z}_\omega + \mathbf{c}^\intercal\mathbf{x}_\omega \tag{3.61}$$

$$
\begin{aligned}
\text{s.t.} \quad & \mathbf{P}\mathbf{x}_\omega \leq \mathbf{S}_\omega \mathbf{y}, && (\boldsymbol{\pi}_\omega), && (3.62)\\
& \mathbf{U}\mathbf{u}_\omega + \mathbf{V}\mathbf{v}_\omega - \mathbf{T}\mathbf{x}_\omega = \mathbf{0}, && (\boldsymbol{\mu}_\omega), && (3.63)\\
& \mathbf{K}\mathbf{u}_\omega + \mathbf{J}\mathbf{z}_\omega \geq \mathbf{D}_\omega, && (\boldsymbol{\nu}_\omega), && (3.64)\\
& \mathbf{M}\mathbf{v}^r_\omega \leq \mathbf{C}^r, && (\boldsymbol{\lambda}_\omega), && (3.65)\\
& \mathbf{G}\mathbf{v}^e_\omega \leq \mathbf{C}^e, && (\boldsymbol{\tau}_\omega), && (3.66)\\
& \mathbf{H}\mathbf{v}^k_\omega \leq \mathbf{C}^k \bar{\mathbf{r}}, && (\boldsymbol{\rho}_\omega), && (3.67)\\
& \mathbf{E}\mathbf{x}^r_\omega \leq \mathbf{C}^r, && (\boldsymbol{\delta}_\omega), && (3.68)\\
& \mathbf{R}\mathbf{x}^e_\omega \leq \mathbf{C}^e, && (\boldsymbol{\iota}_\omega), && (3.69)\\
& \mathbf{L}\mathbf{x}^k_\omega \leq \mathbf{C}^k \bar{\mathbf{r}}, && (\boldsymbol{\kappa}_\omega), && (3.70)\\
& \mathbf{v}_\omega, \mathbf{u}_\omega, \mathbf{z}_\omega, \mathbf{x}_\omega \geq \mathbf{0}.
\end{aligned}
$$

where the Greek bold-faced terms in parenthesis denote the corresponding vectors of the optimal dual solution with appropriate dimensions. Let $\mathrm{Q}(\bar{\mathbf{y}}, \bar{\mathbf{r}}, \zeta^\omega)$ denote the optimal objective value of SUB($\omega$) with first-stage variables $\bar{\mathbf{y}}$, and $\bar{\mathbf{r}}$ under the scenario $\omega$. The optimal objective value and solutions can be derived easily because every SUB($\omega$) is a simple linear programming model. Furthermore, the optimal primal solution SUB($\omega$) for each $\omega \in \Omega$ is feasible for the original problem. Hence, the following equation provides the upper bound, $Z_{ub}$, for the original problem:

$$
Z_{ub} := \mathbf{f}^\intercal \mathbf{y} + \mathbf{e}^\intercal \mathbf{g} + \sum_{\omega \in \Omega} p_\omega Q(\bar{\mathbf{y}}, \bar{\mathbf{r}}, \zeta^\omega) \tag{3.71}
$$

If for every scenario $\omega \in \Omega$, $Q(\bar{\mathbf{y}}, \bar{\mathbf{r}}, \omega)$ is less than or equal to $\bar{\theta}_\omega$ from MP, then the current solution is optimal to the original problem ( i.e.,$Z_{ub} = Z_{lb}$). Otherwise, if the

SUB($\omega$) corresponding to some $\omega$ has $Q(\bar{\mathbf{y}}, \bar{\mathbf{r}}, \omega)$ greater than $\bar{\theta}_\omega$, the corresponding optimality cuts are added to the MP. An optimality cut for scenario $\omega$ is generated as follows:

$$\theta_\omega \geq (\mathbf{a}_\omega^{itr+1})^\intercal \mathbf{y} + (\mathbf{c}_\omega^{itr+1})^\intercal \mathbf{r} + d_\omega^{itr+1} \tag{3.72}$$

Coefficients of the optimality cut are calculated as below:

$$(\mathbf{a}_\omega^{itr+1})^\intercal := \boldsymbol{\pi}_\omega{}^\intercal \mathbf{S}_\omega \tag{3.73}$$

$$(\mathbf{c}_\omega^{itr+1})^\intercal := (\boldsymbol{\rho}_\omega + \boldsymbol{\kappa}_\omega)^\intercal \mathbf{C}^k \tag{3.74}$$

$$d_\omega^{itr+1} := \boldsymbol{\nu}_\omega{}^\intercal \mathbf{D}_\omega + (\boldsymbol{\lambda}_\omega + \boldsymbol{\delta}_\omega)^\intercal \mathbf{C}^r + (\boldsymbol{\tau}_\omega + \boldsymbol{\iota}_\omega)^\intercal \mathbf{C}^e \tag{3.75}$$

This procedure is implemented iteratively until the condition $(Z_{ub} - Z_{lb})/Z_{lb} < \epsilon_{BD}$ is satisfied, where $\epsilon_{BD}$ is the pre-determined control parameter. It is worth mentioning that because the proposed stochastic model has a relatively complete recourse, we do not consider the feasibility cut, which is necessary for the case in which some SUB($\omega$) are infeasible according to the optimal solution of MP.

### 3.4.3 Acceleration method

At the beginning of the typical BD algorithm (TBD), MP is initially solved with an empty set of optimality cuts. Then, based on the optimal dual solution of SUB($\omega$), optimality cuts at the first iteration are added to the MP. We refer to these optimality cuts as *initial optimality cuts*, which are generated in the first iteration.

However, it is obvious that the MP with an empty set of optimality cuts could provide a poor feasible solution (e.g., $\bar{\mathbf{y}}$, $\bar{\mathbf{g}}$, and $\bar{\mathbf{r}}$ are zero and $\bar{\theta}_\omega$ are negative in value).

In this case, initial optimality cuts cannot contribute to creating a better lower bound because poor solutions tend to generate ineffective cuts [78]. Consequently, the TBD algorithm could incur a lot of iterations until the termination condition and naturally increase the total computation time. Therefore, we devise a simple method for accelerating the convergence of bounds in the BD algorithm and for reducing the number of required iterations by generating effective initial optimality cuts at the first step.

Prior to presenting the acceleration method, let us first introduce the *expected value problem* (EVP). This simple model is obtained by replacing all stochastic parameters with their expected values, and the optimal solution to the EVP is called the *expected value solution* (EVS) [22]. The EVS can sometimes be a high-quality solution to the true problem. By utilizing this property, we utilize the EVS to generate better initial cuts than the typical one at the initialization step of the BD algorithm. The detailed procedure for the acceleration method is as follows:

**Acceleration method**

1. Obtain the expected value solution $\bar{\mathbf{y}}, \bar{\mathbf{g}}, \bar{\mathbf{r}}$, by solving the EVP with a commercial solver until the computation time falls within 30 seconds or until the gap between the best solution and the best bound falls within 5%.

2. Solve SUB($\omega$) for each $\omega$ based on the obtained expected value solution in Step 1. Then, obtain the optimal objective value $Q(\bar{\mathbf{y}}, \bar{\mathbf{g}}, \omega)$ and optimal dual solution $\boldsymbol{\pi}_\omega, \boldsymbol{\mu}_\omega, \boldsymbol{\nu}_\omega, \boldsymbol{\lambda}_\omega, \boldsymbol{\tau}_\omega, \boldsymbol{\rho}_\omega, \boldsymbol{\delta}_\omega, \boldsymbol{\iota}_\omega$, and $\boldsymbol{\kappa}_\omega$ for all $\omega$.

3. Generate initial optimality cuts with the obtained objective value and optimal dual solution in Step 2. After generating the initial optimality cuts, the

subsequent procedure is the same as the BD algorithm.

In Step 1, we set the stopping criteria as 30 seconds and the gap within 5% because it costs a computational burden to solve the EVP to get the optimal solution costs in a large-sized problem. Moreover, by implementing a lot of computational experiments, we observed that there was no obvious performance difference between the optimal solution and the sub-optimal solution of the EVP for improving the final computation time of the BD algorithm. Finally, the BD algorithm with the acceleration method (ABD) is presented in Algorithm 2.

**Algorithm 2** Benders decomposition algorithm (Acceleration method)

**Initialization:**

$Z_{ub} \leftarrow \infty,\ Z_{lb} \leftarrow -\infty, itr \leftarrow 1$

solve $EVP$ and get $\bar{\mathbf{y}}, \bar{\mathbf{g}}, \bar{\mathbf{r}}$

**for** $\omega \in \Omega$ **do**

  solve $SUB(\omega)$ based on $\bar{\mathbf{y}}, \bar{\mathbf{g}}$, and $\bar{\mathbf{r}}$

  get $(\mathbf{a}_\omega^{itr})^\intercal, (\mathbf{c}_\omega^{itr})^\intercal, d_\omega$ with optimal dual solutions

  add initial optimality cuts to $MP$

**end**

**while** $Z_{ub} - Z_{lb} \geq \epsilon_{BD} \times Z_{lb}$ **do**

  solve $MP$ and get $\bar{\mathbf{y}}, \bar{\mathbf{g}}, \bar{\mathbf{r}}, \bar{\theta}_\omega, \forall \omega \in \Omega$

  $Z_{lb} \leftarrow \max\left\{Z_{lb}, \mathbf{f}^\intercal\mathbf{y} + \mathbf{e}^\intercal\mathbf{g} + \sum_{\omega\in\Omega} p_\omega \bar{\theta}_\omega\right\}$

  **for** $\omega \in \Omega$ **do**

    solve $SUB(\omega)$ and get dual solution

    get $(\mathbf{a}_\omega^{itr})^\intercal, (\mathbf{c}_\omega^{itr})^\intercal, d_\omega$ with optimal dual solutions

    store the optimal objective value $Q(\bar{\mathbf{y}}, \bar{\mathbf{r}}, \zeta^\omega)$

    **if** $\bar{\theta}_\omega < Q(\bar{\mathbf{y}}, \bar{\mathbf{r}}, \zeta^\omega)$ **then**

      add an optimality cut to $MP$

    **end**

  **end**

  **if** $Z_{ub} > \mathbf{f}^\intercal\mathbf{y} + \mathbf{e}^\intercal\mathbf{g} + \sum_{\omega\in\Omega} p_\omega Q(\bar{\mathbf{y}}, \bar{\mathbf{r}}, \zeta^\omega)$ **then**

    $Z_{ub} \leftarrow \mathbf{f}^\intercal\mathbf{y} + \mathbf{e}^\intercal\mathbf{g} + \sum_{\omega\in\Omega} p_\omega Q(\bar{\mathbf{y}}, \bar{\mathbf{r}}, \zeta^\omega)$

    $\mathbf{y}^* \leftarrow \bar{\mathbf{y}}, \mathbf{g}^* \leftarrow \bar{\mathbf{g}}, \mathbf{r}^* \leftarrow \bar{\mathbf{r}}$

    $\mathbf{u}_\omega^* \leftarrow \bar{\mathbf{u}}_\omega, \mathbf{v}_\omega^* \leftarrow \bar{\mathbf{v}}_\omega, \mathbf{x}_\omega^* \leftarrow \bar{\mathbf{x}}_\omega, \forall \omega \in \Omega$

  **end**

  $itr \leftarrow itr + 1$

**end**

**Return:** $Z_{ub}, Z_{lb}, \mathbf{y}^*, \mathbf{g}^*, \mathbf{r}^*, \mathbf{u}_\omega^*, \mathbf{x}_\omega^*, \mathbf{v}_\omega^*, \forall \omega \in \Omega;$

## 3.5 Computational experiments

In this section, we conducted three types of computational experiments to answer the research questions in Section 3.1. Research question 1 is answered by the results of experiments in Sections 3.5.2 and 3.5.3. Four types of computational experiments

were implemented in Section 3.5.4. The first experiment result answers the Research question 2, and the second and third experiments answer the Research question 3. Research question 4 is answered by the results of the fourth experiment. We suggest several managerial insights in Section 3.5.5 based on the computational results. All the experiments were conducted on a PC with an AMD Ryzen 2700X 8-Core CP, 3.60 GHz processor, and 16GB of RAM with a Windows 10 64-bit system. Test instances were generated using Python 3.8, and every solution approach was developed with FICO Xpress 8.5 and Xpress-Optimizer version 33.01.02.

### 3.5.1 Description of the test instances

To validate the performance of the proposed algorithms, we need benchmark instances. However, as far as we know, there are no existing benchmark instances corresponding to our problem. Therefore, we rely on real-world information for determining the values of the parameters. At first, inventory holding costs, $h_i^r$ and $h_i^k$, were generated on the basis of the article by Hass [67]. The cost of delivery, $b_i$, was determined based on the cost of the parcel delivery service in South Korea. To cover various cases, other deterministic parameters were randomly generated with the range of uniform distributions detailed in Table B.1.

In order to estimate the distributions of stochastic parameters, we used the e-commerce public dataset [74], which consists of demand data for 614 time periods, from September 4, 2016 to September 3, 2018. Then, we fitted the normal distribution to this dataset to estimate the distributions of demands, $D_{it}^{\omega}$. We set the negative-value of realized demands or supplies to zero and adopted the same distribution of $D_{it}^{\omega}$ for $S_{ijt}^{\omega}$. Consequently, every random sample of $N$ scenarios is realized

based on the estimated distribution of stochastic parameters shown in Table B.2.

The locations of suppliers and warehouses are uniformly distributed over the pre-specified width and height of the XY plane. Moreover, the unit transportation costs, $c_j^r, c_j^k$, and $c_j^e$, are assumed to be proportional to the Euclidean distance in the XY plane. Because of the assumption that it is expensive to use the emergency warehouse, the values of $c_j^e$ and $h_i^e$ are significantly larger than the cost of the retailer or provider warehouses.

Based on the model given, the size of a problem is determined by $|\mathcal{I}|$, $|\mathcal{J}|$, $|\mathcal{T}|$, $|\mathcal{K}|$, and $|\mathcal{M}|$. We produced test instances ranging from small to large sizes. In particular, we classified the mathematical model using the test instances 13$\sim$15 for input as the large-sized problem. Every test instance is generated randomly according to the uniform distribution in Table B.1. The detailed characteristics of test instances are indicated in Table 3.2. The columns labeled 'XY' represents the width and height of the XY plane. The number of variables (Vars) and constraints (Cons) are calculated for a scenario size $N = 40$.

We conducted every computational experiment considering one retailer warehouse, one emergency warehouse, and multiple provider warehouses according to assumptions in Section 3.3.1. Through implementing a lot of experiments, we observed that the emergency warehouse was rarely used because of the high operational cost compared to the stockout cost. Hence, in Sections 3.5.2 and 3.5.3, we accommodated the problem in which the emergency warehouse has unlimited capacity for storage and transportation. However, in Section 3.5.4, we considered the limited capacity of the emergency warehouse in order to analyze the effects of lead time and the lost sales cost parameter. The two types of lead time, $L_s$ and $L_d$, were set to

zero in Sections 3.5.2 and 3.5.3, but we evaluated the impacts of these two types of lead time by varying values in Section 3.5.4.

Table 3.2: Test instances specifications ($N = 40$)

| No. | XY | Total Vars | Binary Vars | Cont Vars | Cons | $\lvert\mathcal{I}\rvert$ | $\lvert\mathcal{J}\rvert$ | $\lvert\mathcal{T}\rvert$ | $\lvert\mathcal{K}\rvert$ | $\lvert\mathcal{M}\rvert$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 100×100 | 29,103 | 153 | 28,950 | 14,000 | 2 | 3 | 10 | 5 | 3 |
| 2 | | 32,899 | 131 | 32,768 | 15,104 | 2 | 3 | 8 | 8 | 2 |
| 3 | | 73,844 | 324 | 73,520 | 26,000 | 3 | 4 | 10 | 8 | 4 |
| 4 | | 88,204 | 304 | 87,900 | 30,000 | 3 | 4 | 10 | 10 | 3 |
| 5 | | 88,803 | 483 | 88,320 | 34,800 | 3 | 3 | 12 | 10 | 4 |
| 6 | 300×300 | 146,859 | 471 | 146,388 | 51,168 | 4 | 3 | 12 | 13 | 3 |
| 7 | | 149,404 | 544 | 148,860 | 48,480 | 3 | 4 | 12 | 15 | 3 |
| 8 | | 289,355 | 680 | 288,675 | 76,200 | 4 | 5 | 15 | 15 | 3 |
| 9 | | 434,705 | 1,355 | 433,350 | 109,080 | 5 | 5 | 18 | 15 | 5 |
| 10 | | 467,405 | 1,205 | 466,200 | 112,200 | 5 | 5 | 15 | 20 | 4 |
| 11 | 500×500 | 532,806 | 906 | 531,900 | 114,600 | 5 | 6 | 15 | 20 | 3 |
| 12 | | 561,605 | 1,805 | 559,800 | 135,360 | 5 | 5 | 18 | 20 | 5 |
| 13 | | 1,046,606 | 2,506 | 1,044,100 | 210,800 | 6 | 6 | 20 | 25 | 5 |
| 14 | | 1,049,606 | 4,006 | 1,045,600 | 213,800 | 6 | 6 | 20 | 25 | 8 |
| 15 | | 1,808,006 | 4,506 | 1,803,500 | 345,500 | 7 | 6 | 25 | 30 | 6 |

### 3.5.2 Performance analysis of the proposed algorithms

As mentioned in Section 3.4.2, the SAA problem with the moderate size $N$ could suffer from the computational burden. In this section, we conducted computational experiments to compare the three solution approaches: TBD, ABD, and Solver (solving the given problem with an Xpress-Optimizer). Test instances with different sizes of $N$ were employed to evaluate the performance of the proposed algorithms. For each size of $N$ and solution approach, ten experiments were conducted with different samples of $N$ scenarios. An average of ten experiments has been reported in Table B.3 with comparison results among Solver, TBD, and ABD. The columns labeled 'CPUs' and 'Itr' represent the computation times in seconds and the number

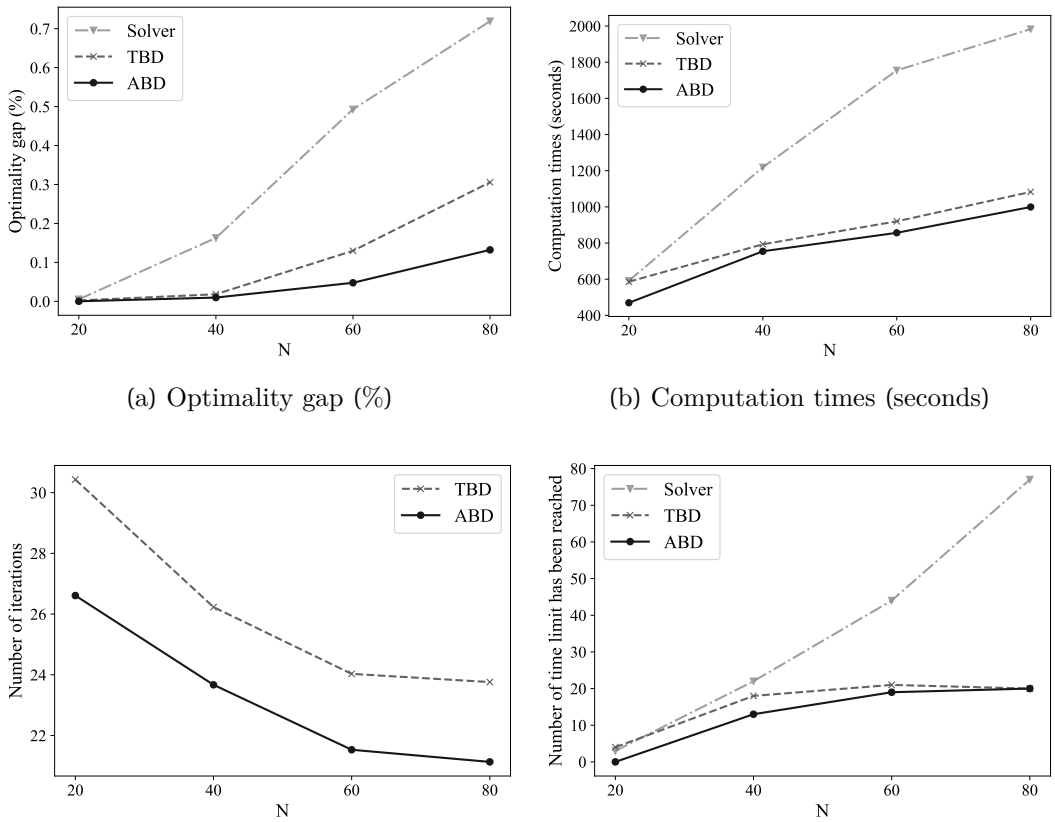
of iterations required to make the optimality gap of BD algorithms (TBD and ABD) less than the pre-determined threshold, $\epsilon_{BD}$. The 'Gap' is defined as follows:

$$\text{Gap} := \left( \frac{\text{Best solution (OBJ by each approach)}}{\text{Best bound} \left( \max \{ Z_{lb} \text{ by ABD}, Z_{lb} \text{ by TBD} \} \right)} - 1 \right) \times 100(\%) \tag{3.76}$$

If the maximum time limit (i.e., 3,600 seconds) was reached, algorithms were terminated, and they output the Gap, CPUs, and Itr obtained so far. We set the $\epsilon_{BD}$ for $10^{-4}$ for both TBD and ABD.

The computational results of all test instances in Table B.3 were averaged in terms of 'Gap', 'CPUs', 'Itr', and the number of times each algorithm reached the time limit are depicted in Figure 3.3. Figure 3.3 indicates that the ABD outperformed the TBD and Solver in terms of every evaluation measure. Furthermore, Figure 3.3 shows that more computation time was required to solve the problem as the size of $N$ increased. However, the computation time of the BD algorithms increased more slowly when compared to the Solver. On the other hand, a small number of iterations was required as the size of $N$ increased for both BD algorithms. The results appeared because as the $N$ increased, it required more time to implement one iteration compared to the smaller size of $N$.

In order to analyze the effects of the initial optimality cuts of ABD, we compared the convergence of bounds for the TBD and ABD. We used test instances 12∼15 with $N = 40$ and set the $\epsilon_{BD}$ to 0.03 for visualizing the apparent convergence. Figure 3.4 represents a comparison between TBD and ABD concerning the $Z_{ub}$, and the $Z_{lb}$. As the number of iterations increased, the upper bound decreased, and the lower bound increased for both algorithms until the value of $(Z_{ub} - Z_{lb})/Z_{lb}$ within




(a) Optimality gap (%)

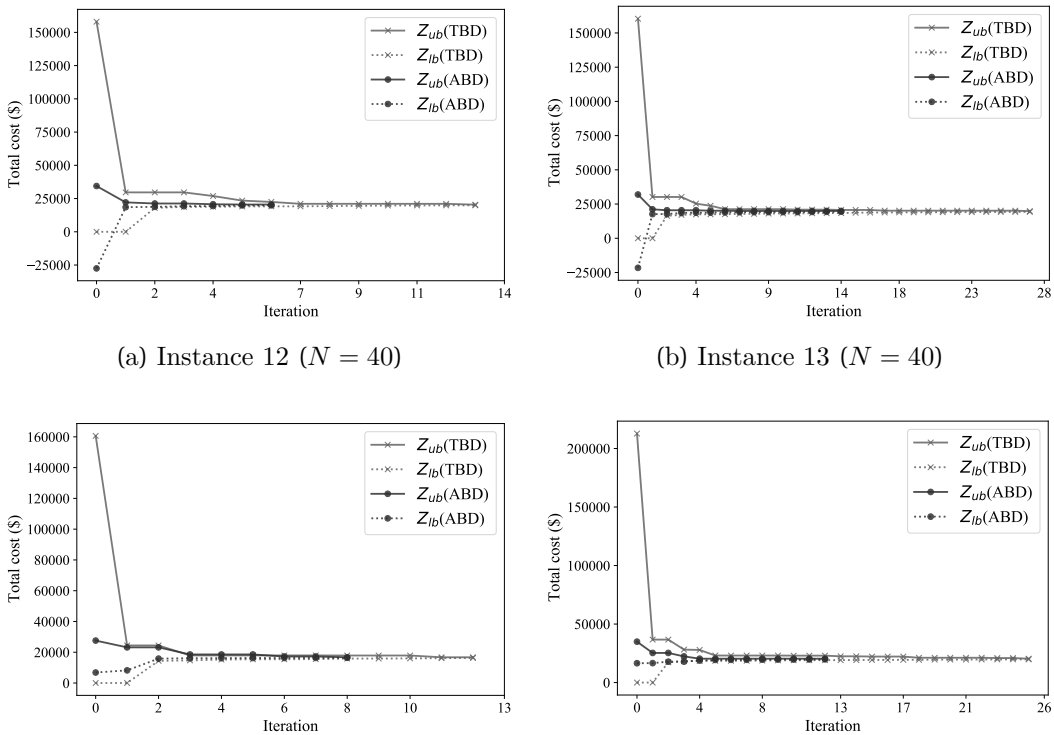
(b) Computation times (seconds)

(c) Number of iterations (%)

(d) Number of time limit has been reached

Figure 3.3: Comparisons between algorithms in terms of four performance measures.

$\epsilon_{BD}$. At the first iteration, the gap between the upper and lower bound of ABD was clearly smaller than the gap of TBD, which meant that ABD created effective initial optimality cuts. Finally, ABD converged faster than TBD with a small number of iterations. Even though the results in Figure 3.4 correspond to test instances 12∼15 with $N = 40$, similar behavior could be observed for other instances with different sizes of $N$.




(a) Instance 12 ($N = 40$)

(b) Instance 13 ($N = 40$)

(c) Instance 14 ($N = 40$)

(d) Instance 15 ($N = 40$)

Figure 3.4: Comparison between TBD and ABD in terms of upper and lower bound.

### 3.5.3 Performance analysis of the stochastic solution

In this section, the quality of the stochastic solution is evaluated through several performance metrics. As mentioned in Section 3.4.1, the stochastic solution is derived from that which has the lowest upper bound, $\bar{f}_{N'}(\hat{\mathbf{y}}, \hat{\mathbf{g}}, \hat{\mathbf{r}})$, value among the number of $M$ SAA replications. We utilize the *value of stochastic solution* (VSS) ,which is well-known performance metrics in stochastic programming research area [22]. VSS can be calculated as:

$$VSS = EEV - RP \tag{3.77}$$

where the EEV is the expected result of the EVP optimal solution and RP is the optimal objective value of the recourse problem.

For every test instance, we carried out the SAA algorithm in Section 3.4.1 with $N' = 3,000, M = 20$, and $\epsilon_{SAA} = 1$. Therefore, the SAA algorithm terminates whenever the relative SAA gap, $Gap^{rel}_{MNN'}$, is within 1%. To compute the statistical lower bound, we progressively increased the number of scenarios in samples from 20 to 200 until the predetermined threshold $\epsilon_{SAA}$ was satisfied, $N \in \{20, 40, 60, 80, 100, 200\}$. Every SAA problem was computed by ABD with $\epsilon_{BD} = 10^{-3}$.

Table 3.3 presents the experiment results from the SAA algorithm. The upper bound values equal to RP, EEV, and WS were derived from the same sample with $N'$ scenarios. By checking the results of the EEV, we could know that the EVS incurred a much higher total cost than the stochastic solution. The values of the VSS showed the performance of stochastic solutions compared with the EVS, which indicated the importance of capturing the stochastic nature of demands and supply for designing the supply chain. For every test instance, a size of $N$ less than 100 was necessary to obtain the stochastic solution with the SAA gap less than 1%. In particular, we could get the high quality stochastic solution only with $N = 20$ for test instances 8∼15.

In Table 3.4, we present the SAA gap estimates from the stochastic solution derived from the SAA algorithm and the EVS. As anticipated, the SAA gap estimates of the stochastic solution were less than 1%. On the other hand, the SAA gap of the EVS was much greater compared to the stochastic solution. In addition, the stochastic solution showed better performance than the EVS in terms of the standard deviation of the SAA gap, $\sigma_{Gap}$. For test instances 2∼8 and 11, the SAA

Table 3.3: Experiment results and statistics of the SAA algorithm

| No. | N | LB | $\sigma_{LB}$ | UB | $\sigma_{UB}$ | EEV | $\sigma_{EEV}$ | WS | VSS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 80 | 13,384.3 | 110.7 | 13,449.1 | 69.3 | 16,720.6 | 127.1 | 12,044.5 | 3,271.5 |
| 2 | 40 | 10,923.4 | 89.3 | 11,023.1 | 53.4 | 20,079.1 | 143.5 | 10,099.1 | 9,056.0 |
| 3 | 40 | 11,147.4 | 74.3 | 11,238.7 | 42.2 | 23,088.7 | 138.0 | 10,459.0 | 11,850.0 |
| 4 | 80 | 16,346.0 | 77.7 | 16,433.5 | 67.0 | 29,244.1 | 159.6 | 15,316.3 | 12,810.6 |
| 5 | 40 | 10,643.5 | 82.0 | 10,685.9 | 40.0 | 23,951.2 | 148.4 | 10,109.7 | 13,265.3 |
| 6 | 80 | 13,820.6 | 57.0 | 13,835.7 | 40.5 | 26,936.8 | 132.4 | 13,108.1 | 13,101.1 |
| 7 | 60 | 13,472.3 | 74.8 | 13,520.1 | 41.7 | 24,377.1 | 129.1 | 12,622.3 | 10,857.0 |
| 8 | 20 | 16,515.2 | 97.6 | 16,663.7 | 36.4 | 36,252.4 | 172.7 | 15,602.0 | 19,588.7 |
| 9 | 20 | 13,189.4 | 50.2 | 13,200.8 | 27.8 | 14,202.4 | 47.9 | 12,703.1 | 1,001.6 |
| 10 | 20 | 14,883.5 | 106.8 | 15,017.9 | 32.4 | 16,158.7 | 60.8 | 14,396.0 | 1,140.8 |
| 11 | 20 | 15,672.7 | 91.5 | 15,771.3 | 31.0 | 26,593.1 | 97.1 | 15,284.8 | 10,821.8 |
| 12 | 20 | 20,379.2 | 73.5 | 20,384.4 | 35.9 | 21,901.6 | 59.7 | 19,913.7 | 1,517.2 |
| 13 | 20 | 19,762.1 | 70.7 | 19,765.8 | 28.5 | 19,888.2 | 33.9 | 19,401.7 | 122.4 |
| 14 | 20 | 16,568.0 | 96.5 | 16,641.7 | 30.6 | 16,740.0 | 34.0 | 16,329.5 | 98.3 |
| 15 | 20 | 20,224.4 | 67.8 | 20,228.8 | 23.8 | 20,585.9 | 27.7 | 20,110.6 | 357.1 |

gap of the EVS was greater than 50%, which meant the provided EVS could not accommodate uncertainty for decision-making. In comparing the cost components derived from the EVS in Table 3.5, we found that the stockout costs absorbed a larger share of the total cost when the SAA gap estimate of the EVS was relatively high. Of special note, the stockout costs accounted for more than 50% of the total cost for test instances 2∼8 and 11.

### 3.5.4 Effects of the ODWS on the supply chain

In this section, we conducted four types of experiments to explore the effects of the ODWS on the supply chain by solving the test instance 10 with $N = 40$. In the first experiment, we investigated the impact that available provider warehouses had on the resulting supply chain. We analyzed the total cost and utilization of provider warehouses by varying the number of available provider warehouses $K_{max}$, which indicates the size of set $\mathcal{K}$. We use the term 'utilization' to refer to the utilization

Table 3.4: SAA gap estimates from stochastic and EVS

| No. | Stochastic solution | | | EVS | | |
|---|---|---|---|---|---|---|
| | $Gap_{MNN'}$ | $Gap^{rel}_{MNN'}$ | $\sigma_{Gap_{MNN'}}$ | $Gap_{MNN'}$ | $Gap^{rel}_{MNN'}$ | $\sigma_{Gap_{MNN'}}$ |
| 1 | 64.8 | 0.48 | 130.60 | 3336.3 | 24.93 | 168.51 |
| 2 | 99.7 | 0.91 | 104.08 | 9155.7 | 83.82 | 169.06 |
| 3 | 91.3 | 0.82 | 85.46 | 11941.3 | 107.12 | 156.71 |
| 4 | 87.5 | 0.54 | 102.61 | 12898.1 | 78.91 | 177.50 |
| 5 | 42.4 | 0.40 | 91.22 | 13307.7 | 125.03 | 169.52 |
| 6 | 15.1 | 0.11 | 69.88 | 13116.2 | 94.90 | 144.10 |
| 7 | 47.7 | 0.35 | 85.69 | 10904.8 | 80.94 | 149.20 |
| 8 | 148.5 | 0.90 | 104.14 | 19737.2 | 119.51 | 198.34 |
| 9 | 11.4 | 0.09 | 57.40 | 1013.0 | 7.68 | 69.41 |
| 10 | 134.4 | 0.90 | 111.56 | 1275.2 | 8.57 | 122.87 |
| 11 | 98.7 | 0.63 | 96.56 | 10920.4 | 69.68 | 133.41 |
| 12 | 5.2 | 0.03 | 81.81 | 1522.4 | 7.47 | 94.69 |
| 13 | 3.7 | 0.02 | 6.29 | 126.1 | 0.64 | 78.45 |
| 14 | 73.7 | 0.44 | 101.24 | 172.0 | 1.04 | 102.33 |
| 15 | 4.4 | 0.02 | 71.83 | 361.5 | 1.79 | 73.22 |

Table 3.5: Cost components derived from EVP solution

| No. | Delivery | | Commitment | | Stockout | | Supplier investment | | Transportation | | Inventory holding | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Cost ($) | % | Cost ($) | % | Cost ($) | % | Cost ($) | % | Cost ($) | % | Cost ($) | % |
| 1 | 4,869.9 | 29.13 | 1,875.9 | 11.22 | 8,076.3 | 48.30 | 1,526.2 | 9.13 | 309.6 | 1.85 | 62.7 | 0.37 |
| 2 | 4,079.1 | 20.32 | 1,453.2 | 7.24 | 13,572.8 | 67.60 | 525.2 | 2.62 | 397.0 | 1.98 | 51.8 | 0.26 |
| 3 | 5,219.6 | 22.61 | 1,488.6 | 6.45 | 15,210.7 | 65.88 | 606.1 | 2.62 | 484.8 | 2.10 | 79.0 | 0.34 |
| 4 | 4,654.6 | 15.92 | 3,096.1 | 10.59 | 20,458.9 | 69.96 | 572.7 | 1.96 | 385.5 | 1.32 | 76.3 | 0.26 |
| 5 | 5,337.5 | 22.28 | 788.3 | 3.29 | 16,838.7 | 70.30 | 517.2 | 2.16 | 351.0 | 1.47 | 118.4 | 0.49 |
| 6 | 5,759.9 | 21.38 | 2,491.6 | 9.25 | 16,568.2 | 61.51 | 532.5 | 1.98 | 1,467.6 | 5.45 | 116.8 | 0.43 |
| 7 | 4,168.7 | 17.10 | 2,961.2 | 12.15 | 15,285.8 | 62.71 | 694.6 | 2.85 | 1,152.5 | 4.73 | 114.3 | 0.47 |
| 8 | 6,901.8 | 19.04 | 2,126.1 | 5.86 | 24,009.7 | 66.23 | 889.1 | 2.45 | 2,154.6 | 5.94 | 171.1 | 0.47 |
| 9 | 7,726.1 | 54.40 | 945.8 | 6.66 | 2,514.1 | 17.70 | 1,197.9 | 8.43 | 1,679.8 | 11.83 | 138.8 | 0.98 |
| 10 | 7,592.7 | 46.99 | 2,129.4 | 13.18 | 3,603.6 | 22.30 | 1,134.8 | 7.02 | 1,587.0 | 9.82 | 111.2 | 0.69 |
| 11 | 7,094.0 | 26.68 | 3,042.4 | 11.44 | 14,002.4 | 52.65 | 990.8 | 3.73 | 1,283.6 | 4.83 | 179.9 | 0.68 |
| 12 | 9,560.4 | 43.65 | 2,278.0 | 10.40 | 4,337.7 | 19.81 | 1,556.2 | 7.11 | 4,044.9 | 18.47 | 124.5 | 0.57 |
| 13 | 11,459.6 | 57.62 | 2,378.7 | 11.96 | 1,976.1 | 9.94 | 1,469.9 | 7.39 | 2,504.7 | 12.59 | 99.2 | 0.50 |
| 14 | 9,907.7 | 59.19 | 838.8 | 5.01 | 3,101.3 | 18.53 | 1,159.8 | 6.93 | 1,612.7 | 9.63 | 119.8 | 0.72 |
| 15 | 12,178.9 | 59.16 | 2,565.0 | 12.46 | 1,761.2 | 8.56 | 1,342.8 | 6.52 | 2,597.5 | 12.62 | 140.4 | 0.68 |

of provider warehouses within the entire time horizon, and it is defined as:

$$\text{Utilization} := \sum_{k \in \mathcal{K}} \sum_{m \in \mathcal{M}} \sum_{t \in \mathcal{T}} r^k_{mt} \tag{3.78}$$

Table 3.6 represents the utilization of provider warehouses and the total cost varying $K_{max}$. As the $K_{max}$ increased until eight, the utilization of provider warehouses also increased. However, the utilization decreased from 30 to 29 when the $K_{max}$ was bigger than eight. On the other hand, the total cost decreased when the $K_{max}$ was increased. In the case where only one provider warehouse was available, it incurred the highest total cost because satisfying demands with only one provider warehouse capacity was challenging. Note that in the case in which $K_{max}$ was bigger than nine, the utilization and total cost did not change. It meant that utilizing provider warehouses from 10 to 20 could not contribute to better solutions for reducing the total cost.

Table 3.6: Impact of different number of available provider warehouses on utilization and total cost

| $K_{max}$ | $\mathcal{K}$ | Utilization | Warehouses | Total cost ($) |
|---|---|---|---|---|
| 1 | $\{1\}$ | 15 | 1 | 32,406.9 |
| 2 | $\{1,2\}$ | 25 | 1,2 | 16,360.0 |
| 3 | $\{1,2,3\}$ | 30 | 1,2,3 | 15,532.5 |
| 4 | $\{1,2,\cdots,4\}$ | 30 | 1,2,3,4 | 15,259.9 |
| 5 | $\{1,2,\cdots,5\}$ | 30 | 1,2,3,4 | 15,259.9 |
| 6 | $\{1,2,\cdots,6\}$ | 30 | 1,2,3,4 | 15,259.9 |
| 7 | $\{1,2,\cdots,7\}$ | 30 | 1,2,3,4,7 | 15,098.2 |
| 8 | $\{1,2,\cdots,8\}$ | 30 | 1,2,3,4,7 | 15,098.2 |
| 9 | $\{1,2,\cdots,9\}$ | 29 | 1,2,3,4,7,9 | 14,986.9 |
| 15 | $\{1,2,\cdots,15\}$ | 29 | 1,2,3,4,7,9 | 14,986.9 |
| 20 | $\{1,2,\cdots,20\}$ | 29 | 1,2,3,4,7,9 | 14,986.9 |

In the second experiment, a sensitivity analysis on the commitment cost of parameter $\alpha$ was conducted to explore the effects on solutions. Figure 3.5 represents the changes of utilization and total cost brought about by varying the value of $\alpha$. As the $\alpha$ increased, utilization decreased and total cost increased. Because of the expensive cost of commitment, utilizing provider warehouses for the SCND was avoided.

When the $\alpha$ was larger than 7,500, the total cost did not vary, and the utilization became zero, which meant provider warehouses were not used.

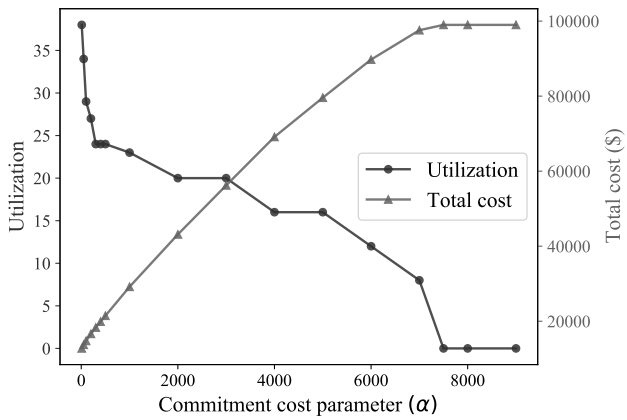



Figure 3.5: Changes of total cost and utilization varying the commitment cost parameter $\alpha$.

Figure 3.6 shows the share of the total cost according to different cost components. Increasing the $\alpha$ resulted in decreasing the percentage of transportation, delivery, and supplier investment costs. Because inventory holding cost parameters, $h_i^r, h_i^r$, and $h_i^e$, were much smaller than other cost parameters, the percentage of inventory holding cost was negligible. The percentage of commitment cost increased and then decreased at the point when the $\alpha$ was larger than 5,000. On the contrary, the percentage of the stockout cost decreased and then increased at the same point for the commitment cost. Like the utilization and total cost in Figure 3.5, the percentage of each cost component did not change when the $\alpha$ was larger than 7,500. Furthermore, the stockout cost accounted for a disproportionately large share of the total cost. Based on this result, we could observe that allowing for the condition of stockout for most of the demands is a better cost-saving strategy compared to using provider warehouses when the $\alpha$ is significantly higher than the stockout cost

parameter $\beta_i$.

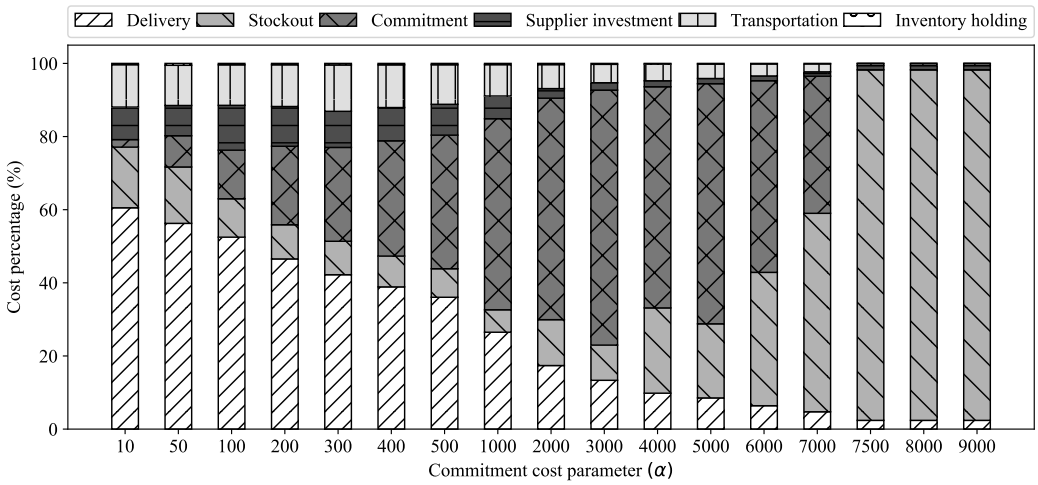



Figure 3.6: Share of total cost for each cost component varying the commitment cost parameter $\alpha$.

