Robust Inventory Management under Supply and Demand Uncertainties

ROBUST INVENTORY MANAGEMENT UNDER SUPPLY AND DEMAND UNCERTAINTIES

BY

JIE CHU, M.Eng., M.S., B.Eng.

A Thesis
Submitted to the DeGroote School of Business
and School of Graduate Studies
of McMaster University
in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

Ph.D. Dissertation (2018)

DeGroote School of Business, McMaster University

Hamilton, Ontario, Canada

TITLE:	Robust Inventory Management under Supply and Demand Uncertainties
AUTHOR:	Jie Chu DeGroote School of Business McMaster University, Hamilton, Canada
SUPERVISOR:	Dr. Kai Huang
COMMITTEE MEMBERS:	Dr. Kai Huang Dr. Mahmut Parlar Dr. Elkafi Hassini
NUMBER OF PAGES:	xi, 113

Abstract

In this thesis, we study three periodic-review, finite-horizon inventory systems in the presence of supply and demand uncertainties. In the first part of the thesis, we study a multi-period single-station problem in which supply uncertainty is modeled by partial supply. Formulating the problem under a robust optimization (RO) framework, we show that solving the robust counterpart is equivalent to solving a nominal problem with a modified deterministic demand sequence. In particular, in the stationary case the optimal robust policy follows the *quasi*-(s, S) form and the corresponding s and S levels are theoretically computable. In the second part of the thesis, we extend the RO framework to a multi-period multi-echelon problem. We show that for a tree structure network, decomposition applies so that the optimal single-station robust policy remains valid for each echelon in the tree. Furthermore, if there are no setup costs in the network, then the problem can be decomposed into several uncapacitated single-station problems with new cost parameters subject to the deterministic demands. In the last part of the thesis, we consider a periodic-review Assemble-To-Order (ATO) system with multiple components and multiple products, where the inventory replenishment for each component follows an independent base-stock policy and product demands are satisfied according to a First-Come-First-Served (FCFS) rule. We jointly consider the inventory replenishment and component allocation problems in the ATO system under stochastic component replenishment lead times and stochastic product demands. The problems are formulated under the stochastic programming (SP) framework, which are difficult to solve exactly due to a large number of scenarios. We use the sample average approximation (SAA) algorithms to find near-optimal solutions, which accuracy is verified by the numerical experiment results.

Acknowledgements

I would like to express my sincere gratitude to all those who have contributed to this thesis and supported me throughout my Ph.D. journey.

First and foremost, I am extremely grateful to my supervisor, Dr. Kai Huang, for his valuable guidance, support, inspiration, and encouragement over the past five years. He has guided me in selecting the thesis topic and continuously helped me develop my research skills. It has been a great honor and pleasure to work and study under his supervision. His advices on both research and career have been invaluable to me. I feel incredibly lucky having him as my doctoral supervisor.

Besides my supervisor, I would like to express my deep gratitude to my committee members, Dr. Mahmut Parlar and Dr. Elkafi Hassini. I truly appreciate all of the time and attention that they have put into this thesis. With their insightful comments and stimulating suggestions, the quality of this thesis has been improved significantly. I would also like to thank Dr. Parkash Abad and Dr. Manish Verma for their high quality Ph.D. courses in the first two years. I have benefited greatly from their classes. I also thank Dr. Yun Zhou for sharing his experience of job searching with me. Thank you my fellow doctoral students, especially Ali, Davod and those who have graduated (Alireza, Hangfei, Majid, and Sophia), for having always been so friendly and supportive. Thanks go to all other DeGroote School of Business members for their kindness and help.

Finally, I would like to express my profound gratitude to my family and friends for their unconditional love and support throughout my years of study. In particular, I would like to thank my loving wife Lulu for her support, great patience, and encouragement at all these years. I am so fortunate to meet her at the beginning of this journey.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation

Inventory management under uncertainty has been extensively studied in the past. There is a large body of literature on finding optimal inventory policies for systems under demand uncertainty while assuming no uncertainty on the supply side. For example, the base-stock policy (i.e., order-up-to level) has been proven to be optimal for a serial inventory system subject to uncertain demand with a known distribution (Clark and Scarf, 1960). The optimality of base-stock policy for more general inventory systems can be found in Zipkin (2000).