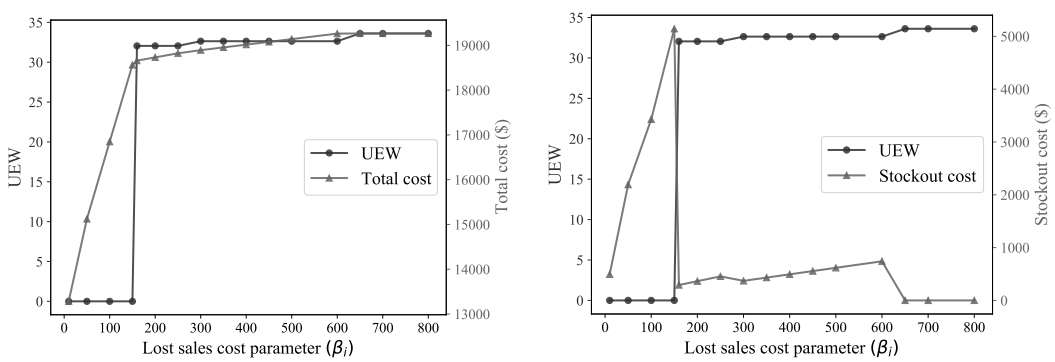
In the third experiment, a sensitivity analysis on the lost sales cost parameter, $\beta_i$, was conducted to observe the relationship between utilization of the emergency warehouse and stockout. In the third and fourth experiments, we assumed that the emergency warehouse is capacitated ($C^e = 70$). The average number of items delivered from the emergency warehouse to customers within the entire time horizon is used to refer to the utilization of the emergency warehouse (UEW), and it is defined as:

$$\text{UEW} := \frac{1}{|\Omega|} \sum_{i \in \mathcal{I}} \sum_{e \in \mathcal{E}} \sum_{t \in \mathcal{T}} \sum_{\omega \in \Omega} u_{it}^{e\omega} \tag{3.79}$$

Figure 3.7 represents the changes in UEW, total cost, and stockout cost brought about by varying the value of $\beta_i$. Until the value of $\beta_i$ was 150, UEW was zero, which meant the emergency warehouse was not used. Instead, every unsatisfied demand

was addressed through the lost sales. At the point $\beta_i$ was 160, UEW increased dramatically from zero to about 32, which means the emergency warehouse was used to satisfy demand. However, UEW slightly increased when $\beta_i$ was bigger than 160.

The total cost and stockout cost increased rapidly until the value of $\beta_i$ was 150. When the $\beta_i$ was bigger than 160, the total cost increased slightly. On the other hand, the stockout cost decreased steeply at the point $\beta_i$ was 160. After that, when the $\beta_i$ was bigger than 160 and smaller than 600, the stockout cost increased slightly. The stockout cost became zero when the $\beta_i$ was bigger than 650, which means that every demand was satisfied. In addition, when the $\beta_i$ was bigger than 650, the UEW, total cost, and stockout cost did not vary.




(a) Total cost

(b) Stockout cost

Figure 3.7: Changes of cost and UEW varying the lost sales cost parameter $\beta_i$.

Figure 3.8 depicts the share of the total cost according to different cost components for the lost sales cost parameter, $\beta_i$. The percentage of inventory holding cost was negligible in the same manner as is shown in Figure 3.6. Depending on the value of $\beta_i$, the percentage of stockout cost and the percentage of transportation

cost tended to move into the opposite directions. In detail, as the $\beta_i$ increased to 150, the percentage of stockout cost increased, and the percentage of delivery and transportation cost decreased. At the point when the $\beta_i$ was 160, the percentage of stockout cost decreased rapidly, and the percentage of transportation cost increased significantly. This result means that the emergency warehouse was used to satisfy demand as much as possible to avoid stockouts because of the high cost of lost sales. When the $\beta_i$ was 160 to 600, the percentage of stockout cost increased slightly, but the stockout cost did not account for any share of the total cost when the $\beta_i$ was larger than 650. In addition, when the $\beta_i$ was larger than 650, the percentage of supplier cost increased, which meant that every demand was satisfied by adopting additional suppliers.

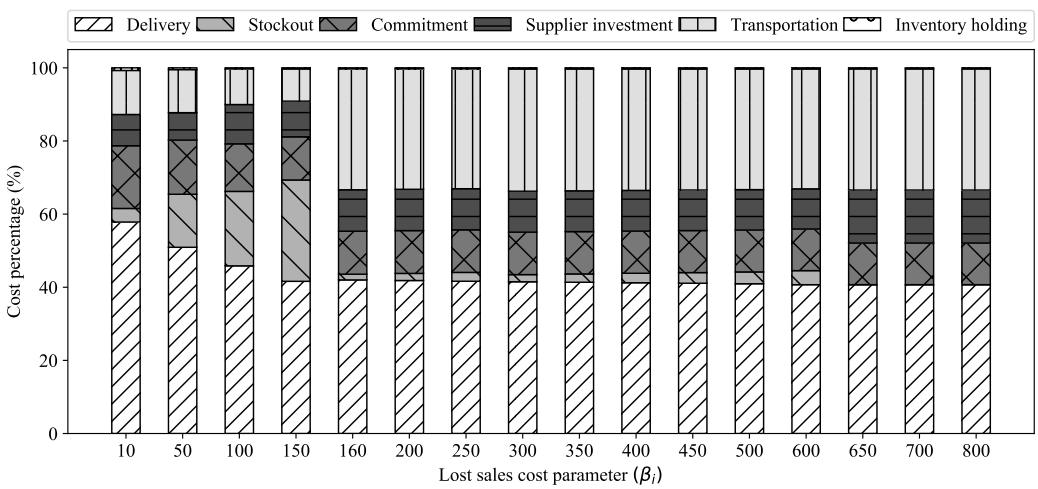



Figure 3.8: Share of total cost for each cost component varying the stockout cost parameter $\beta_i$.

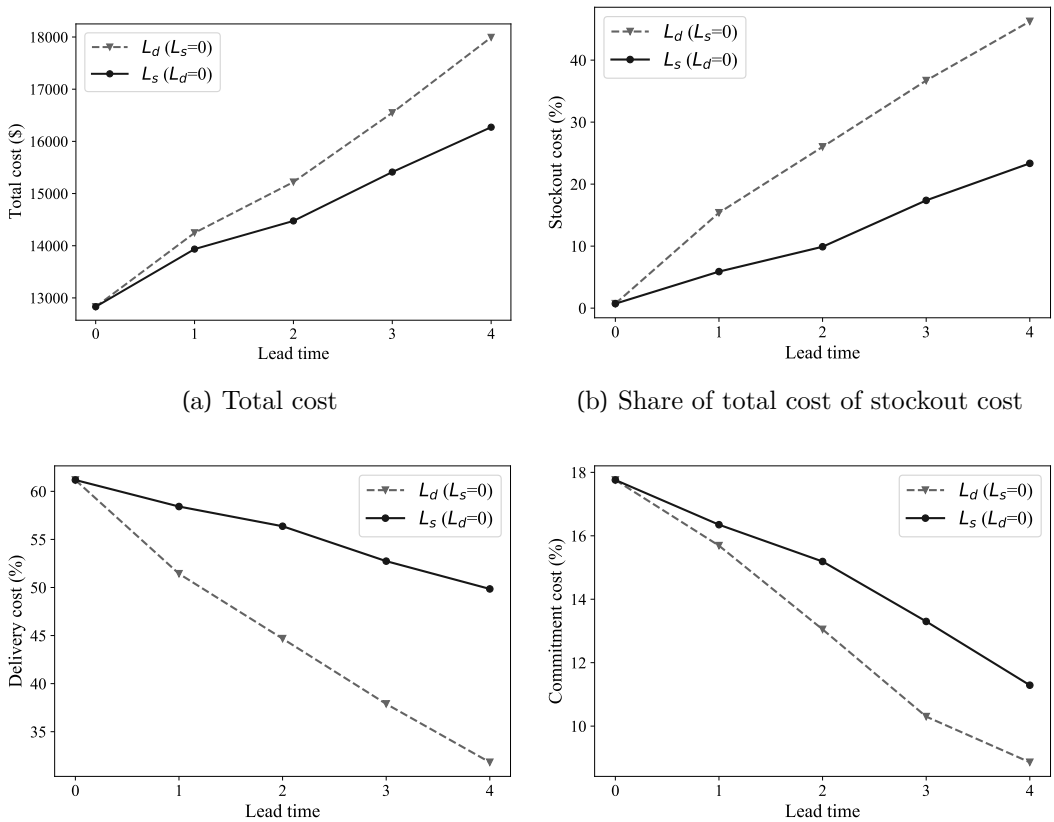
In the fourth experiment, we evaluated the effects of lead times when utilizing the ODWS in the supply chain by varying the values of lead times between suppliers

and warehouses, $L_s$, and between warehouses and customers, $L_d$. In cases in which the lead time exists in the supply chain, a lot of stockout costs can be incurred at the beginning of the planning horizon if the retailer does not hold initial inventory. Therefore, in this experiment, we assumed that the initial inventory is equal to the expected value of demand. Based on the detailed results in Table B.4, we presented in Figure 3.9 the impacts of each type of lead time on total cost, stockout cost, delivery cost, and commitment cost.

In Figure 3.9, we varied the value of one type of lead time, and the other one was fixed to zero to compare the impacts of each type of lead time. For both types of lead time, total cost and stockout cost increased as the value of lead time increased. On the other hand, because the total amount of stockout increased, the percentage of delivery cost decreased. Commitment cost also decreased as the lead time increased, which meant that the utilization of the provider warehouse decreased as well. Finally, by observing that the total cost increased more rapidly when increasing the value of $L_d$ than when increasing the value of $L_s$, we could know that the length of lead time between warehouses and customers severely affected the cost incurred in the supply chain.

### 3.5.5 Managerial insights

The proposed model and stochastic solution approach could contribute to e-commerce retailers who plan to build the supply chain network flexibly during the COVID-19 pandemic. After analyzing the computational results, we can offer several managerial insights that could be instructive to e-commerce retailers who suffer from the limited space of warehouses. The proposed managerial insights are as follows:




(a) Total cost

(b) Share of total cost of stockout cost

(c) Share of total cost of delivery cost

(d) Share of total cost of commitment cost

Figure 3.9: Comparisons between two types of lead time in terms of cost.

1. Utilizing the ODWS can save on the total cost of the supply chain because it has a similar effect as expanding capacity flexibly. Even though most of the demands can be satisfied with enough provider warehouses, we could observe that using a moderate number of provider warehouses is a good strategy for minimizing total cost. Hence, considering the locations of suppliers and provider warehouses and the appropriate number of provider warehouses would be helpful to retailers when constructing an efficient supply chain with the ODWS.

2. Our proposed model is very sensitive to uncertainty because frequent stockouts could occur when insufficient provider warehouses are committed to being used. Even though simple solution approaches could solve the problem (e.g., EVP), most obtained solutions are imprecise for acceptable decision-making. By analyzing the value of the VSS, we could observe that it is important to deal with uncertainty accurately regarding the SCND problem with the ODWS. Therefore, we suggest that retailers who need to address frequent stockouts because of limited capacity should develop an efficient way for accommodating the uncertainty of demand and supply.

3. In the actual case, the value of $\alpha$ is determined by the warehouse operators or the ODWS platform company. However, the value of $\beta_i$ can be estimated by the e-commerce retailers. As shown in Figures 3.6, 3.7, and 3.8, the estimated value of $\beta_i$ has much influence on the quality of solutions. In terms of obtained solutions, if the $\beta_i$ is estimated to be larger than the true value, it results in excess utilization of provider and emergency warehouses. Otherwise, when $\beta_i$ is estimated to be smaller, it could incur a lot of stockouts because of insufficient utilization of provider warehouses and the expensive cost of utilizing the emergency warehouse. Hence, we recommend that retailers conduct an accurate estimation for the value of $\beta_i$ beforehand and then implement our proposed approach.

4. Through several computational experiments, we observed that the lead time increased the total amounts of stockout, which incurred additional costs. The lead time between warehouses and customers was more significant than be-

tween suppliers and warehouses in terms of costs incurred in the supply chain. Therefore, when e-commerce retailers design the supply chain with the ODWS, we recommend choosing a 3PL company operating with short lead times even though the delivery cost is slightly higher. This strategy would be helpful to retailers in minimizing the total cost incurred in the supply chain.

## 3.6 Summary

With e-commerce set to expand rapidly in the coming decades, the ODWS has emerged as a new alternative for satisfying growing demand. By utilizing the ODWS in the supply chain, e-commerce retailers can flexibly respond to demand changes because this service makes short-term rent of warehouses available. However, a high degree of uncertainty regarding demand and supply exists in the e-commerce marketplace, which influences decision-making for the SCND. To the best of our knowledge, there is no existing research dealing with the problem of the SCND with the ODWS under uncertainty. Therefore, we propose the two-stochastic programming model, which reflects the supply chain network of the e-commerce marketplace in South Korea.

Because of the high computational complexity of the proposed model, a solution approach combining the SAA and BD algorithms was presented to solve the proposed model. Of special note, a method to accelerate the convergence of bounds in the BD algorithm, referred to as ABD, was developed. The ABD outperforms the typical version of the BD algorithm and Xpress-Optimizer with regard to the optimality gap and computation times. In addition, the quality of stochastic solutions derived from the SAA algorithm is better than the solutions from the EVP.

Through conducting computational experiments, we could observe that utilizing the ODWS for the SCND saves on the total cost compared to using a small number of warehouses with limited capacity. Furthermore, through our sensitivity analysis, we could see the relationship between parameters of commitment cost and stockout cost for a decision about using the provider and emergency warehouses. We observed the impacts of two types of lead time on the cost incurred in the supply chain considering the ODWS. At last, we present several managerial insights that are helpful for e-commerce retailers who aim to design their supply chain networks with the ODWS.

# Chapter 4

# A decomposition approach for robust omnichannel retail operations considering the third-party platform channel

## 4.1 Introduction

In recent years, several retail companies have sold their products on 3PPs, such as *Amazon* and *Coupang*, despite having their own offline and online channels [151]. In real business, Coupang launched a service called the *C.AVENUE*, and many omnichannel companies, such as *Nike* and *Adidas*, have participated in this service and sold their products using 3PP. From the perspective of retailers, there are distinct advantages to adopting the 3PP channel as one of their sales channels. First, the 3PP companies could implement logistics of fulfillment on behalf of the retailer by using their *self-supporting logistics service system* (SLSS). For example, Amazon has provided a fulfillment service called *Fulfillment by Amazon* (FBA), and it allows retailers to use Amazon to store, pick, pack, and ship customer orders [83]. Second, the retailer could absorb the additional demand of 3PP. A significant number of customers use 3PP to buy products online. Specifically, as of 2022, more than 197 million monthly active users use the Amazon app, and more than 27 million monthly active users use the Coupang app [35]. Therefore, in addition to customers

who want to buy a specific product from a retailer, other users of 3PP could also buy that product while looking around the platform.

Motivated by observing the advantages retailers obtain by using 3PP, we study omnichannel retail operations that have adopted the 3PP channel as one of the sales channels. Moreover, we address decision and optimization problems considering demand uncertainty, which jointly determine the replenishment, allocation, transshipment, and fulfillment of products over a multi-period planning horizon. We assume that the retailer's objective is to minimize the expected total cost over the planning horizon. Of special note, we consider in this study the following two features, which are generally considered in real business: (1) the binary decision for replenishment to accommodate fixed order costs and (2) the constraint restricting replenishment quantity depending on the production capacity of each supplier (i.e., the *production capacity constraint*).

However, there are four issues that make the problem of omnichannel retailer operations challenging. First, the retailer has to make binary replenishment decisions adaptively after demand unfolds over periods (i.e., the *adjustable binary decision*), which increases the complexity of the problem [66]. Second, according to the common assumption in retail environments, the replenishment, allocation, and transshipment of products are decided before the demand is realized (*anticipative manner*), and the fulfillment is decided after demand is realized (*reactive manner*) [72]. Thus, the solution approach providing a good quality solution with integrating anticipative and reactive decisions is necessary. Third, the existence of the 3PP channel makes the problem larger than it would be without this channel. In addition to the retailer's supply chain for online and offline channels, the supply chain for

the 3PP channel (i.e., the *3PP supply chain*) should also be considered if the 3PP channel is adopted. Fourth, the insufficient production capacity of suppliers makes the problem quickly become intractable. To the best of our knowledge, no existing study addresses the above four issues simultaneously, even though Lim et al. [93] and Jiu [72] dealt with the first and second issues.

In order to fill these research gaps, our study deals with a multi-period stochastic optimization model that takes into account the logistics operations of an omnichannel retailer's supply chain and the supply chain of the 3PP simultaneously. Additionally, we propose a novel decomposition method, which is called *DECOM*, to enhance computational efficiency. We present the main contributions of our study from the following two perspectives:

- *Modeling:* As far as we know, this is the first study to develop the stochastic optimization model addressing both the retailer's supply chain and the supply chain of the 3PP for omnichannel retail operations. Furthermore, we deal with the production capacity of suppliers and transshipment between logistics centers, which are two elements that have not been addressed in related existing studies. Finally, our model can jointly determine every decision, and the anticipative and reactive manners are implemented seamlessly as the demand unfolds over periods.

- *Solution approach:* We propose a DECOM based on the TPA based on RO approach, which is the state-of-the-art method to deal with adjustable binary decisions [93]. We first utilize the original TPA to solve our problem, but it requires a significant computational burden to solve the realistic problem instances. In addition, the TPA could not solve our problem within acceptable

times when the production capacity of suppliers is insufficient. To alleviate these issues, we decompose the total supply chain into two streams, one for the retailer's supply chain and the other for the 3PP supply chain, by introducing artificial variables. Through extensive computational experiments, we evaluate the performance of DECOM by comparing it with several approaches from existing literature. The experimental results suggest that DECOM could provide high-quality solutions similar to solutions derived from the TPA. Furthermore, in terms of computational efficiency, DECOM outperforms the TPA by solving large-scale problems within a reasonable time. Finally, even though the production capacity becomes insufficient, the computation time of DECOM does not increase significantly compared to that of the TPA.

## 4.2 Literature review

The literature review will focus on three streams of research in operations management: omnichannel retail operations, the 3PP channel, and RO.

### 4.2.1 Omnichannel retail operations

The last few years have seen a huge growth in the number of papers published on the topic of omnichannel leverage in retail operations [27]. Many researchers have empirically studied this topic to find the effects of adopting the omnichannel in retail operations [47, 48, 9, 81]. Instead of reviewing all existing studies related to the omnichannel topic, we present a detailed review of recent literature regarding the optimization problem in the omnichannel from the retailer's perspective.

Govindarajan et al. [58] considered the inventory and fulfillment decisions in the

omnichannel network with multiple stores and fulfillment centers for the omnichannel retailer. They developed scalable heuristic solutions for joint decisions, including the pooling of online demands across locations using a hindsight-optimal bound. Park et al. [105] studied the problem to create an efficient showcase inventory, allowing different desired products to be experienced by as many customers as possible. They presented a MILP model to maximize the expected customer showcasing utility and analyzed the effects of the proposed model through the case study of dealerships in the US. Pichka et al. [109] dealt with the problem of jointly deciding fulfillment and pricing decisions for omnichannel retailers. They first presented customer demands using the multinomial logit choice model. Using the developed demand model, they proposed two MINLP models to optimize decisions for fulfillment, pricing, and inventory. These two MINLP models were transformed into MILP models to be solved efficiently. Abouelrous et al. [2] addressed the multi-location inventory problem, aiming to determine the initial inventory at each location within a given planning horizon. They also simultaneously considered the stochastic online and in-store demands, which were general assumptions in omnichannel retail operations. In order to enhance computational efficiency, they approximated the problem by developing a two-stage stochastic optimization with a scenario reduction technique.

We present two relevant studies that utilize the RO for omnichannel retailing. First, Qiu et al. [114] addressed the problem for pricing and ordering optimization considering full-refund and no-refund policies. They also defined the demand as a linear function of the price and refund to accommodate the general case that demands depend on the prices and available return policies. Using historical data, they presented a nonlinear robust omnichannel pricing and ordering optimization

model to cope with demand uncertainty. The robust counterpart of the proposed model was transformed into the tractable MILP model by using the duality theory. However, the presented approach is challenging to apply in the multi-period problem, and the computational efficiency was not analyzed.

On the other hand, Jiu [72], which is the most relevant study to our research, addressed the multi-period problem for robust omnichannel retailing. The study used the TPA, developed by Lim et al. [93], to solve the problem. The TPA could provide high-quality solutions compared to existing approaches. In addition, through computational experiments on large-scale problems, the study indicated that the TPA was scalable to the problem. Our study has several differences compared with the study by Jiu [72]. One of these differences is that both transshipment decisions and production capacity are considered in our model. However, the most apparent contribution of our study is that we adopt the 3PP channel in our model. In other words, when optimizing the proposed problem, the retailer's supply chain and the supply chain of the 3PP should be considered simultaneously. In the following section, we present several studies that analyze the effects of adopting the 3PP channel for retailing.

### 4.2.2 Third-party platform channel

By investigating the existing studies considering 3PP in retail, we observe that 3PP companies can be classified into two types depending on the existence of SLSS in those companies. For 3PP companies without the SLSS, the retailer or manufacturer who participates in 3PP can only sell their products using the platform, but the logistics of products must be implemented by themselves. On the other hand, for

3PP companies with the SLSS, the retailer can sell their products on 3PPs. Adding to that, the 3PP company implements every logistics and fulfillment procedure on behalf of the retailer. Our study considers the latter type for a 3PP company by reflecting real cases of Coupang and Amazon.

First, we present several previous studies considering the 3PP company without the SLSS. Ryan et al. [117] addressed a research question of whether the retailer that has its own sales channel should expand the sales channel by using 3PP. They considered the participation fee for the 3PP channel and a revenue-sharing requirement. Using game theory, they derived the optimal decision and system equilibrium for both the retailer and the 3PP company. Xiao and Xu [146] studied commission contract design between a 3PP company and sellers who have superior demand information to achieve two goals: (1) to incentivize the seller to install optimal capacity and (2) to extract full surplus. To achieve these two goals, they applied the lost-sale penalty contract, which charges a penalty cost to sellers if a stockout occurs. Zhen et al. [150] considered a model with a financial capital constraint from the perspective of the manufacturer. The manufacturer was assumed to sell its products through a retailer and the 3PP channel. Also, the setup dictated that the manufacturer could borrow financing from the 3PP, the retailer, or the bank. The authors derived the best financing option for the manufacturer considering the channel competition, the revenue sharing rate, and the unit production cost.

From this point onward in this section, we will introduce literature that considers the 3PP company operating with SLSS. Qin et al. [113] addressed the SLSS of 3PP, which is provided to retailers that participate in the 3PP channel. They analyzed the strategic and economic impacts of logistic service sharing and examined the

equilibrium mode between 3PP and the retailer considering the logistics service level and the market potential. Zhen and Xu [151] dealt with a research question of whether the retailer who has online and offline channels should adopt 3PP for the sales channel. In order to answer the research question, they developed a game-theoretical model. Furthermore, they explored the impact of the direction of the spillover effect between sales channels by varying the degree of channel competition and assuming the agency fee for using 3PP. Lai et al. [83] investigated the effects of FBA, which is a fulfillment service offered by Amazon, on both Amazon itself and on retailers that use this service. They developed a strategic competition model and found that FBA could alleviate price competition between Amazon and the retailer. In addition, FBA could improve the service level of retailers, and Amazon also benefits because the sales of Amazon's products increased because of the FBA.

The abovementioned literature only investigated whether the retailer who owns its offline and online channels should expand sales channels by utilizing the 3PP channel. Also, the effects of utilizing 3PP on both the retailer and the 3PP company were examined. However, in a setting where the retailer has determined to utilize the 3PP channel in advance, there is a lack of research investigating the optimal way to operate both the retailer's supply chain and the supply chain of the 3PP. To fill these gaps, our study addresses the problem that the retailer has determined to utilize the 3PP channel in advance. Furthermore, we aim to provide efficient logistics operations by minimizing the total expected cost from the perspective of the retailer. We adopt RO as our solution approach, and several key papers in the RO research area will be presented in the following section.

### 4.2.3 Robust optimization

RO is one of the approaches that deals with uncertainty in optimization problems. In contrast to other approaches (e.g., stochastic programming and dynamic programming), RO does not need any knowledge about the probability distribution. But instead, it assumes that the uncertainty value belongs within a predetermined set, called the *uncertainty set*. RO aims to find the optimal solution under the worst-case scenario, and the obtained solution should be guaranteed to be feasible for any realizations of uncertain parameters in the uncertainty set [10].

In order to make the RO model tractable, the uncertainty set is generally defined as a convex set [114]. Soyster et al. [128] first addressed a box shape of the uncertainty set for the RO formulation. Even though the solution was feasible for all perturbations in an interval, a conservative solution was obtained. To reduce the level of conservatism of the robust solutions, Ben-Tal and Nemirovski [12] developed the RO model for the ellipsoidal uncertainty set. Berstimas and Sim [20] developed a family of polyhedral uncertainty sets in which cardinally constraints were considered using a budget of uncertainty.

Two types of decisions can be utilized for the multi-period decisions problem: (1) *here-and-now* and (2) *wait-and-see*. For the here-and-now scheme, every decision is determined before the planning horizon starts (i.e., before every uncertain parameter is revealed). In contrast, for the wait-and-see scheme, we can postpone making decisions until some of the uncertain parameters are revealed. Therefore, the wait-and-see decision is less conservative than the here-and-now decision because it can be adjusted flexibly according to the realized portion of uncertain parameters at each stage [147]. However, it is complex to deal with the wait-and-see decision because

of the large feasible space of adjustable variables.

The ARO is developed to deal with multi-stage problems, which commonly assume the multi-period setting and consider adjustable variables to implement the wait-and-see decision. Because of tractability reasons, it is typical to restrict feasible space by optimizing a certain type of parameterized function. This function is usually called the *decision rule*. Several researches have used nonlinear functions for the decision rule [19, 50]. However, a broad body of literature has adopted the linear function for the decision rule, which is called the *linear decision rule* (LDR). Ben-Tal et al. [11] first presented the LDR for a production inventory problem. Because the LDR could lead problems to be reformulated to be tractable, it has attracted considerable interest in many domains, and in particular, it has been widely utilized in inventory management [21, 122, 126]. The simplest version of the decision rule is the *static rule*, in which decisions are fixed regardless of the realization of uncertain parameters. For some cases, the static rule has proved to be optimal [122, 18, 96].

The solution approaches of the abovementioned studies have focused on adjustable continuous variables; thus, they cannot apply to adjustable binary variables. Only a few studies developed solution approaches to deal with adjustable binary variables: the K-adaptability approach [66], the finite adaptability approach (FA) [16, 111], and the binary decision rule (BDR) [17]. In particular, Lim et al. [94] developed the target-oriented robust optimization (TRO) method to address the adjustable binary and continuous variables at the same time. The TRO aims to maximize the chance of fulfilling a prespecified target [30]. Lim et al. [94] proved that TRO could provide a static rule that was optimal for a multi-product, multi-period inventory problem. By utilizing the strength of TRO, Lim et al. [93] developed the

TPA. In the TPA, they decoupled adjustable binary variables and adjustable continuous variables for making decisions. TPA decided the adjustable binary variables by a static rule of TRO and resorted to the LDR for determining the adjustable continuous variables. The experimental results showed that the TPA outperformed existing approaches, BDR and FA, for both solution quality and computational efficiency.

Even though the TPA has shown outstanding performance compared to existing approaches, it could not be scalable to our problem. The TPA has required a significant computational burden for large-scale instances because our problem considers the retailer's supply chain (online and offline channels) and the supply chain of the 3PP (3PP channel) simultaneously. Therefore, our study develops the DECOM approach, which could be scalable to large-scale problems.

## 4.3 Problem description and mathematical model

### 4.3.1 Problem description

We consider a model in which a retailer sells products, $i \in \mathcal{I}$, to customers through several sales channels. By following the assumption of Jiu [72], we also assume that a retailer replenishes the inventory of each individual product from a single supplier (i.e., each product $i$ can only be provided from the corresponding supplier $i$). Also, each supplier $i$ has a limited production capacity, $s^{it}$. Furthermore, we assume that each product $i$ can only be provided from the corresponding supplier $i$. There are three types of sales channels (1) a retailer's offline channel, (2) a retailer's online channel, and (3) the 3PP channel. The supply chain network consists of multiple capacitated logistics centers, $j \in \mathcal{J}$, and offline stores, $k \in \mathcal{K}_O$, several logistics

centers, $j \in \mathcal{J}_D$, operated by the retailer, which is called *DC*, and the others, $j \in \mathcal{J}_F$, operated by the 3PP, which is called *FC*. In the case of the retailer's offline channel, we assume that the offline store $k$ is located at each offline demand zone $k$. Therefore, each demand zone is fulfilled by the corresponding offline store. For the retailer's online channel, there are multiple online demand zones for DCs. On the other hand, for the 3PP channel, we consider the aggregate demand for FCs because the 3PP company can deliver products from FCs to customers using its SLSS. It should be noted that our model can be easily extended to the general case, the multiple online demand zones for FCs, by defining the set of online demand zones for FCs. We assume that each demand type should be fulfilled by the corresponding channel, and we do not anticipate any customer switching between channels if there is a stockout.

We consider a multi-period problem with a finite planning horizon divided into period $t \in \mathcal{T}$. For each period $t$, the replenishment, transshipment, allocation, and fulfillment decisions are made, and the following sequence of an event is repeated:

1. At the start of period $t$, the quantity of product $i$ replenished at $t - L_j^i$ period arrives at the logistics center $j$. The retailer decides the replenishment quantity for each logistics center $j$ from each supplier $i$ (i.e., replenishment decision, $\delta_j^{it}$ and $q_j^{it}$).

2. The retailer then decides the transshipment quantity between DCs and how many products to allocate from DCs to offline stores (i.e., transshipment and allocation decisions, $u_{l,jj'}^{it}$ and $u_{e,jk}^{it}$).

3. At the end of period $t$, each type of demand is realized. The retailer determines

how many products to fulfill for each type of demand, and from which DCs, FCs, and offline stores to fulfill it (i.e., fulfillment decision, $v^{it}_{\rho,k}$, $v^{it}_{\eta,k}$ and $r^{it}_{jk}$). If customers face a stockout, the demand gets lost, which is a general assumption in retail environments [54].

Figure 4.1 describes the retailer's supply chain and four types of decisions.

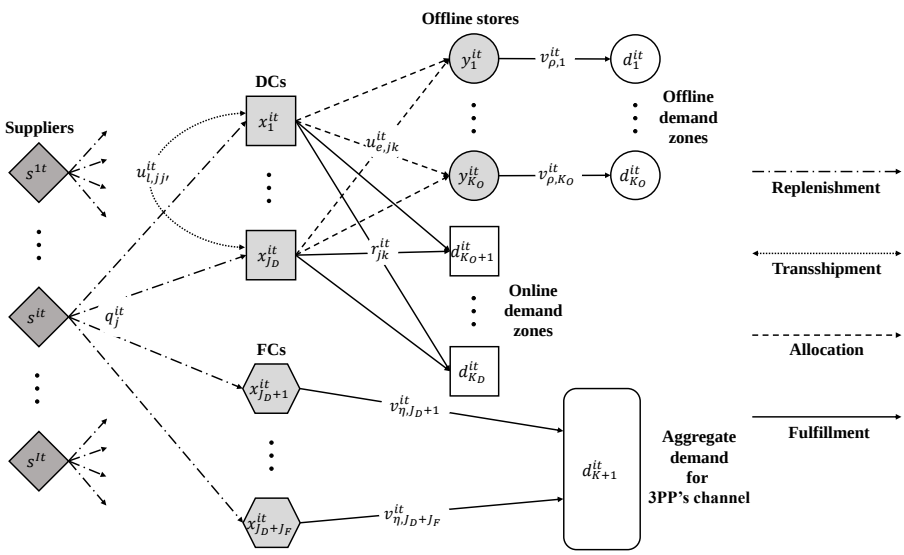



Figure 4.1: Supply chain network of the proposed problem.

We utilize the following notations to formulate the mathematical model:

**Indices and sets:**

| | |
|---|---|
| $\mathcal{T}$ | set of time periods, $t \in \mathcal{T} = \{1, 2, \ldots, T\}$ |
| $\mathcal{T}^+$ | $t \in \mathcal{T}^+ = \{1, 2, \ldots, T+1\}$ |
| $\mathcal{I}$ | set of products (=suppliers), $i \in \mathcal{I} = \{1, 2, \ldots, I\}$ |
| $\mathcal{K}_O$ | set of offline demand zones (=offline stores), $k \in \mathcal{K}_O = \{1, 2, \ldots, K_O\}$ |
| $\mathcal{K}_D$ | set of online demand zones for DCs, $k \in \mathcal{K}_D = \{K_O + 1, \ldots, K_O + K_D\}$ |

| | |
|---|---|
| $\mathcal{K}^-$ | set of online and offline demand zones for DCs, $k \in \mathcal{K}^- = \{1, \ldots, K\}$ $(K = K_O + K_D)$ |
| $\mathcal{K}$ | set of demand zones for DCs and FCs, $k \in \mathcal{K} = \{1, \ldots, K+1\}$ |
| $\mathcal{J}_D$ | set of capacitated DCs, $j \in \mathcal{J}_D = \{1, 2, \ldots, J_D\}$ |
| $\mathcal{J}_F$ | set of capacitated FCs, $j \in \mathcal{J}_F = \{J_D + 1, \ldots, J_D + J_F\}$ |
| $\mathcal{J}$ | $j \in \mathcal{J} = \{1, \ldots, J\}$ $(J = J_D + J_F)$ |

**Parameters:**

| | |
|---|---|
| $S_j^{it}$ | fixed cost to order product $i$ for the logistics center $j$ from supplier $i$ at period $t$ |
| $h_{x,j}^{it}$ | unit inventory holding cost for the logistics center $j$ per product $i$ at period $t$ |
| $h_{y,k}^{it}$ | unit inventory holding cost for the offline store $k$ per product $i$ at period $t$ |
| $c_o^{ij}$ | distance between the supplier $i$ and the logistics center $j$ |
| $c_l^{jj'}$ | distance between the DC $j$ and the other DC $j'$ |
| $c_e^{jk}$ | distance between the DC $j$ and offline store $k$ |
| $c_g^{jk}$ | distance between the DC $j$ and the online demand zone $k$ |
| $\lambda_o^{it}$ | transportation cost per 1km for the replenishment of product $i$ at period $t$ |
| $\lambda_l^{it}$ | transportation cost per 1km for the transshipment of product $i$ at period $t$ |
| $\lambda_e^{it}$ | transportation cost per 1km for allocation from DCs to offline stores for product $i$ at period $t$ |
| $\lambda_g^{it}$ | transportation cost per 1km for fulfillment from DCs to online demand zones of DCs for product $i$ at period $t$ |
| $p_k^{it}$ | lost sales cost for demand type $k$ per product $i$ at period $t$ |
| $\rho_k^{it}$ | fulfillment cost for the offline demand zone $k$ per product $i$ at period $t$ |
| $\eta_j^{it}$ | fulfillment cost for the aggregate demand for FC $j$ per product $i$ at period $t$ |
| $s^{it}$ | production capacity of supplier $i$ at period $t$ |
| $L_j^i$ | replenishment lead time of product $i$ from supplier $i$ to the logistics center $j$ |
| $\bar{x}_j$ | storage capacity of the logistics center $j$ |

| | |
|---|---|
| $\bar{y}_k$ | storage capacity of the offline store $k$ |
| $d_k^{it}$ | realized value of demand type $k$ for product $i$ at period $t$ |

**Decision variables:**

| | |
|---|---|
| $\delta_j^{it}$ | 1 if product $i$ is replenished at period $t$ from supplier $i$ to the logistics center $j$, 0 otherwise |
| $q_j^{it}$ | replenishment quantity of the product $i$ at period $t$ from supplier $i$ to the logistics center $j$ |
| $x_j^{it}$ | on-hand level of product $i$ from the logistics center $j$ at period $t$ |
| $y_k^{it}$ | on-hand level of product $i$ from the offline store $k$ at period $t$ |
| $u_{l,jj'}^{it}$ | transshipment quantity of the product $i$ from the DC $j$ to the other DC $j'$ |
| $u_{e,jk}^{it}$ | allocation quantity of the product $i$ at period $t$ from the DC $j$ to offline store $k$ at period $t$ |
| $v_{\rho,k}^{it}$ | fulfillment quantity of the product $i$ to satisfy the offline demand zone $k$ at period $t$ |
| $v_{\eta,j}^{it}$ | fulfillment quantity of the product $i$ from FC $j$ to satisfy aggregate demand for the 3PP channel at period $t$ |
| $r_{jk}^{it}$ | fulfillment quantity of the product $i$ from the DC $j$ to the online demand zone $k$ at period $t$ |
| $z_k^{it}$ | lost sales of product $i$ for the demand type $k$ at period $t$ |

The total cost incurred in the supply chain consists of ten cost components: (1) the fixed cost to place an order, $S_j^{it}\delta_j^{it}$, (2) the per-unit ordering cost, $\lambda_o^{it}c_o^{ij}q_j^{it}$, (3) the inventory holding cost for DCs and FCs, $h_{x,j}^{it}x_j^{i,t+1}$, (4) the inventory holding cost for offline stores, $h_{y,k}^{it}y_k^{i,t+1}$, (5) the stockout cost, $p_k^{it}z_k^{it}$, (6) the transshipment cost between DCs, $\lambda_l^{it}c_l^{jj'}u_{l,jj'}^{it}$, (7) the allocation cost from DCs to offline stores, $\lambda_e^{it}c_e^{jk}u_{e,jk}^{it}$, (8) the fulfillment cost for online demand zones for DCs, $\lambda_g^{it}c_g^{jk}r_{jk}^{it}$, (9)

the fulfillment cost for offline demand zones, $\rho_k^{it} v_{\rho,k}^{it}$, and (10) the fulfillment cost for the aggregate demand for the 3PP channel, $\eta_j^{it} v_{\eta,j}^{it}$. It should be noted that some expenses (e.g., a fixed participation fee) could be incurred when the retailer uses the 3PP channel [117]. Even though our study simply defines the same fixed cost parameter, $S_j^{it}$, for DCs ($j \in \mathcal{J}_D$) and for FCs ($j \in \mathcal{J}_F$), these expenses could be accommodated easily by revising the value of $S_j^{it}$, depending on whether logistics center $j$ is included in $\mathcal{J}_D$ or $\mathcal{J}_F$.

We first present a deterministic model in which all demand information within the entire planning horizon is known at the start of period $t = 1$. The deterministic model ($\mathrm{P_{DET}}$) is formulated as follows:

($\mathrm{P_{DET}}$)

$$\begin{aligned}
\min \quad & \sum_{i\in\mathcal{I}} \sum_{t\in\mathcal{T}} \Bigg( \sum_{j\in\mathcal{J}} S_j^{it} \delta_j^{it} + \sum_{j\in\mathcal{J}} \lambda_o^{it} c_o^{ij} q_j^{it} + \sum_{j\in\mathcal{J}} h_{x,j}^{it} x_j^{i,t+1} + \sum_{k\in\mathcal{K}_O} h_{y,k}^{it} y_k^{i,t+1} + \sum_{k\in\mathcal{K}} p_k^{it} z_k^{it} \\
& + \sum_{j\in\mathcal{J}_D} \sum_{j'\in\mathcal{J}_D} \lambda_l^{it} c_l^{jj'} u_{l,jj'}^{it} + \sum_{j\in\mathcal{J}_D} \sum_{k\in\mathcal{K}_O} \lambda_e^{it} c_e^{jk} u_{e,jk}^{it} + \sum_{j\in\mathcal{J}_D} \sum_{k\in\mathcal{K}_D} \lambda_g^{it} c_g^{jk} r_{jk}^{it} \qquad (4.1) \\
& + \sum_{k\in\mathcal{K}_O} \rho_k^{it} v_{\rho,k}^{it} + \sum_{j\in\mathcal{J}_F} \eta_j^{it} v_{\eta,j}^{it} \Bigg)
\end{aligned}$$

$$\text{s.t.} \quad q_j^{it} \le \bar{q}_j^i \delta_j^{it}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \qquad (4.2)$$

$$\sum_{j\in\mathcal{J}} q_j^{it} \le s^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T} \qquad (4.3)$$

$$\sum_{i\in\mathcal{I}} \left( x_j^{it} + q_j^{i,t-L_j^i} \right) \le \bar{x}_j, \qquad \forall j \in \mathcal{J}_F, t \in \mathcal{T} \qquad (4.4)$$

$$\sum_{i\in\mathcal{I}}\left(x_j^{it}+q_j^{i,t-L_j^i}+\sum_{j'\in\mathcal{J}_D\setminus\{j\}}u_{l,j'j}^{it}-\sum_{k\in\mathcal{K}_O}u_{e,jk}^{it}-\sum_{j'\in\mathcal{J}_D\setminus\{j\}}u_{l,jj'}^{it}\right)\leq\bar{x}_j,$$

$$\forall j\in\mathcal{J}_D,t\in\mathcal{T} \tag{4.5}$$

$$x_j^{it}+q_j^{i,t-L_j^i}\geq\sum_{j'\in\mathcal{J}_D\setminus\{j\}}u_{l,jj'}^{it},\qquad\forall i\in\mathcal{I},j\in\mathcal{J}_D,t\in\mathcal{T} \tag{4.6}$$

$$\sum_{i\in\mathcal{I}}\left(y_k^{it}+\sum_{j\in\mathcal{J}_D}u_{e,jk}^{it}\right)\leq\bar{y}_k,\qquad\forall k\in\mathcal{K}_O,t\in\mathcal{T} \tag{4.7}$$

$$\sum_{j\in\mathcal{J}_D}r_{jk}^{it}+z_k^{it}=d_k^{it},\qquad\forall i\in\mathcal{I},t\in\mathcal{T},k\in\mathcal{K}_D \tag{4.8}$$

$$v_{\rho,k}^{it}+z_k^{it}=d_k^{it},\qquad\forall i\in\mathcal{I},t\in\mathcal{T},k\in\mathcal{K}_O \tag{4.9}$$

$$\sum_{j\in\mathcal{J}_F}v_{\eta,j}^{it}+z_{K+1}^{it}=d_{K+1}^{it},\qquad\forall i\in\mathcal{I},t\in\mathcal{T} \tag{4.10}$$

$$x_j^{i,t+1}=x_j^{it}+q_j^{i,t-L_j^i}+\sum_{j'\in\mathcal{J}_D\setminus\{j\}}u_{l,j'j}^{it}-\sum_{j'\in\mathcal{J}_D\setminus\{j\}}u_{l,jj'}^{it}$$

$$-\sum_{k\in\mathcal{K}_O}u_{e,jk}^{it}-\sum_{k\in\mathcal{K}_D}r_{jk}^{it},\qquad\forall i\in\mathcal{I},j\in\mathcal{J}_D,t\in\mathcal{T} \tag{4.11}$$

$$x_j^{i,t+1}=x_j^{it}+q_j^{i,t-L_j^i}-v_{\eta,j}^{it},\qquad\forall i\in\mathcal{I},j\in\mathcal{J}_F,t\in\mathcal{T} \tag{4.12}$$

$$y_k^{i,t+1}=y_k^{it}+\sum_{j\in\mathcal{J}_D}u_{e,jk}^{it}-v_{\rho,k}^{it},\qquad\forall i\in\mathcal{I},k\in\mathcal{K}_O,t\in\mathcal{T} \tag{4.13}$$

$$q_j^{it}\geq0,\delta_j^{it}\in\{0,1\},\qquad\forall i\in\mathcal{I},j\in\mathcal{J},t\in\mathcal{T} \tag{4.14}$$

$$x_j^{it}\geq0,\qquad\forall i\in\mathcal{I},j\in\mathcal{J},t\in\mathcal{T}^+ \tag{4.15}$$

$$y_k^{it}\geq0,\qquad\forall i\in\mathcal{I},k\in\mathcal{K}_O,t\in\mathcal{T}^+ \tag{4.16}$$

$$u_{l,jj'}^{it}\geq0,\qquad\forall j,j'\in\mathcal{J}_D,i\in\mathcal{I},t\in\mathcal{T} \tag{4.17}$$

$$u_{e,jk}^{it}\geq0,\qquad\forall j\in\mathcal{J}_D,k\in\mathcal{K}_O,i\in\mathcal{I},t\in\mathcal{T} \tag{4.18}$$

$$v_{\rho,k}^{it}\geq0,\qquad\forall i\in\mathcal{I},k\in\mathcal{K}_O,t\in\mathcal{T} \tag{4.19}$$

$$v_{\eta,j}^{it}\geq0,\qquad\forall i\in\mathcal{I},j\in\mathcal{J}_F,t\in\mathcal{T} \tag{4.20}$$

$$r_{jk}^{it}\geq0,\qquad\forall i\in\mathcal{I},j\in\mathcal{J}_D,k\in\mathcal{K}_D,t\in\mathcal{T} \tag{4.21}$$

$$z_k^{it}\geq0,\qquad\forall i\in\mathcal{I},k\in\mathcal{K},t\in\mathcal{T} \tag{4.22}$$

The objective function (4.1) minimizes the total cost incurred within the supply chain. In the objective function, the first and second terms are the ordering cost, the third and fourth terms are the inventory holding cost, and the fifth term is the lost sales cost. The sixth term is the transshipment cost between DCs, and the seventh term is the allocation cost from DCs to offline stores. The eighth, ninth, and tenth terms are the fulfillment cost to the demand zones. Constraint (4.2) represents that if products are ordered, a fixed ordering cost is incurred. Constraint (4.3) enforces that the total number of products replenished from supplier $i$ cannot exceed the given production capacity $s^{it}$. Constraint (4.4) enforces that the inventory of the FC $j$ cannot exceed its capacity, $\bar{x}_j$, after products arrive. Constraint (4.5) also represents the storage capacity constraint for the DC $j$ considering the replenishment, transshipment, and allocation quantities. Constraint (4.6) represents that the number of products transshipped from the DC $j$ to other DCs should be less than the inventory of the DC $j$. Constraint (4.7) restricts that the inventory of the offline store $k$ cannot exceed its capacity, $\bar{y}_k$, after products arrive. Constraints (4.8), (4.9), and (4.10) ensure that the demand is satisfied by inventories held in DCs, offline stores, and FCs, respectively. Moreover, these constraints ensure that all unsatisfied demand becomes lost. Constraints (4.11), (4.12), and (4.13) are the balance equations for inventories of DCs, FCs, and offline stores, respectively. Finally, Constraints (4.14)−(4.22) ensure that decision variables are non-negative real variables, except for $\delta_j^{it}$, which are binary variables.

### 4.3.2 Stochastic optimization model

In this section, we present the stochastic optimization model to accommodate the demand uncertainty. We use $\tilde{d}_k^{it}$ to denote random demand $k$ for product $i$ at period $t$ for all $i \in \mathcal{I}, k \in \mathcal{K}, t \in \mathcal{T}$. The mean values of demand $\tilde{d}_k^{it}$ are denoted as $\hat{d}_k^{it}$, and the realization of the demand is denoted as $d_k^{it}$. For ease of the exposition, we utilize $\tilde{\mathbf{d}}^t = \left(\tilde{d}_k^{i\tau}, \forall i \in \mathcal{I}, k \in \mathcal{K}, \tau \in \{1, \ldots, t\}\right)$ to denote a collection of all demands from period 1 to period $t$, and $\tilde{\mathbf{d}}$ denotes $\tilde{\mathbf{d}}^T$. The realization of the demand $\tilde{\mathbf{d}}^t$ and $\tilde{\mathbf{d}}$ are denoted as $\mathbf{d}^t$ and $\mathbf{d}$, respectively.

In the proposed stochastic optimization model, we consider the *adjustable decision variables* to accommodate two different types of decisions (i.e., anticipative and reactive manners). The adjustable decision variables can postpone the decision until some portion of the demand is realized (i.e., wait-and-see decisions), which is different from the process that every decision should be made at the start of period 1 (i.e., here-and-now decisions). We define the following adjustable decision variables based on the information of the decision variables in the deterministic model:

**Adjustable decision variables:**

| | |
|---|---|
| $\delta_j^{it}(\tilde{\mathbf{d}}^{t-1})$ | 1 if product $i$ is replenished from supplier $i$ to the logistics center $j$ at the start of period $t$ after $\tilde{\mathbf{d}}^{t-1}$ is realized, 0 otherwise |
| $q_j^{it}(\tilde{\mathbf{d}}^{t-1})$ | quantity of the product $i$ replenished from supplier $i$ to the logistics center $j$ at the start of the period $t$ after $\tilde{\mathbf{d}}^{t-1}$ is realized |
| $x_j^{it}(\tilde{\mathbf{d}}^{t-1})$ | on-hand level of product $i$ in the logistics center $j$ at the start of period $t$ after $\tilde{\mathbf{d}}^{t-1}$ is realized |
| $y_k^{it}(\tilde{\mathbf{d}}^{t-1})$ | on-hand level of product $i$ in the offline store $k$ at the start of period $t$ after $\tilde{\mathbf{d}}^{t-1}$ is realized |
| $u_{l,jj'}^{it}(\tilde{\mathbf{d}}^{t-1})$ | quantity of the product $i$ transshipped from the DC $j$ to the other DC $j'$ at the start of the period $t$ after $\tilde{\mathbf{d}}^{t-1}$ is realized |

| | |
|---|---|
| $u^{it}_{e,jk}(\tilde{\mathbf{d}}^{t-1})$ | quantity of the product $i$ allocated from the DC $j$ to the offline store $k$ at the start of period $t$ after $\tilde{\mathbf{d}}^{t-1}$ is realized |
| $v^{it}_{\rho,k}(\tilde{\mathbf{d}}^{t})$ | quantity of the product $i$ fulfilled to satisfy the offline demand zone $k$ at the end of period $t$ after $\tilde{\mathbf{d}}^{t}$ is realized |
| $v^{it}_{\eta,j}(\tilde{\mathbf{d}}^{t})$ | quantity of the product $i$ from the FC $j$ fulfilled to satisfy the aggregate demand for FCsat the end of period $t$ after $\tilde{\mathbf{d}}^{t}$ is realized |
| $r^{it}_{jk}(\tilde{\mathbf{d}}^{t})$ | quantity of the product $i$ from the DC $j$ fulfilled to satisfy the online demand zone for DCs $k$ at the end of period $t$ after $\tilde{\mathbf{d}}^{t}$ is realized |
| $z^{it}_{k}(\tilde{\mathbf{d}}^{t})$ | lost sales of product $i$ for the demand type $k$ at the end of period $t$ after $\tilde{\mathbf{d}}^{t}$ is realized |

It should be noted that among the above adjustable decision variables, only the $\delta^{it}_j(\tilde{\mathbf{d}}^{t-1})$ are the *adjustable binary variables*, and the others are the *adjustable continuous variables*. In addition, because $\delta^{it}_j(\tilde{\mathbf{d}}^{t-1}), q^{it}_j(\tilde{\mathbf{d}}^{t-1}), x^{it}_j(\tilde{\mathbf{d}}^{t-1}), u^{it}_{l,jj'}(\tilde{\mathbf{d}}^{t-1})$, and $u^{it}_{e,jk}(\tilde{\mathbf{d}}^{t-1})$ are decided at the start of period $t$, these decisions are determined based on the anticipative manner. On the other hand, because $v^{it}_{\rho,k}(\tilde{\mathbf{d}}^{t}), v^{it}_{\eta,j}(\tilde{\mathbf{d}}^{t}), r^{it}_{jk}(\tilde{\mathbf{d}}^{t})$, and $z^{it}_k(\tilde{\mathbf{d}}^{t})$ are decided at the end of period $t$, these decisions are determined based on the reactive manner. For ease of exposition, let $\boldsymbol{\delta}(\tilde{\mathbf{d}}) = \left(\delta^{it}_j(\tilde{\mathbf{d}}^{t-1}), \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}\right)$ denote a collection of the adjustable binary variables. We use notations $\boldsymbol{\pi}(\tilde{\mathbf{d}})$ and $\boldsymbol{\mu}(\tilde{\mathbf{d}})$ to denote a collection of the adjustable continuous variables determined based on the anticipative and reactive manners, respectively:

$$\boldsymbol{\pi}(\tilde{\mathbf{d}}) = \left(q^{it}_j(\tilde{\mathbf{d}}^{t-1}), x^{it}_j(\tilde{\mathbf{d}}^{t-1}), y^{it}_k(\tilde{\mathbf{d}}^{t-1}), u^{it}_{l,jj'}(\tilde{\mathbf{d}}^{t-1}), u^{it}_{e,jk}(\tilde{\mathbf{d}}^{t-1}),\ \forall i \in \mathcal{I}, j \in \mathcal{J}, j' \in \mathcal{J}, k \in \mathcal{K}_O, t \in \mathcal{T}\right)$$

$$\boldsymbol{\mu}(\tilde{\mathbf{d}}) = \left(v^{it}_{\rho,k}(\tilde{\mathbf{d}}^{t}), v^{it}_{\eta,j}(\tilde{\mathbf{d}}^{t}), r^{it}_{jk}(\tilde{\mathbf{d}}^{t}), z^{it}_k(\tilde{\mathbf{d}}^{t}),\ \forall i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K}, t \in \mathcal{T}\right)$$

If the demand is given as $\mathbf{d}$, the total cost incurred in the supply chain is defined

as follows:

$$\boldsymbol{\Psi}\left(\boldsymbol{\delta}(\mathbf{d}), \boldsymbol{\pi}(\mathbf{d}), \boldsymbol{\mu}(\mathbf{d})\right) =$$

$$\sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}}\Bigg(\sum_{j\in\mathcal{J}} S_j^{it}\delta_j^{it}(\mathbf{d}^{t-1}) + \sum_{j\in\mathcal{J}}\lambda_o^{it}c_o^{ij}q_j^{it}(\mathbf{d}^{t-1}) + \sum_{j\in\mathcal{J}} h_{x,j}^{it}x_j^{i,t+1}(\mathbf{d}^t) + \sum_{k\in\mathcal{K}_O} h_{y,k}^{it}y_k^{i,t+1}(\mathbf{d}^t)$$

$$+\sum_{k\in\mathcal{K}} p_k^{it}z_k^{it}(\mathbf{d}^t) + \sum_{j\in\mathcal{J}_D}\sum_{j'\in\mathcal{J}_D}\lambda_l^{it}c_l^{jj'}u_{l,jj'}^{it}(\mathbf{d}^{t-1}) + \sum_{j\in\mathcal{J}_D}\sum_{k\in\mathcal{K}_O}\lambda_e^{it}c_e^{jk}u_{e,jk}^{it}(\mathbf{d}^{t-1})$$

$$+\sum_{j\in\mathcal{J}_D}\sum_{k\in\mathcal{K}_D}\lambda_g^{it}c_g^{jk}r_{jk}^{it}(\mathbf{d}^t) + \sum_{k\in\mathcal{K}_O}\rho_k^{it}v_{\rho,k}^{it}(\mathbf{d}^t) + \sum_{j\in\mathcal{J}_F}\eta_j^{it}v_{\eta,j}^{it}(\mathbf{d}^t)\Bigg)$$

We propose the following stochastic optimization model ($\mathrm{P_{STOC}}$) by accommodating the demand uncertainty:

($\mathrm{P_{STOC}}$)

$$\min \quad \mathbb{E}_{\tilde{\mathbf{d}}}\left[\boldsymbol{\Psi}\left(\boldsymbol{\delta}(\tilde{\mathbf{d}}), \boldsymbol{\pi}(\tilde{\mathbf{d}}), \boldsymbol{\mu}(\tilde{\mathbf{d}})\right)\right] \tag{4.23}$$

$$\text{s.t.} \quad q_j^{it}(\tilde{\mathbf{d}}^{t-1}) \le \bar{q}_j^i\delta_j^{it}(\tilde{\mathbf{d}}^{t-1}), \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}, t\in\mathcal{T} \tag{4.24}$$

$$\sum_{j\in\mathcal{J}} q_j^{it}(\tilde{\mathbf{d}}^{t-1}) \le s^{it}, \qquad \forall i\in\mathcal{I}, t\in\mathcal{T} \tag{4.25}$$

$$\sum_{i\in\mathcal{I}}\left(x_j^{it}(\tilde{\mathbf{d}}^{t-1}) + q_j^{i,t-L_j^i}(\tilde{\mathbf{d}}^{t-L_j^i-1})\right) \le \bar{x}_j, \qquad \forall j\in\mathcal{J}_F, t\in\mathcal{T} \tag{4.26}$$

$$\sum_{i\in\mathcal{I}}\Bigg(x_j^{it}(\tilde{\mathbf{d}}^{t-1}) + q_j^{i,t-L_j^i}(\tilde{\mathbf{d}}^{t-L_j^i-1}) + \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,j'j}^{it}(\tilde{\mathbf{d}}^{t-1}) - \sum_{k\in\mathcal{K}_O} u_{e,jk}^{it}(\tilde{\mathbf{d}}^{t-1})$$

$$-\sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,jj'}^{it}(\tilde{\mathbf{d}}^{t-1})\Bigg) \le \bar{x}_j, \qquad \forall j\in\mathcal{J}_D, t\in\mathcal{T} \tag{4.27}$$

$$x_j^{it}(\tilde{\mathbf{d}}^{t-1}) + q_j^{i,t-L_j^i}(\tilde{\mathbf{d}}^{t-L_j^i-1}) \ge \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,jj'}^{it}(\tilde{\mathbf{d}}^{t-1}), \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}_D, t\in\mathcal{T} \tag{4.28}$$

$$\sum_{i\in\mathcal{I}}\left(y_k^{it}(\tilde{\mathbf{d}}^{t-1}) + \sum_{j\in\mathcal{J}_D} u_{e,jk}^{it}(\tilde{\mathbf{d}}^{t-1})\right) \le \bar{y}_k, \qquad \forall k\in\mathcal{K}_O, t\in\mathcal{T} \tag{4.29}$$

$$\sum_{j\in\mathcal{J}_D} r_{jk}^{it}(\tilde{\mathbf{d}}^t) + z_k^{it}(\tilde{\mathbf{d}}^t) = \tilde{d}_k^{it}, \qquad \forall i\in\mathcal{I}, k\in\mathcal{K}_D, t\in\mathcal{T} \tag{4.30}$$

$$v_{\rho,k}^{it}(\tilde{\mathbf{d}}^t) + z_k^{it}(\tilde{\mathbf{d}}^t) = \tilde{d}_k^{it}, \qquad \forall i\in\mathcal{I}, k\in\mathcal{K}_O, t\in\mathcal{T} \tag{4.31}$$

$$\sum_{j\in\mathcal{J}_F} v_{\eta,j}^{it}(\tilde{\mathbf{d}}^t) + z_{K+1}^{it}(\tilde{\mathbf{d}}^t) = \tilde{d}_{K+1}^{it}, \qquad \forall i\in\mathcal{I}, t\in\mathcal{T} \tag{4.32}$$

$$x_j^{i,t+1}(\tilde{\mathbf{d}}^t) = x_j^{it}(\tilde{\mathbf{d}}^{t-1}) + q_j^{i,t-L_j^i}(\tilde{\mathbf{d}}^{t-L_j^i-1}) + \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,j'j}^{it}(\tilde{\mathbf{d}}^{t-1}) - \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,jj'}^{it}(\tilde{\mathbf{d}}^{t-1})$$
$$- \sum_{k\in\mathcal{K}_O} u_{e,jk}^{it}(\tilde{\mathbf{d}}^{t-1}) - \sum_{k\in\mathcal{K}_D} r_{jk}^{it}(\tilde{\mathbf{d}}^t), \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}_D, t\in\mathcal{T} \tag{4.33}$$

$$x_j^{i,t+1}(\tilde{\mathbf{d}}^t) = x_j^{it}(\tilde{\mathbf{d}}^{t-1}) + q_j^{i,t-L_j^i}(\tilde{\mathbf{d}}^{t-L_j^i-1}) - v_{\eta,j}^{it}(\tilde{\mathbf{d}}^t), \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}_F, t\in\mathcal{T} \tag{4.34}$$

$$y_k^{i,t+1}(\tilde{\mathbf{d}}^t) = y_k^{it}(\tilde{\mathbf{d}}^{t-1}) + \sum_{j\in\mathcal{J}_D} u_{e,jk}^{it}(\tilde{\mathbf{d}}^{t-1}) - v_{\rho,k}^{it}(\tilde{\mathbf{d}}^t), \qquad \forall i\in\mathcal{I}, k\in\mathcal{K}_O, t\in\mathcal{T} \tag{4.35}$$

$$q_j^{it}(\tilde{\mathbf{d}}^{t-1}) \geq 0, q_j^{it}(\tilde{\mathbf{d}}^{t-1}) \in \mathcal{R}^{t-1}, \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}, t\in\mathcal{T} \tag{4.36}$$

$$x_j^{it}(\tilde{\mathbf{d}}^{t-1}) \geq 0, x_j^{it}(\tilde{\mathbf{d}}^{t-1}) \in \mathcal{R}^{t-1}, \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}, t\in\mathcal{T}^+ \tag{4.37}$$

$$y_k^{it}(\tilde{\mathbf{d}}^{t-1}) \geq 0, y_k^{it}(\tilde{\mathbf{d}}^{t-1}) \in \mathcal{R}^{t-1}, \qquad \forall i\in\mathcal{I}, k\in\mathcal{K}_O, t\in\mathcal{T}^+ \tag{4.38}$$

$$u_{l,jj'}^{it}(\tilde{\mathbf{d}}^{t-1}) \geq 0, u_{l,jj'}^{it}(\tilde{\mathbf{d}}^{t-1}) \in \mathcal{R}^{t-1}, \qquad \forall j,j'\in\mathcal{J}_D, i\in\mathcal{I}, t\in\mathcal{T} \tag{4.39}$$

$$u_{e,jk}^{it}(\tilde{\mathbf{d}}^{t-1}) \geq 0, u_{e,jk}^{it}(\tilde{\mathbf{d}}^{t-1}) \in \mathcal{R}^{t-1}, \qquad \forall j\in\mathcal{J}_D, k\in\mathcal{K}_O, i\in\mathcal{I}, t\in\mathcal{T} \tag{4.40}$$

$$v_{\rho,k}^{it}(\tilde{\mathbf{d}}^t) \geq 0, v_{\rho,k}^{it}(\tilde{\mathbf{d}}^t) \in \mathcal{R}^t, \qquad \forall i\in\mathcal{I}, k\in\mathcal{K}_O, t\in\mathcal{T} \tag{4.41}$$

$$v_{\eta,j}^{it}(\tilde{\mathbf{d}}^t) \geq 0, v_{\eta,j}^{it}(\tilde{\mathbf{d}}^t) \in \mathcal{R}^t, \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}_F, t\in\mathcal{T} \tag{4.42}$$

$$r_{jk}^{it}(\tilde{\mathbf{d}}^t) \geq 0, r_{jk}^{it}(\tilde{\mathbf{d}}^t) \in \mathcal{R}^t, \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}_D, k\in\mathcal{K}_D, t\in\mathcal{T} \tag{4.43}$$

$$z_k^{it}(\tilde{\mathbf{d}}^t) \geq 0, z_k^{it}(\tilde{\mathbf{d}}^t) \in \mathcal{R}^t, \qquad \forall i\in\mathcal{I}, k\in\mathcal{K}, t\in\mathcal{T} \tag{4.44}$$

$$\delta_j^{it}(\tilde{\mathbf{d}}^{t-1}) \in \mathcal{B}^{t-1}, \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}, t\in\mathcal{T} \tag{4.45}$$

where $\mathcal{R}^\tau$ and $\mathcal{B}^\tau$ functions are mapping from $\mathbb{R}^{I\times\tau\times(K+1)}$ to $\mathbb{R}$ and $\{0,1\}$, respectively. The objective of the $\mathrm{P}_{\mathrm{STOC}}$ is to minimize the total expected cost, and

every constraint must be satisfied for all demand realizations. The $\mathrm{P}_{\mathrm{STOC}}$ is the multistage stochastic optimization problem that is generally computationally intractable to solve [123]. Traditionally, dynamic programming or multistage stochastic programming methods are used to solve the stochastic optimization problem by characterizing demand uncertainty with a known probability distribution. However, assumptions about demand distribution could be unrealistic if a decision maker has insufficient demand data. If the gap between true demand and assumed distributions is large, solutions derived from these methods could show poor performance in practice. Furthermore, the computational complexity to solve $\mathrm{P}_{\mathrm{STOC}}$ is increased significantly due to the existence of the adjustable binary variables $\boldsymbol{\delta}(\tilde{\mathbf{d}})$. Through numerical experiments, Lim et al. [93] and Jiu [72] showed that existing approaches, specifically the BDR [15] and the FA [17], require significant computational burdens to solve the problem with the adjustable binary variables.

Lim et al. [93] proposed a TPA that does not require any assumptions about demand distribution and could reduce computational burdens. Because of these distinct advantages of the TPA, Jiu [72] also extended the applicability of the TPA to robust omnichannel retail operations. While the TPA performs well in certain problems, it shows poor performance in our problem because of large-scale issues incurred by the supply chain of the 3PP. In addition, Constraint (4.25) that restricts the replenishment quantity with suppliers' production capacity increases the computational complexity. These issues motivate us to develop a suitable approach to our problem (i.e., DECOM). Before explaining the proposed approach, we briefly introduce how we customize the TPA for our problem in the following section.

## 4.4 A two-phase approach (TPA) based on robust optimization

A TPA solves the proposed problem by decoupling binary decision variables and the continuous decision variables. In Phase 1, the binary decisions are determined with the static rule (i.e., $\boldsymbol{\delta}(\tilde{\mathbf{d}}^{t-1}) = \boldsymbol{\delta}$) by utilizing a TRO [94]. In Phase 2, we adaptively decide the continuous variables by utilizing the LDR with an objective of minimizing the worst-case expected total cost [11]. In order to adopt a TPA, it is assumed that the demand $\tilde{d}_k^{it}$ is $\hat{d}_k^{it}$ mean random variables and fall in a support set $\left[\underline{d}_k^{it}, \bar{d}_k^{it}\right], \forall i \in \mathcal{I}, k \in \mathcal{K}, t \in \mathcal{T}$. Considering this assumption, the uncertainty set for each $\tilde{d}_k^{it}$ is defined as $D_k^{it} := \left\{d_k^{it} \mid \underline{d}_k^{it} \leq d_k^{it} \leq \bar{d}_k^{it}\right\}$ where $\underline{\zeta}_k^{it} = \hat{d}_k^{it} - \underline{d}_k^{it}$ and $\bar{\zeta}_k^{it} = \bar{d}_k^{it} - \hat{d}_k^{it}$, $\forall i \in \mathcal{I}, k \in \mathcal{K}, t \in \mathcal{T}$.

### 4.4.1 Phase 1 of TPA

In Phase 1, we determine the binary decisions by utilizing the TRO that maximizes the sizes of the uncertainty sets and makes a total cost lower than a predetermined cost target. Lim and Wang [94] proved that a static rule is optimal for TRO formulation and showed that the computational burden could be reduced significantly. In order to reformulate $\mathrm{P}_{\mathrm{STOC}}$ into the TRO model, we define the *adjustable uncertainty set* for each $\tilde{d}_k^{it}$ as $D_k^{it}(\gamma) := \left\{d_k^{it} \mid \hat{d}_k^{it} - \gamma\underline{\zeta}_k^{it} \leq d_k^{it} \leq \hat{d}_k^{it} + \gamma\bar{\zeta}_k^{it}\right\}$ where $\underline{\zeta}_k^{it} = \hat{d}_k^{it} - \underline{d}_k^{it}$ and $\bar{\zeta}_k^{it} = \bar{d}_k^{it} - \hat{d}_k^{it}$. For notational convenience, let $\mathbf{D}^t(\gamma) = \left(D_k^{i\tau}(\gamma), \forall k \in \mathcal{K}, i \in \mathcal{I}, \tau \in \{1, \ldots, t\}\right)$ and $\mathbf{D}(\gamma) = \mathbf{D}^T(\gamma)$. In addition, we define a *cost target* $\psi$ to restrict total cost to be no more than a predetermined value $\psi$ under any demand realizations. We present the TRO model, $\mathrm{P}_{\mathrm{TRO}}$, as follows:

$(\mathrm{P}_{\mathrm{TRO}})$

$$
\begin{aligned}
\gamma^* = \ & \max\ \gamma \\
\text{s.t.} \quad & \boldsymbol{\Psi}\left(\boldsymbol{\delta}(\mathbf{d}), \boldsymbol{\pi}(\mathbf{d}), \boldsymbol{\mu}(\mathbf{d})\right) \leq \psi, \qquad \forall \mathbf{d} \in \mathbf{D}(\gamma) \\
& \text{Constraints } (4.24) - (4.45), \qquad \forall \mathbf{d}^t \in \mathbf{D}^t(\gamma) \\
& 0 \leq \gamma \leq 1
\end{aligned}
$$

The objective of the model is to absorb as much uncertainty as by maximizing the sizes of the adjustable uncertainty set. We control the sizes of adjustable uncertainty set by adopting the new decision variable $\gamma$ ($0 \leq \gamma \leq 1$). Simultaneously, the total cost must be lower than a cost target $\psi$ as indicated in the first constraint. The other constraints are the same as $\mathrm{P_{STOC}}$. However, the equality constraints (4.30)−(4.32) could cause an infeasibility issue if the static rule is adopted. Fortunately, we can overcome this issue by allowing Constraints (4.30)−(4.32) to be relaxed from equality to inequality as follows [93]:

$(\mathrm{P_{TRO-R}})$

$$
\begin{aligned}
\gamma' = \ & \max\ \gamma && (4.46) \\
\text{s.t.} \quad & \boldsymbol{\Psi}\left(\boldsymbol{\delta}(\mathbf{d}), \boldsymbol{\pi}(\mathbf{d}), \boldsymbol{\mu}(\mathbf{d})\right) \leq \psi, \qquad \forall \mathbf{d} \in \mathbf{D}(\gamma) && (4.47) \\
& \sum_{j \in \mathcal{J}_D} r_{jk}^{it}(\mathbf{d}^t) + z_k^{it}(\mathbf{d}^t) \geq d_k^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_D, \qquad \forall \mathbf{d}^t \in \mathbf{D}^t(\gamma) && (4.48) \\
& v_{\rho,k}^{it}(\mathbf{d}^t) + z_k^{it}(\mathbf{d}^t) \geq d_k^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_O, \qquad \forall \mathbf{d}^t \in \mathbf{D}^t(\gamma) && (4.49) \\
& \sum_{j \in \mathcal{J}_F} v_{\eta,j}^{it}(\mathbf{d}^t) + z_{K+1}^{it}(\mathbf{d}^t) \geq d_{K+1}^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, \qquad \forall \mathbf{d}^t \in \mathbf{D}^t(\gamma) && (4.50) \\
& \text{Constraints } (4.24) - (4.29), (4.33) - (4.45) \qquad \forall \mathbf{d}^t \in \mathbf{D}^t(\gamma) && (4.51) \\
& 0 \leq \gamma \leq 1 && (4.52)
\end{aligned}
$$

It should be noted that Constraints (4.48)−(4.50) lead to $\gamma^* \leq \gamma'$.

We define *uncertainty variables* $f_k^{it}$ and $g_k^{it}$ falling in $F_k^{it}(\gamma) := \left\{ f_k^{it} \mid 0 \leq f_k^{it} \leq \gamma \zeta_{\underline{k}}^{it} \right\}$

and $G_k^{it}(\gamma) := \left\{g_k^{it} \mid 0 \le g_k^{it} \le \gamma\bar{\zeta}_k^{it}\right\}$, respectively. By adopting uncertainty variables, we can tighten constraints in $\mathrm{P}_{\mathrm{TRO-R}}$, and each demand can be represented as $d_k^{it} = \hat{d}_k^{it} - f_k^{it} + g_k^{it},\ \forall i \in \mathcal{I}, k \in \mathcal{K}, t \in \mathcal{T}$. For convenience, we define boldface notation to denote collections of $f_k^{it}, g_k^{it}, F_k^{it}(\gamma)$, and $G_k^{it}(\gamma)$ as

$$\mathbf{f}^t = \left(f_k^{i\tau}(\gamma),\ \forall i \in \mathcal{I}, k \in \mathcal{K}, \tau \in \{1,\ldots,t\}\right),\ \mathbf{g}^t = \left(g_k^{i\tau}(\gamma),\ \forall i \in \mathcal{I}, k \in \mathcal{K}, \tau \in \{1,\ldots,t\}\right),$$
$$\mathbf{F}^t(\gamma) = \left(F_k^{i\tau}(\gamma),\ \forall i \in \mathcal{I}, k \in \mathcal{K}, \tau \in \{1,\ldots,t\}\right),\ \mathbf{G}^t(\gamma) = \left(G_k^{i\tau}(\gamma),\ \forall i \in \mathcal{I}, k \in \mathcal{K}, \tau \in \{1,\ldots,t\}\right).$$

By replacing $d_k^{it}$ with $\hat{d}_k^{it} + g_k^{it}$ in Constraints (4.48)−(4.50), $\mathrm{P}_{\mathrm{TRO-R}}$ can be approximated as follows:

$(\mathrm{P}_{\mathrm{TRO-A}})$

$$\begin{aligned}
\gamma'' = \max\ & \gamma \\
\text{s.t.}\quad & \boldsymbol{\Psi}\left(\boldsymbol{\delta}(\mathbf{d}), \boldsymbol{\pi}(\mathbf{d}), \boldsymbol{\mu}(\mathbf{d})\right) \le \psi, \quad \forall \mathbf{f}^t \in \mathbf{F}^t(\gamma), \mathbf{g}^t \in \mathbf{G}^t(\gamma) \\
& \sum_{j \in \mathcal{J}_D} r_{jk}^{it}(\mathbf{d}^t) + z_k^{it}(\mathbf{d}^t) \ge \hat{d}_k^{it} + g_k^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_D, \quad \forall \mathbf{f}^t \in \mathbf{F}^t(\gamma), \mathbf{g}^t \in \mathbf{G}^t(\gamma) \\
& v_{\rho,k}^{it}(\mathbf{d}^t) + z_k^{it}(\mathbf{d}^t) \ge \hat{d}_k^{it} + g_k^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_O, \quad \forall \mathbf{f}^t \in \mathbf{F}^t(\gamma), \mathbf{g}^t \in \mathbf{G}^t(\gamma) \\
& \sum_{j \in \mathcal{J}_F} v_{\eta,j}^{it}(\mathbf{d}^t) + z_{K+1}^{it}(\mathbf{d}^t) \ge \hat{d}_{K+1}^{it} + g_{K+1}^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, \quad \mathbf{f}^t \in \mathbf{F}^t(\gamma), \mathbf{g}^t \in \mathbf{G}^t(\gamma) \\
& \text{Constraints } (4.24) - (4.29), (4.33) - (4.45), \quad \mathbf{f}^t \in \mathbf{F}^t(\gamma), \mathbf{g}^t \in \mathbf{G}^t(\gamma) \\
& 0 \le \gamma \le 1
\end{aligned}$$

Because constraints in $\mathrm{P}_{\mathrm{TRO-A}}$ are tighter than those of Problem $\mathrm{P}_{\mathrm{TRO-R}}$, it is obvious that $\gamma'' \le \gamma'$.

We consider a static rule; thus, decisions are fixed regardless of the revealed uncertainties. Therefore, every adjustable variable is replaced with the decision variables of the deterministic problem (e.g., $\delta_j^{it}(\mathbf{d}^{t-1}) \to \delta_j^{it}$ and $\boldsymbol{\delta}(\mathbf{d}) \to \boldsymbol{\delta}$). We

define the total cost for the static rule as follows:

$$\boldsymbol{\Psi}^{\dagger}(\boldsymbol{\delta},\boldsymbol{\pi},\boldsymbol{\mu}) =$$

$$\sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}}\Bigg(\sum_{j\in\mathcal{J}} S_j^{it}\delta_j^{it} + \sum_{j\in\mathcal{J}}\lambda_o^{it}c_o^{ij}q_j^{it} + \sum_{j\in\mathcal{J}} h_{x,j}^{it}x_j^{i,t+1} + \sum_{k\in\mathcal{K}_O} h_{y,k}^{it}y_k^{i,t+1} + \sum_{k\in\mathcal{K}} p_k^{it}z_k^{it}$$

$$+\sum_{j\in\mathcal{J}_D}\sum_{j'\in\mathcal{J}_D}\lambda_l^{it}c_l^{jj'}u_{l,jj'}^{it} + \sum_{j\in\mathcal{J}_D}\sum_{k\in\mathcal{K}_O}\lambda_e^{it}c_e^{jk}u_{e,jk}^{it} + \sum_{j\in\mathcal{J}_D}\sum_{k\in\mathcal{K}_D}\lambda_g^{it}c_g^{jk}r_{jk}^{it} + \sum_{k\in\mathcal{K}_O}\rho_k^{it}v_{\rho,k}^{it} + \sum_{j\in\mathcal{J}_F}\eta_j^{it}v_{\eta,j}^{it}\Bigg)$$

The static rule can be derived by solving the following $\mathrm{P}_{\mathrm{TRO-S}}$:

$(\mathrm{P}_{\mathrm{TRO-S}})$

$$\gamma^s = \max \quad \gamma \tag{4.53}$$

$$\text{s.t.} \quad \boldsymbol{\Psi}^{\dagger}(\boldsymbol{\delta},\boldsymbol{\pi},\boldsymbol{\mu}) \le \psi \tag{4.54}$$

$$\sum_{j\in\mathcal{J}_D} r_{jk}^{it} + z_k^{it} \ge \hat{d}_k^{it} + g_k^{it}, \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}, k\in\mathcal{K}_D, \quad \forall \mathbf{f}^t\in\mathbf{F}^t(\gamma), \mathbf{g}^t\in\mathbf{G}^t(\gamma) \tag{4.55}$$

$$v_{\rho,k}^{it} + z_k^{it} \ge \hat{d}_k^{it} + g_k^{it}, \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}, k\in\mathcal{K}_O, \quad \forall \mathbf{f}^t\in\mathbf{F}^t(\gamma), \mathbf{g}^t\in\mathbf{G}^t(\gamma) \tag{4.56}$$

$$\sum_{j\in\mathcal{J}_F} v_{\eta,j}^{it} + z_{K+1}^{it} \ge \hat{d}_{K+1}^{it} + g_{K+1}^{it}, \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}, \quad \forall \mathbf{f}^t\in\mathbf{F}^t(\gamma), \mathbf{g}^t\in\mathbf{G}^t(\gamma) \tag{4.57}$$

$$\text{Constraints } (4.2)-(4.7), (4.11)-(4.22) \tag{4.58}$$

$$0 \le \gamma \le 1 \tag{4.59}$$

We could know that $\gamma^s \le \gamma''$ because decisions with the static rule are more restrictive than adjustable decisions. Before presenting an approach to derive an optimal static rule for $\mathrm{P}_{\mathrm{TRO-S}}$, we use the notation $\boldsymbol{\theta}$ to denote a collection of uncertainty variables $f_k^{it}$ and $g_k^{it}$ (i.e., $\boldsymbol{\theta} = \big(f_k^{it}, g_k^{it}, \forall i\in\mathcal{I}, k\in\mathcal{K}, t\in\mathcal{T}\big)$). Given $\gamma$, let $\boldsymbol{\Theta}(\gamma)$ denote the support set of $\boldsymbol{\theta}$. For ease of exposition, we represent $\mathrm{P}_{\mathrm{TRO-S}}$

as the following simple form:

$$
\begin{aligned}
\gamma^s = \max \quad & \gamma \\
\text{s.t.} \quad & \mathbf{C}(\boldsymbol{\theta})\boldsymbol{\kappa} \leq \mathbf{e}(\boldsymbol{\theta}), && \forall \boldsymbol{\theta} \in \boldsymbol{\Theta}(\gamma) \\
& \boldsymbol{\kappa} \in \boldsymbol{\Pi}, && \forall \boldsymbol{\theta} \in \boldsymbol{\Theta}(\gamma)
\end{aligned}
$$

where $\mathbf{C}(\boldsymbol{\theta})$ and $\mathbf{e}(\boldsymbol{\theta})$ represent all coefficients, and $\boldsymbol{\kappa}$ and $\boldsymbol{\Pi}$ represent decision variables for the static rule and the feasible set, respectively. We present the definition of the *worst-case scenario of uncertainty* as follows:

**Definition 4.1** (Worst-case scenario of uncertainty)**.** *Given the coefficients* $\mathbf{C}(\boldsymbol{\theta})$ *and* $\mathbf{e}(\boldsymbol{\theta})$ *in* $\mathrm{P}_{\mathrm{TRO-S}}$*, an element* $\breve{\boldsymbol{\theta}}(\gamma) \in \boldsymbol{\Theta}(\gamma)$ *is called the worst-case scenario of uncertainty if for each* $\boldsymbol{\kappa} \in \boldsymbol{\Pi}$ *that satisfies* $\mathbf{C}(\breve{\boldsymbol{\theta}}(\gamma))\boldsymbol{\kappa} \leq \mathbf{e}(\breve{\boldsymbol{\theta}}(\gamma))$*, it also satisfies* $\mathbf{C}(\boldsymbol{\theta})\boldsymbol{\kappa} \leq \mathbf{e}(\boldsymbol{\theta})$*,* $\forall \boldsymbol{\theta} \in \boldsymbol{\Theta}(\gamma)$*.*

We reformulate $\mathrm{P}_{\mathrm{TRO-S}}$ with the worst-case scenario of uncertainty $\breve{\boldsymbol{\theta}}(\gamma)$ by replacing the right-hand side inequality Constraints (4.55)−(4.57) from $\hat{d}_k^{it} + g_k^{it}$ to $\hat{d}_k^{it} + \gamma\bar{\zeta}_k^{it}$. Finally, the problem with the worst-case scenario of uncertainty is defined as the following deterministic problem:

$$
\begin{aligned}
&(\mathrm{P}_{\mathrm{STATIC}}) \\
&\quad \gamma^\dagger = \max \quad \gamma \\
&\quad \text{s.t.} \quad \boldsymbol{\Psi}^\dagger(\boldsymbol{\delta}, \boldsymbol{\pi}, \boldsymbol{\mu}) \leq \psi \\
&\qquad \sum_{j \in \mathcal{J}_D} r_{jk}^{it} + z_k^{it} \geq \hat{d}_k^{it} + \gamma\bar{\zeta}_k^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_D \\
&\qquad v_{\rho,k}^{it} + z_k^{it} \geq \hat{d}_k^{it} + \gamma\bar{\zeta}_k^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_O \\
&\qquad \sum_{j \in \mathcal{J}_F} v_{\eta,j}^{it} + z_{K+1}^{it} \geq \hat{d}_{K+1}^{it} + \gamma\bar{\zeta}_{K+1}^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}
\end{aligned}
$$

$$\text{Constraints } (4.2) - (4.7), (4.11) - (4.22)$$

$$0 \leq \gamma \leq 1$$

Let $\bar{\boldsymbol{\delta}}$ denote the optimal solution of $\boldsymbol{\delta}$ obtained by solving the $\mathrm{P}_{\mathrm{STATIC}}$. Because the constraints of $\mathrm{P}_{\mathrm{STATIC}}$ are more restrictive than those of $\mathrm{P}_{\mathrm{TRO-S}}$, we have $\gamma^{\dagger} \leq \gamma^{s} \leq \gamma^{''}$. Interestingly, Lim et al. [93] shows that $\gamma^{\dagger} \geq \gamma^{s} \geq \gamma^{''}$ in Theorem 1. Therefore, we have $\gamma^{\dagger} = \gamma^{''}$; thus, the optimal solution of the deterministic Problem $\mathrm{P}_{\mathrm{STATIC}}$ is also optimal for $\mathrm{P}_{\mathrm{TRO-A}}$.