As global supply chains have been growing considerably, supply uncertainty that often arises from higher variability in overseas suppliers performance can adversely impact the overall performance of a supply chain, which suggests a need to simultaneously incorporate supply and demand uncertainties into decision-making. Supply uncertainty has been modeled in several different forms in the production and inventory management literature. *Yield* uncertainty, as one major form of supply uncertainty that occurs in the production process, typically models the situations where partial supply frequently happens. The reader is referred to Yano and Lee (1995) for a thorough literature review on yield uncertainty.

In contrast, *supply disruption* and *lead time* uncertainties are closely related to variabilities in the process of inventory replenishment. Supply disruption refers to partial (or complete) inoperation of the entities within a supply chain (e.g., road, truck and facility) due to terrorist attacks or major natural disasters (e.g., earthquake, tornado and flood). Such uncertainty is highly unpredictable and may significantly influence the entire supply chain when it occurs, so many stochastic programming models have been proposed to allow for recourse actions (see e.g., Ozbay and Ozguven, 2007).

Lead time uncertainty is usually concerned with unexpected shipment (or production) delay. It is noteworthy that analyzing inventory problems under uncertain lead time could be cumbersome because of a phenomenon called "*order crossing*". That is, orders may arrive in a different sequence than the one in which they were initially placed (He et al., 1998). To avoid such complexity, most research papers on this issue have assumed no order cross when there are stochastic lead times. Nevertheless, Srinivasan et al. (2011) pointed out that in practice the order crossing is inevitable due to widely employed multi-sourcing and just-in-time (JIT) strategies.

Our goal in this thesis is to provide general modeling frameworks and efficiently compute optimal policies for the periodic-review, finite-horizon inventory systems in the presence of supply and demand uncertainties. Specifically, we study three periodic-review, finite-horizon inventory systems subject to uncertain supply and uncertain demand, they are: (1) single-station system, (2) multi-echelon system, and (3) Assemble-To-Order (ATO) system.

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

1.2 Contributions and Organization of the Thesis

In this thesis, we formulate single-station and multi-echelon inventory problems subject to uncertain supply and uncertain demand using the RO formulation, while the ATO system under stochastic replenishment lead times and stochastic product demands is formulated under the SP formulation.

In Chapter 2, we review advances in the RO and SP literatures with their applications in the inventory management. Also, we briefly describe a RO framework for linear programming (LP) problems with row-wise uncertainty as well as a general form of a two-stage SP, which will be used in this thesis.

In Chapter 3, we first study multi-period single-station problem subject to uncertain supply and uncertain demand, where supply uncertainty is modeled by partial supply. By restricting supply ratio and demand to budget (polyhedral) uncertain sets, we formulate the problem under the RO framework of Bertsimas and Sim (2003, 2004). We show that the computational burden of the robust counterpart is not much higher than the corresponding nominal problem. We also provide theoretical results regarding the optimal inventory policy of the robust counterparts. Specifically, we show that solving the robust counterpart amounts to solving a nominal problem with a modified deterministic demand sequence. In the stationary case, the optimal robust policy is *quasi*-(s, S), where s and S levels are theoretically computable. In addition, we consider the capacitated single-station problems and investigate how the optimal robust policy is affected. The numerical results show that the proposed robust policy could significantly outperform the nominal policy and the robust policy of Bertsimas and Thiele (2006) in the average performance when different cost parameters and demand distributions are considered.

In Chapter 4, we extend the RO framework to a multi-period multi-echelon inventory problem with a tree structure, in which we assume that the supply uncertainty only affects orders placed by main storage hubs. Although the proposed model is more complicated than that of the single-station problem in terms of the problem size, it still can be solved within polynomial time to optimality if there are no setup costs. We provide theoretical insights into the optimal robust policy. We show that the optimal robust policy is decomposable into those of robust single-station problems. Specifically, the problem can be decomposed into several interconnected single-station problems with (or without) time-varying capacity on orders. If there are no setup costs, then the problem can be decomposed into several uncapacitated single-station problems with the deterministic demands. The numerical results suggest that the significant benefits in terms of cost savings and performance stability can be realized by incorporating both supply and demand uncertainties.