By controlling the cost target $\psi$, a decision maker could choose the degree of conservativeness for the obtained solution. To determine the proper value for $\psi$, we utilize the following affine function of $\phi$, which is called the *target coefficient*:

$$\psi(\phi) := (1-\phi)\nu(1) + \phi\nu(0)$$

where $\nu(1)$ and $\nu(0)$ is the optimal objective of the following deterministic problem:

$$(\mathrm{P}_{\mathrm{TPA}-\nu(\gamma)})$$

$$\nu(\gamma) = \min \boldsymbol{\Psi}^{\dagger}(\boldsymbol{\delta}, \boldsymbol{\pi}, \boldsymbol{\mu})$$

$$\text{s.t.} \quad \sum_{j \in \mathcal{J}_D} r_{jk}^{it} + z_k^{it} \geq \hat{d}_k^{it} + \gamma \bar{\zeta}_k^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_D$$

$$v_{\rho,k}^{it} + z_k^{it} \geq \hat{d}_k^{it} + \gamma \bar{\zeta}_k^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_O$$

$$\sum_{j \in \mathcal{J}_F} v_{\eta,j}^{it} + z_{K+1}^{it} \geq \hat{d}_{K+1}^{it} + \gamma \bar{\zeta}_{K+1}^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}$$

$$\text{Constraints } (4.2) - (4.7), (4.11) - (4.22)$$

If the $\phi$ is close to zero, the conservativeness of solutions is increased; otherwise, it is decreased. Until now, we briefly introduced the principle of the TRO approach in

this section. For further information, we recommend readers refer to Lim and Wang [94] and Lim et al. [93].

### 4.4.2 Phase 2 of TPA

In Phase 2, we determine the adjustable continuous variables with fixed binary decisions $\bar{\boldsymbol{\delta}}$ obtained in Phase 1. As mentioned in Section 4.4.1, without any knowledge on the true demand distribution, only the mean of $\tilde{d}_k^{it}$ (i.e., $\hat{d}_k^{it}$) and the support set $\left[\underline{d}_k^{it}, \bar{d}_k^{it}\right]$ are given. In order to deal with distributional ambiguity, we adopt the solution approach proposed by Giloba and Schmeidler [53]. We first consider $\mathcal{F}$ as a family of distributions of $\tilde{\mathbf{d}}$, and the mean support set $\hat{\mathbf{D}} = \left(\hat{D}_k^{it}, \forall k \in \mathcal{K}, i \in \mathcal{I}, t \in \mathcal{T}\right)$. Let $\mathcal{P}$ denote any distribution of $\tilde{\mathbf{d}}$ included in $\mathcal{F}$, $\mathcal{P} \in \mathcal{F}$; thus, we have $\mathbb{E}_{\mathcal{P}}\left[\tilde{\mathbf{d}}\right] \in \hat{\mathbf{D}}$. We solve the following problem with the objective of minimizing the worst-case expected total cost over a family of distributions $\mathcal{F}$:

$$
\begin{aligned}
&(\mathrm{P}_{\mathrm{ARO}}) \\
\min \max_{\mathcal{P} \in \mathcal{F}} \quad & \mathbb{E}_{\mathcal{P}}\left[\boldsymbol{\Psi}\left(\boldsymbol{\pi}(\tilde{\mathbf{d}}), \boldsymbol{\mu}(\tilde{\mathbf{d}})\right)\right] \\
\text{s.t.} \quad & q_j^{it}(\mathbf{d}^{t-1}) \leq \bar{q}_j^i \bar{\delta}_j^{it}, \qquad i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}, \qquad \forall \mathbf{d}^{t-1} \in \mathbf{D}^{t-1} \\
& \text{Constraints (4.25)} - \text{(4.44)}, \qquad \forall \mathbf{d}^t \in \mathbf{D}^t
\end{aligned}
$$

with the fixed binary decisions $\bar{\delta}_j^{it}$ in the first constraint. Because it is generally intractable to solve $\mathrm{P}_{\mathrm{ARO}}$, we rely on optimizing parameterized functions, where the feasible space is restricted to linear functions (i.e., the LDR [147]). For each adjustable continuous variable, we define the following LDR:

$$
\begin{aligned}
q_j^{it}\left(\mathbf{d}^{t-1}\right) &= q_j^{it,0} + \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} q_j^{it,\sigma\tau} d_\sigma^{i\tau},, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \\
x_j^{it}\left(\mathbf{d}^{t-1}\right) &= x_j^{it,0} + \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} x_j^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}^+
\end{aligned}
$$

$$y_k^{it}\left(\mathbf{d}^{t-1}\right) = y_k^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1} y_k^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}^+$$

$$u_{l,jj'}^{it}\left(\mathbf{d}^{t-1}\right) = u_{l,jj'}^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1} u_{l,jj'}^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, j' \in \mathcal{J}_D, t \in \mathcal{T}$$

$$u_{e,jk}^{it}\left(\mathbf{d}^{t-1}\right) = u_{e,jk}^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1} u_{e,jk}^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_O, t \in \mathcal{T}$$

$$v_{\rho,k}^{it}\left(\mathbf{d}^{t}\right) = v_{\rho,k}^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t} v_{\rho,k}^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}$$

$$v_{\eta,j}^{it}\left(\mathbf{d}^{t}\right) = v_{\eta,j}^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t} v_{\eta,j}^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}$$

$$r_{jk}^{it}\left(\mathbf{d}^{t}\right) = r_{jk}^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t} r_{jk}^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_D, t \in \mathcal{T}$$

$$z_k^{it}\left(\mathbf{d}^{t}\right) = z_k^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t} z_k^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}, t \in \mathcal{T}$$

Note that each product is independent of other products by following the assumption of Lim et al. [93]. If the coefficient of the LDR is given, each type of decision is determined as demand unveiled.

We present $\mathrm{P_{LDR}}$ in Appendix C.1. We could obtain the coefficient of the LDR by solving $\mathrm{P_{LDR}}$ considering coefficients as decision variables. We develop the $\mathrm{P_{LDR}}$ based on the Theorem 2 in Lim et al. [93]. $\mathrm{P_{LDR}}$ can be transformed to the linear deterministic model by duality theory [10]. Consequently, the coefficient can be obtained by solving the linear deterministic model with a commercial solver. We present the linear deterministic model transformed from the $\mathrm{P_{LDR}}$ in Appendix C.2.

## 4.5 A decomposition approach (DECOM)

Given cost target $\psi$, three MILP models ($\mathrm{P_{TPA-\nu(0)}}, \mathrm{P_{TPA-\nu(1)}}, \mathrm{P_{STATIC}}$) and one linear programming (LP) model ($\mathrm{P_{LDR}}$) must be solved for applying the TPA. However, the existence of the supply chain of 3PP and the production capacity constraint

increases the complexity of the problem because two supply chains, one for the retailer and the other for the 3PP, should be considered simultaneously. Therefore, it requires a significant computational burden to solve the three MILP models. To alleviate this issue, we develop a DECOM approach which can be regarded as an extended version of the TPA. The key idea of DECOM is to adopt the *artificial variable* $w^{it}$. Let $\boldsymbol{w} = \left(w^{it}, \forall i \in \mathcal{I}, t \in \mathcal{T}\right)$ denote a collection of the artificial variable. The production capacity constraint (4.3) in $\mathrm{P_{DET}}$ is reformulated as the following constraints by introducing decision variables $\boldsymbol{w}$:

$$\sum_{j \in \mathcal{J}_D} q_j^{it} \leq s^{it} w^{it}, \qquad i \in \mathcal{I}, t \in \mathcal{T} \tag{4.60}$$

$$\sum_{j \in \mathcal{J}_F} q_j^{it} \leq s^{it} \left(1 - w^{it}\right), \qquad i \in \mathcal{I}, t \in \mathcal{T} \tag{4.61}$$

$$w^{it} \geq 0, \qquad i \in \mathcal{I}, t \in \mathcal{T} \tag{4.62}$$

There are two advantages to using variables $\boldsymbol{w}$. First, given $\boldsymbol{w}$, the feasible region for variables $q_j^{it}, \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}$ can be reduced. Figure 4.2 presents the feasible region reduced by Constraints (4.60)−(4.62) for three cases: $w^{it} = 0.25, 0.50,$ and 0.75. Second, $\mathrm{P_{TPA-\nu(0)}}$ and $\mathrm{P_{STATIC}}$ can be solved separately for a retailer's supply chain and the supply chain of 3PP. Consequently, these two advantages could significantly reduce the computational burden, and experimental results will be presented in Section 4.6.

### 4.5.1 Phase 1 of DECOM

Phase 1 of DECOM aims to determine the binary decision $\boldsymbol{\delta}$, which is similar to Phase 1 of the TPA. Of special note, we also determine the artificial variable $\boldsymbol{w}$ in Phase 1. We use the $\boldsymbol{\delta}_D, \boldsymbol{\pi}_D$, and $\boldsymbol{\mu}_D$ to denote a collection of variables for

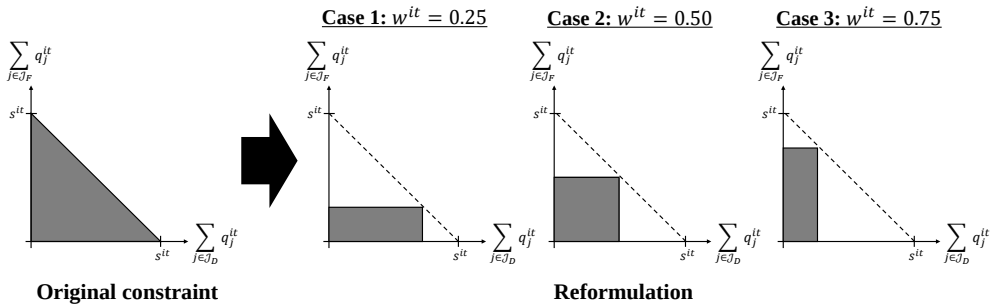



Figure 4.2: Effects of introducing artificial variables $w^{it}$. The shaded area is the feasible region for $q_j^{it}$.

the retailer's supply chain and the $\boldsymbol{\delta}_F, \boldsymbol{\pi}_F$, and $\boldsymbol{\mu}_F$ for the supply chain of 3PP as follows:

$$\boldsymbol{\delta}_D = \left(\delta_j^{it}, \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}\right), \quad \boldsymbol{\delta}_F = \left(\delta_j^{it}, \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}\right),$$

$$\boldsymbol{\pi}_D = \left(q_j^{it}, x_j^{it}, y_k^{it}, u_{l,jj'}^{it}, u_{e,jk}^{it}, \ \forall i \in \mathcal{I}, j \in \mathcal{J}_D, j' \in \mathcal{J}_D, k \in \mathcal{K}_O, t \in \mathcal{T}\right),$$

$$\boldsymbol{\pi}_F = \left(q_j^{it}, x_j^{it}, \ \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}\right),$$

$$\boldsymbol{\mu}_D = \left(v_{\rho,k}^{it}, r_{jk}^{it}, z_k^{it}, \ \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}^-, t \in \mathcal{T}\right), \quad \boldsymbol{\mu}_D = \left(v_{\eta,j}^{it}, , z_{K+1}^{it}, \ \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}\right).$$

Given $\boldsymbol{\delta}_D, \boldsymbol{\delta}_F, \boldsymbol{\pi}_D, \boldsymbol{\pi}_F, \boldsymbol{\mu}_D$, and $\boldsymbol{\mu}_F$, the total cost for retailer's supply chain is defined as

$$\boldsymbol{\Psi}_D^{\dagger}\left(\boldsymbol{\delta}_D, \boldsymbol{\pi}_D, \boldsymbol{\mu}_D\right) =$$

$$\sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}}\Bigg(\sum_{j\in\mathcal{J}_D} S_j^{it}\delta_j^{it} + \sum_{j\in\mathcal{J}_D}\lambda_o^{it}c_o^{ij}q_j^{it} + \sum_{j\in\mathcal{J}_D} h_{x,j}^{it}x_j^{i,t+1} + \sum_{k\in\mathcal{K}_O} h_{y,k}^{it}y_k^{i,t+1} + \sum_{k\in\mathcal{K}^-} p_k^{it}z_k^{it}$$

$$+ \sum_{j\in\mathcal{J}_D}\sum_{j'\in\mathcal{J}_D}\lambda_l^{it}c_l^{jj'}u_{l,jj'}^{it} + \sum_{j\in\mathcal{J}_D}\sum_{k\in\mathcal{K}_O}\lambda_e^{it}c_e^{jk}u_{e,jk}^{it} + \sum_{j\in\mathcal{J}_D}\sum_{k\in\mathcal{K}_D}\lambda_g^{it}c_g^{jk}r_{jk}^{it} + \sum_{k\in\mathcal{K}_O}\rho_k^{it}v_{\rho,k}^{it}\Bigg),$$

and the total cost for the 3PP supply chain is defined as

$$\boldsymbol{\Psi}_F^{\dagger}\left(\boldsymbol{\delta}_F, \boldsymbol{\pi}_F, \boldsymbol{\mu}_F\right) = \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}}\left(\sum_{j\in\mathcal{J}_F} S_j^{it}\delta_j^{it} + \sum_{j\in\mathcal{J}_F}\lambda_o^{it}c_o^{ij}q_j^{it} + \sum_{j\in\mathcal{J}_F} h_{x,j}^{it}x_j^{i,t+1} + p_{K+1}^{it}z_{K+1}^{it} + \sum_{j\in\mathcal{J}_F}\eta_j^{it}v_{\eta,j}^{it}\right).$$

In Phase 1, we first solve the following MILP problem to determine $\boldsymbol{w}$:

$$
\begin{aligned}
&(\mathrm{P}_{\mathrm{DECOM}-\nu(1)})\\
&\qquad \nu(1) = \min \boldsymbol{\Psi}_D^{\dagger}\left(\boldsymbol{\delta}_D, \boldsymbol{\pi}_D, \boldsymbol{\mu}_D\right) + \boldsymbol{\Psi}_F^{\dagger}\left(\boldsymbol{\delta}_F, \boldsymbol{\pi}_F, \boldsymbol{\mu}_F\right)\\
&\qquad \text{s.t.} \quad \sum_{j\in\mathcal{J}_D} q_j^{it} \le s^{it}w^{it}, \qquad i\in\mathcal{I}, t\in\mathcal{T}\\
&\qquad\qquad \sum_{j\in\mathcal{J}_F} q_j^{it} \le s^{it}\left(1-w^{it}\right), \qquad i\in\mathcal{I}, t\in\mathcal{T}\\
&\qquad\qquad \sum_{j\in\mathcal{J}_D} r_{jk}^{it} + z_k^{it} \ge \hat{d}_k^{it} + \bar{\zeta}_k^{it}, \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}, k\in\mathcal{K}_D\\
&\qquad\qquad v_{\rho,k}^{it} + z_k^{it} \ge \hat{d}_k^{it} + \bar{\zeta}_k^{it}, \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}, k\in\mathcal{K}_O\\
&\qquad\qquad \sum_{j\in\mathcal{J}_F} v_{\eta,j}^{it} + z_{K+1}^{it} \ge \hat{d}_{K+1}^{it} + \bar{\zeta}_{K+1}^{it}, \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}\\
&\qquad\qquad w^{it} \ge 0, \qquad i\in\mathcal{I}, t\in\mathcal{T}\\
&\qquad\qquad \text{Constraints } (4.2), (4.4)-(4.7), (4.11)-(4.22)
\end{aligned}
$$

Let $\bar{\boldsymbol{w}}$ denote the optimal solution of $\boldsymbol{w}$. We use $\mathrm{P}_{\mathrm{DECOM}-\nu(1)}$ to determine $\boldsymbol{w}$ because of the following two reasons. First, we utilize $\mathrm{P}_{\mathrm{DECOM}-\nu(1)}$ to obtain the robust solution of $\boldsymbol{w}$. Because $\mathrm{P}_{\mathrm{DECOM}-\nu(1)}$ considers the worst-case scenario of uncertainty with $\gamma = 1$, it is obvious that the robust solution of $\boldsymbol{w}$ could be obtained. Second, because the optimal value $\nu(1)$ of $\mathrm{P}_{\mathrm{DECOM}-\nu(1)}$ is used to get the cost target for applying the TRO approach, it is not mandatory to implement another unnecessary scheme to determine $\boldsymbol{w}$, which could save computational time. Note that the

$\boldsymbol{w}$ is not used for actual decisions (i.e., replenishment, transshipment, allocation, and fulfillment). The $\boldsymbol{w}$ is only used to decompose the proposed problem and reduce computational times.

Let $\bar{\boldsymbol{\delta}}_D^1, \bar{\boldsymbol{\pi}}_D^1, \bar{\boldsymbol{\mu}}_D^1, \bar{\boldsymbol{\delta}}_F^1, \bar{\boldsymbol{\pi}}_F^1$, and, $\bar{\boldsymbol{\mu}}_F^1$ are optimal solutions of Problem $\mathrm{P}_{\mathrm{DECOM}-\nu(1)}$. We define $\nu_D(1) = \boldsymbol{\Psi}_D^\dagger\left(\bar{\boldsymbol{\delta}}_D^1, \bar{\boldsymbol{\pi}}_D^1, \bar{\boldsymbol{\mu}}_D^1\right)$ and $\nu_F(1) = \boldsymbol{\Psi}_F^\dagger\left(\bar{\boldsymbol{\delta}}_F^1, \bar{\boldsymbol{\pi}}_F^1, \bar{\boldsymbol{\mu}}_F^1\right)$, and the sum of $\nu_D(1)$ and $\nu_F(1)$ is equal to $\nu(1)$. Then, we solve the following problem to get value $\nu(0)$ with fixed value $\bar{\boldsymbol{w}}$:

$$
\begin{aligned}
&(\mathrm{P}_{\mathrm{DECOM}-\nu(0)}) \\
&\qquad \nu(0) = \min \boldsymbol{\Psi}_D^\dagger\left(\boldsymbol{\delta}_D, \boldsymbol{\pi}_D, \boldsymbol{\mu}_D\right) + \boldsymbol{\Psi}_F^\dagger\left(\boldsymbol{\delta}_F, \boldsymbol{\pi}_F, \boldsymbol{\mu}_F\right) \\
&\qquad \text{s.t.} \quad \sum_{j \in \mathcal{J}_D} q_j^{it} \leq s^{it}\bar{w}^{it}, \qquad i \in \mathcal{I}, t \in \mathcal{T} \\
&\qquad\qquad \sum_{j \in \mathcal{J}_F} q_j^{it} \leq s^{it}\left(1 - \bar{w}^{it}\right), \qquad i \in \mathcal{I}, t \in \mathcal{T} \\
&\qquad\qquad \sum_{j \in \mathcal{J}_D} r_{jk}^{it} + z_k^{it} \geq \hat{d}_k^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_D \\
&\qquad\qquad v_{\rho,k}^{it} + z_k^{it} \geq \hat{d}_k^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_O \\
&\qquad\qquad \sum_{j \in \mathcal{J}_F} v_{\eta,j}^{it} + z_{K+1}^{it} \geq \hat{d}_{K+1}^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T} \\
&\qquad\qquad \text{Constraints } (4.2), (4.4) - (4.7), (4.11) - (4.22)
\end{aligned}
$$

The first and second constraints use the fixed value $\bar{\boldsymbol{w}}$, which is obtained by solving the $\mathrm{P}_{\mathrm{DECOM}-\nu(1)}$. Let $\bar{\boldsymbol{\delta}}_D^0, \bar{\boldsymbol{\pi}}_D^0, \bar{\boldsymbol{\mu}}_D^0, \bar{\boldsymbol{\delta}}_F^0, \bar{\boldsymbol{\pi}}_F^0$, and, $\bar{\boldsymbol{\mu}}_F^0$ are optimal solutions of Problem $\mathrm{P}_{\mathrm{DECOM}-\nu(0)}$. We have $\nu_D(0) = \boldsymbol{\Psi}_D^\dagger\left(\bar{\boldsymbol{\delta}}_D^0, \bar{\boldsymbol{\pi}}_D^0, \bar{\boldsymbol{\mu}}_D^0\right)$ and $\nu_F(0) = \boldsymbol{\Psi}_F^\dagger\left(\bar{\boldsymbol{\delta}}_F^0, \bar{\boldsymbol{\pi}}_F^0, \bar{\boldsymbol{\mu}}_F^0\right)$, and the sum of $\nu_D(0)$ and $\nu_F(0)$ is equal to $\nu(0)$.

In contrast to the procedure of the TPA, DECOM adopts two cost targets, $\psi_D$ and $\psi_D$, one for the retailer and the other for the 3PP. The $\psi_D$ and $\psi_F$ are

determined with the following two affine functions of $\phi$, respectively:

$$\psi_D(\phi) := (1-\phi)\nu_D(1) + \phi\nu_D(0)$$
$$\psi_F(\phi) := (1-\phi)\nu_F(1) + \phi\nu_F(0).$$

Given $\bar{\boldsymbol{w}}$, the stochastic optimization model $\mathrm{P}_{\mathrm{STOC}}$ can be decomposed into two models, one for the retailer's supply chain and the other for the supply chain of the 3PP. For each model, we could derive two MILP models, $\mathrm{P}_{\mathrm{STATIC-D}}$ and $\mathrm{P}_{\mathrm{STATIC-F}}$, by applying the TRO approach presented in Section 4.4.1. $\mathrm{P}_{\mathrm{STATIC-D}}$ and $\mathrm{P}_{\mathrm{STATIC-F}}$ are formulated for the retailer's and the 3PP supply chains, respectively. In the case of the TPA, the $\gamma^\dagger$, which is for maximizing the adjustable uncertainty set, is the same for the retailer's and the 3PP supply chains. On the other hand, in DECOM, we define $\gamma_D^\dagger$ for the objective value of $\mathrm{P}_{\mathrm{STATIC-D}}$, and $\gamma_F^\dagger$ for the objective value of $\mathrm{P}_{\mathrm{STATIC-F}}$. $\mathrm{P}_{\mathrm{STATIC-D}}$ and $\mathrm{P}_{\mathrm{STATIC-F}}$ are presented as follows:

$$
\begin{aligned}
&(\mathrm{P}_{\mathrm{STATIC-D}}) \\
&\gamma_D^\dagger = \max \quad \gamma \\
&\qquad \text{s.t.} \quad \boldsymbol{\Psi}_D^\dagger\left(\boldsymbol{\delta}, \boldsymbol{\pi}, \boldsymbol{\mu}\right) \le \psi_D \\
&\qquad\qquad q_j^{it} \le \bar{q}_j^i \delta_j^{it}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T} \\
&\qquad\qquad \sum_{j \in \mathcal{J}_D} q_j^{it} \le s^{it}\bar{w}^{it}, \qquad i \in \mathcal{I}, t \in \mathcal{T} \\
&\qquad\qquad \sum_{j \in \mathcal{J}_D} r_{jk}^{it} + z_k^{it} \ge \hat{d}_k^{it} + \gamma\bar{\zeta}_k^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_D \\
&\qquad\qquad v_{\rho,k}^{it} + z_k^{it} \ge \hat{d}_k^{it} + \gamma\bar{\zeta}_k^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_O \\
&\qquad\qquad q_j^{it} \ge 0, \delta_j^{it} \in \{0,1\}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T} \\
&\qquad\qquad x_j^{it} \ge 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}^+
\end{aligned}
$$

$$
\begin{aligned}
& z_k^{it} \geq 0, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}^-, t \in \mathcal{T} \\
& \text{Constraints } (4.5) - (4.7), (4.11), (4.13), (4.16) - (4.19), (4.21) \\
& 0 \leq \gamma \leq 1
\end{aligned}
$$

$(\mathrm{P}_{\mathrm{STATIC-F}})$

$$
\begin{aligned}
\gamma_F^{\dagger} = \max \quad & \gamma \\
\text{s.t.} \quad & \boldsymbol{\Psi}_F^{\dagger}(\boldsymbol{\delta}, \boldsymbol{\pi}, \boldsymbol{\mu}) \leq \psi_F \\
& q_j^{it} \leq \bar{q}_j^i \delta_j^{it}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T} \\
& \sum_{j \in \mathcal{J}_D} q_j^{it} \leq s^{it}(1 - \bar{w}^{it}), \qquad i \in \mathcal{I}, t \in \mathcal{T} \\
& \sum_{j \in \mathcal{J}_F} v_{\eta,j}^{it} + z_{K+1}^{it} \geq \hat{d}_{K+1}^{it} + \gamma \bar{\zeta}_{K+1}^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T} \\
& q_j^{it} \geq 0, \delta_j^{it} \in \{0, 1\}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T} \\
& x_j^{it} \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}^+ \\
& z_{K+1}^{it} \geq 0, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T} \\
& \text{Constraints } (4.4), (4.12), (4.20) \\
& 0 \leq \gamma \leq 1
\end{aligned}
$$

Let $\bar{\boldsymbol{\delta}}_D$ and $\bar{\boldsymbol{\delta}}_F$ be optimal solutions for $\boldsymbol{\delta}_D$ and $\boldsymbol{\delta}_F$ obtained by solving the Problems $\mathrm{P}_{\mathrm{STATIC-D}}$ and $\mathrm{P}_{\mathrm{STATIC-F}}$, respectively. Consequently, the $\bar{\boldsymbol{\delta}}_D$ and $\bar{\boldsymbol{\delta}}_F$ will be used for binary replenishment decisions in Phase 2 of DECOM.

### 4.5.2 Phase 2 of DECOM

The goal of Phase 2 of DECOM is to determine the adjustable continuous variables, which is similar to the goal of Phase 2 of the TPA. However, a key difference between these two approaches is that Phase 2 of DECOM utilizes the solution for the artificial variable $\bar{\boldsymbol{w}}$ obtained in Phase 1. In addition, by using the fixed $\bar{\boldsymbol{w}}$, we

can decompose the $\mathrm{P}_{\mathrm{ARO}}$ into the following two problems $\mathrm{P}_{\mathrm{ARO-D}}$ and $\mathrm{P}_{\mathrm{ARO-F}}$:

$(\mathrm{P}_{\mathrm{ARO-D}})$

$$\begin{aligned}
\min \max_{\mathcal{P}\in\mathcal{F}} \quad & \mathbb{E}_{\mathcal{P}}\left[\boldsymbol{\Psi}_D\left(\boldsymbol{\pi}_D(\tilde{\mathbf{d}}), \boldsymbol{\mu}_D(\tilde{\mathbf{d}})\right)\right] \\
\text{s.t.} \quad & q_j^{it}(\mathbf{d}^{t-1}) \leq \bar{q}_j^i \bar{\delta}_j^{it}, \qquad i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, \qquad \forall \mathbf{d}^{t-1} \in \mathbf{D}^{t-1} \\
& \sum_{j\in\mathcal{J}_D} q_j^{it}(\mathbf{d}^{t-1}) \leq s^{it}\bar{w}^{it}, \qquad i \in \mathcal{I}, t \in \mathcal{T}, \qquad \forall \mathbf{d}^{t-1} \in \mathbf{D}^{t-1} \\
& q_j^{it}(\tilde{\mathbf{d}}^{t-1}) \geq 0, q_j^{it}(\tilde{\mathbf{d}}^{t-1}) \in \mathcal{R}^{t-1}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T} \\
& x_j^{it}(\tilde{\mathbf{d}}^{t-1}) \geq 0, x_j^{it}(\tilde{\mathbf{d}}^{t-1}) \in \mathcal{R}^{t-1}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}^+ \\
& z_k^{it}(\tilde{\mathbf{d}}^{t}) \geq 0, z_k^{it}(\tilde{\mathbf{d}}^{t}) \in \mathcal{R}^{t}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}^-, t \in \mathcal{T} \\
& \text{Constraints } (4.27)-(4.31), (4.33), (4.35), (4.38)-(4.41), (4.43) \qquad \forall \mathbf{d}^t \in \mathbf{D}^t
\end{aligned}$$

$(\mathrm{P}_{\mathrm{ARO-F}})$

$$\begin{aligned}
\min \max_{\mathcal{P}\in\mathcal{F}} \quad & \mathbb{E}_{\mathcal{P}}\left[\boldsymbol{\Psi}_F\left(\boldsymbol{\pi}_F(\tilde{\mathbf{d}}), \boldsymbol{\mu}_F(\tilde{\mathbf{d}})\right)\right] \\
\text{s.t.} \quad & q_j^{it}(\mathbf{d}^{t-1}) \leq \bar{q}_j^i \bar{\delta}_j^{it}, \qquad i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}, \qquad \forall \mathbf{d}^{t-1} \in \mathbf{D}^{t-1} \\
& \sum_{j\in\mathcal{J}_F} q_j^{it}(\mathbf{d}^{t-1}) \leq s^{it}(1-\bar{w}^{it}), \qquad i \in \mathcal{I}, t \in \mathcal{T}, \qquad \forall \mathbf{d}^{t-1} \in \mathbf{D}^{t-1} \\
& q_j^{it}(\tilde{\mathbf{d}}^{t-1}) \geq 0, q_j^{it}(\tilde{\mathbf{d}}^{t-1}) \in \mathcal{R}^{t-1}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T} \\
& x_j^{it}(\tilde{\mathbf{d}}^{t-1}) \geq 0, x_j^{it}(\tilde{\mathbf{d}}^{t-1}) \in \mathcal{R}^{t-1}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}^+ \\
& z_{K+1}^{it}(\tilde{\mathbf{d}}^{t}) \geq 0, z_{K+1}^{it}(\tilde{\mathbf{d}}^{t}) \in \mathcal{R}^{t}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T} \\
& \text{Constraints } (4.26), (4.32), (4.34), (4.42) \qquad \forall \mathbf{d}^t \in \mathbf{D}^t
\end{aligned}$$

where given $\mathbf{d}$

$$\boldsymbol{\Psi}_D\left(\boldsymbol{\pi}_D(\mathbf{d}), \boldsymbol{\mu}_D(\mathbf{d})\right) =$$

$$\sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}}\left(\sum_{j\in\mathcal{J}_D}\lambda_o^{it}c_o^{ij}q_j^{it}(\mathbf{d}^{t-1}) + \sum_{j\in\mathcal{J}_D}h_{x,j}^{it}x_j^{i,t+1}(\mathbf{d}^t) + \sum_{k\in\mathcal{K}_O}h_{y,k}^{it}y_k^{i,t+1}(\mathbf{d}^t) + \sum_{k\in\mathcal{K}^-}p_k^{it}z_k^{it}(\mathbf{d}^t)\right.$$

$$+ \sum_{j \in \mathcal{J}_D} \sum_{j' \in \mathcal{J}_D} \lambda_l^{it} c_l^{jj'} u_{l,jj'}^{it}(\mathbf{d}^{t-1}) + \sum_{j \in \mathcal{J}_D} \sum_{k \in \mathcal{K}_O} \lambda_e^{it} c_e^{jk} u_{e,jk}^{it}(\mathbf{d}^{t-1}) + \sum_{j \in \mathcal{J}_D} \sum_{k \in \mathcal{K}_D} \lambda_g^{it} c_g^{jk} r_{jk}^{it}(\mathbf{d}^{t})$$

$$+ \sum_{k \in \mathcal{K}_O} \rho_k^{it} v_{\rho,k}^{it}(\mathbf{d}^{t}) \Bigg),$$

$$\mathbf{\Psi}_F\left(\boldsymbol{\pi}_F(\mathbf{d}), \boldsymbol{\mu}_F(\mathbf{d})\right) =$$

$$\sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} \left( \sum_{j \in \mathcal{J}_F} \lambda_o^{it} c_o^{ij} q_j^{it}(\mathbf{d}^{t-1}) + \sum_{j \in \mathcal{J}_F} h_{x,j}^{it} x_j^{i,t+1}(\mathbf{d}^{t}) + p_{K+1}^{it} z_{K+1}^{it}(\mathbf{d}^{t}) + \sum_{j \in \mathcal{J}_F} \eta_j^{it} v_{\eta,j}^{it}(\mathbf{d}^{t}) \right).$$

In order to restrict feasible space to linear functions, we also utilize the LDR for each adjustable continuous variable. The LDR for a retailer's supply chain is defined as

$$q_j^{it}\left(\mathbf{d}^{t-1}\right) = q_j^{it,0} + \sum_{\sigma \in \mathcal{K}^-} \sum_{\tau=1}^{t-1} q_j^{it,\sigma\tau} d_\sigma^{i\tau},, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}$$

$$x_j^{it}\left(\mathbf{d}^{t-1}\right) = x_j^{it,0} + \sum_{\sigma \in \mathcal{K}^-} \sum_{\tau=1}^{t-1} x_j^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}^+$$

$$y_k^{it}\left(\mathbf{d}^{t-1}\right) = y_k^{it,0} + \sum_{\sigma \in \mathcal{K}^-} \sum_{\tau=1}^{t-1} y_k^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}^+$$

$$u_{l,jj'}^{it}\left(\mathbf{d}^{t-1}\right) = u_{l,jj'}^{it,0} + \sum_{\sigma \in \mathcal{K}^-} \sum_{\tau=1}^{t-1} u_{l,jj'}^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, j' \in \mathcal{J}_D, t \in \mathcal{T}$$

$$u_{e,jk}^{it}\left(\mathbf{d}^{t-1}\right) = u_{e,jk}^{it,0} + \sum_{\sigma \in \mathcal{K}^-} \sum_{\tau=1}^{t-1} u_{e,jk}^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_O, t \in \mathcal{T}$$

$$v_{\rho,k}^{it}\left(\mathbf{d}^{t}\right) = v_{\rho,k}^{it,0} + \sum_{\sigma \in \mathcal{K}^-} \sum_{\tau=1}^{t} v_{\rho,k}^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}$$

$$r_{jk}^{it}\left(\mathbf{d}^{t}\right) = r_{jk}^{it,0} + \sum_{\sigma \in \mathcal{K}^-} \sum_{\tau=1}^{t} r_{jk}^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_D, t \in \mathcal{T}$$

$$z_k^{it}\left(\mathbf{d}^{t}\right) = z_k^{it,0} + \sum_{\sigma \in \mathcal{K}^-} \sum_{\tau=1}^{t} z_k^{it,\sigma\tau} d_\sigma^{i\tau}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}^-, t \in \mathcal{T}$$

, and for the 3PP supply chain is defined as

$$q_j^{it}\left(\mathbf{d}^{t-1}\right) = q_j^{it,0} + \sum_{\tau=1}^{t-1} q_j^{it,K+1,\tau} d_{K+1}^{i\tau},, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}$$

$$x_j^{it}\left(\mathbf{d}^{t-1}\right) = x_j^{it,0} + \sum_{\tau=1}^{t-1} x_j^{it,K+1,\tau} d_{K+1}^{i\tau}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}^+$$

$$v_{\eta,j}^{it}\left(\mathbf{d}^{t}\right) = v_{\eta,j}^{it,0} + \sum_{\tau=1}^{t} v_{\eta,j}^{it,K+1,\tau} d_{K+1}^{i\tau}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}$$

$$z_{K+1}^{it}\left(\mathbf{d}^{t}\right) = z_{K+1}^{it,0} + \sum_{\tau=1}^{t} z_{K+1}^{it,K+1,\tau} d_{K+1}^{i\tau}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}.$$

Based on the above LDR, we present $\mathrm{P}_{\mathrm{LDR-D}}$ for the retailer's supply chain and $\mathrm{P}_{\mathrm{LDR-F}}$ for the 3PP supply chain to obtain the coefficient of LDR. $\mathrm{P}_{\mathrm{LDR-D}}$ and $\mathrm{P}_{\mathrm{LDR-F}}$ are defined as follows:

($\mathrm{P}_{\mathrm{LDR-D}}$)

$$\begin{aligned}
\min \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}} \Bigg( & \sum_{j\in\mathcal{J}_D} \lambda_o^{it} c_o^{ij} \left( q_j^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t-1} q_j^{it,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) + \sum_{j\in\mathcal{J}_D} h_{x,j}^{it} \left( x_j^{i,t+1,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t} x_j^{i,t+1,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) \\
& + \sum_{k\in\mathcal{K}_O} h_{y,k}^{it} \left( y_k^{i,t+1,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t} y_k^{i,t+1,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) + \sum_{k\in\mathcal{K}^-} p_k^{it} \left( z_k^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t} z_k^{it,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) \\
& + \sum_{j\in\mathcal{J}_D}\sum_{j'\in\mathcal{J}_D} \lambda_l^{it} c_l^{jj'} \left( u_{l,jj'}^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t-1} u_{l,jj'}^{it,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) + \sum_{j\in\mathcal{J}_D}\sum_{k\in\mathcal{K}_O} \lambda_e^{it} c_e^{jk} \left( u_{e,jk}^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t-1} u_{e,jk}^{it,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) \\
& + \sum_{j\in\mathcal{J}_D}\sum_{k\in\mathcal{K}_D} \lambda_g^{it} c_g^{jk} \left( r_{jk}^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t} r_{jk}^{it,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) + \sum_{k\in\mathcal{K}_O} \rho_k^{it} \left( v_{\rho,k}^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t} v_{\rho,k}^{it,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) \Bigg)
\end{aligned}$$

$$\begin{aligned}
\text{s.t.} \quad & q_j^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t-1} q_j^{it,\sigma\tau} d_\sigma^{i\tau} \le \bar{q}_j^i \bar{\delta}_j^{it}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1} \\
& \sum_{j\in\mathcal{J}_D} \left( q_j^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t-1} q_j^{it,\sigma\tau} d_\sigma^{i\tau} \right) \le s^{it}\bar{w}^{it}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1} \\
& \sum_{i\in\mathcal{I}} \Bigg( x_j^{it,0} + q_j^{i,t-L_j^i,0} + \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,j'j}^{it,0} - \sum_{k\in\mathcal{K}_O} u_{e,jk}^{it,0} - \sum_{j\in\mathcal{J}_D\setminus\{j\}} u_{l,jj'}^{it,0} \\
& + \sum_{\sigma\in\mathcal{K}^-} \Bigg( \sum_{\tau=1}^{t-1} \left( x_j^{it,\sigma\tau} + \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,j'j}^{it,\sigma\tau} - \sum_{k\in\mathcal{K}_O} u_{e,jk}^{it,\sigma\tau} - \sum_{j\in\mathcal{J}_D\setminus\{j\}} u_{l,jj'}^{it,\sigma\tau} \right) d_\sigma^{i\tau} \\
& + \sum_{\tau=1}^{t-L_j^i-1} q_j^{i,t-L_j^i,\sigma\tau} d_\sigma^{i\tau} \Bigg) \Bigg) \le \bar{x}_j, \qquad \forall j \in \mathcal{J}_D, t \in \mathcal{T}, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}
\end{aligned}$$

$$x_j^{it,0} + q_j^{i,t-L_j^i,0} - \sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,jj'}^{it,0} + \sum_{\sigma \in \mathcal{K}^-} \left( \sum_{\tau=1}^{t-1} \left( x_j^{it,\sigma\tau} - \sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,jj'}^{it,\sigma\tau} \right) d_\sigma^{i\tau} \right.$$

$$\left. + \sum_{\tau=1}^{t-L_j^i-1} q_j^{i,t-L_j^i,\sigma\tau} d_\sigma^{i\tau} \right) \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}$$

$$\sum_{i \in \mathcal{I}} \left( y_k^{it,0} + \sum_{j \in \mathcal{J}_D} u_{e,jk}^{it,0} + \sum_{\sigma \in \mathcal{K}^-} \sum_{\tau=1}^{t-1} \left( y_k^{it,\sigma\tau} + \sum_{j \in \mathcal{J}_D} u_{e,jk}^{it,\sigma\tau} \right) d_\sigma^{i\tau} \right) \leq \bar{y}_k, \qquad \forall k \in \mathcal{K}_O, t \in \mathcal{T}, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}$$

$$\sum_{j \in \mathcal{J}_D} r_{jk}^{it,0} + z_k^{it,0} = 0, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_D,$$

$$\sum_{j \in \mathcal{J}_D} r_{jk}^{it,\sigma\tau} + z_k^{it,\sigma\tau} = \begin{cases} 1 & \text{if) } \tau = t, \sigma = k \\ \\ 0 & \text{otherwise)} \end{cases}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_D, \sigma \in \mathcal{K}^-, \tau \in \{1, \ldots, t\}$$

$$v_{\rho,k}^{it,0} + z_k^{it,0} = 0, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_O,$$

$$v_{\rho,k}^{it,\sigma\tau} + z_k^{it,\sigma\tau} = \begin{cases} 1 & \text{if) } \tau = t, \sigma = k \\ \\ 0 & \text{otherwise)} \end{cases}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_O, \sigma \in \mathcal{K}^-, \tau \in \{1, \ldots, t\}$$

$$x_j^{i,t+1,0} = x_j^{it,0} + q_j^{i,t-L_j^i,0} +$$

$$\sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,j'j}^{it,0} - \sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,jj'}^{it,0} - \sum_{k \in \mathcal{K}_O} u_{e,jk}^{it,0} - \sum_{k \in \mathcal{K}_D} r_{jk}^{it,0}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}$$

$$x_j^{i,t+1,\sigma\tau} = \begin{cases} x_j^{it,\sigma\tau} + q_j^{i,t-L_j^i,\sigma\tau} + \sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,j'j}^{it,\sigma\tau} - \sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,jj'}^{it,\sigma\tau} & \text{if) } \tau = 1, \ldots, t - L_j^i - 1 \\ \\ - \sum_{k \in \mathcal{K}_O} u_{e,jk}^{it,\sigma\tau} - \sum_{k \in \mathcal{K}_D} r_{jk}^{it,\sigma\tau} & \\ \\ x_j^{it,\sigma\tau} + \sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,j'j}^{it,\sigma\tau} - \sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,jj'}^{it,\sigma\tau} & \text{if) } \tau = t - L_j^i, \ldots, t - 1 \\ \\ - \sum_{k \in \mathcal{K}_O} u_{e,jk}^{it,\sigma\tau} - \sum_{k \in \mathcal{K}_D} r_{jk}^{it,\sigma\tau} & \\ \\ - \sum_{k \in \mathcal{K}_D} r_{jk}^{it,\sigma\tau} & \text{if) } \tau = t \end{cases}$$

$$, \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}^-, \tau \in \{1, \ldots, t\}$$

$$y_k^{i,t+1,0} = y_k^{it,0} + \sum_{j \in \mathcal{J}_D} u_{e,jk}^{it,0} - v_{\rho,k}^{it,0}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}$$

$$y_k^{i,t+1,\sigma\tau} = \begin{cases} y_k^{it,\sigma\tau} + \sum_{j\in\mathcal{J}_D} u_{e,jk}^{it,\sigma\tau} - v_{\rho,k}^{it,\sigma\tau}, & \text{if) } \tau = 1,\dots,t-1 \\ -v_{\rho,k}^{it,\sigma\tau}, & \text{if) } \tau = t \end{cases}$$

$$,\forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}, \sigma \in \mathcal{K}^-, t \in \mathcal{T}, \tau \in \{1,\dots,t\}$$

$$q_j^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t-1} q_j^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}$$

$$x_j^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t-1} x_j^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}^+, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}$$

$$y_k^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t-1} y_k^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}^+, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}$$

$$u_{l,jj'}^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t-1} u_{l,jj'}^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, j, j' \in \mathcal{J}_D, t \in \mathcal{T}, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}$$

$$u_{e,jk}^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t-1} u_{e,jk}^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, k \in \mathcal{K}_O, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}$$

$$v_{\rho,k}^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t} v_{\rho,k}^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}, \mathbf{d}^{t} \in \mathbf{D}^{t}$$

$$r_{jk}^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t} r_{jk}^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_D, t \in \mathcal{T}, \mathbf{d}^{t} \in \mathbf{D}^{t}$$

$$z_k^{it,0} + \sum_{\sigma\in\mathcal{K}^-}\sum_{\tau=1}^{t} z_k^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}^-, \mathbf{d}^{t} \in \mathbf{D}^{t}$$

$$q_j^{it,0}, q_j^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}^-, \tau \in \{1,\dots,t-1\}$$

$$x_j^{it,0}, x_j^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}^+, \sigma \in \mathcal{K}^-, \tau \in \{1,\dots,t-1\}$$

$$y_k^{it,0}, y_k^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}^+, \sigma \in \mathcal{K}^-, \tau \in \{1,\dots,t-1\}$$

$$u_{l,jj'}^{it,0}, u_{l,jj'}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, j' \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}^-, \tau \in \{1,\dots,t-1\}$$

$$v_{\rho,k}^{it,0}, v_{\rho,k}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}, \sigma \in \mathcal{K}^-, \tau \in \{1,\dots,t\}$$

$$r_{jk}^{it,0}, r_{jk}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}^-, \tau \in \{1,\dots,t\}$$

$$z_k^{it,0}, z_k^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}^-, t \in \mathcal{T}, \sigma \in \mathcal{K}^-, \tau \in \{1,\dots,t\}$$

$(\mathrm{P}_{\mathrm{LDR-F}})$

$$\min \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}}\left(\sum_{j\in\mathcal{J}_F}\lambda_o^{it}c_o^{ij}\left(q_j^{it,0}+\sum_{\tau=1}^{t-1}q_j^{it,K+1,\tau}\hat{d}_{K+1}^{i\tau}\right)+\sum_{j\in\mathcal{J}_F}h_{x,j}^{it}\left(x_j^{i,t+1,0}+\sum_{\tau=1}^{t}x_j^{i,t+1,K+1,\tau}\hat{d}_{K+1}^{i\tau}\right)\right.$$
$$\left.+p_{K+1}^{it}\left(z_{K+1}^{it,0}+\sum_{\tau=1}^{t}z_{K+1}^{it,K+1,\tau}\hat{d}_{K+1}^{i\tau}\right)+\sum_{j\in\mathcal{J}_F}\eta_j^{it}\left(v_{\eta,j}^{it,0}+\sum_{\tau=1}^{t}v_{\eta,j}^{it,K+1,\tau}\hat{d}_{K+1}^{i\tau}\right)\right)$$

$$\text{s.t.}\quad q_j^{it,0}+\sum_{\tau=1}^{t-1}q_j^{it,K+1,\tau}d_{K+1}^{i\tau}\le\bar{q}_j^i\bar{\delta}_j^{it},\qquad\forall i\in\mathcal{I},j\in\mathcal{J}_F,t\in\mathcal{T},\mathbf{d}^{t-1}\in\mathbf{D}^{t-1}$$

$$\sum_{j\in\mathcal{J}_F}\left(q_j^{it,0}+\sum_{\tau=1}^{t-1}q_j^{it,K+1,\tau}d_{K+1}^{i\tau}\right)\le s^{it}(1-\bar{w}^{it}),\qquad\forall i\in\mathcal{I},t\in\mathcal{T},\mathbf{d}^{t-1}\in\mathbf{D}^{t-1}$$

$$\sum_{i\in\mathcal{I}}\left(x_j^{it,0}+q_j^{i,t-L_j^i,0}+\left(\sum_{\tau=1}^{t-1}x_j^{it,K+1,\tau}d_{K+1}^{i\tau}+\sum_{\tau=1}^{t-L_j^i-1}q_j^{i,t-L_j^i,K+1,\tau}d_{K+1}^{i\tau}\right)\right)\le\bar{x}_j,$$

$$\forall j\in\mathcal{J}_F,t\in\mathcal{T},\mathbf{d}^{t-1}\in\mathbf{D}^{t-1}$$

$$\sum_{j\in\mathcal{J}_F}v_{\eta,j}^{it,0}+z_{K+1}^{it,0}=0,\qquad\forall i\in\mathcal{I},t\in\mathcal{T},$$

$$\sum_{j\in\mathcal{J}_F}v_{\eta,j}^{it,K+1,\tau}+z_{K+1}^{it,K+1,\tau}=\begin{cases}1 & \text{if) } \tau=t\\ 0 & \text{otherwise)}\end{cases},\qquad\forall i\in\mathcal{I},t\in\mathcal{T},\tau\in\{1,\dots,t\}$$

$$x_j^{i,t+1,0}=x_j^{it,0}+q_j^{i,t-L_j^i,0}-v_{\eta,j}^{it,0},\qquad\forall i\in\mathcal{I},j\in\mathcal{J}_F,t\in\mathcal{T}$$

$$x_j^{i,t+1,K+1,\tau}=\begin{cases}x_j^{it,K+1,\tau}+q_j^{i,t-L_j^i,K+1,\tau}-v_{\eta,j}^{it,K+1,\tau}, & \text{if) } \tau=1,\dots,t-L_j^i-1\\ x_j^{it,K+1,\tau}-v_{\eta,j}^{it,K+1,\tau}, & \text{if) } \tau=t-L_j^i,\dots,t-1\\ -v_{\eta,j}^{it,K+1,\tau}, & \text{if) } \tau=t\end{cases}$$

$$,\forall i\in\mathcal{I},j\in\mathcal{J}_F,t\in\mathcal{T},\tau\in\{1,\dots,t\}$$

$$q_j^{it,0}+\sum_{\tau=1}^{t-1}q_j^{it,K+1,\tau}d_{K+1}^{i\tau}\ge 0,\qquad\forall i\in\mathcal{I},j\in\mathcal{J}_F,t\in\mathcal{T},\mathbf{d}^{t-1}\in\mathbf{D}^{t-1}$$

$$x_j^{it,0}+\sum_{\tau=1}^{t-1}x_j^{it,K+1,\tau}d_{K+1}^{i\tau}\ge 0,\qquad\forall i\in\mathcal{I},j\in\mathcal{J}_F,t\in\mathcal{T}^+,\mathbf{d}^{t-1}\in\mathbf{D}^{t-1}$$

$$v_{\eta,j}^{it,0}+\sum_{\tau=1}^{t}v_{\eta,j}^{it,K+1,\tau}d_{K+1}^{i\tau}\ge 0,\qquad\forall i\in\mathcal{I},j\in\mathcal{J}_F,t\in\mathcal{T},\mathbf{d}^{t}\in\mathbf{D}^{t}$$

$$z_{K+1}^{it,0}+\sum_{\tau=1}^{t}z_{K+1}^{it,K+1,\tau}d_{K+1}^{i\tau}\ge 0,\qquad\forall i\in\mathcal{I},t\in\mathcal{T},\mathbf{d}^{t}\in\mathbf{D}^{t}$$

$$q_j^{it,0}, q_j^{it,K+1,\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}, \tau \in \{1, \ldots, t-1\}$$

$$x_j^{it,0}, x_j^{it,K+1,\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}^+, \tau \in \{1, \ldots, t-1\}$$

$$v_{\eta,j}^{it,0}, v_{\eta,j}^{it,K+1,\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}, \tau \in \{1, \ldots, t\}$$

$$z_{K+1}^{it,0}, z_{K+1}^{it,K+1,\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, \tau \in \{1, \ldots, t\}$$

By following the same logic outlined in Section 4.4.2 and in Appendix C.2, $\mathrm{P}_{\mathrm{LDR-D}}$ and $\mathrm{P}_{\mathrm{LDR-F}}$ also can be reformulated to the linear deterministic model using the duality theory. In summary, we must solve four MILP models (i.e., $\mathrm{P}_{\mathrm{DECOM}-\nu(1)}$, $\mathrm{P}_{\mathrm{DECOM}-\nu(0)}$, $\mathrm{P}_{\mathrm{STATIC-D}}$, and $\mathrm{P}_{\mathrm{STATIC-F}}$) and two LP models (i.e., $\mathrm{P}_{\mathrm{LDR-D}}$ and $\mathrm{P}_{\mathrm{LDR-F}}$) to implement the DECOM approach. Figure 4.3 presents frameworks of the TPA and DECOM.

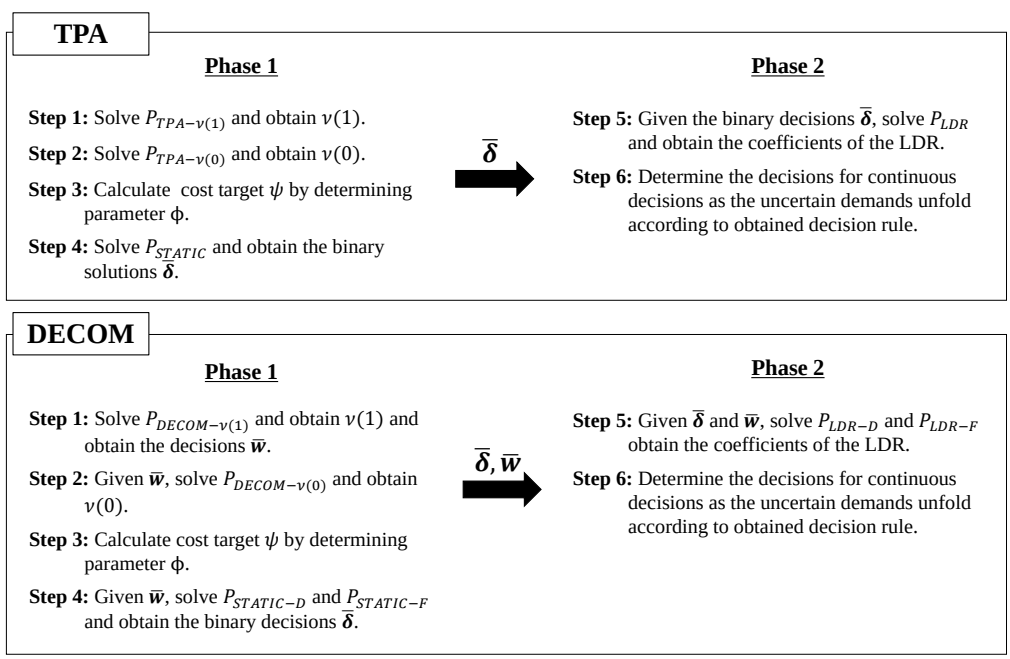



Figure 4.3: Frameworks of TPA and DECOM.

## 4.6 Computational experiments

In this section, we implement three types of computational experiments. In Section 4.6.1, we evaluate the performance of DECOM with respect to various demand distributions. We compare our developed approach with the two benchmark algorithms: TPA and an alternative two-phase approach (DTPA). The DTPA determines the adjustable binary variables $\boldsymbol{\delta}(\tilde{\mathbf{d}})$ with the static rule by solving the EVP, i.e., the deterministic model $\mathrm{P}_{\mathrm{DET}}$ with mean demands [93]. On the other hand, the adjustable continuous variables are determined by applying Phase 2 of the TPA. In Section 4.6.2, we examine the advantages of DECOM in terms of computational efficiency for large-scale instances. In Section 4.6.3, we compare the performance of DECOM and TPA by varying the production capacity. In addition, a cost analysis is also performed with different values of the target coefficient $\phi$. In Section 4.6.4, we present several managerial insights on the basis of the experimental results.

All the experiments were conducted on a PC with an AMD Ryzen 2700X 7-Core CP, 3.60GHz processor, and 16GB of RAM with a Windows 10 64-bit system. In addition, every test instance is generated using Python 3.8 with the libraries *SciPy* and *Numpy*. The DTPA, TPA, and DECOM were developed with FICO Xpress 8.6, and we solved every model by utilizing the Xpress-Optimizer version 33.01.02 with its default parameter settings. In addition, we set the integrality gap tolerance in Xpress to one percent by following the setting of Lim et al. [93].

We benchmark Jiu [72], the most relevant models to our study, to determine constant parameters in the mathematical model. We generated parameters randomly according to the uniform distributions in Table 4.1. The replenishment lead time $L_j^i$ was generated by the discrete uniform distribution. The continuous uniform distri-

bution was used to determine the rest of the parameters. The locations of logistics centers and offline stores were uniformly distributed over the pre-specified size of the XY plane. We determined $c_o^{ij}, c_l^{jj'}, c_e^{jk}$, and $c_g^{jk}$ based on the Euclidean distance between each location.

We assumed that the sum of every demand for item $i$ at period $t$ falls in $[80, 120]$ (i.e., $\sum_{k\in\mathcal{K}} d_k^{it} \in [80, 120]\,,\ \forall i \in \mathcal{I}, t \in \mathcal{T}$). To determine each demand $k \in \mathcal{K}$, we define the share of each distribution channel for $\sum_{k\in\mathcal{K}} d_k^{it}$ as: (1) $\alpha_1$ for the retailer's offline channel, (2) $\alpha_2$ for the retailer's online channel, and (3) $\alpha_3$ for the 3PP channel. We set $\alpha_1 = 0.2, \alpha_2 = 0.3$, and $\alpha_3 = 0.5$. Each channel's demand is generated by the assumed demand distributions, which fall in the corresponding support set represented in Table 4.2. The mean demand $\hat{d}_k^{it}$ is determined according to the assumed demand distribution. Even though we assume the demand distribution to generate random demand, every algorithm is implemented without any knowledge about the demand distribution.

Table 4.1: Ranges of the parameters.

| $S_j^{it}$ | $h_{x,j}^{it}$ | $h_{y,k}^{it}$ | $\lambda_o^{it}$ & $\lambda_e^{it}$ | $\lambda_l^{it}$ | $\lambda_g^{it}$ | $p_k^{it}$ | $\rho_k^{it}$ & $\eta_k^{it}$ | $L_j^i$ |
|---|---|---|---|---|---|---|---|---|
| $\mathcal{U}\,(50, 80)$ | $\mathcal{U}\,(0.2, 0.5)$ | $\mathcal{U}\,(0.3, 0.6)$ | $\mathcal{U}\,(0.05, 0.1)$ | $\mathcal{U}\,(0.02, 0.03)$ | $\mathcal{U}\,(0.08, 0.13)$ | $\mathcal{U}\,(60, 80)$ | $\mathcal{U}\,(2, 3)$ | $\mathcal{U}\,\{0, 1\}$ |

Table 4.2: Support set of each channel for item $i$ and period $t$

| Retailer's offline channel ($d_k^{it}$, $\forall k \in K_O$) | Retailer's online channel ($d_k^{it}$, $\forall k \in K_D$) | 3PP's online channel ($d_{K+1}^{it}$) |
|---|---|---|
| $\left[\frac{\alpha_1}{K_O} \times \sum_{k\in\mathcal{K}} \underline{d_k^{it}},\ \frac{\alpha_1}{K_O} \times \sum_{k\in\mathcal{K}} \bar{d_k^{it}}\right]$ | $\left[\frac{\alpha_2}{K_D} \times \sum_{k\in\mathcal{K}} \underline{d_k^{it}},\ \frac{\alpha_2}{K_D} \times \sum_{k\in\mathcal{K}} \bar{d_k^{it}}\right]$ | $\left[\alpha_3 \times \sum_{k\in\mathcal{K}} \underline{d_k^{it}},\ \alpha_3 \times \sum_{k\in\mathcal{K}} \bar{d_k^{it}}\right]$ |

In order to analyze the effects of the production capacity constraint, we define

the following affine function of $\xi$ to determine the $s^{it}$:

$$s^{it}(\xi) := \xi \times \sum_{k \in \mathcal{K}} \hat{d}_k^{it} + (1 - \xi) \times 2 \sum_{k \in \mathcal{K}} \bar{d}_k^{it}.$$

where $0 \leq \xi \leq 1$. According to the above affine function, the production capacity becomes insufficient as the $\xi$ is close to one. Otherwise, there is sufficient production capacity when the $\xi$ is close to zero.

### 4.6.1 Experiment 1: Performance analysis in small problems under symmetric and asymmetric demand distributions

Experiment 1 is conducted for the following two purposes. First, we validate the obtained decision rule through Monte Carlo (MC) simulation. Every MC simulation is implemented with 500 samples. Second, we evaluate our approach for symmetric and asymmetric demand distributions. We utilize the beta distribution by referring to Jiu [72]. In this section, we set $I = 3,\ K_O = 3,\ K_D = 3,\ J_D = 2,\ J_F = 2$, and conduct experiments on a $50 \times 50$ XY plane. Also, we set $\xi = 0$ to assume sufficient production capacity. We have tested on this setup three different planning horizons: $T = 4$, 7, and 10. Furthermore, we define the set of candidate parameters $\Phi$ to find the best cost target. We use the notation $\phi^*$ to denote the target coefficient, which shows the best performance. We consider six candidate values for $\phi$ as $\Phi = \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$. Note that we use notation '$a \sim b$' to indicate that multiple values of $\phi^* \in \Phi$ between $a$ and $b$ show the same best performances (i.e., $a \leq \phi^* \leq b$).

First, we have conducted experiments on three types of symmetric distribution, $Beta(0.3, 0.3)$, $Beta(1, 1)$, and $Beta(4, 4)$, and experimental results are reported in Table 4.3. We provide shapes of symmetric and asymmetric distributions in

Appendix C.4. In Table 4.3, "LDR" means the objective value of $\mathrm{P_{LDR}}$ with the fixed order cost $S_j^{it}\delta_j^{it}$ for the TPA, and the sum of objective values of $\mathrm{P_{LDR-D}}$ and $\mathrm{P_{LDR-F}}$ with the fixed order cost for DECOM. The "SIM" indicates the expected total cost implemented by MC simulation utilizing the obtained decision rule, and the "Std" is the standard deviation of the total cost for 500 samples. The "CPU(s)" means the computation times in seconds. We adopt the *expected value of perfect information* (EVPI) to evaluate the solution quality of each algorithm. To derive the EVPI, we solve the deterministic model $\mathrm{P_{DET}}$ under the perfect information setting (i.e., the deterministic demand setting). We use the "Gap(%)" to measure the solution quality, which is calculated by $(\mathrm{SIM} - \mathrm{EVPI}) \times 100/\mathrm{EVPI}$.

Every experimental result of the TPA and DECOM was reported by adopting the best target coefficient $\phi^*$. The values of LDR and SIM were indifferent, which meant that the obtained decision rule achieved our goal (i.e., minimizing the expected total cost). In terms of solution quality, the Gap of the TPA and DECOM was smaller than 10 percent, except for a result for $Beta(0.3, 0.3)$ with $T = 4$. However, the Gap of the DTPA was bigger than 20 percent, except for a result for $Beta(1, 1)$ with $T = 7$. As shown in Figure 4.4, the total cost of the TPA and DECOM was similar and significantly lower than the total cost of the DTPA. Also, the standard deviation of the TPA and DECOM was relatively small compared to that of the DTPA. For symmetric distributions, there is a tendency for the best solutions of the TPA and DECOM to be derived when the value of $\phi^*$ is small. This tendency meant that conservative binary decisions were necessary when the demand distribution was symmetric. Concerning CPU(s), it takes the least time to implement the DTPA because the TPA and DECOM were implemented for $|\Phi|$ times to find the best

target coefficient $\phi^*$.

Table 4.3: Experimental results on symmetric demand distributions.

| | | $Beta(0.3, 0.3)$ | | | $Beta(1, 1)$ | | | $Beta(4, 4)$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $T = 4$ | $T = 7$ | $T = 10$ | $T = 4$ | $T = 7$ | $T = 10$ | $T = 4$ | $T = 7$ | $T = 10$ |
| DTPA | LDR($\times 10^2$) | 73.01 | 149.65 | 174.94 | 64.98 | 121.33 | 203.46 | 69.66 | 119.22 | 168.13 |
| | SIM($\times 10^2$) | 72.86 | 149.66 | 174.57 | 65.01 | 121.49 | 203.56 | 69.60 | 119.27 | 168.12 |
| | Gap(%) | 24.84 | 48.93 | 24.46 | 22.20 | 19.55 | 44.59 | 26.04 | 30.97 | 24.16 |
| | Std($\times 10^2$) | 3.94 | 8.45 | 7.79 | 4.05 | 4.39 | 6.17 | 1.99 | 2.61 | 2.92 |
| | CPU(s) | 1.95 | 4.98 | 10.89 | 1.94 | 6.97 | 15.50 | 1.84 | 7.10 | 13.00 |
| TPA | LDR($\times 10^2$) | 64.64 | 108.84 | 150.50 | 58.44 | 108.91 | 151.80 | 60.63 | 99.19 | 145.57 |
| | SIM($\times 10^2$) | 64.67 | 108.86 | 150.48 | 58.44 | 108.97 | 151.81 | 60.63 | 99.19 | 145.58 |
| | Gap(%) | 10.80 | 8.33 | 7.28 | 9.85 | 7.24 | 7.83 | 9.79 | 8.92 | 7.51 |
| | Std($\times 10^2$) | 0.86 | 1.22 | 1.32 | 0.60 | 0.92 | 1.04 | 0.29 | 0.48 | 0.57 |
| | CPU(s) | 10.87 | 27.75 | 75.17 | 8.73 | 47.84 | 112.10 | 8.60 | 35.03 | 104.51 |
| | $\phi^*$ | 0.0 | 0.0 | 0.0 | 0.2∼0.8 | 0.0 | 0.2 | 0.0 | 0.4 | 0.2 |
| DECOM | LDR($\times 10^2$) | 64.78 | 108.26 | 150.70 | 58.44 | 108.83 | 151.74 | 59.90 | 99.30 | 146.42 |
| | SIM($\times 10^2$) | 64.83 | 108.27 | 150.66 | 58.44 | 108.89 | 151.75 | 59.90 | 99.31 | 146.43 |
| | Gap(%) | 11.08 | 7.74 | 7.41 | 9.85 | 7.16 | 7.78 | 8.47 | 9.05 | 8.14 |
| | Std($\times 10^2$) | 0.83 | 1.22 | 1.30 | 0.60 | 0.92 | 1.06 | 0.34 | 0.49 | 0.56 |
| | CPU(s) | 6.42 | 19.12 | 41.33 | 6.25 | 26.91 | 61.87 | 7.29 | 16.09 | 46.91 |
| | $\phi^*$ | 0.0 | 0.0 | 0.2 | 0.0∼0.8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2∼0.6 |
| EVPI | ($\times 10^2$) | 58.36 | 100.49 | 140.27 | 53.20 | 101.62 | 140.79 | 55.22 | 91.07 | 135.41 |

Second, we have conducted experiments on four types of asymmetric distribution, $Beta(2, 5)$, $Beta(5, 2)$, $Beta(1, 6)$, and $Beta(6, 1)$, and the experimental results were reported in Table C.1 in the Appendix C.3. As in the case of symmetric distributions, the values of LDR and SIM were indifferent when the demand distributions were asymmetric. However, when the beta distributions were skewed to the right ($Beta(a, b),\ a < b$), the Gap was bigger than 10 percent. The binary decisions with the static rule $\boldsymbol{\delta}$ could be too conservative for the beta distribution with $a < b$ because the realized demand was usually smaller than the mean value. Also, because the realized demand was relatively small, the $\phi^*$ value was high compared to the symmetric distributions. On the other hand, when the beta distributions were skewed to the left ($Beta(a, b),\ b < a$), the GAP was smaller than 10 percent

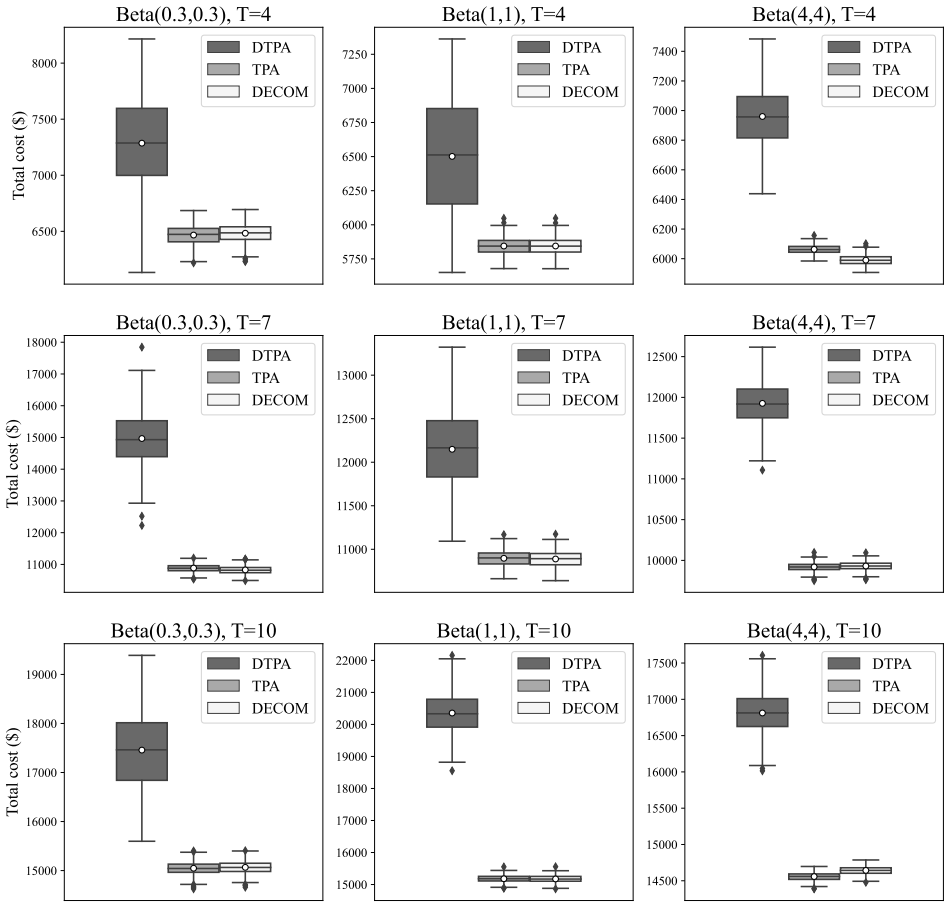



Figure 4.4: Box plots for the total cost for 500 samples for every algorithm.

except for a result implemented by DECOM for $Beta(5, 2)$ with $T = 10$. Because the realized demand was usually bigger than the mean value, there was no doubt that robust solutions were necessary; thus, the $\phi^*$ value was small.

Figure 4.5 is presented to compare the performance of the TPA and DECOM in terms of solution quality (Gap) and computational efficiency (CPU(s)). Among 21 results of experiments (9 for the symmetric distribution and 12 for the asymmetric distribution), the number of wins of DECOM and TPA was the same regarding the

Gap. However, for CPU(s), DECOM outperformed TPA except for one result.

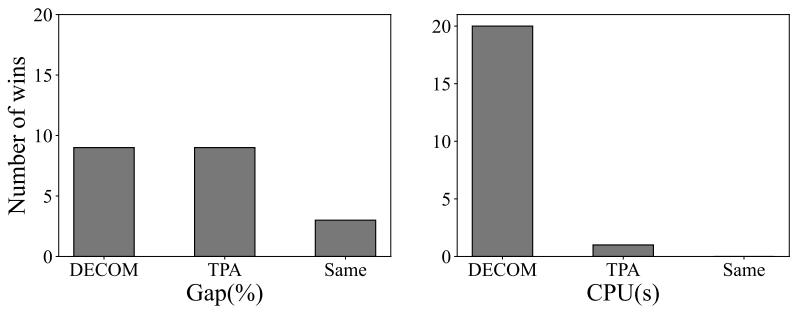



Figure 4.5: Comparison of performance between TPA and DECOM in terms of Gap and CPU(s).

### 4.6.2 Experiment 2: Computational efficiency of DECOM in large-scale problems

In this section, we have conducted several experiments to examine the computational efficiency of DECOM in large-scale problems. For every experiment, we fix the value of $T$, $K_O$, and $K_D$ as 10, 5, and 5, respectively. In addition, we vary with the value of $I, J_D$, and $J_F$ to change the problem scales, in which the $(I, J_D, J_F)$ vary from $(3, 3, 3)$ to $(10, 10, 10)$. We set the size of the XY plane as $100 \times 100$, and we assume that the demand distribution follows $Beta(1, 1)$. Based on the experimental result for $Beta(1, 1)$ in Table 4.3, we set $\phi = 0.0$ for DECOM and TPA. Furthermore, we set $\xi = 0$ following the setting of Experiment 1. When solving the MILP models, we terminate the commercial solver if the time limit, i.e., 3600 seconds, is reached and output the feasible solution obtained so far.

We present experimental results for large-scale problems in Table 4.4, which presents Gap, Std, CPU(s), and EVPI. We keep in mind that $\mathrm{P_{DET}}$ is a MILP model; thus, significant computational power is necessary to solve it 500 times to

obtain the EVPI in large-scale problems. Therefore, we utilize the "alternative" EVPI.

In the alternative EVPI, we first obtain the optimal binary solution $\bar{\boldsymbol{\delta}}$ by using Phase 1 of DECOM. Then, we fix the binary variable with the value $\bar{\boldsymbol{\delta}}$ to make $\mathrm{P}_{\mathrm{DET}}$ as an LP model. Therefore, $\mathrm{P}_{\mathrm{DET}}$ can be solved 500 times with perfect information within a reasonable time. To avoid confusion, the obtained value from the alternative EVPI is also indicated by the term "EVPI" in Table 4.4. The DTPA has the largest value for Gap and Std compared to other approaches, which meant the solution quality of the DTPA was poor.

In particular, we present the following five types of CPU(s): "$\mathrm{P}_{\nu(1)}$", "$\mathrm{P}_{\nu(0)}$", "$\mathrm{P}_{\mathrm{STATIC}}$", "Phase 1", and "Phase 2". "$\mathrm{P}_{\nu(1)}$" means the computation time to solve the $\mathrm{P}_{\mathrm{TPA}-\nu(1)}$ for the TPA, and $\mathrm{P}_{\mathrm{DECOM}-\nu(1)}$ for the DECOM. "$\mathrm{P}_{\nu(0)}$" means the computation time to solve the $\mathrm{P}_{\mathrm{TPA}-\nu(0)}$ for the TPA, and $\mathrm{P}_{\mathrm{DECOM}-\nu(0)}$ for the DECOM. "$\mathrm{P}_{\mathrm{STATIC}}$" means the sum of computation time to solve both the $\mathrm{P}_{\mathrm{STATIC-D}}$ and the $\mathrm{P}_{\mathrm{STATIC-F}}$ for the DECOM. "Phase 1" is computed by summing values in $\mathrm{P}_{\nu(1)}, \mathrm{P}_{\nu(0)}$, and $\mathrm{P}_{\mathrm{STATIC}}$ for the TPA and DECOM. On the other hand, for the DTPA, "Phase 1" is the computation time to solve the $\mathrm{P}_{\mathrm{DET}}$ with the mean demand. "Phase 2" is the computation time to solve both the $\mathrm{P}_{\mathrm{LDR-D}}$ and the $\mathrm{P}_{\mathrm{LDR-F}}$ for DECOM, and the $\mathrm{P}_{\mathrm{LDR}}$ for the TPA and DTPA. Finally, "Total" can be computed by summing values in "Phase 1" and "Phase 2", which is the total computation time to implement each approach.

Figure 4.6 represents the Total, Phase 1, and Phase 2 CPU(s) for three approaches. In terms of Total CPU(s), it required the shortest time to implement the DECOM compared to the TPA and DTPA except for the case of $(I, J_D, J_F) =$

$(8, 8, 8)$. On the contrary, it took the highest computation time to implement the TPA, except for $(I, J_D, J_F) = (3, 3, 3)$. The DTPA could finish Phase 1 within a relatively short computation time compared to the TPA and DECOM. The DTPA and TPA required similar computation times to conduct Phase 2. However, DECOM required less of a computational burden compared to the DTPA and TPA to conduct Phase 2.

Figure 4.7 depicts the CPU(s) of $\mathrm{P}_{\nu(1)}$,$\mathrm{P}_{\nu(0)}$, and $\mathrm{P}_{\mathrm{STATIC}}$. Because these three procedures did not require implementing the DTPA, only DECOM and TPA are represented in Figure 4.7. For $\mathrm{P}_{\nu(1)}$, the performance of the TPA and DECOM was indifferent. However, for $\mathrm{P}_{\nu(0)}$, DECOM requires a much shorter time to solve the problem than TPA. Specifically, we could observe that the TPA could not solve the problem until the time limit when $(I, J_D, J_F) = (10, 10, 10)$. In contrast, DECOM required only about 12 seconds to solve the problem in the same experimental setting. The DECOM could finish the procedure for $\mathrm{P}_{\nu(0)}$ within a short time because the feasible region was substantially reduced by fixing the value for $\bar{\boldsymbol{w}}$. Furthermore, DECOM also had high computational efficiency for $\mathrm{P}_{\mathrm{STATIC}}$. When $I, J_D, J_F \geq 6$, the TPA could not solve Problem $\mathrm{P}_{\mathrm{STATIC}}$ until the time limit. However, DECOM could solve both Problems $\mathrm{P}_{\mathrm{STATIC-D}}$ and $\mathrm{P}_{\mathrm{STATIC-F}}$ within the time limit except for $(I, J_D, J_F) = (10, 10, 10)$. For $(I, J_D, J_F) = (10, 10, 10)$, DECOM could not solve Problem $\mathrm{P}_{\mathrm{STATIC-D}}$ within the time limit; but, Problem $\mathrm{P}_{\mathrm{STATIC-F}}$ could be solved in less than a second.

It can be clearly seen from the above experiments that DECOM could solve the large-scale problem within a reasonable time and derive promising solutions. Even though the TPA could also derive high-quality solutions, it required a significant

computational burden to solve the large-scale problem.

Table 4.4: Experimental results on large size problems.

| | | | $(I, J_D, J_F)$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | $(3,3,3)$ | $(4,4,4)$ | $(5,5,5)$ | $(6,6,6)$ | $(7,7,7)$ | $(8,8,8)$ | $(9,9,9)$ | $(10,10,10)$ |
| DTPA | Gap(%) | | 11.64 | 22.94 | 27.31 | 17.98 | 36.70 | 19.47 | 39.70 | 29.53 |
| | Std($\times 10^2$) | | 2.04 | 5.25 | 8.60 | 3.18 | 13.11 | 5.47 | 12.26 | 5.60 |
| | CPU(s) | Phase 1 | 0.91 | 0.93 | 3.36 | 36.40 | 390.09 | 121.99 | 173.57 | 3600* |
| | | Phase 2 | 84.05 | 143.85 | 279.07 | 473.40 | 1210.38 | 1667.72 | 3515.68 | 7501.68 |
| | | Total | 84.96 | 144.78 | 282.43 | 509.79 | 1600.47 | 1789.71 | 3689.24 | 11101.68 |
| TPA | Gap(%) | | 6.47 | 5.12 | 5.95 | 6.04 | 13.94 | 6.78 | 6.65 | 7.37 |
| | Std($\times 10^2$) | | 1.31 | 1.33 | 1.72 | 1.42 | 8.50 | 1.49 | 1.47 | 1.68 |
| | CPU(s) | $P_{\nu(1)}$ | 0.89 | 1.14 | 1.52 | 19.70 | 166.59 | 347.44 | 562.47 | 3600* |
| | | $P_{\nu(0)}$ | 1.26 | 0.90 | 2.68 | 68.92 | 3541.60 | 239.68 | 300.31 | 3600* |
| | | $P_{\text{STATIC}}$ | 2.00 | 8.75 | 19.16 | 3600* | 3600* | 3600* | 3600* | 3600* |
| | | Phase 1 | 4.14 | 10.79 | 23.36 | 3688.63 | 7308.19 | 4187.12 | 4462.77 | 10800.00 |
| | | Phase 2 | 80.40 | 135.10 | 318.40 | 447.90 | 1035.43 | 1544.60 | 3772.61 | 6629.06 |
| | | Total | 84.54 | 145.89 | 341.76 | 4136.52 | 8343.63 | 5731.72 | 8235.38 | 17429.06 |
| DECOM | Gap(%) | | 6.59 | 5.03 | 4.75 | 6.20 | 12.50 | 5.96 | 7.51 | 6.45 |
| | Std($\times 10^2$) | | 1.29 | 1.34 | 1.72 | 1.39 | 7.32 | 1.45 | 3.42 | 1.66 |
| | CPU(s) | $P_{\nu(1)}$ | 1.01 | 1.17 | 1.53 | 20.01 | 140.46 | 133.47 | 573.35 | 3600* |
| | | $P_{\nu(0)}$ | 0.42 | 0.26 | 0.79 | 2.26 | 5.35 | 2.85 | 8.50 | 12.71 |
| | | $P_{\text{STATIC}}$ | 1.25 | 1.48 | 2.45 | 52.31 | 122.13 | 775.11 | 43.43 | 3600* + 0.44[a] |
| | | Phase 1 | 2.67 | 2.92 | 4.77 | 74.58 | 267.94 | 911.43 | 625.28 | 7212.71 |
| | | Phase 2 | 51.42 | 86.26 | 183.98 | 265.70 | 659.95 | 901.27 | 3671.85 | 2862.29 |
| | | Total | 54.09 | 89.18 | 188.75 | 340.28 | 927.88 | 1812.71 | 4297.13 | 10075.00 |
| EVPI | ($\times 10^2$) | | 180.43 | 208.28 | 293.84 | 284.69 | 403.73 | 338.81 | 378.90 | 472.49 |

* Time limit was reached.

[a] $P_{\text{STATIC-D}}$ could not be solved within the time limit; however, it required 0.44 seconds to solve the $P_{\text{STATIC-F}}$.

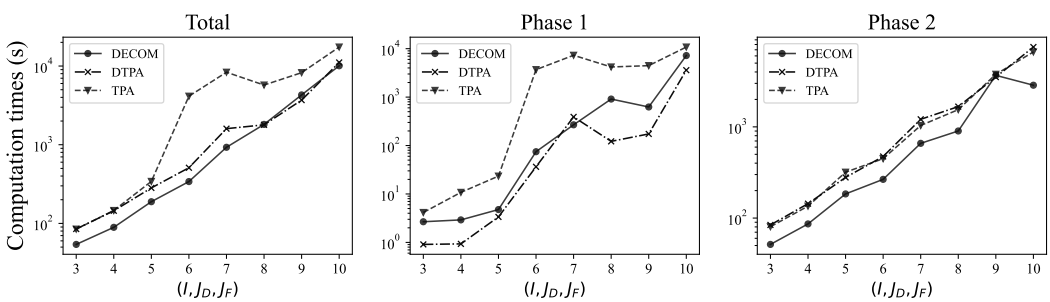



Figure 4.6: Computation times of Total, Phase1, and Phase 2 for three approaches.

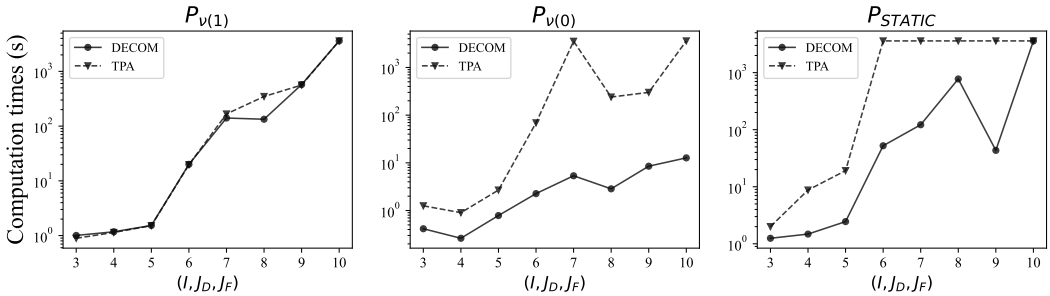



Figure 4.7: Computation times of $\mathrm{P}_{\nu(1)}$, $\mathrm{P}_{\nu(0)}$, and $\mathrm{P}_{\mathrm{STATIC}}$ for DECOM and TPA.

### 4.6.3 Experiment 3: Performance analysis by varying the production capacity

In this section, we have conducted two types of experiments to analyze the effects of production capacity on the problem complexity and the total cost. In the first experiment, we analyzed the complexity by varying the production capacity with different values of $\xi \in \{0.0, 0.1, \ldots, 0.9, 1.0\}$. We set $T = 7, K_O = 5, K_D = 5$, and locations were distributed randomly on a $50 \times 50$ XY plane. Also, demand distribution follows $Beta(1, 1)$ as in the setting of Experiment 2. We have conducted experiments on two different sizes of problems: $(I, J_D, J_F) = (4, 4, 4)$ and $(5, 5, 5)$.

Table C.2 in Appendix C.3 shows the experimental results with different production capacities. We excluded the DTPA in this experiment because of poor solution quality when the production capacity was insufficient. For the $\xi = 0.9$ and 1.0, the Gaps of DECOM and TPA were bigger than 10 percent because stockout frequently occurred owing to insufficient production capacity. We reported CPU(s) for Phase 1, Phase 2, and Total in Table C.2, and Total CPU(s) for TPA and DECOM was represented in Figure 4.8.