In Chapter 5, we consider a multi-product, multi-component, periodic-review ATO system that simultaneously incorporates stochastic replenishment lead times and stochastic product demands. The system enforces an independent base-stock policy and a First-Come-First-Served (FCFS) allocation rule. We first consider the circumstance where the decision maker has full knowledge of the realized lead times and propose a two-stage stochastic integer program to jointly optimize the base-stock levels and component allocation. Subsequently, we advance the methodology by considering a more general situation where the decision maker only has full distributional knowledge of the random lead times and propose a multi-stage stochastic integer program for the joint optimization. To solve the proposed models, we use the sample average approximation (SAA) algorithms. The effectiveness of the SAA solutions is demonstrated through the tightness of the gaps between lower and upper bounds yielded by the algorithms. The benefit of incorporating lead time uncertainty into the base-stock optimization is evaluated by simulation in a comparison with the base-stock levels of Akçay and Xu (2004) where the deterministic lead times are assumed.

We summarize the major contributions of this thesis, as well as some directions for the future research in Chapter 6.

Chapter 2

Literature Review

This chapter introduces robust optimization (RO) and stochastic programming (SP) approaches, and discusses their applications in the production and inventory management literature. In particular, we present a RO framework specifically tailored for linear programming (LP) problems subject to row-wise uncertainty and a general form of a two-stage SP model, as they will be used in this thesis.

2.1 Robust Optimization Literature

There are two major RO modeling methodologies in the literature, i.e., the scenario-based RO and set-based RO. It is worth noting that these two RO methods are very distinct in nature. The scenario-based RO was developed in Mulvey et al. (1995), in which a finite set of scenarios was considered and the probability distribution of the scenarios was known in advance, and it focuses on the expected performance. The formulation explicitly allows for the constraint violations under some of the scenarios, and such violations are penalized in the objective function. Thus, the trade-off between solution and model robustness can be quantitatively measured. However, this approach suffers from tractability issues as the size of the problem increases.

As an alternative, the set-based RO provides a framework to address the issue of data uncertainty when incomplete distributional information is provided. This RO assumes a bounded, convex uncertainty set for uncertain parameters. The formulation aims to minimize the maximum cost (or maximize the minimum revenue), where the maximum cost (or minimum revenue) is computed over the uncertainty set. As opposed to the scenario-based RO, set-based RO enforces feasibility of the optimal solution for all values of the uncertain parameters within the uncertainty set.

The set-based RO was pioneered by Soyster (1973), who considered column-wise convex uncertainty sets for a LP problem and sought a feasible solution for all realizations in the sets. The author showed that the problem is equivalent to another LP problem where each uncertain parameter equals its worst-case value within the set. This leads to an extremely conservative solution, which largely impedes its practical implementation. To tackle the over-conservative issue, Ben-Tal and Nemirovski (1999) proposed a RO formulation with row-wise ellipsoidal uncertainty sets. Although the RO with ellipsoidal uncertainty sets provides less conservative solutions than Soyster (1973), solving the resulting conic quadratic programs, that is, the robust counterparts of LPs, could be computationally burdensome.

More recently, Bertsimas and Sim (2003, 2004) developed a RO framework with row-wise polyhedral uncertainty sets in the context of LP problems, when limited distributional information is provided (i.e., mean and standard deviation). In the RO, the uncertain parameters are assumed to belong to the symmetric intervals centered at their mean (or nominal) values. Realizing that it is unlikely that all uncertainty parameters will take their worst-case value within the intervals, a pre-determined parameter called *budget-of-uncertainty* is imposed to rule out large cumulative deviations. Therefore, this RO has also been referred to as the RO with budget (polyhedral) uncertainty sets. The authors further proved the existence of probabilistic bounds against the constraint violations. A most attractive feature of the RO is that the computational burden of RO is usually not much higher than the original problem. For example, the robust counterpart of a LP remains a LP (Bertsimas and Sim, 2003, 2004).