For both TPA and DECOM, the computation time to implement Phase 2 was

indifferent, although the value of $\xi$ was changed. In other words, production capacity had insignificant effects on the complexity of Phase 2. On the other hand, the production capacity significantly affected the TPA's computational efficiency for implementing Phase 1; thus, the Total CPU(s) of the TPA was also affected by the value of $\xi$. However, in the experiment setting of $(I, J_D, J_F) = (4, 4, 4)$, we required less than a minute to implement DECOM, except for in the case of $\xi = 0.9$. Furthermore, in the experiment setting of $(I, J_D, J_F) = (5, 5, 5)$, it takes less than 80 seconds to implement DECOM, except for in the cases of $\xi = 0.9$ and 1.0. Consequently, despite the production capacity changes, we observe that DECOM showed steady performance in terms of computational efficiency as represented in Figure 4.8.

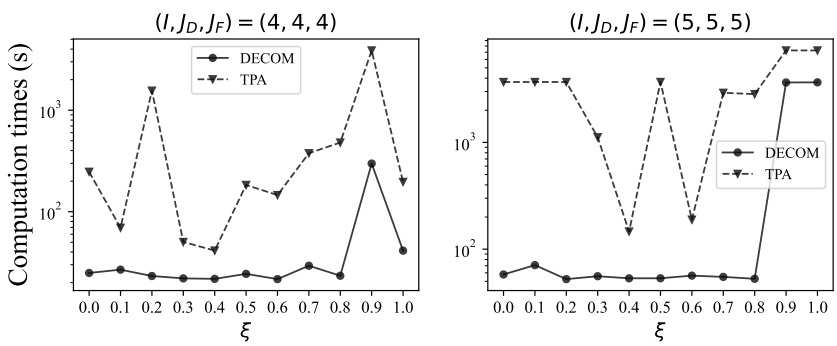



Figure 4.8: Total computation times of DECOM and TPA with different values of $\xi$.

In the second experiment, we implemented cost analysis by varying the production capacities. Figure 4.9 presents bar plots for six cost components: (1) the total cost for the whole supply chain (TC), (2) the total cost for the retailer's supply chain (TJD), (3) the total cost for the 3PP supply chain (TJF), (4) the stockout cost for the whole supply chain (PC), (5) the stockout cost for the retailer's supply chain

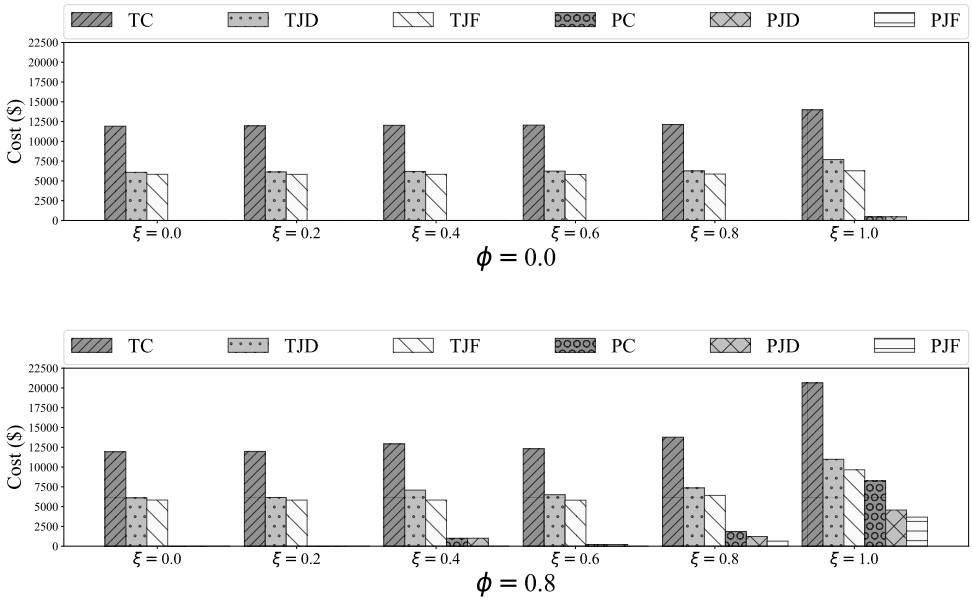



Figure 4.9: Cost analysis by varying the production capacity.

(PJD), and (6) the stockout cost for the supply chain of the 3PP (PJF). We conducted experiments for two decision rules; one was obtained from the DECOM with $\phi = 0.0$, and the other was obtained from the DECOM with $\phi = 0.8$. Because the DECOM with $\phi = 0.0$ could derive the conservative decision rule to the uncertainty, the stockout only occurred when the $\xi = 1.0$, which was the case in which suppliers had the smallest production capacities. However, because the DECOM with $\phi = 0.8$ output the aggressive decision rule to the uncertainty, the stockout occurred when $\xi \geq 0.4$. Of special note, when sufficient production capacity existed ($\xi = 0.0$ and 0.2), two decision rules incurred the same total cost.

### 4.6.4 Managerial insights

On the basis of the experimental results, we present the following managerial insights which could be instructive to practitioners who are concerned about setting up an effective supply chain in an omnichannel environment considering demand uncertainty:

- Even though our problem consists of only two types of supply chains, significant computational power is needed to optimize the flow of products when considering both supply chains simultaneously. In real business, the supply chain becomes more complicated and faces high demand uncertainty. Therefore, more computing power is no doubt required to optimize that complicated supply chain. Because of this, we recommend that practitioners seek the method that most efficiently decomposes and optimizes each supply chain, which is the DECOM approach. An efficient decomposition method could lead to promising policies within much shorter computation times compared to the method that simultaneously considers every supply chain component.

- Because of logistics disruption and production delays during COVID-19, the production capacity in the supply chain could be unstable. Hence, it is important to seek a promising strategy to operate the supply chain depending on the status of supplies. The DECOM can derive the decision rule which is conservative to the demand uncertainty if the value of $\phi$ is zero. On the contrary, the decision rule that is aggressive to the demand uncertainty could be obtained if the $\phi$ is close to one. Therefore, if practitioners are willing to utilize our approach, they should derive the policy appropriate to the production

capacity in their supply chain by setting the best value $\phi$.

- Our approach can be applied in practice without accurate demand distribution and only requires three pieces of information about demand: lower bound and upper bound of demand support set and mean value. However, because retail companies have been able to obtain lots of historical demand data with the rapid development of computational technology, these immense demand data should be utilized for improved decisions. Even if the accurate demand distribution is not estimated using historical data, it is necessary to determine the appropriate value $\phi$ after identifying how the distribution is roughly shaped (e.g., examine the skewness or variance of demand distributions). For example, as presented in Section 4.6.1, if the demand distribution is skewed to the right, it is necessary to derive a conservative solution by setting a small value for $\phi$. Otherwise, if the distribution is skewed to the left, the $\phi$ should be set with a large value to obtain an aggressive solution to demand uncertainty.

## 4.7 Summary

We studied the optimization problem considering demand uncertainty in a setting where the omnichannel retailer determined to utilize the 3PP channel in advance. In the proposed problem, the retailer's online and offline channels were operated by the retailer's supply chain, and the 3PP channel was operated by the supply chain of the 3PP. Moreover, we considered joint replenishment, allocation, transshipment, and fulfillment decisions over a multi-period planning horizon. To minimize the expected total cost, we presented the stochastic optimization model from the perspective of a retailer.

However, there were four challenges in our problem. First, the adjustable binary decisions for replenishment should be considered, which incurs a fixed order cost. Second, we should integrate anticipative and reactive decisions when solving the problem. Third, the existence of the 3PP channel increased the problem size because the retailer's supply chain and the supply chain of the 3PP should be considered simultaneously. Fourth, the production capacity constraint made the problem more intractable.

Even though the TPA developed by Lim et al. [93] could mitigate the first and second challenges, TPA often required a high computational burden to solve the proposed problem because of the third and fourth challenges. In particular, performance of the TPA was aggravated when the production capacity was insufficient. As a way to overcome these challenges, we proposed a DECOM approach by utilizing artificial variables, and it can solve the problem separately according to the retailer's supply chain and the 3PP supply chain.

Through conducting computational experiments, we observed that DECOM and TPA provided solutions with similar quality in various demand distributions. However, DECOM outperformed TPA in terms of computational efficiency. In particular, DECOM was scalable to large-scale problems while maintaining its high solution quality. In addition, despite insufficient production capacity, DECOM showed steady performance compared to the sufficient case, while the TPA suffered from high computational complexity. Based on the experimental results, we presented several managerial insights that could be instructive to the omnichannel retailer who needs to operate both a retailer's supply chain and a 3PP supply chain effectively.

# Chapter 5

# Conclusions

## 5.1 Summary and contributions

Because of the high flexibility and low risk of sharing economy, logistics practitioners have started to embrace a sharing economy with logistics to bring efficiency to fulfillment services. This thesis aimed to alleviate several challenges of e-commerce retailers by sharing logistics resources, and we proposed three operation problems. However, the adoption of sharing logistics resources increased the uncertainty and complexity of the proposed problems. Therefore, it was required to develop advanced solution methodologies which could consider uncertainty systematically and be scalable to realistic problem instances.

In Chapter 2, we proposed the lateral transshipment model for fresh food by accommodating the key attributes of the OOCS: heterogeneous shelf life, proactive transshipment, and non-negligible transshipment time. We developed the hybrid DRL approach by combining the SAC algorithm with SQLT policy and RS. The proposed approach had the following three advantages. First, the proposed approach greatly alleviated the curse of dimensionality, which is incurred due to the perishable nature of fresh produce. Second, because the hybrid DRL could derive policy by directly utilizing data, it does not need any knowledge or assumption about demand

distribution. Third, the hybrid DRL was stable during the training process compared to the original SAC algorithm because it could mitigate issues incurred due to large action spaces. Experimental results showed that the hybrid DRL could outperform existing approaches developed by Haijema and Minner [64] and Dehghani et al. [38]. In addition, we found that transshipment substantially reduces the outdating cost by allowing the offline channel to make good use of the old products that will be discarded in the online channel, which is new to the literature.

In Chapter 3, we proposed the SCND problem considering ODWS under demand and yield uncertainty. We considered the commitment variables and uncertainty in ODWS, which is new to the literature. The proposed problem was formulated by the TSSP model, and we solved it by utilizing the method combined with SAA and BD algorithms. Of special note, we developed the ABD, which could generate effective initial cuts for improving the convergence speed of the BD algorithm. The ABD could outperform the typical version of the BD algorithm and Xpress-Optimizer with regard to optimality gap and computation times. Under various experiment settings, we could observe cost-saving effects when ODWS was used for designing supply chain networks. Through a sensitivity analysis, the parameter values for commitment and stockout affected decisions for whether to utilize ODWS or not.

In Chapter 4, we proposed an omnichannel retail operations problem considering production capacity constraint and the 3PP channel. We adopted the adjustable binary and continuous variables as wait-and-see decisions, and the state-of-art-method TPA [93] was utilized to solve the problem. However, the existence of the 3PP channel and production capacity constraints rapidly increased the computational complexity; thus, TPA suffered from the high computational burden. Therefore,

we proposed a novel approach, DECOM, and it had two distinctive advantages: (1) reduce the feasible region and (2) decompose the original problem into two small problems, one for the retailer's supply chain and the other for the supply chain of the 3PP. Through computational experiments, DECOM could be scalable to large-scale problems while maintaining high solution quality. In contrast, TPA could not solve the same problems within acceptable computation times.

## 5.2 Future research

Based on the several limitations of this thesis, we suggested some lines of future research for each chapter. Chapter 2 has two limitations. The first limitation of Chapter 2 is that we only considered two outlets. On the other hand, leading companies with the OOCS commonly operate multiple online distribution centers and offline retail stores. The second limitation is that DRL requires massive computational efforts to train neural networks for one instance. In order to systematically analyze the impacts of the proposed model, the trained neural networks of DRL need to be evaluated in different instances by varying values of parameters. However, it requires several weeks or months to train neural networks from scratch for every different combination of parameters. The third limitation is that we only considered the transshipment to deal with outdated products. In real business, transshipment could incur high operational costs for allocating and packaging products. Based on the above three limitations, several future studies could be suggested. Reflecting the distances of multiple outlet locations in cost parameters and developing the appropriate DRL approach could be important lines of future research. Also, instead of training neural networks for DRL from scratch for each instance, future studies

should target developing methods that reutilize neural networks completed training for another instance. In addition, adopting the *promotion* to sell outdated products at lower sale prices could be interesting topics for future studies. However, if promotion is considered one of our model's decisions, the action space should be larger than our proposed model. In order to consider promotion, future studies should target overcoming challenges incurred from the large action space by developing an appropriate MDP model or utilizing other solution approaches, such as stochastic programming and robust optimization.

For further research of Chapter 3, we intend to extend our study by using multi-stage stochastic programming, which has an advantage for dealing with uncertainty under a multi-period setting. The nature of TSSP enables the stochastic parameters to become known in a single moment. However, regarding the problem with a planning horizon with multiple periods, the uncertainty can be dealt with more accurately when the stochastic parameters have been realized progressively in each period. Therefore, through utilizing the above scheme, some decisions will be made before the realization of uncertainty, and other decisions will be made after the realization in each period [57].

In Chapter 4, the first limitation of our study is that we failed to acquire real-world data and have just utilized the benchmark data provided by Jiu [72] to validate the proposed approach. The second limitation is that we determine the value of cost target $\psi$, which affects the performance of DECOM, by simply utilizing the affine function of target coefficient $\phi$. Based on the above two limitations, validating the DECOM by defining the demand uncertainty set with real-world data could be an important line of future research, which could contribute more insight to real

problems. In addition, instead of finding the best value of $\psi$ by conducting extensive experiments, future studies should target developing methods that could provide the best $\psi$ appropriate to the given problem. Furthermore, future studies could provide the compact search space to tune $\psi$ for reducing the workload.

# Appendices

# Appendix A

# Supplementary materials for Chapter 2

## A.1 Information about hyperparameters of the hybrid DRL approach

Table A.1: Hyperparameters used for the hybrid DRL approach

| Hyperparameter name | Hyperparameter used |
|---|---|
| Size of the replay buffer $N_{\mathcal{D}}$ | 200,000 |
| Size of the minibatch $\|\mathcal{B}\|$ | 128 |
| Soft update factor $\psi$ | 0.002 |
| Prioritization factor $\eta$ | 0.6 |
| Compensation factor $\beta$ | 0.4 |
| Discount factor $\gamma$ | 0.99 |
| Learning rate $\lambda$ | 0.0001 |
| Hidden layer of neural networks | [64,64] |
| Optimizer | Adam |
| Activation function hidden layers | Relu |
| Activation function output layers | Softmax |

## A.2 Pseudocode for SACDPE

**Algorithm 3** SACDPE

Initialize $Q^{soft}_{\theta_1} : \mathcal{S} \to \mathbb{R}^{|\mathcal{A}|},\ Q^{soft}_{\theta_2} : \mathcal{S} \to \mathbb{R}^{|\mathcal{A}|},\ \pi_\phi : \mathcal{S} \to [0,1]^{|\mathcal{A}|}$
Initialize $Q^{soft}_{\bar{\theta}_1} : \mathcal{S} \to \mathbb{R}^{|\mathcal{A}|},\ Q^{soft}_{\bar{\theta}_2} : \mathcal{S} \to \mathbb{R}^{|\mathcal{A}|},\ \mathcal{D} \leftarrow \emptyset$
$\bar{\theta}_1 \leftarrow \theta_1,\ \bar{\theta}_2 \leftarrow \theta_2$
Declare the environment for SACDPE ($ENV_{RL}$)
$e \leftarrow 1$
**for** *each episode* $e = 1, \cdots, E$ **do**
  $t \leftarrow 1$
  **for** *each timestep* $t = 1, \cdots, T$ **do**
    Observe $s_t$ and choose action $a_t \sim \pi_\phi(\cdot|s_t)$
    Observe $r_t = PF^{RL}_t$ and $s_{t+1}$ from $ENV_{RL}$
    $\mathcal{D} \leftarrow \mathcal{D} \cup \{(s_t, a_t, r_t, s_{t+1})\}$ with maximal priority $p_t = \max_{i<t} p_i$
    Sample a mini-batch $\mathcal{B}$ from $\mathcal{D}$ according to probability $P(d) = p^\eta_d / \sum_{k=1}^{N_\mathcal{D}} p^\eta_k, \forall d \in \mathcal{D}$
    $\Delta\theta_1, \Delta\theta_2, \Delta\phi, \Delta\alpha = 0$
    **for** $b \in \mathcal{B}$ **do**
      $w_b = \left(\frac{1}{N_\mathcal{D}} \times \frac{1}{P(d)}\right)^\beta / \max_{i\in\mathcal{B}} w_i$
      $|\delta_b| = \min\left\{\left(Q^{soft}_{\theta_1}(s) - (r + \gamma V_{\bar{\theta}}(s'))\right)^2, \left(Q^{soft}_{\theta_2}(s) - (r + \gamma V_{\bar{\theta}}(s'))\right)^2\right\}$
      $p_b \leftarrow |\delta_b| + \epsilon_{per}$
      $\Delta\theta_i \leftarrow \Delta\theta_i + w_b \nabla_{\theta_i} J_{Q^{soft}}(\theta_i)$, for $i \in \{1, 2\}$
      $\Delta\phi \leftarrow \Delta\phi + w_b \nabla_\phi J_\pi(\phi)$
      $\Delta\alpha \leftarrow \Delta\alpha + \nabla_\alpha J(\alpha)$
    **end**
    Update soft Q networks $\theta_i \leftarrow \theta_i - \lambda\Delta\theta_i$, for $i \in \{1, 2\}$
    Update policy network $\phi \leftarrow \phi - \lambda\Delta\phi$
    Adjust temperature $\alpha \leftarrow \alpha - \lambda\Delta\alpha$
    Update target soft Q networks $\bar{\theta}_i \leftarrow \psi\theta_i + (1-\psi)\bar{\theta}_i$, for $i \in \{1, 2\}$
    $t \leftarrow t + 1$
  **end**
  $e \leftarrow e + 1$
**end**
**Return:** $\theta_1,\ \theta_2,\ \pi_\phi$

## A.3 The reasons for using the existing data

Because RL is a learning-based algorithm, the training data significantly affects the policy derived by training RL. In addition, because we deal with the novel problem and the application research for RL, it is very important to utilize the appropriate data to the proposed problem. It is ideal to utilize the real data of *Oasis Market*, but we could not secure that data. To overcome these challenges, we investigated literature related to the RL for inventory problems. However, most existing studies could not obtain the appropriate data for their problems or secure immense real data for training the proposed RL algorithms. They usually organize the training data with the following two schemes:

1. Fit a probability distribution to small-size real data and generate demand data for training [52, 104].

2. Assume a probability distribution and generate demand data under various population parameters [36, 139].

In our problem, we consider the inventory model with a single item and two different channels (i.e., OOCS with single item). In real practice, it is obvious that the demand distributions of online and offline channels are different. As illustrated in Figure A.1, the OOCS with a single item can be interpreted as the single channel model with two types of items. In this model, the type of item can be transformed to the other type by implementing transshipment. Therefore, the demand distribution of each channel in OOCS can be interpreted as the demand distribution of each type of item in a single channel.

Oroojlooyjadid et al. [104] presented demand data sets for three different items.

Also, the demand distributions of the three items were different. Therefore, we adopted the data provided by Oroojlooyjadid et al. [104] in Section 2.5.1. At last, to validate the performance of the hybrid DRL, we have generated six test instances by adopting the different demand data sets for each channel in OOCS.

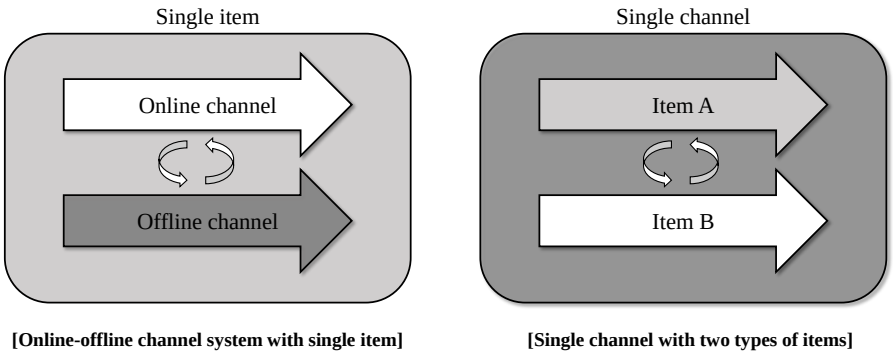



Figure A.1: OOCS with a single item and single channel with two types of item.

## A.4 Improvement effects of transshipment varying the unit transshipment cost parameter

Table A.2: Improvement effects of utilizing transshipment varying the unit transshipment cost parameter $c_l$

| | $c_l$ | Measure | RV | HC | SC | WC | TC | OC | PF |
|---|---|---|---|---|---|---|---|---|---|
| No-transshipment | - | Average | 162.53 | 16.62 | 9.40 | 13.88 | 0.00 | 58.49 | 64.14 |
| Transshipment | 0 | Average | 163.62 | 15.04 | 8.69 | 10.83 | 0.00 | 58.00 | 71.06 |
| | | Improvement[a] | 1.09 | 1.58 | 0.71 | 3.05 | 0.00 | 0.49 | 6.91 |
| | 1 | Average | 163.25 | 15.80 | 8.87 | 10.56 | 2.10 | 57.81 | 68.11 |
| | | Improvement[a] | 0.72 | 0.81 | 0.53 | 3.32 | -2.10 | 0.68 | 3.97 |
| | 2 | Average | 162.23 | 16.44 | 9.38 | 10.41 | 2.22 | 57.46 | 66.33 |
| | | Improvement[a] | -0.29 | 0.18 | 0.02 | 3.47 | -2.22 | 1.03 | 2.18 |
| | 3 | Average | 162.63 | 16.49 | 9.22 | 11.22 | 2.46 | 57.80 | 65.43 |
| | | Improvement[a] | 0.10 | 0.13 | 0.18 | 2.65 | -2.46 | 0.69 | 1.29 |
| | 4 | Average | 162.12 | 16.70 | 9.51 | 11.80 | 1.51 | 57.81 | 64.79 |
| | | Improvement[a] | -0.41 | -0.08 | -0.11 | 2.07 | -1.51 | 0.69 | 0.64 |
| | 5 | Average | 162.30 | 16.69 | 9.43 | 12.24 | 1.43 | 57.98 | 64.53 |
| | | Improvement[a] | -0.22 | -0.07 | -0.03 | 1.63 | -1.43 | 0.51 | 0.39 |
| | 6 | Average | 162.39 | 16.69 | 9.39 | 12.39 | 1.51 | 58.05 | 64.37 |
| | | Improvement[a] | -0.13 | -0.07 | 0.01 | 1.49 | -1.51 | 0.44 | 0.23 |
| | 7 | Average | 162.46 | 16.68 | 9.36 | 12.54 | 1.53 | 58.11 | 64.23 |
| | | Improvement[a] | -0.07 | -0.07 | 0.04 | 1.34 | -1.53 | 0.38 | 0.09 |
| | 8 | Average | 162.53 | 16.62 | 9.40 | 13.88 | 0.00 | 58.49 | 64.14 |
| | | Improvement[a] | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 9 | Average | 162.53 | 16.62 | 9.40 | 13.88 | 0.00 | 58.49 | 64.14 |
| | | Improvement[a] | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

[a] Improvement: (Transshipment−No-transshipment) for revenue and profit, and (No-transshipment−Transshipment) for cost components

## A.5 Improvement effects of transshipment varying the shelf life of online and offline channels

Table A.3: Improvement effects of utilizing transshipment varying the shelf life of product held in online and offline channels, $M^{ON}$ and $M^{OF}$

| | | Short shelf life | | | Long shelf life | | |
|---|---|---|---|---|---|---|---|
| | $M^{ON}$ | 3 | 3 | 3 | 5 | 5 | 5 |
| | $M^{OF}$ | 5 | 6 | 7 | 7 | 8 | 9 |
| | $M^{OF} - M^{ON}$ | 2 | 3 | 4 | 2 | 3 | 4 |
| | | Average value per period | | | | | |
| RV | No-transshipment | 161.40 | 162.31 | 162.84 | 168.34 | 167.79 | 167.79 |
| | Transshipment | 162.15 | 162.78 | 163.41 | 169.77 | 169.76 | 169.54 |
| | Improvement[a] | 0.75 | 0.47 | 0.57 | 1.42 | 1.97 | 1.76 |
| HC | No-transshipment | 15.99 | 17.24 | 17.96 | 23.43 | 23.08 | 23.10 |
| | Transshipment | 14.04 | 16.03 | 16.69 | 20.30 | 20.19 | 20.12 |
| | Improvement[a] | 1.95 | 1.21 | 1.26 | 3.12 | 2.90 | 2.98 |
| SC | No-transshipment | 10.03 | 9.46 | 9.12 | 6.37 | 6.65 | 6.65 |
| | Transshipment | 9.48 | 9.08 | 8.69 | 5.63 | 5.69 | 5.68 |
| | Improvement[a] | 0.55 | 0.37 | 0.43 | 0.74 | 0.96 | 0.97 |
| WC | No-transshipment | 13.23 | 11.91 | 11.36 | 3.02 | 2.53 | 2.48 |
| | Transshipment | 10.13 | 7.46 | 7.50 | 1.43 | 1.12 | 1.04 |
| | Improvement[a] | 3.10 | 4.45 | 3.86 | 1.60 | 1.41 | 1.44 |
| OC | No-transshipment | 57.92 | 57.87 | 57.90 | 57.05 | 56.73 | 56.71 |
| | Transshipment | 57.32 | 56.76 | 57.00 | 57.01 | 56.89 | 56.86 |
| | Improvement[a] | 0.60 | 1.11 | 0.90 | 0.03 | -0.16 | -0.15 |
| PF | No-transshipment | 64.24 | 65.84 | 66.51 | 78.48 | 78.79 | 78.85 |
| | Transshipment | 71.19 | 73.46 | 73.53 | 85.39 | 85.88 | 85.84 |
| | Improvement[a] | 6.95 | 7.61 | 7.02 | 6.92 | 7.08 | 7.00 |

[a] Improvement: (Transshipment−No-transshipment) for revenue and profit, and (No-transshipment−Transshipment) for cost components

# Appendix B

# Supplementary materials for Chapter 3

## B.1 Parameter information

Table B.1: Ranges of the deterministic parameters

| $F_j$ | $\alpha$ | $\beta_i$ | $\gamma$ | $C^k$ & $C^r$ |
|---|---|---|---|---|
| $\mathcal{U}(500, 1000)$ | $\mathcal{U}(100, 200)$ | $\mathcal{U}(30, 70)$ | $\mathcal{U}(0.80, 0.99)$ | $\mathcal{U}(50, 100)$ |

Table B.2: Probability distributions for stochastic parameters

| $D_{it}^w$ | $S_{ijt}^w$ |
|---|---|
| $\mathcal{N}\left(\frac{179.06}{\lvert\mathcal{I}\rvert}, (\frac{91.18}{\lvert\mathcal{I}\rvert})^2\right)$ | $\mathcal{N}\left(\frac{179.06}{\lvert\mathcal{I}\rvert}, (\frac{91.18}{\lvert\mathcal{I}\rvert})^2\right)$ |

## B.2 Comparison of performance for solving SAA problems and computational results about the two types of lead time

Table B.3: Comparison of performance between Xpress Solver, TBD, and ABD for solving SAA problems

| No. | Method | $N$ | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 20 | | | 40 | | | 60 | | | 80 | | |
| | | Gap | CPUs | Itr | Gap | CPUs | Itr | Gap | CPUs | Itr | Gap | CPUs | Itr |
| 1 | Solver | 0.00 | 11.86 | - | 0.00 | 18.39 | - | 0.00 | 30.47 | - | 0.00 | 44.40 | - |
| | TBD | 0.00 | 4.39 | 13.7 | 0.00 | 3.87 | 12.1 | 0.00 | 5.35 | 12.1 | 0.00 | 6.90 | 11.9 |
| | ABD | 0.00 | 4.46 | 11 | 0.00 | 3.46 | 9.4 | 0.00 | 4.68 | 9.7 | 0.00 | 6.29 | 9.6 |
| 2 | Solver | 0.00 | 8.88 | - | 0.00 | 16.04 | - | 0.00 | 22.96 | - | 0.00 | 33.82 | - |
| | TBD | 0.00 | 4.30 | 13.4 | 0.00 | 3.72 | 13.3 | 0.00 | 4.59 | 12.5 | 0.00 | 7.45 | 13.5 |
| | ABD | 0.00 | 3.96 | 11.2 | 0.00 | 3.56 | 11.6 | 0.00 | 3.74 | 9.3 | 0.00 | 7.01 | 11.4 |
| 3 | Solver | 0.00 | 37.08 | - | 0.00 | 77.97 | - | 0.00 | 112.47 | - | 0.00 | 191.44 | - |
| | TBD | 0.00 | 28.05 | 20.0 | 0.00 | 39.45 | 18.1 | 0.00 | 34.47 | 15.7 | 0.00 | 64.91 | 17.4 |
| | ABD | 0.00 | 24.64 | 17.4 | 0.00 | 30.81 | 14.7 | 0.00 | 31.46 | 13.6 | 0.00 | 56.24 | 14.3 |
| 4 | Solver | 0.00 | 30.21 | - | 0.00 | 54.99 | - | 0.00 | 110.47 | - | 0.00 | 138.46 | - |
| | TBD | 0.00 | 16.57 | 19.3 | 0.00 | 21.74 | 16.4 | 0.00 | 32.91 | 17.5 | 0.00 | 44.50 | 17.7 |
| | ABD | 0.00 | 14.80 | 16.2 | 0.00 | 18.04 | 13.4 | 0.00 | 33.00 | 16.2 | 0.00 | 38.51 | 15.0 |
| 5 | Solver | 0.00 | 61.87 | - | 0.00 | 91.22 | - | 0.00 | 180.89 | - | 0.00 | 301.62 | - |
| | TBD | 0.00 | 30.64 | 18.3 | 0.00 | 32.89 | 15.9 | 0.00 | 62.89 | 17.8 | 0.00 | 71.48 | 16.5 |
| | ABD | 0.00 | 27.51 | 15.1 | 0.00 | 31.38 | 13.8 | 0.00 | 54.98 | 14.1 | 0.00 | 56.07 | 12.7 |
| 6 | Solver | 0.00 | 38.60 | - | 0.00 | 94.68 | - | 0.00 | 153.23 | - | 0.00 | 232.66 | - |
| | TBD | 0.00 | 18.68 | 16.0 | 0.00 | 31.89 | 15.5 | 0.00 | 47.82 | 15.9 | 0.00 | 63.51 | 16.1 |
| | ABD | 0.00 | 16.16 | 12.9 | 0.00 | 24.08 | 11.3 | 0.00 | 38.06 | 12.2 | 0.00 | 48.28 | 12.2 |
| 7 | Solver | 0.00 | 66.07 | - | 0.00 | 137.47 | - | 0.00 | 242.08 | - | 0.00 | 342.36 | - |
| | TBD | 0.00 | 79.86 | 27.3 | 0.00 | 62.83 | 22.4 | 0.00 | 102.06 | 23.4 | 0.00 | 162.87 | 24.7 |
| | ABD | 0.00 | 77.83 | 25.5 | 0.00 | 69.86 | 20.9 | 0.00 | 102.76 | 21.2 | 0.00 | 167.03 | 22.5 |
| 8 | Solver | 0.00 | 578.07 | - | 0.00 | 1496.01 | - | 0.20 | 3231.62 | - | 0.38 | 3599.15 | - |
| | TBD | 0.00 | 576.96 | 39.6 | 0.00 | 746.29 | 42.1 | 0.00 | 838.18 | 32.5 | 0.00 | 1531.10 | 36.5 |
| | ABD | 0.00 | 613.06 | 38.6 | 0.00 | 755.93 | 38.7 | 0.00 | 839.66 | 29.9 | 0.00 | 1372.98 | 33.9 |
| 9 | Solver | 0.00 | 837.11 | - | 0.00 | 1995.72 | - | 0.11 | 3340.58 | - | 1.45 | 3600* | - |
| | TBD | 0.00 | 730.23 | 37.2 | 0.00 | 935.83 | 31.5 | 0.00 | 1260.52 | 30.8 | 0.00 | 1445.09 | 29.6 |
| | ABD | 0.00 | 870.76 | 34.5 | 0.00 | 862.95 | 27.9 | 0.00 | 1057.39 | 26.7 | 0.00 | 1187.31 | 25.4 |
| 10 | Solver | 0.00 | 438.08 | - | 0.02 | 1628.80 | - | 0.02 | 2453.07 | - | 0.25 | 3419.30 | - |
| | TBD | 0.00 | 323.73 | 33.9 | 0.00 | 598.13 | 31.7 | 0.00 | 1055.44 | 31.5 | 0.00 | 1318.84 | 32.6 |
| | ABD | 0.00 | 324.92 | 33.3 | 0.00 | 559.46 | 29.4 | 0.00 | 842.94 | 28.5 | 0.00 | 1193.71 | 29.3 |
| 11 | Solver | 0.00 | 488.98 | - | 0.00 | 981.57 | - | 0.00 | 2046.29 | - | 0.02 | 3441.95 | - |
| | TBD | 0.00 | 151.46 | 31.0 | 0.00 | 212.41 | 25.6 | 0.00 | 306.67 | 25.3 | 0.00 | 439.46 | 26.2 |
| | ABD | 0.00 | 147.37 | 27.7 | 0.00 | 198.26 | 22.3 | 0.00 | 303.81 | 22.6 | 0.00 | 455.59 | 24.4 |
| 12 | Solver | 0.00 | 732.77 | - | 0.00 | 1806.96 | - | 0.03 | 3600* | - | 0.26 | 3600* | - |
| | TBD | 0.00 | 733.32 | 39.6 | 0.00 | 651.79 | 30.6 | 0.00 | 1043.33 | 29.8 | 0.00 | 1290.18 | 29.1 |
| | ABD | 0.00 | 576.05 | 34.5 | 0.00 | 611.11 | 26.8 | 0.00 | 825.19 | 25.5 | 0.00 | 1129.70 | 24.9 |
| 13 | Solver | 0.05 | 1938.19 | - | 0.34 | 3600* | - | 4.07 | 3600* | - | 4.20 | 3600* | - |
| | TBD | 0.00 | 2422.31 | 65.8 | 0.05 | 3578.91 | 53.6 | 0.43 | 3600* | 43.8 | 1.72 | 3600* | 37.2 |
| | ABD | 0.00 | 1747.20 | 52.7 | 0.04 | 3313.95 | 51.5 | 0.34 | 3600* | 42.4 | 1.44 | 3600* | 33.6 |
| 14 | Solver | 0.00 | 905.16 | - | 0.00 | 2682.87 | - | 0.32 | 3600* | - | 1.34 | 3600* | - |
| | TBD | 0.00 | 617.21 | 28.1 | 0.00 | 1375.98 | 25.6 | 0.00 | 1806.77 | 24.0 | 0.00 | 2594.65 | 24.6 |
| | ABD | 0.00 | 492.04 | 21.8 | 0.00 | 1234.17 | 21.2 | 0.00 | 1511.06 | 19.1 | 0.00 | 2071.29 | 19.3 |
| 15 | Solver | 0.01 | 2697.94 | - | 2.07 | 3600* | - | 2.62 | 3600* | - | 2.87 | 3600* | - |
| | TBD | 0.01 | 3053.00 | 53.3 | 0.20 | 3600* | 39.0 | 1.50 | 3600* | 27.9 | 2.85 | 3600* | 22.8 |
| | ABD | 0.00 | 2104.27 | 46.7 | 0.08 | 3600* | 42.1 | 0.35 | 3600* | 31.9 | 0.53 | 3600* | 28.4 |

* Time limit was reached

Table B.4: Impacts of lead time on cost incurred in supply chain with the ODWS

| $L_s$ | $L_d$ | Total cost ($) | Delivery | | Commitment | | Stockout | | Supplier investment | | Transportation | | Inventory holding | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Cost ($) | % | Cost ($) | % | Cost ($) | % | Cost ($) | % | Cost ($) | % | Cost ($) | % |
| 0 | 0 | 12,830.6 | 7,849.4 | 61.18 | 2,278.4 | 17.76 | 93.0 | 0.72 | 1,134.8 | 8.84 | 1,368.0 | 10.66 | 107.1 | 0.83 |
| 0 | 1 | 14,246.8 | 7,326.8 | 51.43 | 2,235.7 | 15.69 | 2,191.4 | 15.38 | 1,134.8 | 7.97 | 1,271.4 | 8.92 | 86.8 | 0.61 |
| 0 | 2 | 15,220.2 | 6,799.9 | 44.68 | 1,986.2 | 13.05 | 3,958.3 | 26.01 | 1,134.8 | 7.46 | 1,266.2 | 8.32 | 74.9 | 0.49 |
| 0 | 3 | 16,548.1 | 6,270.3 | 37.89 | 1,704.8 | 10.30 | 6,072.4 | 36.70 | 1,134.8 | 6.86 | 1,285.0 | 7.77 | 80.8 | 0.49 |
| 0 | 4 | 17,992.7 | 5,725.5 | 31.82 | 1,593.4 | 8.86 | 8,321.8 | 46.25 | 1,134.8 | 6.31 | 1,157.7 | 6.43 | 59.5 | 0.33 |
| 1 | 0 | 13,935.3 | 8,140.5 | 58.42 | 2,278.4 | 16.35 | 820.1 | 5.89 | 1,134.8 | 8.14 | 1,482.9 | 10.64 | 78.7 | 0.56 |
| 1 | 1 | 15,287.0 | 7,618.1 | 49.83 | 2,166.9 | 14.17 | 2,894.1 | 18.93 | 1,134.8 | 7.42 | 1,409.0 | 9.22 | 64.0 | 0.42 |
| 1 | 2 | 16,284.1 | 7,075.4 | 43.45 | 1,900.2 | 11.67 | 4,720.8 | 28.99 | 1,134.8 | 6.97 | 1,400.1 | 8.60 | 52.8 | 0.32 |
| 1 | 3 | 17,627.9 | 6,527.6 | 37.03 | 1,656.4 | 9.40 | 6,898.1 | 39.13 | 1,134.8 | 6.44 | 1,350.3 | 7.66 | 60.7 | 0.34 |
| 1 | 4 | 19,163.1 | 6,000.0 | 31.31 | 1,450.2 | 7.57 | 9,233.1 | 48.18 | 1,134.8 | 5.92 | 1,302.4 | 6.80 | 42.7 | 0.22 |
| 2 | 0 | 14,475.1 | 8,159.9 | 56.37 | 2,198.1 | 15.19 | 1,433.5 | 9.90 | 1,134.8 | 7.84 | 1,484.6 | 10.26 | 64.2 | 0.44 |
| 2 | 1 | 16,009.7 | 7,611.5 | 47.54 | 1,900.2 | 11.87 | 3,819.7 | 23.86 | 1,134.8 | 7.09 | 1,482.2 | 9.26 | 61.3 | 0.38 |
| 2 | 2 | 16,989.9 | 7,081.0 | 41.68 | 1,693.9 | 9.97 | 5,589.3 | 32.90 | 1,134.8 | 6.68 | 1,436.6 | 8.46 | 54.3 | 0.32 |
| 2 | 3 | 18,262.0 | 6,542.5 | 35.83 | 1,545.0 | 8.46 | 7,690.3 | 42.11 | 1,134.8 | 6.21 | 1,304.0 | 7.14 | 45.3 | 0.25 |
| 2 | 4 | 19,733.2 | 6,021.5 | 30.51 | 1,353.1 | 6.86 | 9,844.6 | 49.89 | 1,134.8 | 5.75 | 1,324.7 | 6.71 | 54.3 | 0.28 |
| 3 | 0 | 15,412.1 | 8,127.7 | 52.74 | 2,049.2 | 13.30 | 2,677.9 | 17.38 | 1,134.8 | 7.36 | 1,353.7 | 8.78 | 68.9 | 0.45 |
| 3 | 1 | 16,782.2 | 7,599.4 | 45.28 | 1,774.2 | 10.57 | 4,837.7 | 28.83 | 1,134.8 | 6.76 | 1,374.1 | 8.19 | 62.0 | 0.37 |
| 3 | 2 | 17,884.8 | 7,055.7 | 39.45 | 1,550.7 | 8.67 | 6,782.4 | 37.92 | 1,134.8 | 6.35 | 1,306.5 | 7.31 | 54.7 | 0.31 |
| 3 | 3 | 18,924.5 | 6,547.9 | 34.60 | 1,359.1 | 7.18 | 8,500.7 | 44.92 | 1,134.8 | 6.00 | 1,324.1 | 7.00 | 58.1 | 0.31 |
| 3 | 4 | 20,331.2 | 6,030.3 | 29.66 | 1,290.3 | 6.35 | 10,614.5 | 52.21 | 1,134.8 | 5.58 | 1,215.6 | 5.98 | 45.8 | 0.23 |
| 4 | 0 | 16,271.7 | 8,112.0 | 49.85 | 1,837.2 | 11.29 | 3,799.9 | 23.35 | 1,134.8 | 6.97 | 1,320.2 | 8.11 | 67.6 | 0.42 |
| 4 | 1 | 17,830.6 | 7,574.2 | 42.48 | 1,725.8 | 9.68 | 6,091.1 | 34.16 | 1,134.8 | 6.36 | 1,243.5 | 6.97 | 61.3 | 0.34 |
| 4 | 2 | 18,646.1 | 7,056.9 | 37.85 | 1,482.0 | 7.95 | 7,721.5 | 41.41 | 1,134.8 | 6.09 | 1,195.9 | 6.41 | 55.1 | 0.30 |
| 4 | 3 | 19,696.8 | 6,542.8 | 33.22 | 1,349.6 | 6.85 | 9,451.2 | 47.98 | 1,134.8 | 5.76 | 1,163.1 | 5.90 | 55.4 | 0.28 |
| 4 | 4 | 20,991.6 | 6,022.4 | 28.69 | 1,221.0 | 5.82 | 11,488.2 | 54.73 | 1,134.8 | 5.41 | 1,076.3 | 5.13 | 49.0 | 0.23 |

# Appendix C

# Supplementary materials for Chapter 4

## C.1 $\mathrm{P}_{\mathrm{LDR}}$ for Phase 2 of TPA

We propose the following robust optimization Problem $\mathrm{P}_{\mathrm{LDR}}$ based on the Theorem 2 in [93]:

$(\mathrm{P}_{\mathrm{LDR}})$

$$\begin{aligned}
\min \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}}\Bigg(&\sum_{j\in\mathcal{J}}\lambda_o^{it}c_o^{ij}\left(q_j^{it,0}+\sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1}q_j^{it,\sigma\tau}\hat{d}_\sigma^{i\tau}\right)+\sum_{j\in\mathcal{J}}h_{x,j}^{it}\left(x_j^{i,t+1,0}+\sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t}x_j^{i,t+1,\sigma\tau}\hat{d}_\sigma^{i\tau}\right)\\
&+\sum_{k\in\mathcal{K}_O}h_{y,k}^{it}\left(y_k^{i,t+1,0}+\sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t}y_k^{i,t+1,\sigma\tau}\hat{d}_\sigma^{i\tau}\right)+\sum_{k\in\mathcal{K}}p_k^{it}\left(z_k^{it,0}+\sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t}z_k^{it,\sigma\tau}\hat{d}_\sigma^{i\tau}\right)\\
&+\sum_{j\in\mathcal{J}_D}\sum_{j'\in\mathcal{J}_D}\lambda_l^{it}c_l^{jj'}\left(u_{l,jj'}^{it,0}+\sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1}u_{l,jj'}^{it,\sigma\tau}\hat{d}_\sigma^{i\tau}\right)+\sum_{j\in\mathcal{J}_D}\sum_{k\in\mathcal{K}_O}\lambda_e^{it}c_e^{jk}\left(u_{e,jk}^{it,0}+\sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1}u_{e,jk}^{it,\sigma\tau}\hat{d}_\sigma^{i\tau}\right)\\
&+\sum_{j\in\mathcal{J}_D}\sum_{k\in\mathcal{K}_D}\lambda_g^{it}c_g^{jk}\left(r_{jk}^{it,0}+\sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t}r_{jk}^{it,\sigma\tau}\hat{d}_\sigma^{i\tau}\right)+\sum_{k\in\mathcal{K}_O}\rho_k^{it}\left(v_{\rho,k}^{it,0}+\sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t}v_{\rho,k}^{it,\sigma\tau}\hat{d}_\sigma^{i\tau}\right)\\
&+\sum_{j\in\mathcal{J}_F}\eta_j^{it}\left(v_{\eta,j}^{it,0}+\sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t}v_{\eta,j}^{it,\sigma\tau}\hat{d}_\sigma^{i\tau}\right)\Bigg)
\end{aligned}$$

$$\begin{aligned}
\text{s.t.}\quad & q_j^{it,0}+\sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1}q_j^{it,\sigma\tau}d_\sigma^{i\tau}\le\bar{q}_j^i\bar{\delta}_j^{it}, \qquad \forall i\in\mathcal{I},j\in\mathcal{J},t\in\mathcal{T},\mathbf{d}^{t-1}\in\mathbf{D}^{t-1}\\
& \sum_{j\in\mathcal{J}}\left(q_j^{it,0}+\sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1}q_j^{it,\sigma\tau}d_\sigma^{i\tau}\right)\le s^{it}, \qquad \forall i\in\mathcal{I},t\in\mathcal{T},\mathbf{d}^{t-1}\in\mathbf{D}^{t-1}
\end{aligned}$$

$$\sum_{i\in\mathcal{I}}\left(x_j^{it,0}+q_j^{i,t-L_j^i,0}+\sum_{\sigma\in\mathcal{K}}\left(\sum_{\tau=1}^{t-1}x_j^{it,\sigma\tau}d_\sigma^{i\tau}+\sum_{\tau=1}^{t-L_j^i-1}q_j^{i,t-L_j^i,\sigma\tau}d_\sigma^{i\tau}\right)\right)\leq\bar{x}_j,\qquad\forall j\in\mathcal{J}_F,t\in\mathcal{T},\mathbf{d}^{t-1}\in\mathbf{D}^{t-1}$$

$$\sum_{i\in\mathcal{I}}\left(x_j^{it,0}+q_j^{i,t-L_j^i,0}+\sum_{j'\in\mathcal{J}_D\setminus\{j\}}u_{l,j'j}^{it,0}-\sum_{k\in\mathcal{K}_O}u_{e,jk}^{it,0}-\sum_{j\in\mathcal{J}_D\setminus\{j\}}u_{l,jj'}^{it,0}\right.$$
$$+\sum_{\sigma\in\mathcal{K}}\left(\sum_{\tau=1}^{t-1}\left(x_j^{it,\sigma\tau}+\sum_{j'\in\mathcal{J}_D\setminus\{j\}}u_{l,j'j}^{it,\sigma\tau}-\sum_{k\in\mathcal{K}_O}u_{e,jk}^{it,\sigma\tau}-\sum_{j\in\mathcal{J}_D\setminus\{j\}}u_{l,jj'}^{it,\sigma\tau}\right)d_\sigma^{i\tau}\right.$$
$$\left.\left.+\sum_{\tau=1}^{t-L_j^i-1}q_j^{i,t-L_j^i,\sigma\tau}d_\sigma^{i\tau}\right)\right)\leq\bar{x}_j,\qquad\forall j\in\mathcal{J}_D,t\in\mathcal{T},\mathbf{d}^{t-1}\in\mathbf{D}^{t-1}$$

$$x_j^{it,0}+q_j^{i,t-L_j^i,0}-\sum_{j'\in\mathcal{J}_D\setminus\{j\}}u_{l,jj'}^{it,0}+\sum_{\sigma\in\mathcal{K}}\left(\sum_{\tau=1}^{t-1}\left(x_j^{it,\sigma\tau}-\sum_{j'\in\mathcal{J}_D\setminus\{j\}}u_{l,jj'}^{it,\sigma\tau}\right)d_\sigma^{i\tau}\right.$$
$$\left.+\sum_{\tau=1}^{t-L_j^i-1}q_j^{i,t-L_j^i,\sigma\tau}d_\sigma^{i\tau}\right)\geq 0,\qquad\forall i\in\mathcal{I},j\in\mathcal{J}_D,t\in\mathcal{T},\mathbf{d}^{t-1}\in\mathbf{D}^{t-1}$$

$$\sum_{i\in\mathcal{I}}\left(y_k^{it,0}+\sum_{j\in\mathcal{J}_D}u_{e,jk}^{it,0}+\sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1}\left(y_k^{it,\sigma\tau}+\sum_{j\in\mathcal{J}_D}u_{e,jk}^{it,\sigma\tau}\right)d_\sigma^{i\tau}\right)\leq\bar{y}_k,\qquad\forall k\in\mathcal{K}_O,t\in\mathcal{T},\mathbf{d}^{t-1}\in\mathbf{D}^{t-1}$$

$$\sum_{j\in\mathcal{J}_D}r_{jk}^{it,0}+z_k^{it,0}=0,\qquad\forall i\in\mathcal{I},t\in\mathcal{T},k\in\mathcal{K}_D,$$

$$\sum_{j\in\mathcal{J}_D}r_{jk}^{it,\sigma\tau}+z_k^{it,\sigma\tau}=\begin{cases}1 & \text{if) } \tau=t,\sigma=k\\ 0 & \text{otherwise)}\end{cases},\qquad\forall i\in\mathcal{I},t\in\mathcal{T},k\in\mathcal{K}_D,\sigma\in\mathcal{K},\tau\in\{1,\ldots,t\}$$

$$v_{\rho,k}^{it,0}+z_k^{it,0}=0,\qquad\forall i\in\mathcal{I},t\in\mathcal{T},k\in\mathcal{K}_O,$$

$$v_{\rho,k}^{it,\sigma\tau}+z_k^{it,\sigma\tau}=\begin{cases}1 & \text{if) } \tau=t,\sigma=k\\ 0 & \text{otherwise)}\end{cases},\qquad\forall i\in\mathcal{I},t\in\mathcal{T},k\in\mathcal{K}_O,\sigma\in\mathcal{K},\tau\in\{1,\ldots,t\}$$

$$\sum_{j\in\mathcal{J}_F}v_{\eta,j}^{it,0}+z_{K+1}^{it,0}=0,\qquad\forall i\in\mathcal{I},t\in\mathcal{T},$$

$$\sum_{j\in\mathcal{J}_F}v_{\eta,j}^{it,\sigma\tau}+z_{K+1}^{it,\sigma\tau}=\begin{cases}1 & \text{if) } \tau=t,\sigma=K+1\\ 0 & \text{otherwise)}\end{cases},\qquad\forall i\in\mathcal{I},t\in\mathcal{T},\sigma\in\mathcal{K},\tau\in\{1,\ldots,t\}$$

$$x_j^{i,t+1,0}=x_j^{it,0}+q_j^{i,t-L_j^i,0}+$$

$$\sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,j'j}^{it,0} - \sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,jj'}^{it,0} - \sum_{k \in \mathcal{K}_O} u_{e,jk}^{it,0} - \sum_{k \in \mathcal{K}_D} r_{jk}^{it,0}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}$$

$$x_j^{i,t+1,\sigma\tau} = \begin{cases} x_j^{it,\sigma\tau} + q_j^{i,t-L_j^i,\sigma\tau} + \sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,j'j}^{it,\sigma\tau} - \sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,jj'}^{it,\sigma\tau} & \text{if) } \tau = 1, \ldots, t - L_j^i - 1 \\ - \sum_{k \in \mathcal{K}_O} u_{e,jk}^{it,\sigma\tau} - \sum_{k \in \mathcal{K}_D} r_{jk}^{it,\sigma\tau} & \\ x_j^{it,\sigma\tau} + \sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,j'j}^{it,\sigma\tau} - \sum_{j' \in \mathcal{J}_D \setminus \{j\}} u_{l,jj'}^{it,\sigma\tau} & \text{if) } \tau = t - L_j^i, \ldots, t-1 \\ - \sum_{k \in \mathcal{K}_O} u_{e,jk}^{it,\sigma\tau} - \sum_{k \in \mathcal{K}_D} r_{jk}^{it,\sigma\tau} & \\ - \sum_{k \in \mathcal{K}_D} r_{jk}^{it,\sigma\tau} & \text{if) } \tau = t \end{cases}$$

$$, \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$x_j^{i,t+1,0} = x_j^{it,0} + q_j^{i,t-L_j^i,0} - v_{\eta,j}^{it,0}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}$$

$$x_j^{i,t+1,\sigma\tau} = \begin{cases} x_j^{it,\sigma\tau} + q_j^{i,t-L_j^i,\sigma\tau} - v_{\eta,j}^{it,\sigma\tau}, & \text{if) } \tau = 1, \ldots, t - L_j^i - 1 \\ x_j^{it,\sigma\tau} - v_{\eta,j}^{it,\sigma\tau}, & \text{if) } \tau = t - L_j^i, \ldots, t-1 \\ -v_{\eta,j}^{it,\sigma\tau}, & \text{if) } \tau = t \end{cases}$$

$$, \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$y_k^{i,t+1,0} = y_k^{it,0} + \sum_{j \in \mathcal{J}_D} u_{e,jk}^{it,0} - v_{\rho,k}^{it,0}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}$$

$$y_k^{i,t+1,\sigma\tau} = \begin{cases} y_k^{it,\sigma\tau} + \sum_{j \in \mathcal{J}_D} u_{e,jk}^{it,\sigma\tau} - v_{\rho,k}^{it,\sigma\tau}, & \text{if) } \tau = 1, \ldots, t-1 \\ -v_{\rho,k}^{it,\sigma\tau}, & \text{if) } \tau = t \end{cases}$$

$$, \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}, \sigma \in \mathcal{K}, t \in \mathcal{T}, \tau \in \{1, \ldots, t\}$$

$$q_j^{it,0} + \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} q_j^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}$$

$$x_j^{it,0} + \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} x_j^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}^+, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}$$

$$y_k^{it,0} + \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} y_k^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}^+, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}$$

$$u_{l,jj'}^{it,0} + \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} u_{l,jj'}^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, j, j' \in \mathcal{J}_D, t \in \mathcal{T}, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}$$

$$u_{e,jk}^{it,0} + \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} u_{e,jk}^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, k \in \mathcal{K}_O, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}$$

$$v_{\rho,k}^{it,0} + \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t} v_{\rho,k}^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}, \mathbf{d}^{t} \in \mathbf{D}^{t}$$

$$v_{\eta,j}^{it,0} + \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t} v_{\eta,j}^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}, \mathbf{d}^{t} \in \mathbf{D}^{t}$$

$$r_{jk}^{it,0} + \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t} r_{jk}^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_D, t \in \mathcal{T}, \mathbf{d}^{t} \in \mathbf{D}^{t}$$

$$z_{k}^{it,0} + \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t} z_{k}^{it,\sigma\tau} d_\sigma^{i\tau} \geq 0, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}, \mathbf{d}^{t} \in \mathbf{D}^{t}$$

$$q_j^{it,0}, q_j^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$x_j^{it,0}, x_j^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}^+, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$y_k^{it,0}, y_k^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}^+, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$u_{l,jj'}^{it,0}, u_{l,jj'}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, j' \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$v_{\rho,k}^{it,0}, v_{\rho,k}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$v_{\eta,j}^{it,0}, v_{\eta,j}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$r_{jk}^{it,0}, r_{jk}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$z_{k}^{it,0}, z_{k}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

## C.2 Linear deterministic model of $\mathrm{P}_{\mathrm{LDR}}$

The constraints of $\mathrm{P}_{\mathrm{LDR}}$ hold $\mathbf{d}^t \in \mathbf{D}^t$. Therefore, $\mathrm{P}_{\mathrm{LDR}}$ cannot be solved directly using the commercial solver. By using the inner optimization and strong duality theory, we will transform $\mathrm{P}_{\mathrm{LDR}}$ to a linear deterministic model, which is tractable. In order to ease understanding for readers, we present how we transform the first constraint of $\mathrm{P}_{\mathrm{LDR}}$ to the tractable form in detail.

First, we present the first contraint of $\mathrm{P}_{\mathrm{LDR}}$ as

$$q_j^{it,0} + \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} q_j^{it,\sigma\tau} d_\sigma^{i\tau} \leq \bar{q}_j^i \bar{\delta}_j^{it}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}, \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}. \tag{C.1}$$

Constraint (C.1) can be equivalently to

$$q_j^{it,0} + \max_{d_\sigma^{i\tau}} \left( \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} q_j^{it,\sigma\tau} d_\sigma^{i\tau} \right) \leq \bar{q}_j^i \bar{\delta}_j^{it}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}, \forall \mathbf{d}^{t-1} \in \mathbf{D}^{t-1}. \tag{C.2}$$

Remember that $\mathrm{P}_{\mathrm{LDR}}$ is a minimization problem. Therefore, $\mathrm{P}_{\mathrm{LDR}}$ with Constraint (C.2) is the min-max problem. We convert it to the min-min problem by taking the dual of the inner optimization problem as follows:

$$\begin{aligned}
\max_{d_\sigma^{it}} \quad & \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} q_j^{it,\sigma\tau} d_\sigma^{i\tau} && \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \\
\text{s.t.} \quad & d_\sigma^{i\tau} \leq \bar{d}_\sigma^{i\tau} \qquad \forall \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\} \\
& -d_\sigma^{i\tau} \leq -\underline{d}_\sigma^{i\tau} \qquad \forall \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\} \\
\Longleftrightarrow & \\
\min_{\alpha_{a,jt}^{i\sigma\tau}, \beta_{a,jt}^{i\sigma\tau}} \quad & \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} \left( \alpha_{a,jt}^{i\sigma\tau} \bar{d}_\sigma^{i\tau} - \beta_{a,jt}^{i\sigma\tau} \underline{d}_\sigma^{i\tau} \right) && \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \\
\text{s.t.} \quad & \alpha_{a,jt}^{i\sigma\tau} - \beta_{a,jt}^{i\sigma\tau} = q_j^{it,\sigma\tau} \qquad \forall \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}
\end{aligned}$$

By strong duality, the above two problems have the same optimal objective value. Finally, Constraint (C.1) is transformed as the following two constraints:

$$q_j^{it,0} + \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} \left( \alpha_{a,jt}^{i\sigma\tau} \bar{d}_\sigma^{i\tau} - \beta_{a,jt}^{i\sigma\tau} \underline{d}_\sigma^{i\tau} \right) \leq \bar{q}_j^i \bar{\delta}_j^{it}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}$$

$$\alpha_{a,jt}^{i\sigma\tau} - \beta_{a,jt}^{i\sigma\tau} = q_j^{it,\sigma\tau} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}.$$

All inequality constraints of $\mathrm{P}_{\mathrm{LDR}}$ will be transformed in the same manner of the above procedure. We use notations $\alpha$ and $\beta$ to denote dual variables of inequality constraints of $\mathrm{P}_{\mathrm{LDR}}$. Finally, the linear deterministic model is presented as follows:

**Linear deterministic model**

$$
\begin{aligned}
\min \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}} \Bigg( & \sum_{j\in\mathcal{J}} \lambda_o^{it} c_o^{ij} \left( q_j^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1} q_j^{it,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) + \sum_{j\in\mathcal{J}} h_{x,j}^{it} \left( x_j^{i,t+1,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t} x_j^{i,t+1,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) \\
& + \sum_{k\in\mathcal{K}_O} h_{y,k}^{it} \left( y_k^{i,t+1,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t} y_k^{i,t+1,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) + \sum_{k\in\mathcal{K}} p_k^{it} \left( z_k^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t} z_k^{it,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) \\
& + \sum_{j\in\mathcal{J}_D}\sum_{j'\in\mathcal{J}_D} \lambda_l^{it} c_l^{jj'} \left( u_{l,jj'}^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1} u_{l,jj'}^{it,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) + \sum_{j\in\mathcal{J}_D}\sum_{k\in\mathcal{K}_O} \lambda_e^{it} c_e^{jk} \left( u_{e,jk}^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1} u_{e,jk}^{it,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) + \sum_{j\in\mathcal{J}_D}\sum_{k\in\mathcal{K}_D} \lambda_g^{it} c_g^{jk} \left( r_{jk}^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t} r_{jk}^{it,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) \\
& + \sum_{k\in\mathcal{K}_O} \rho_k^{it} \left( v_{\rho,k}^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t} v_{\rho,k}^{it,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) + \sum_{j\in\mathcal{J}_F} \eta_j^{it} \left( v_{\eta,j}^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t} v_{\eta,j}^{it,\sigma\tau} \hat{d}_\sigma^{i\tau} \right) \Bigg)
\end{aligned}
$$



$$
\begin{aligned}
\text{s.t.} \quad & q_j^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1} \left( \alpha_{a,jt}^{i\sigma\tau} \bar{d}_\sigma^{i\tau} - \beta_{a,jt}^{i\sigma\tau} \underline{d}_\sigma^{i\tau} \right) \le \bar{q}_j^i \bar{\delta}_j^{it}, \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}, t\in\mathcal{T} \\
& q_j^{it,\sigma\tau} - \alpha_{a,jt}^{i\sigma\tau} + \beta_{a,jt}^{i\sigma\tau} = 0, \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}, t\in\mathcal{T}, \sigma\in\mathcal{K}, \tau\in\{1,\dots,t-1\} \\
& \sum_{j\in\mathcal{J}} q_j^{it,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1} \left( \alpha_{b,t}^{i\sigma\tau} \bar{d}_\sigma^{i\tau} - \beta_{b,t}^{i\sigma\tau} \underline{d}_\sigma^{i\tau} \right) \le s^{it}, \qquad \forall i\in\mathcal{I}, t\in\mathcal{T} \\
& \sum_{j\in\mathcal{J}} q_j^{it,\sigma\tau} - \alpha_{b,t}^{i\sigma\tau} + \beta_{b,t}^{i\sigma\tau} = 0 \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}, \sigma\in\mathcal{K}, \tau\in\{1,\dots,t-1\} \\
& \sum_{i\in\mathcal{I}} \left( x_j^{it,0} + q_j^{i,t-L_j^i,0} + \sum_{\sigma\in\mathcal{K}}\sum_{\tau=1}^{t-1} \left( \alpha_{d,jt}^{i\sigma\tau} \bar{d}_\sigma^{i\tau} - \beta_{d,jt}^{i\sigma\tau} \underline{d}_\sigma^{i\tau} \right) \right) \le \bar{x}_j, \qquad \forall j\in\mathcal{J}_F, t\in\mathcal{T}
\end{aligned}
$$

$$x_j^{it,\sigma\tau} - \alpha_{d,jt}^{i\sigma\tau} + \beta_{d,jt}^{i\sigma\tau} = \begin{cases} -q_j^{i,t-L_j^i,\sigma\tau} & \text{if) } \tau = 1,\dots,t-L_j^i-1 \\ 0 & \text{if) } \tau = t-L_j^i,\dots,t-1 \end{cases} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1,\dots,t-1\}$$

$$\sum_{i\in\mathcal{I}} \left( x_j^{it,0} + q_j^{i,t-L_j^i,0} + \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,j'j}^{it,0} - \sum_{k\in\mathcal{K}_O} u_{e,jk}^{it,0} - \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,jj'}^{it,0} + \sum_{\sigma\in\mathcal{K}} \sum_{\tau=1}^{t-1} \left( \alpha_{e,jt}^{i\sigma\tau} \bar{d}_\sigma^{i\tau} - \beta_{e,jt}^{i\sigma\tau} \underline{d}_\sigma^{i\tau} \right) \right) \le \bar{x}_j,$$

$$\forall j \in \mathcal{J}_D, t \in \mathcal{T}$$

$$x_j^{it,\sigma\tau} + \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,j'j}^{it,\sigma\tau} - \sum_{k\in\mathcal{K}_O} u_{e,jk}^{it,\sigma\tau} - \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,jj'}^{it,\sigma\tau} - \alpha_{e,jt}^{i\sigma\tau} + \beta_{e,jt}^{i\sigma\tau} = \begin{cases} -q_j^{it-L_j^i,\sigma\tau} & \text{if) } \tau = 1,\dots,t-L_j^i-1 \\ 0 & \text{if) } \tau = t-L_j^i,\dots,t-1 \end{cases}$$

$$\forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1,\dots,t-1\}$$

$$x_j^{it,0} + q_j^{i,t-L_j^i,0} - \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,jj'}^{it,0} - \sum_{\sigma\in\mathcal{K}} \sum_{\tau=1}^{t-1} \left( \alpha_{f,jt}^{i\sigma\tau} \bar{d}_\sigma^{i\tau} - \beta_{f,jt}^{i\sigma\tau} \underline{d}_\sigma^{i\tau} \right) \ge 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}$$

$$x_j^{it,\sigma\tau} - \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,jj'}^{it,\sigma\tau} + \alpha_{f,jt}^{i\sigma\tau} - \beta_{f,jt}^{i\sigma\tau} = \begin{cases} -q_j^{i,t-Lj^i,\sigma\tau} & \text{if) } \tau = 1,\dots,t-L_j^i-1 \\ 0 & \text{if) } \tau = t-L_j^i,\dots,t-1 \end{cases}$$

$$\forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1,\dots,t-1\}$$

$$\sum_{i\in\mathcal{I}} \left( y_k^{it,0} + \sum_{j\in\mathcal{J}_D} u_{e,jk}^{it,0} + \sum_{\sigma\in\mathcal{K}} \sum_{\tau=1}^{t-1} \left( \alpha_{g,kt}^{i\sigma\tau} \bar{d}_\sigma^{i\tau} - \beta_{g,kt}^{i\sigma\tau} \underline{d}_\sigma^{i\tau} \right) \right) \le \bar{y}_k, \qquad \forall k \in \mathcal{K}_O, t \in \mathcal{T},$$

$$y_k^{it,\sigma\tau} + \sum_{j\in\mathcal{J}_D} u_{e,jk}^{it,\sigma\tau} - \alpha_{g,kt}^{i\sigma\tau} + \beta_{g,kt}^{i\sigma\tau} = 0, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1,\dots,t-1\}$$

$$\sum_{j\in\mathcal{J}_D} r_{jk}^{it,0} + z_k^{it,0} = 0, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_D,$$

$$\sum_{j\in\mathcal{J}_D} r_{jk}^{it,\sigma\tau} + z_k^{it,\sigma\tau} = \begin{cases} 1 & \text{if) } \tau = t, \sigma = k \\ \\ 0 & \text{otherwise)} \end{cases}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_D, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$v_{\rho,k}^{it,0} + z_k^{it,0} = 0, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_O$$

$$v_{\rho,k}^{it,\sigma\tau} + z_k^{it,\sigma\tau} = \begin{cases} 1 & \text{if) } \tau = t, \sigma = k \\ \\ 0 & \text{otherwise)} \end{cases}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}_O, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$\sum_{j\in\mathcal{J}_F} v_{\eta,j}^{it,0} + z_{K+1}^{it,0} = 0, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}$$

$$\sum_{j\in\mathcal{J}_F} v_{\eta,j}^{it,\sigma\tau} + z_{K+1}^{it,\sigma\tau} = \begin{cases} 1 & \text{if) } \tau = t, \sigma = K+1 \\ \\ 0 & \text{otherwise)} \end{cases}, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$x_j^{i,t+1,0} = x_j^{it,0} + q_j^{i,t-L_j^i,0} + \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,j'j}^{it,0} - \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,jj'}^{it,0} - \sum_{k\in\mathcal{K}_O} u_{e,jk}^{it,0} - \sum_{k\in\mathcal{K}_D} r_{jk}^{it,0}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}$$

$$x_j^{i,t+1,\sigma\tau} = \begin{cases} x_j^{it,\sigma\tau} + q_j^{i,t-L_j^i,\sigma\tau} + \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,j'j}^{it,\sigma\tau} - \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,jj'}^{it,\sigma\tau} - \sum_{k\in\mathcal{K}_O} u_{e,jk}^{it,\sigma\tau} - \sum_{k\in\mathcal{K}_D} r_{jk}^{it,\sigma\tau}, & \text{if) } \tau = 1, \ldots, t - L_j^i - 1 \\ \\ x_j^{it,\sigma\tau} + \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,j'j}^{it,\sigma\tau} - \sum_{j'\in\mathcal{J}_D\setminus\{j\}} u_{l,jj'}^{it,\sigma\tau} - \sum_{k\in\mathcal{K}_O} u_{e,jk}^{it,\sigma\tau} - \sum_{k\in\mathcal{K}_D} r_{jk}^{it,\sigma\tau} & \text{if) } \tau = t - L_j^i, \ldots, t-1 \\ \\ -\sum_{k\in\mathcal{K}_D} r_{jk}^{it,\sigma\tau}, & \text{if) } \tau = t \end{cases}$$

$$\forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$x_j^{i,t+1,0} = x_j^{it,0} + q_j^{i,t-L_j^i,0} - v_{\eta,j}^{it,0}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}$$

$$x_j^{i,t+1,\sigma\tau} = \begin{cases} x_j^{it,\sigma\tau} + q_j^{i,t-L_j^i,\sigma\tau} - v_{\eta,j}^{it,\sigma\tau}, & \text{if) } \tau = 1,\dots,t-L_j^i-1 \\ x_j^{it,\sigma\tau} - v_{\eta,j}^{it,\sigma\tau}, & \text{if) } \tau = t-L_j^i,\dots,t-1 \\ -v_{\eta,j}^{it,\sigma\tau}, & \text{if) } \tau = t \end{cases}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1,\dots,t\}$$

$$y_k^{i,t+1,0} = y_k^{it,0} + \sum_{j \in \mathcal{J}_D} u_{e,jk}^{it,0} - v_{\rho,k}^{it,0}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}$$

$$y_k^{i,t+1,\sigma\tau} = \begin{cases} y_k^{it,\sigma\tau} + \sum_{j \in \mathcal{J}_D} u_{e,jk}^{it,\sigma\tau} - v_{\rho,k}^{it,\sigma\tau}, & \text{if) } \tau = 1,\dots,t-1 \\ -v_{\rho,k}^{it,\sigma\tau}, & \text{if) } \tau = t \end{cases}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}, \sigma \in \mathcal{K}, t \in \mathcal{T}, \tau \in \{1,\dots,t\}$$

$$q_j^{it,0} - \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} \left( \alpha_{q,jt}^{i\sigma\tau} \bar{d}_\sigma^{i\tau} - \beta_{q,jt}^{i\sigma\tau} \underline{d}_\sigma^{i\tau} \right) \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}$$

$$q_j^{it,\sigma\tau} + \alpha_{q,jt}^{i\sigma\tau} - \beta_{q,jt}^{i\sigma\tau} = 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1,\dots,t-1\}$$

$$x_j^{it,0} - \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} \left( \alpha_{x,jt}^{i\sigma\tau} \bar{d}_\sigma^{i\tau} - \beta_{x,jt}^{i\sigma\tau} \underline{d}_\sigma^{i\tau} \right) \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}^+,$$

$$x_j^{it,\sigma\tau} + \alpha_{x,jt}^{i\sigma\tau} - \beta_{x,jt}^{i\sigma\tau} = 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}^+, \sigma \in \mathcal{K}, \tau \in \{1,\dots,t-1\}$$

$$y_k^{it,0} - \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} \left( \alpha_{y,kt}^{i\sigma\tau} \bar{d}_\sigma^{i\tau} - \beta_{y,kt}^{i\sigma\tau} \underline{d}_\sigma^{i\tau} \right) \geq 0, \qquad \forall k \in \mathcal{K}_O, i \in \mathcal{I}, t \in \mathcal{T}^+$$

$$y_k^{it,\sigma\tau} + \alpha_{y,kt}^{i\sigma\tau} - \beta_{y,kt}^{i\sigma\tau} = 0, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}^+, \sigma \in \mathcal{K}, \tau \in \{1,\dots,t-1\}$$

$$u_{l,jj'}^{it,0} - \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} \left( \alpha_{u,jj't}^{i\sigma\tau} \bar{d}_\sigma^{i\tau} - \beta_{u,jj't}^{i\sigma\tau} \underline{d}_\sigma^{i\tau} \right) \geq 0, \qquad \forall i \in \mathcal{I}, j, j' \in \mathcal{J}_D, t \in \mathcal{T},$$

$$u_{l,jj'}^{it,\sigma\tau} + \alpha_{u,jj't}^{i\sigma\tau} - \beta_{u,jj't}^{i\sigma\tau} = 0, \qquad \forall i \in \mathcal{I}, j, j' \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1,\dots,t-1\}$$

$$u_{e,jk}^{it,0} - \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t-1} \left( \alpha_{v,jkt}^{i\sigma\tau} \bar{d}_\sigma^{i\tau} - \beta_{v,jkt}^{i\sigma\tau} \underline{d}_\sigma^{i\tau} \right) \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_O, t \in \mathcal{T},$$

$$u_{e,jk}^{it,\sigma\tau} + \alpha_{v,jkt}^{i\sigma\tau} - \beta_{v,jkt}^{i\sigma\tau} = 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_O, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$v_{\rho,k}^{it,0} - \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t} \left( \alpha_{\rho,kt}^{i\sigma\tau} \bar{d}_{\sigma}^{i\tau} - \beta_{\rho,kt}^{i\sigma\tau} \underline{d}_{\sigma}^{i\tau} \right) \geq 0, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}$$

$$v_{\rho,k}^{it,\sigma\tau} + \alpha_{\rho,kt}^{i\sigma\tau} - \beta_{\rho,kt}^{i\sigma\tau} = 0, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$v_{\eta,j}^{it,0} - \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t} \left( \alpha_{\eta,jt}^{i\sigma\tau} \bar{d}_{\sigma}^{i\tau} - \beta_{\eta,jt}^{i\sigma\tau} \underline{d}_{\sigma}^{i\tau} \right) \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}$$

$$v_{\eta,j}^{it,\sigma\tau} + \alpha_{\eta,jt}^{i\sigma\tau} - \beta_{\eta,jt}^{i\sigma\tau} = 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$r_{jk}^{it,0} - \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t} \left( \alpha_{r,jkt}^{i\sigma\tau} \bar{d}_{\sigma}^{i\tau} - \beta_{r,jkt}^{i\sigma\tau} \underline{d}_{\sigma}^{i\tau} \right) \geq 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_D, t \in \mathcal{T},$$

$$r_{jk}^{it,\sigma\tau} + \alpha_{r,jkt}^{i\sigma\tau} - \beta_{r,jkt}^{i\sigma\tau} = 0, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$z_{k}^{it,0} - \sum_{\sigma \in \mathcal{K}} \sum_{\tau=1}^{t} \left( \alpha_{z,kt}^{i\sigma\tau} \bar{d}_{\sigma}^{i\tau} - \beta_{z,kt}^{i\sigma\tau} \underline{d}_{\sigma}^{i\tau} \right) \geq 0, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}$$

$$z_{k}^{it,\sigma\tau} + \alpha_{z,kt}^{i\sigma\tau} - \beta_{z,kt}^{i\sigma\tau} = 0, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, k \in \mathcal{K}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$q_{j}^{it,0}, q_{j}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$x_{j}^{it,0}, x_{j}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}^{+}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$y_{k}^{it,0}, y_{k}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}^{+}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$u_{l,jj'}^{it,0}, u_{l,jj'}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j, j' \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$u_{e,jk}^{it,0}, u_{e,jk}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_O, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$v_{\rho,k}^{it,0}, v_{\rho,k}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$v_{\eta,j}^{it,0}, v_{\eta,j}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$r_{jk}^{it,0}, r_{jk}^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$z_k^{it,0}, z_k^{it,\sigma\tau} \in \mathbb{R}, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$\alpha_{a,jt}^{i\sigma\tau}, \beta_{a,jt}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$\alpha_{b,t}^{i\sigma\tau}, \beta_{b,t}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$\alpha_{d,jt}^{i\sigma\tau}, \beta_{d,jt}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$\alpha_{e,jt}^{i\sigma\tau}, \beta_{e,jt}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$\alpha_{f,jt}^{i\sigma\tau}, \beta_{f,jt}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$\alpha_{g,kt}^{i\sigma\tau}, \beta_{g,kt}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$\alpha_{q,jt}^{i\sigma\tau}, \beta_{q,jt}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$\alpha_{x,jt}^{i\sigma\tau}, \beta_{x,jt}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}^+, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$\alpha_{y,kt}^{i\sigma\tau}, \beta_{y,kt}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}^+, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$\alpha_{u,jj't}^{i\sigma\tau}, \beta_{u,jj't}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, j, j' \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$\alpha_{v,jkt}^{i\sigma\tau}, \beta_{v,jkt}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_D, k \in \mathcal{K}_O, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t-1\}$$

$$\alpha_{\rho,kt}^{i\sigma\tau}, \beta_{\rho,kt}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_O, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$\alpha_{\eta,jt}^{i\sigma\tau}, \beta_{\eta,jt}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}_F, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$\alpha_{r,jkt}^{i\sigma\tau}, \beta_{r,jkt}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}_D, j \in \mathcal{J}_D, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

$$\alpha_{z,kt}^{i\sigma\tau}, \beta_{z,kt}^{i\sigma\tau} \in \mathbb{R}^+, \qquad \forall i \in \mathcal{I}, k \in \mathcal{K}, t \in \mathcal{T}, \sigma \in \mathcal{K}, \tau \in \{1, \ldots, t\}$$

## C.3 Experimental results on asymmetric demand distributions and different production capacities

Table C.1: Experimental results on asymmetric demand distributions.

| | | $Beta(2,5)$ | | | $Beta(5,2)$ | | | $Beta(1,6)$ | | | $Beta(6,1)$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $T=4$ | $T=7$ | $T=10$ | $T=4$ | $T=7$ | $T=10$ | $T=4$ | $T=7$ | $T=10$ | $T=4$ | $T=7$ | $T=10$ |
| DTPA | LDR($\times 10^2$) | 70.48 | 122.28 | 185.67 | 58.29 | 124.53 | 155.14 | 62.94 | 100.75 | 146.78 | 63.18 | 110.66 | 158.88 |
| | SIM($\times 10^2$) | 70.59 | 122.55 | 185.47 | 58.31 | 124.66 | 155.23 | 62.78 | 100.20 | 146.64 | 63.19 | 110.62 | 158.89 |
| | Gap(%) | 61.93 | 47.54 | 38.45 | 18.21 | 34.24 | 14.18 | 38.49 | 43.53 | 30.28 | 22.17 | 5.77 | 1.20 |
| | Std($\times 10^2$) | 3.66 | 3.60 | 5.32 | 1.26 | 1.99 | 1.83 | 2.20 | 4.41 | 5.06 | 0.87 | 0.69 | 0.43 |
| | CPU(s) | 2.00 | 5.49 | 21.06 | 2.06 | 8.27 | 17.55 | 2.00 | 5.39 | 13.05 | 1.95 | 4.90 | 19.25 |
| TPA | LDR($\times 10^2$) | 51.20 | 98.66 | 163.37 | 52.33 | 97.55 | 141.33 | 54.87 | 87.99 | 144.14 | 53.26 | 107.22 | 158.83 |
| | SIM($\times 10^2$) | 51.20 | 98.69 | 163.41 | 52.36 | 97.56 | 141.32 | 54.84 | 87.97 | 144.02 | 53.27 | 107.21 | 158.83 |
| | Gap(%) | 17.45 | 18.82 | 21.98 | 6.14 | 5.05 | 3.95 | 20.98 | 26.01 | 27.95 | 3.01 | 2.52 | 1.17 |
| | Std($\times 10^2$) | 0.26 | 0.41 | 3.44 | 0.30 | 0.43 | 0.48 | 0.45 | 0.39 | 4.51 | 0.17 | 0.34 | 0.43 |
| | CPU(s) | 8.05 | 28.33 | 128.54 | 9.93 | 83.20 | 681.72 | 7.93 | 30.56 | 74.80 | 8.95 | 30.70 | 127.49 |
| | $\phi^*$ | 0.8 | 0.4 | 0.6 | 0.0∼0.4 | 0.0∼0.4 | 0.0 | 0.8 | 0.4 | 0.8 | 0.0∼0.2 | 0.0∼0.6 | 0.0∼1.0 |
| DECOM | LDR($\times 10^2$) | 51.00 | 93.47 | 148.51 | 52.33 | 97.48 | 158.36 | 54.95 | 86.20 | 130.72 | 53.26 | 107.31 | 158.83 |
| | SIM($\times 10^2$) | 51.01 | 93.49 | 148.49 | 52.36 | 97.49 | 158.59 | 54.91 | 86.14 | 130.73 | 53.27 | 107.30 | 158.84 |
| | Gap(%) | 17.00 | 12.56 | 10.84 | 6.14 | 4.98 | 16.65 | 21.15 | 23.40 | 16.14 | 3.01 | 2.60 | 1.17 |
| | Std($\times 10^2$) | 0.27 | 0.40 | 0.61 | 0.30 | 0.46 | 2.75 | 0.45 | 2.55 | 0.80 | 0.17 | 0.35 | 0.43 |
| | CPU(s) | 7.41 | 21.34 | 66.52 | 6.50 | 19.86 | 46.14 | 7.12 | 17.35 | 64.40 | 7.18 | 30.94 | 70.16 |
| | $\phi^*$ | 0.6 | 0.8 | 0.6 | 0.0∼0.6 | 0.2 | 0.0 | 0.8 | 0.8 | 0.8 | 0.0∼0.8 | 0.0∼0.6 | 0.0∼1.0 |
| EVPI | ($\times 10^2$) | 43.59 | 83.06 | 133.97 | 49.33 | 92.87 | 135.96 | 45.33 | 69.81 | 112.56 | 51.72 | 104.58 | 157.00 |

Table C.2: Experimental results on different production capacities.

| | | | $(I, J_D, J_F) = (4, 4, 4)$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $\xi = 0.0$ | $\xi = 0.1$ | $\xi = 0.2$ | $\xi = 0.3$ | $\xi = 0.4$ | $\xi = 0.5$ | $\xi = 0.6$ | $\xi = 0.7$ | $\xi = 0.8$ | $\xi = 0.9$ | $\xi = 1.0$ |
| TPA | Gap(%) | | 7.90 | 8.48 | 8.58 | 8.44 | 8.29 | 8.39 | 8.60 | 8.13 | 7.66 | 11.97 | 20.16 |
| | Std($\times 10^2$) | | 0.80 | 0.76 | 0.75 | 0.76 | 0.78 | 0.83 | 0.82 | 0.87 | 0.91 | 1.10 | 1.85 |
| | CPU(s) | Phase1 | 216.45 | 39.44 | 1528.14 | 19.99 | 9.13 | 150.13 | 114.43 | 346.35 | 450.53 | 3852.61 | 163.15 |
| | | Phase2 | 30.88 | 30.02 | 30.92 | 30.28 | 32.25 | 33.32 | 31.30 | 30.46 | 29.62 | 32.08 | 33.41 |
| | | Total | 247.33 | 69.46 | 1559.07 | 50.27 | 41.38 | 183.45 | 145.73 | 376.81 | 480.15 | 3884.68 | 196.56 |
| DECOM | Gap(%) | | 8.18 | 7.97 | 8.28 | 8.31 | 7.77 | 8.02 | 7.75 | 7.89 | 7.45 | 10.65 | 19.92 |
| | Std($\times 10^2$) | | 0.79 | 0.78 | 0.76 | 0.76 | 0.78 | 0.82 | 0.89 | 0.89 | 0.91 | 0.99 | 1.74 |
| | CPU(s) | Phase1 | 3.52 | 7.13 | 3.43 | 2.92 | 2.17 | 4.23 | 3.47 | 9.56 | 4.61 | 278.74 | 20.57 |
| | | Phase2 | 21.39 | 19.71 | 19.80 | 19.09 | 19.61 | 20.15 | 18.18 | 19.77 | 18.75 | 19.43 | 20.73 |
| | | Total | 24.90 | 26.83 | 23.23 | 22.01 | 21.78 | 24.38 | 21.65 | 29.33 | 23.37 | 298.17 | 41.31 |
| EVPI | ($\times 10^2$) | | 113.63 | 113.74 | 113.79 | 113.85 | 114.08 | 114.25 | 114.41 | 115.08 | 115.89 | 116.46 | 117.56 |

| | | | $(I, J_D, J_F) = (5, 5, 5)$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $\xi = 0.0$ | $\xi = 0.1$ | $\xi = 0.2$ | $\xi = 0.3$ | $\xi = 0.4$ | $\xi = 0.5$ | $\xi = 0.6$ | $\xi = 0.7$ | $\xi = 0.8$ | $\xi = 0.9$ | $\xi = 1.0$ |
| TPA | Gap(%) | | 9.06 | 9.21 | 9.55 | 9.27 | 8.32 | 8.23 | 8.06 | 7.39 | 6.93 | 10.66 | 20.95 |
| | Std($\times 10^2$) | | 0.91 | 0.94 | 0.96 | 0.94 | 1.02 | 1.07 | 1.04 | 1.10 | 1.11 | 1.46 | 2.47 |
| | CPU(s) | Phase1 | 3604.53 | 3616.16 | 3606.61 | 1049.77 | 79.11 | 3620.62 | 118.20 | 2849.45 | 2769.65 | 7209.42 | 7200.74 |
| | | Phase2 | 76.75 | 65.27 | 75.73 | 69.06 | 67.04 | 63.71 | 70.76 | 67.27 | 66.01 | 65.87 | 74.38 |
| | | Total | 3681.28 | 3681.43 | 3682.34 | 1118.83 | 146.15 | 3684.33 | 188.96 | 2916.72 | 2835.66 | 7275.29 | 7275.12 |
| DECOM | Gap(%) | | 8.68 | 9.22 | 9.13 | 8.33 | 8.70 | 7.83 | 7.75 | 7.28 | 6.88 | 9.54 | 20.41 |
| | Std($\times 10^2$) | | 0.90 | 0.94 | 0.91 | 1.04 | 0.95 | 1.08 | 1.08 | 1.09 | 1.09 | 1.24 | 2.62 |
| | CPU(s) | Phase1 | 6.88 | 18.50 | 6.69 | 6.57 | 6.39 | 8.41 | 10.23 | 6.32 | 4.29 | 3601.00 | 3601.40 |
| | | Phase2 | 51.05 | 52.37 | 45.72 | 49.22 | 46.94 | 44.97 | 46.26 | 48.68 | 48.49 | 43.25 | 53.58 |
| | | Total | 57.93 | 70.86 | 52.41 | 55.78 | 53.33 | 53.38 | 56.49 | 54.99 | 52.77 | 3644.25 | 3654.97 |
| EVPI | ($\times 10^2$) | | 134.21 | 134.34 | 134.45 | 134.78 | 135.26 | 135.83 | 136.16 | 136.88 | 137.85 | 138.32 | 139.27 |

## C.4 Shapes of symmetric and asymmetric demand distributions

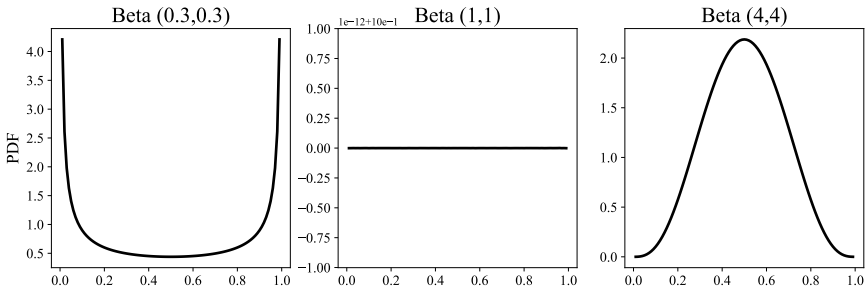



Figure C.1: Shapes of symmetric demand distributions utilized in Experiment 1

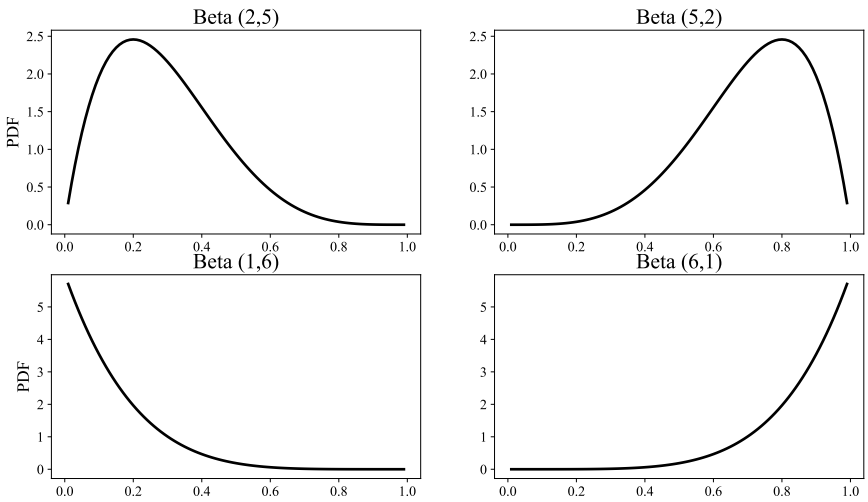



Figure C.2: Shapes of asymmetric demand distributions utilized in Experiment 1

# Bibliography


[1] H. Abouee-Mehrizi, O. Berman, and S. Sharma, *Optimal joint replenishment and transshipment policies in a multi-period inventory system with lost sales*, Operations Research, 63 (2015), pp. 342–350.