Let's consider the following LP problem:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}_i'\mathbf{x} \le b_i, \qquad \forall i, \\ & \mathbf{x} \ge 0, \end{aligned} \tag{2.1}$$

where $\mathbf{c}$, $\mathbf{A}_i$ are row vectors of size n, $(\cdot)'$ is the vector transpose, b_i is a constant, and $\mathbf{x}$ is a vector of n non-negative variables. We assume that the entries a_{ij}, $j = 1, \ldots, n$, of the vector $\mathbf{A}_i$ are uncertain parameters and may vary in the interval $[\bar{a}_{ij} - \hat{a}_{ij}, \bar{a}_{ij} + \hat{a}_{ij}]$. Let $z_{ij} = (a_{ij} - \bar{a}_{ij})/\hat{a}_{ij}$ denote the *scaled deviation* such that $a_{ij} = \bar{a}_{ij} + \hat{a}_{ij}z_{ij}$ and $z_{ij} \in [-1, 1]$ for all i, j. It is assumed that the scaled deviations belong to the following budget uncertainty set:

$$\mathcal{Z}_i = \{-1 \le z_{ij} \le 1,\ \sum_{j=1}^{n} |z_{ij}| \le \Gamma_i,\ \forall j \le n\}, \quad \forall i \tag{2.2}$$

where the parameter Γ_i represents the budget-of-uncertainty for constraint i, which varies in the interval $[0, n]$. If $\Gamma_i = 0$, it means all a_{ij} take their nominal values, namely $\bar{a}_{ij}$; if $\Gamma_i = n$, it allows all a_{ij} to take their worst-case values, namely $\bar{a}_{ij} \pm \hat{a}_{ij}$.

Bertsimas and Sim (2003, 2004) showed that the LP problem (2.1) with the uncertainty set (2.2) is equivalent to another LP problem:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & \sum_{j=1}^{n} \bar{a}_{ij}x_j + p_i\Gamma_i + \sum_{j=1}^{n} q_{ij} \le b_i, && \forall i, \\ & p_i + q_{ij} \ge \hat{a}_{ij}, && \forall i, j, \\ & p_i \ge 0,\ q_{ij} \ge 0, && \forall i, j, \end{aligned} \tag{2.3}$$

$$x_j \geq 0, \qquad \forall j,$$

where p_i and q_{ij} are dual variables associated with the constraints in the uncertainty set (2.2). This is the set-based RO that we will use in Chapter 3 and 4. We adopt this approach because of its computational efficiency and the effectiveness of its solution.

The set-based RO has been used in a wide range of operations management applications such as scheduling (Bohle et al., 2010), humanitarian logistics (Paul and Wang, 2015), capacity expansion (Ordóñez and Zhao, 2007), portfolio management (Bertsimas and Pachamanova, 2008), and pricing (Thiele, 2006). Among others, we are particularly interested in the prior works that used the RO with budget uncertainty sets in the inventory management area.

In Bertsimas and Thiele (2006), the authors first applied the RO of Bertsimas and Sim (2003, 2004) to classical multi-period inventory problems under demand uncertainty. Furthermore, they showed that in the case of stationary costs the robust counterparts are equivalent to the nominal problems with a modified deterministic demand in each time period. They also concluded that the optimal robust policies follow the (s, S) form. The quality of the robust policies was highlighted through numerical studies when compared to those obtained from dynamic programming and a myopic approach when erroneous distributions were assumed.

The methodology is subsequently applied to the sawmill planning problems with yield uncertainty in Alvarez and Vera (2014). Aouam and Brahimi (2013) formulated an integrated production planning with order acceptance problem under the RO framework, where customer demands are distinguished by marginal revenue and subject to uncertainty. We also note applications of the RO in the inventory management literature that involve multi-source uncertainty such as Alem and Morabito (2012) (uncertain costs and demands), Varas et al. (2014) (uncertain raw material supplies and demands), and Wei et al. (2011) (uncertain returns and demands).