[2] A. Abouelrous, A. F. Gabor, and Y. Zhang, *Optimizing the inventory and fulfillment of an omnichannel retailer: a stochastic approach with scenario clustering*, Computers and Industrial Engineering, 173 (2022), p. 108723.

[3] S. C. Allen, *Redistribution of total stock over several user locations*, Naval Research Logistics Quarterly, 5 (1958), pp. 337–345.

[4] K. J. Arrow, T. Harris, and J. Marschak, *Optimal inventory policy*, Econometrica: Journal of the Econometric Society, (1951), pp. 250–272.

[5] A. Azaron, U. Venkatadri, and A. Farhang Doost, *Designing profitable and responsive supply chains under uncertainty*, International Journal of Production Research, 59 (2021), pp. 213–225.

[6] V. Azizi, G. Hu, and M. Mokari, *A two-stage stochastic programming model for multi-period reverse logistics network design with lot-sizing*, Computers and Industrial Engineering, 143 (2020), p. 106397.

[7] H. Badri, M. Bashiri, and T. H. Hejazi, *Integrated strategic and tactical planning in a supply chain network design with a heuristic solution method*, Computers and Operations Research, 40 (2013), pp. 1143–1154.

[8] M. Bashiri, H. Badri, and J. Talebi, *A new approach to tactical and strategic planning in production–distribution networks*, Applied Mathematical Modelling, 36 (2012), pp. 1703–1717.

[9] D. R. Bell, S. Gallino, and A. Moreno, *Offline showrooms in omnichannel retail: Demand and operational benefits*, Management Science, 64 (2018), pp. 1629–1651.

[10] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust optimization*, vol. 28, Princeton university press, 2009.

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

[12] A. Ben-Tal and A. Nemirovski, *Robust solutions of uncertain linear programs*, Operations Research Letters, 25 (1999), pp. 1–13.

[13] J. F. Benders, *Partitioning procedures for solving mixed-variables programming problems*, Numerische Mathematik, 4 (1962), pp. 238–252.

[14] D. Bertsekas, *Dynamic programming and optimal control: Volume I*, vol. 1, Athena scientific, 2012.

[15] D. Bertsimas and C. Caramanis, *Adaptability via sampling*, in 2007 46th IEEE Conference on Decision and Control, IEEE, 2007, pp. 4717–4722.

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

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

[18] D. Bertsimas, V. Goyal, and B. Y. Lu, *A tight characterization of the performance of static solutions in two-stage adjustable robust linear optimization*, Mathematical Programming, 150 (2015), pp. 281–319.

[19] D. Bertsimas, D. A. Iancu, and P. A. Parrilo, *A hierarchy of near-optimal policies for multistage adaptive optimization*, IEEE Transactions on Automatic Control, 56 (2011), pp. 2809–2824.

[20] D. Bertsimas and M. Sim, *The price of robustness*, Operations Research, 52 (2004), pp. 35–53.

[21] D. Bertsimas and A. Thiele, *A robust optimization approach to inventory theory*, Operations Research, 54 (2006), pp. 150–168.

[22] J. R. Birge and F. Louveaux, *Introduction to stochastic programming*, Springer Science & Business Media, 2011.

[23] R. N. Boute, J. Gijsbrechts, W. van Jaarsveld, and N. Vanvuchelen, *Deep reinforcement learning for inventory control: A roadmap*, European Journal of Operational Research, (2021).

[24] R. A. Broekmeulen and K. H. Van Donselaar, *A heuristic to manage perishable inventory with batch ordering, positive lead-times, and time-varying demand*, Computers and Operations Research, 36 (2009), pp. 3013–3018.

[25] E. V. Bulinskaya, *Some results concerning optimum inventory policies*, Theory of Probability & Its Applications, 9 (1964), pp. 389–403.

[26] J. Burton and A. Banerjee, *Cost-parametric analysis of lateral transshipment policies in two-echelon supply chains*, International Journal of Production Economics, 93 (2005), pp. 169–178.

[27] Y.-J. Cai and C. K. Lo, *Omni-channel management in the new retailing era: A systematic review and future research agenda*, International Journal of Production Economics, 229 (2020), p. 107729.

[28] M. C. Carissimi and A. Creazza, *The role of the enabler in sharing economy service triads: A logistics perspective*, Cleaner Logistics and Supply Chain, 5 (2022), p. 100077.

[29] S. Ceschia, M. Gansterer, S. Mancini, and A. Meneghetti, *The on-demand warehousing problem*, International Journal of Production Research, (2022), pp. 1–19.

[30] W. Chen and M. Sim, *Goal-driven optimization*, Operations Research, 57 (2009), pp. 342–357.

[31] X. Chen and J. Zhang, *A stochastic programming duality approach to inventory centralization games*, Operations Research, 57 (2009), pp. 840–851.

[32] T. Cheong, *Joint inventory and transshipment control for perishable products of a two-period lifetime*, The International Journal of Advanced Manufacturing Technology, 66 (2013), pp. 1327–1341.

[33] P. Christodoulou, *Soft actor-critic for discrete action settings*, arXiv preprint arXiv:1910.07207, (2019).

[34] M. A. Cohen, *Analysis of single critical number ordering policies for perishable inventories*, Operations Research, 24 (1976), pp. 726–741.

[35] S. Daniel, *Amazon: The World' s Most Powerful Economic and Cultural Force*. `https://www.investing.com/academy/statistics/amazon-facts/`, 2023. [Online; accessed 05-May-2023].

[36] B. J. De Moor, J. Gijsbrechts, and R. N. Boute, *Reward shaping to improve the performance of deep reinforcement learning in perishable inventory management*, European Journal of Operational Research, 301 (2022), pp. 535–545.

[37] M. Dehghani and B. Abbasi, *An age-based lateral-transshipment policy for perishable items*, International Journal of Production Economics, 198 (2018), pp. 93–103.

[38] M. Dehghani, B. Abbasi, and F. Oliveira, *Proactive transshipment in the blood supply chain: A stochastic programming approach*, Omega, 98 (2021), p. 102112.

[39] Q. Duan and T. W. Liao, *A new age-based replenishment policy for supply chain inventory optimization of highly perishable products*, International Journal of Production Economics, 145 (2013), pp. 658–671.

[40] F. Dunke, I. Heckmann, S. Nickel, and F. Saldanha-da Gama, *Time traps in supply chains: Is optimal still good enough?*, European Journal of Operational Research, 264 (2018), pp. 813–829.

[41] A. Emelogu, S. Chowdhury, M. Marufuzzaman, L. Bian, and B. Eksioglu, *An enhanced sample average approximation method for stochastic optimization*, International Journal of Production Economics, 182 (2016), pp. 230–252.

[42] S. Emma, M. Utpal, and B. Beth, *A Study of 46,000 Shoppers Shows That Omnichannel Retailing Works*. `https://hbr.org/2017/01/a-study-of-46000-shoppers-shows-that-omnichannel-retailing-works`, 2017. [Online; accessed 05-May-2023].

[43] M. Fattahi and K. Govindan, *Integrated forward/reverse logistics network design under uncertainty with pricing for collection of used products*, Annals of Operations Research, 253 (2017), pp. 193–225.

[44] M. Fattahi, M. Mahootchi, and S. Moattar Husseini, *Integrated strategic and tactical supply chain planning with price-sensitive demands*, Annals of Operations Research, 242 (2016), pp. 423–456.

[45] S. S. Fazeli, S. Venkatachalam, R. B. Chinnam, and A. Murat, *Two-Stage Stochastic Choice Modeling Approach for Electric Vehicle Charg-*

*ing Station Network Design in Urban Communities*, IEEE Transactions on Intelligent Transportation Systems, 22 (2021), pp. 3038–3053.

[46] M. Firouz, B. B. Keskin, and S. H. Melouk, *An integrated supplier selection and inventory problem with multi-sourcing and lateral transshipments*, Omega, 70 (2017), pp. 77–93.

[47] S. Gallino, A. Moreno, and I. Stamatopoulos, *Channel integration, sales dispersion, and inventory management*, Management Science, 63 (2017), pp. 2813–2831.

[48] F. Gao and X. Su, *Online and offline information for omnichannel retailing*, Manufacturing & Service Operations Management, 19 (2017), pp. 84–98.

[49] A. M. Geoffrion and G. W. Graves, *Multicommodity distribution system design by benders decomposition*, Management Science, 20 (1974), pp. 822–844.

[50] A. Georghiou, W. Wiesemann, and D. Kuhn, *Generalized decision rule approximations for stochastic programming via liftings*, Mathematical Programming, 152 (2015), pp. 301–338.

[51] S. M. Ghorashi Khalilabadi, S. H. Zegordi, and E. Nikbakhsh, *A multi-stage stochastic programming approach for supply chain risk mitigation via product substitution*, Computers and Industrial Engineering, 149 (2020), p. 106786.

[52] J. Gijsbrechts, R. N. Boute, J. A. Van Mieghem, and D. J. Zhang, *Can deep reinforcement learning improve inventory management? perfor-*

*mance on lost sales, dual-sourcing, and multi-echelon problems*, Manufacturing & Service Operations Management, (2022).

[53] I. Gilboa and D. Schmeidler, *Maxmin expected utility with non-unique prior*, Journal of Mathematical Economics, 18 (1989), pp. 141–153.

[54] J. Goedhart, R. Haijema, and R. Akkerman, *Modelling the influence of returns for an omni-channel retailer*, European Journal of Operational Research, 306 (2023), pp. 1248–1263.

[55] J. Goldberg, *E-Commerce Sales Grew 50% to $ 870 Billion During The Pandemic*. `https://www.forbes.com/sites/jasongoldberg/2022/02/18/e-commerce-sales-grew-50-to-870-billion-during-the-pandemic/`, 2022. "[Online; accessed 04-July-2022]".

[56] A. Gosavi, *Reinforcement learning: A tutorial survey and recent advances*, INFORMS Journal on Computing, 21 (2009), pp. 178–192.

[57] K. Govindan, M. Fattahi, and E. Keyvanshokooh, *Supply chain network design under uncertainty: A comprehensive review and future research directions*, European Journal of Operational Research, 263 (2017), pp. 108–141.

[58] A. Govindarajan, A. Sinha, and J. Uichanco, *Joint inventory and fulfillment decisions for omnichannel retail networks*, Naval Research Logistics (NRL), 68 (2021), pp. 779–794.

[59] M. Gupta, S. Tiwari, and C. K. Jaggi, *Retailer' s ordering policies for time-varying deteriorating items with partial backlogging and permissible delay*

*in payments in a two-warehouse environment*, Annals of Operations Research, 295 (2020), pp. 139–161.

[60] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, *Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor*, in International conference on machine learning, PMLR, 2018, pp. 1861–1870.

[61] T. Haarnoja, A. Zhou, K. Hartikainen, G. Tucker, S. Ha, J. Tan, V. Kumar, H. Zhu, A. Gupta, P. Abbeel, et al., *Soft actor-critic algorithms and applications*, arXiv preprint arXiv:1812.05905, (2018).

[62] R. Haijema, *A new class of stock-level dependent ordering policies for perishables with a short maximum shelf life*, International Journal of Production Economics, 143 (2013), pp. 434–439.

[63] R. Haijema and S. Minner, *Stock-level dependent ordering of perishables: A comparison of hybrid base-stock and constant order policies*, International Journal of Production Economics, 181 (2016), pp. 215–225.

[64] ——, *Improved ordering of perishables: The value of stock-age information*, International Journal of Production Economics, 209 (2019), pp. 316–324.

[65] R. Haijema, J. van der Wal, and N. M. van Dijk, *Blood platelet production: Optimization by dynamic programming and simulation*, Computers and Operations Research, 34 (2007), pp. 760–779.

[66] G. A. Hanasusanto, D. Kuhn, and W. Wiesemann, *K-adaptability in two-stage robust binary programming*, Operations Research, 63 (2015), pp. 877–891.

[67] M. Hass, *On-demand Warehousing 101: Top 10 Locations for Warehousing & eCommerce Fulfillment.* https://www.flexe.com/articles/on-demand-warehousing-101-top-10-warehousing-fulfillment-locations/, 2019. [Online; accessed 23-August-2021].

[68] H. Hasselt, *Double q-learning*, Advances in neural information processing systems, 23 (2010).

[69] S. Hong and J. Lee, *Competition in Korean dawn, same-day delivery market intensifies with new players jumping in.* https://pulsenews.co.kr/view.php?year=2022&no=234432, 2022. [Online; accessed 23-November-2022].

[70] S. Hong and S. Lee, *Overnight delivery race spills over to big and foreign names in Korea.* https://pulsenews.co.kr/view.php?sc=30800028&year=2022&no=537584&mc=, 2022. [Online; accessed 23-November-2022].

[71] S. Huang and S. Ontañón, *A closer look at invalid action masking in policy gradient algorithms*, arXiv preprint arXiv:2006.14171, (2020).

[72] S. Jiu, *Robust omnichannel retail operations with the implementation of ship-from-store*, Transportation Research Part E: Logistics and Transportation Review, 157 (2022), p. 102550.

[73] C. Jonas, *Optimizing a two-warehouse system under shortage backordering, trade credit, and decreasing rental conditions*, International Journal of Production Economics, 209 (2019), pp. 147–155.

[74] Kaggle, *Brazilian E-Commerce Public Dataset by Olist.* https://www.kaggle.com/olistbr/brazilian-ecommerce/, 2018. [Online; accessed 23-August-2021].

[75] W. Kang and J. Lee, *Korea' s overnight fresh food delivery firm Oasis seeks Kosdaq listing.* https://pulsenews.co.kr/view.php?sc=30800028&year=2022&no=381446, 2022. [Online; accessed 23-November-2022].

[76] A. Kara and I. Dogan, *Reinforcement learning approaches for specifying ordering policies of perishable inventory systems*, Expert Systems with Applications, 91 (2018), pp. 150–158.

[77] M. Khatami, M. Mahootchi, and R. Z. Farahani, *Benders' decomposition for concurrent redesign of forward and closed-loop supply chain network with demand and return uncertainties*, Transportation Research Part E: Logistics and Transportation Review, 79 (2015), pp. 1–21.

[78] F. Kiya and H. Davoudpour, *Stochastic programming approach to redesigning a warehouse network under uncertainty*, Transportation Research Part E: Logistics and Transportation Review, 48 (2012), pp. 919–936.

[79] A. J. Kleywegt, A. Shapiro, and T. Homem-de Mello, *The Sample Average Approximation Method for Stochastic Discrete Optimization*, SIAM Journal on Optimization, 12 (2002), pp. 479–502.

[80] A. Klose and A. Drexl, *Facility location models for distribution system design*, European Journal of Operational Research, 162 (2005), pp. 4–29.

[81] R. Kong, L. Luo, L. Chen, and M. F. Keblis, *The effects of bops implementation under different pricing strategies in omnichannel retailing*, Transportation Research Part E: Logistics and Transportation Review, 141 (2020), p. 102014.

[82] A. Kranenburg and G.-J. Van Houtum, *A new partial pooling structure for spare parts networks*, European Journal of Operational Research, 199 (2009), pp. 908–921.

[83] G. Lai, H. Liu, W. Xiao, and X. Zhao, "*fulfilled by amazon*": *A strategic perspective of competition at the e-commerce platform*, Manufacturing & Service Operations Management, 24 (2022), pp. 1406–1420.

[84] J. Lee, C. Ko, and I. Moon, *E-commerce supply chain network design using on-demand warehousing system under uncertainty*, International Journal of Production Research, (2022), pp. 1–27.

[85] J. Lee, J. Park, and I. Moon, *Inventory and commitment decisions for on-demand warehousing system*, in IFIP International Conference on Advances in Production Management Systems,Springer,Cham, 2021, pp. 455–463.

[86] T. Lee, *Retailers ditch dawn deliveries in favor of more profitable options*. `https://koreajoongangdaily.joins.com/2022/10/03/business/industry/korea-delivery-grocery/20221003171548624.html`, 2022. [Online; accessed 23-November-2022].

[87] Y. H. Lee, J. W. Jung, and Y. S. Jeon, *An effective lateral transshipment policy to improve service level in the supply chain*, International Journal of Production Economics, 106 (2007), pp. 115–126.

[88] M. Leonard, *How the pandemic drove retailers to on-demand warehousing.* `https://www.supplychaindive.com/news/pandemic-covid-on-demand-warehouse-flowspace-flexe-inventory/598150/`, 2021. "[Online; accessed 23-August-2021]".

[89] G. Li, X. He, J. Zhou, and H. Wu, *Pricing, replenishment and preservation technology investment decisions for non-instantaneous deteriorating items*, Omega, 84 (2019), pp. 114–126.

[90] Q. Li, P. Yu, and L. Du, *Separation of perishable inventories in offline retailing through transshipment*, Operations Research, (2021).

[91] Y. Li, Y. Liao, X. Hu, and W. Shen, *Lateral transshipment with partial request and random switching*, Omega, 92 (2020), p. 102134.

[92] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, *Continuous control with deep reinforcement learning*, arXiv preprint arXiv:1509.02971, (2015).

[93] Y. F. Lim, S. Jiu, and M. Ang, *Integrating anticipative replenishment allocation with reactive fulfillment for online retailing using robust optimization*, Manufacturing & Service Operations Management, 23 (2021), pp. 1616–1633.

[94] Y. F. Lim and C. Wang, *Inventory management based on target-oriented robust optimization*, Management Science, 63 (2017), pp. 4409–4427.

[95] R. Manzini and E. Gebennini, *Optimization models for the dynamic facility location and allocation problem*, International Journal of Production Research, 46 (2008), pp. 2061–2086.

[96] A. Marandi and D. Den Hertog, *When are static and adjustable robust optimization problems with constraint-wise uncertainty equivalent?*, Mathematical Programming, 170 (2018), pp. 555–568.

[97] J. Meissner and O. V. Senicheva, *Approximate dynamic programming for lateral transshipment problems in multi-location inventory systems*, European Journal of Operational Research, 265 (2018), pp. 49–64.

[98] M. T. Melo, S. Nickel, and F. S. Da Gama, *Dynamic multi-commodity capacitated facility location: a mathematical modeling framework for strategic supply chain planning*, Computers and Operations Research, 33 (2006), pp. 181–208.

[99] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, *Asynchronous methods for deep reinforcement learning*, in International conference on machine learning, PMLR, 2016, pp. 1928–1937.

[100] S. Nahmias, *A comparison of alternative approximations for ordering perishable inventory*, INFOR: Information Systems and Operational Research, 13 (1975), pp. 175–184.

[101] D. Nakandala, H. Lau, and P. K. Shum, *A lateral transshipment model for perishable inventory management*, International Journal of Production Research, 55 (2017), pp. 5341–5354.

[102] V. I. Norkin, G. C. Pflug, and A. Ruszczyński, *A branch and bound method for stochastic global optimization*, Mathematical Programming, 83 (1998), pp. 425–450.

[103] F. Nur, M. Aboytes-Ojeda, K. K. Castillo-Villar, and M. Marufuzzaman, *A two-stage stochastic programming model for biofuel supply chain network design with biomass quality implications*, IISE Transactions, 53 (2021), pp. 845–868.

[104] A. Oroojlooyjadid, M. Nazari, L. V. Snyder, and M. Takáč, *A deep q-network for the beer game: Deep reinforcement learning for inventory optimization*, Manufacturing & Service Operations Management, 24 (2022), pp. 285–304.

[105] J. Park, I. Dayarian, and B. Montreuil, *Showcasing optimization in omnichannel retailing*, European Journal of Operational Research, 294 (2021), pp. 895–905.

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

[107] S. K. Paul, M. A. Moktadir, K. Sallam, T.-M. Choi, and R. K. Chakrabortty, *A recovery planning model for online business operations*

*under the covid-19 outbreak*, International Journal of Production Research, (2021), pp. 1–23.

[108] J. A. Pazour and K. Unnu, *On the unique features and benefits of on-demand distribution models*, 15th IMHRC Proceedings (Savannah, Georgia. USA – 2018), 13 (2018).

[109] K. Pichka, L. C. Alwan, and X. Yue, *Fulfillment and pricing optimization for omni-channel retailers considering shipment of in-store demand*, Transportation Research Part E: Logistics and Transportation Review, 167 (2022), p. 102912.

[110] D. Ponce, I. Contreras, and G. Laporte, *E-commerce shipping through a third-party supply chain*, Transportation Research Part E: Logistics and Transportation Review, 140 (2020), p. 101970.

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

[112] W. B. Powell, *A unified framework for stochastic optimization*, European Journal of Operational Research, 275 (2019), pp. 795–821.

[113] X. Qin, Z. Liu, and L. Tian, *The strategic analysis of logistics service sharing in an e-commerce platform*, Omega, 92 (2020), p. 102153.

[114] R. Qiu, L. Ma, and M. Sun, *A robust omnichannel pricing and ordering optimization approach with return policies based on data-driven support vector*

*clustering*, European Journal of Operational Research, 305 (2023), pp. 1337–1354.

[115] R. Rahmaniani, T. G. Crainic, M. Gendreau, and W. Rei, *The benders decomposition algorithm: A literature review*, European Journal of Operational Research, 259 (2017), pp. 801–817.

[116] S. Ren, T.-M. Choi, K.-M. Lee, and L. Lin, *Intelligent service capacity allocation for cross-border-e-commerce related third-party-forwarding logistics operations: A deep learning approach*, Transportation Research Part E: Logistics and Transportation Review, 134 (2020), p. 101834.

[117] J. K. Ryan, D. Sun, and X. Zhao, *Competition and coordination in online marketplaces*, Production and Operations Management, 21 (2012), pp. 997–1014.

[118] T. Santoso, S. Ahmed, M. Goetschalckx, and A. Shapiro, *A stochastic programming approach for supply chain network design under uncertainty*, European Journal of Operational Research, 167 (2005), pp. 96–115.

[119] T. Schaul, J. Quan, I. Antonoglou, and D. Silver, *Prioritized experience replay*, arXiv preprint arXiv:1511.05952, (2015).

[120] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, *Proximal policy optimization algorithms*, arXiv preprint arXiv:1707.06347, (2017).

[121] P. Schütz, A. Tomasgard, and S. Ahmed, *Supply chain design under uncertainty using sample average approximation and dual decomposition*, European Journal of Operational Research, 199 (2009), pp. 409–419.

[122] C.-T. See and M. Sim, *Robust approximation to multiperiod inventory management*, Operations Research, 58 (2010), pp. 583–594.

[123] A. Shapiro and A. Nemirovski, *On complexity of stochastic programming problems*, Continuous optimization: Current trends and modern applications, (2005), pp. 111–146.

[124] Y. Shi, X. Guo, and Y. Yu, *Dynamic warehouse size planning with demand forecast and contract flexibility*, International Journal of Production Research, 56 (2018), pp. 1313–1325.

[125] Y. Shi, Y. Yu, and Y. Dong, *Warehousing platform' s revenue management: A dynamic model of coordinating space allocation for self-use and rent*, European Journal of Operational Research, 293 (2021), pp. 167–176.

[126] Y. Shin, S. Lee, and I. Moon, *Robust multiperiod inventory model considering trade-in program and refurbishment service: Implications to emerging markets*, Transportation Research Part E: Logistics and Transportation Review, 138 (2020), p. 101932.

[127] L. V. Snyder and Z.-J. M. Shen, *Fundamentals of supply chain theory*, John Wiley & Sons, 2019.

[128] A. L. Soyster, *Convex programming with set-inclusive constraints and applications to inexact linear programming*, Operations Research, 21 (1973), pp. 1154–1157.

[129] N. N. Sultana, H. Meisheri, V. Baniwal, S. Nath, B. Ravindran, and H. Khadilkar, *Reinforcement learning for multi-product multi-node inventory management in supply chains*, arXiv preprint arXiv:2006.04037, (2020).

[130] W. Sun, Y. Zou, X. Zhang, N. Guo, B. Zhang, and G. Du, *High robustness energy management strategy of hybrid electric vehicle based on improved soft actor-critic deep reinforcement learning*, Energy, (2022), p. 124806.

[131] G. Tagaras and D. Vlachos, *Effectiveness of stock transshipment under various demand distributions and nonnegligible transshipment times*, Production and Operations Management, 11 (2002), pp. 183–198.

[132] P. N. Thanh, N. Bostel, and O. Péton, *A dynamic model for facility location in the design of complex supply chains*, International Journal of Production Economics, 113 (2008), pp. 678–693.

[133] Z. Tian and G. Zhang, *Multi-echelon fulfillment warehouse rent and production allocation for online direct selling*, Annals of Operations Research, 304 (2021), pp. 427–451.

[134] S. Tiwari, C. K. Jaggi, A. K. Bhunia, A. A. Shaikh, and M. Goh, *Two-warehouse inventory model for non-instantaneous deteriorating items with stock-dependent demand and inflation using particle swarm optimization*, Annals of Operations Research, 254 (2017), pp. 401–423.

[135] F. Tornese, K. Unnu, M. Gnoni, and J. Pazour, *On-demand warehousing: main features and business models*, In XXV Summer School "Francesco Turo" - Industrial Systems Engineering, (2020).

[136] K. Unnu and J. A. Pazour, *Analyzing varying cost structures of alternative warehouse strategies*, in IIE Annual Conference. Proceedings, Institute of Industrial and Systems Engineers (IISE), 2019, pp. 480–485.

[137] ——, *Evaluating on-demand warehousing via dynamic facility location models*, IISE Transactions, (2022), pp. 1–16.

[138] G. Van der Heide, P. Buijs, K. J. Roodbergen, and I. F. Vis, *Dynamic shipments of inventories in shared warehouse and transportation networks*, Transportation Research Part E: Logistics and Transportation Review, 118 (2018), pp. 240–257.

[139] L. van Hezewijk, N. Dellaert, T. Van Woensel, and N. Gademann, *Using the proximal policy optimisation algorithm for solving the stochastic capacitated lot sizing problem*, International Journal of Production Research, (2022), pp. 1–24.

[140] N. Vanvuchelen, J. Gijsbrechts, and R. Boute, *Use of proximal policy optimization for the joint replenishment problem*, Computers in Industry, 119 (2020), p. 103239.

[141] B. Verweij, S. Ahmed, A. J. Kleywegt, G. Nemhauser, and A. Shapiro, *The sample average approximation method applied to stochastic*

*routing problems: a computational study*, Computational Optimization and Applications, 24 (2003), pp. 289–333.

[142] K.-M. Wang and Z.-J. Ma, *Age-based policy for blood transshipment during blood shortage*, Transportation Research Part E: Logistics and Transportation Review, 80 (2015), pp. 166–183.

[143] Z. Wang, Y. Dai, S.-C. Fang, Z.-Z. Jiang, and Y. Xu, *Inventory transshipment game with limited supply: Trap or treat*, Naval Research Logistics (NRL), 67 (2020), pp. 383–403.

[144] C. J. Watkins and P. Dayan, *Q-learning*, Machine Learning, 8 (1992), pp. 279–292.

[145] Y. Wei, M. Yang, J. Chen, L. Liang, and T. Ding, *Dynamic lateral transshipment policy of perishable foods with replenishment and recycling*, Computers and Industrial Engineering, 172 (2022), p. 108574.

[146] W. Xiao and Y. Xu, *Should an online retailer penalize its independent sellers for stockout?*, Production and Operations Management, 27 (2018), pp. 1124–1132.

[147] İ. Yanikoğlu, B. L. Gorissen, and D. den Hertog, *A survey of adjustable robust optimization*, European Journal of Operational Research, 277 (2019), pp. 799–813.

[148] B. Zahiri, S. A. Torabi, M. Mohammadi, and M. Aghabegloo, *A multi-stage stochastic programming approach for blood supply chain planning*, Computers and Industrial Engineering, 122 (2018), pp. 1–14.

[149] C. Zhang, T. Ayer, C. C. White, J. N. Bodeker, and J. D. Roback, *Inventory sharing for perishable products: Application to platelet inventory management in hospital blood banks*, Operations Research, (2022).

[150] X. Zhen, D. Shi, Y. Li, and C. Zhang, *Manufacturer' s financing strategy in a dual-channel supply chain: Third-party platform, bank, and retailer credit financing,* Transportation Research Part E: Logistics and Transportation Review, 133 (2020), p. 101820.

[151] X. Zhen and S. Xu, *Who should introduce the third-party platform channel under different pricing strategies?*, European Journal of Operational Research, 299 (2022), pp. 168–182.

[152] Q. Zhou, Y. Yang, and S. Fu, *Deep reinforcement learning approach for solving joint pricing and inventory problem with reference price effects*, Expert Systems with Applications, 195 (2022), p. 116564.

[153] Z. Zhu, K. Lin, and J. Zhou, *Transfer learning in deep reinforcement learning: A survey*, arXiv preprint arXiv:2009.07888, (2020).

# 국문초록

통신 기술이 발전하고 비대면 수요가 증가함에 따라 이커머스 시장은 최근 몇 년 동안 크게 성장하였으며, 이커머스 유통업체의 수 또한 증가하였다. 하지만 이커머스 시장의 경쟁과열과 수요의 불확실성으로 발생하는 높은 운영 비용으로 인하여 많은 유통업체가 어려움을 겪고 있다. 이러한 어려움을 극복하는 방안으로, 공유 경제는 유연한 물류 운영을 위한 비즈니스 모델로 주목받고 있다. 본 학위논문에서는 불확실성하에서 물류 자원 공유를 고려한 의사결정 모델을 개발하는 것을 목표로 한다. 또한, 공급망 관리 연구 분야와 관련된 다음의 세 가지 문제를 다룬다: (1) 소멸성 상품의 재고 관리, (2) 공급망 네트워크 설계, (3) 주문, 할당 및 배송 결정. 그리고 제시된 문제들에서 물류 자원을 공유하기 위한 세 가지 서비스 및 전략을 고려한다.

첫 번째로, 환적과 온라인-오프라인 채널 시스템을 동시에 고려한 소멸성 상품의 재고 관리 문제를 다룬다. 온라인-오프라인 채널 시스템의 특성을 고려한 마르코프 의사결정 모델을 제시한다. 또한, 마르코프 의사결정 모델에서 발생할 수 있는 차원의 저주를 극복하기 위해 소프트 액터 크리틱 알고리즘에 기반한 하이브리드 심층 강화학습 알고리즘을 개발한다. 그리고 환적을 통해 제품의 폐기 비용을 줄일 수 있는 효과를 확인한다. 두 번째로, 온디맨드 창고 시스템을 고려한 공급망 네트워크 설계 문제를 다룬다. 불확실성을 고려하고 제안된 문제를 모형화하기 위해 2단계 추계적 수리 모델을 제시한다. 표본 평균 근사법과 벤더스 분해법을 결합하여 제안된 문제를 해결한다. 특히, 효과적인 초기 절단면을 생성하여 벤더스 분해법의 수렴 속도를 증가시키는 방법을 개발한다. 세 번째로, 제3자 플랫폼의 판매 채널을 고려한 주문, 할당 및 배송 문제를 다룬다. 유통업체의 공급망과 제3자 플랫폼 기업의 공급망을 동시에 고려한 추계적 최적화 모형을 고려한다. 추계적 최적화 모형을 다루기 힘든 어려움을

해결하기 위해, 2단계 강건 최적화에 기반한 분해 기법을 제안한다. 실험적 결과를 통해 개발된 분해 기법이 대규모 문제들에서도 좋은 성능을 보이는 것을 확인한다.

**주요어**: 물류 자원 공유, 이커머스, 공급망관리, 강화학습, 추계적 계획법, 강건최적화
**학번**: 2021-34229

# 감사의 글

산업공학에 대한 학문적 궁금증을 해결하고 더 넓은 환경에서 공부하겠다는 일념으로 관악에 온 지 어느덧 4년 반의 시간이 흘렀습니다. 설렘에 가득 차고 아무것도 몰랐던 대학생이, 이제는 한 명의 성인이자 연구자로서 졸업을 앞두고 있습니다. 4년 반의 시간을 돌이켜보면 힘든 시간도 많았고 지칠 때도 많았지만, 제 인생에서 제일 행복했던 시간이었습니다. 행복한 시간을 만들어 주고 박사과정을 무사히 마칠 수 있게 도와주었던 분들에게 감사의 마음을 전하고자 합니다.

먼저, 산업공학에 대한 기초가 부족했던 저를 받아주신 문일경 교수님께 깊이 감사드립니다. 교수님을 만나 공급망관리 연구실의 일원일 될 수 있었던 것은 최고의 행운이었습니다. 제가 하는 연구에 대해 큰 격려와 응원을 해주시고, 지지해 주신 것에 대해 진심으로 감사드립니다. 공급망관리 연구실에 들어오기 전에 저는 예의범절이 약한 어린아이였지만, 교수님께서 해주셨던 따끔한 충고와 진심 어린 조언으로 인해 인성적으로 한층 더 성장할 수 있었습니다. 최고의 자리에 계심에도 불구하고 항상 최선을 다하시고 학생들을 배려해 주시는 교수님의 모습은 저에게 큰 가르침이 되었습니다. 교수님께서 몸소 보여주셨던 주인의식, 책임의식들을 마음속에 새기며 졸업하더라도 교수님께 부끄럽지 않은 제자가 되도록 노력하겠습니다. 항상 존경하고 감사합니다, 교수님.

본 학위 논문의 심사를 맡아 주신 이재욱 교수님, 박건수 교수님, 장윤석 교수님, 김병수 교수님께 깊은 감사의 마음을 전합니다. 이재욱 교수님, 교수님의 수업을 통해 통계 및 확률의 엄밀한 정의에 관해 공부할 수 있었습니다. 교수님의 조언으로 인해 통계적 유의성에 따른 연구 결과 도출에 대해서 한 번 더 고민할 수 있었습니다. 박건수 교수님, 교수님 수업을 통해 배웠던 확률적 재고모형, 환적을 고려한 재고모형을 통해서 본 학위논문의 연구 주제들을 설정할 수 있었습니다. 또한 진심 어린 조언과

세부적인 사항까지 되짚어 주셨던 세심함에 다시 한번 감사드립니다. 장윤석 교수님, 교수님께서 해주신 조언으로 학위논문의 흐름을 개선하고 각 챕터 간의 연관성을 강조할 수 있었습니다. 바쁘신 와중에도 진로에 대해 좋은 말씀 해주신 것에 대해 진심으로 감사드립니다. 김병수 교수님, 교수님께서 해주신 날카로운 조언으로 연구의 가정사항에 대해서 세심히 되짚을 수 있었습니다. 또한 교수님께서 진로와 인생 방향 설정에 대해 말씀해 주셨던 따뜻한 말씀들이 저에게 큰 힘이 되었습니다. 심사위원 교수님들뿐만 아니라 서울대학교 산업공학과 모든 교수님께도 진심으로 감사드립니다. 교수님들의 훌륭한 수업으로 인하여 연구 역량을 차근차근 쌓아 갈 수 있었고, 연구의 의미에 대해서 깊이 고민해 보면서 정진할 수 있었습니다.

저에게 공급망 관리 연구실 동료들은 항상 미안하고 감사했던 소중한 사람들입니다. 연구실에 처음 들어와서 많이 헤맬 때 따뜻이 도와주었던 선배들, 그리고 동기 세원이에게 감사의 마음을 전합니다. 학위 과정 동안 힘든 점, 궁금한 점에 관해서 물어볼 때 자기 일처럼 대답해 줬던 세원이와 선배들이 없었더라면 더 많은 어려움을 겪었을 것입니다. 그리고 선배들께서 연구에 대해 해줬던 소중한 조언들로 인하여 학위 과정을 무사히 마칠 수 있었습니다. 특히, 항상 의지할 수 있었고 귀감으로 삼을 수 있었던 영철이형, 윤제형, 광이형, 종민이형, 민정이에게 깊은 감사를 표합니다. 그리고 부족한 저를 묵묵히 따라와 주고 저의 고민을 들어주던 자랑스러운 후배들에게도 감사한 마음을 전하고 싶습니다. 이제는 연구실의 버팀목이 된 건우, 준석, 지연, 성배가 저보다 더 영리하고 책임감이 있기 때문에 연구실의 미래에 대한 걱정 없이 졸업할 수 있습니다. 남들은 어렵고 힘들다고 하던 대학원 생활을, 좋은 추억과 행복한 기억들만 남기고 연구실을 떠날 수 있게 해준 모든 공급망 관리 연구실 동료들에게 진심으로 감사합니다. 항상 건강하며 모두의 앞날에 행복만 있기를 기원합니다. 자주 연락하고 소중한 인연의 끈을 놓지 않도록 노력하겠습니다.

마지막으로 30년의 긴 세월 동안 저를 지지하고 사랑해 주셨던 가족들에게 감사한 마음을 전합니다. 어머니의 따뜻한 응원과 격려로 힘든 시간을 헤쳐 나가고 이겨낼

수 있었습니다. 주말에 어머니가 해주셨던 식사와 소소한 대화들이 그다음 주를 헤쳐나갈 수 있는 원동력이 되었습니다. 겉으로는 강하게 말씀하시지만 속으로는 응원을 많이을 해주시는 아버지 덕분에 내적으로 단단한 사람이 될 수 있었습니다. 아버지가 우리 가족을 위해 헌신하신 점에 대해 항상 감사히 생각하며, 저도 아버지처럼 훌륭한 사람이 될 수 있도록 노력하겠습니다. 그리고 연구, 진로 등에 대한 유용한 팁들을 주며 저를 배려한 형 덕분에 여러 고민을 잘 해결할 수 있었습니다. 형한테 많이 의지할 수 있었고 어린 동생의 철없는 장난들을 받아준 것에 대해 미안하며 감사합니다. 앞으로도 자랑스러운 아들 그리고 동생이 되도록 노력하며, 우리 가족에게 항상 좋은 일만 있고 건강하기를 기원합니다. 사랑합니다.

2023년 8월

이준혁 올림