Although the derived robust counterparts in the above papers exhibit computational tractability, the hidden over-conservative issue of the RO approach has been investigated in Thiele (2010). To alleviate such the conservativeness, the author suggested to either examine the efficiency of the budget-of-uncertainty through simulation, or use smaller polyhedral sets. Bienstock and ÖZbay (2008) addressed this over-conservativeness by solving a "true" min-max single-station inventory problem of Bertsimas and Thiele (2006), where the base-stock policy was enforced. The proposed Benders' decomposition algorithms perform well under two demand models (risk budget demand and bursty demand). Nevertheless, solving the problem with Benders' decomposition algorithms may require extensive computational efforts.

We refer to the RO discussed thus far as the static RO approach since it does not allow for recourse actions as time evolves. Ben-Tal et al. (2004) developed an adjustable RO (ARO) approach to account for recourse actions, thus the underlying problems are solved in a dynamic fashion. The ARO produces less conservative solutions than the static RO, but it is computationally intractable in general. We notice that many efforts have been made to derive tractable approximations for the ARO in the inventory management literature (e.g., Ben-Tal et al., 2005, 2009; See and Sim, 2010; Solyalı et al., 2015). In addition, researchers have recently investigated the performance of affine policies in multistage adjustable optimization. In particular, Bertsimas et al. (2010) proved the optimality of disturbance-affine control policies for one-dimensional, constrained, multistage robust optimization and Bertsimas et al. (2011) introduced a hierarchy of near-optimal polynomial control policies for linear dynamical systems subject to uncertainty. For more details on the ARO and its recent advances, the reader is referred to Gabrel et al. (2014).

2.2 Stochastic Programming Literature

SP is a scenario-based approach that is well-suited for situations where the probability distribution for the underlying uncertain parameters is known with certainty. For example, the probability for scenario i in which we observe that the realization of demand d_i is equal to p_i for $i = 1, 2, \ldots, I$, where I represents the number of possible scenarios in the system. In contrast to the RO, the methodology is established to evaluate the expected performance and aims to minimize expected cost (or maximize expected revenue) over the finite possible scenarios. Overall, the SP approach produces a safe and useful solution that guarantees the feasibility for all scenarios.

The two-stage (linear) SP, as first introduced in Dantzig (1955), is the most widely used and studied method in the literature. It consists of two groups of decision variables, namely, first-stage variables (also called *here-at-now*) and second-stage variables (also called *wait-and-see*). The decision maker initially decides on the first-stage values, after which the uncertainty parameters are realized, then the recourse actions are taken (i.e., decides on second-stage values) so as to mitigate any negative effects incurred by the first-stage decision.

Suppose we deal with a two-stage SP problem that has the following form:

$$\begin{aligned} \max \quad & \mathbb{E}_{\xi}[Q(\mathbf{x}, \xi)] \\ \text{s.t.} \quad & \mathbf{A}'\mathbf{x} \leq b, \\ & \mathbf{x} \geq 0, \end{aligned} \tag{2.4}$$

where $Q(\mathbf{x}, \xi(\omega))$ is the optimal objective value of the second-stage problem

$$\begin{aligned} Q(\mathbf{x}, \xi(\omega)) = \max \quad & \mathbf{q}(\omega)'\mathbf{y} \\ \text{s.t.} \quad & \mathbf{W}'\mathbf{y} \leq \mathbf{h}(\omega) - \mathbf{T}(\omega)'\mathbf{x}, \end{aligned} \tag{2.5}$$

$$\mathbf{y} \geq 0.$$

The first-stage decision variables are $\mathbf{x}$, which must be made before the realization of random vectors ξ. The second-stage decision variables $\mathbf{y}$ are made with respect to the realized second-stage problem parameters $\xi(\omega) = (\mathbf{q}(\omega), \mathbf{h}(\omega), \mathbf{T}(\omega))$. Note that each realization of the random vector ξ corresponds to a *scenario*. Since there are different realizations of the random vector ξ, the objective of (2.4) aims to maximize the expectation of the second-stage objective function over a set of scenarios (see Birge and Louveaux, 2011 for details).

We see many applications of the two-stage SP in supply chain design (e.g., MirHassani et al., 2000, Tsiakis et al., 2001) and transportation planning (e.g., Barbarosoğlu and Arda, 2004) problems because the immediate recourse actions are required in response to highly unpredictable events.

As a natural extension of the two-stage SP, multi-stage SP is usually adopted in the multi-period environments where the uncertain parameters are revealed sequentially over discrete time. In such a context, the multi-stage SP yields a superior solution to the two-stage SP in the sense that it better characterizes the dynamic behavior of the stochastic process and adds more flexibility into the decision-making. It is worth noting that the evolution of the stochastic parameters are usually represented by a *scenario tree*, in which each node represents a possible realization of a set of stochastic parameters at a certain stage (or period), and a path from the root node to a leaf node in the tree corresponds to a scenario (see Huang and Ahmed, 2009 for details of the scenario tree).

On the other hand, the multi-stage SP problems are notorious for their difficulties to solve, especially when the recourse decisions are required to be integer (Birge and Louveaux, 2011). The complexity of the multi-stage SP problems were discussed in Shapiro and Nemirovski (2005). The authors stated that the multi-stage SP models are in general computational intractable. Some techniques to solve the multi-stage SP models include Benders' decomposition (e.g., Fragniere et al., 2000), augmented Lagrangian relaxation (e.g., Ruszczynski, 1999), and Monte Carlo sampling-based approximation (e.g., Dantzig and Glynn, 1990, Kleywegt et al., 2002), among others. The reader is referred to Birge and Louveaux (2011), Kall and Wallace (1994) for a thorough overview of the SP and the corresponding solution methods.

The application of the SP approach in the production and inventory management literature goes back to the late 1960's when El Agizy (1969) considered a single-item inventory problem subject to demand uncertainty. The author defined the decision variables as the function of the possible demand sequences. He further identified that the resulting LP problem can be reformulated as a network flow problem. In the context of multi-period problems, most existing works formulated the problems within a two-stage SP framework (e.g., Bakir and Byrne, 1998, Hood et al., 2003, Dillon et al., 2017). However, as noted above, the two-stage SP cannot accurately capture the dynamic stochastic process in such problems.

In Escudero et al. (1993), the authors proposed a multi-stage SP model to address a multi-period multi-product production planning problem with demand uncertainty. They computationally showed that the proposed model with a LP structure can be solved with very modest computational efforts. Huang and Ahmed (2009) presented a multi-stage SP model for a slightly different problem than that in El Agizy (1969) and analyzed the dynamic stochastic process using scenario trees. Brandimarte (2006) considered a capacitated lot-sizing problem subject to uncertain demand, in which the proposed multi-stage SP model was compared with a deterministic model of the problem using expected demands. Kazemi Zanjani et al. (2010) proposed a multi-stage SP model to cope with multi-product capacitated inventory problems with uncertain supply (i.e., process yield) and uncertain demand. Moreover, they applied the proposed model to the sawmill production problem of a realistic scale, and the results indicate that the solution performance of the multi-stage SP model significantly outperforms the two-stage SP model.

Chapter 3

Robust Optimization for a Single-Station System

3.1 Introduction

The periodic-review, finite-horizon inventory planning problems under demand uncertainty have been extensively studied in the past, with rich analytical results and applications. For example, Clark and Scarf (1960) has proved the optimality of base-stock policy for a serial inventory system subject to stochastic demand.

As global supply chains have been growing considerably, increasing firms take advantage of global sourcing opportunities to lower their labor and production costs. On the other hand, supply uncertainty that often arises from higher variability in overseas suppliers production performance as well as the shipment delay due to the long-distance transportation, or supply chain disruption, may adversely impact the overall performance of a supply chain, which suggests a practical need to incorporate supply uncertainty into decision-making.

Early works on inventory problems subject to supply and demand uncertainties can be found in Gerchak et al. (1988), Parlar and Berkin (1991), and Ciarallo et al. (1994), among many others. A majority of the prior works heavily relies on the assumption that the distribution for the underlying uncertain parameters is precisely known. Unfortunately, the distributional information in practice might be very difficult to accurately estimate or acquire.

Bertsimas and Sim (2003, 2004) developed a robust optimization (RO) framework specifically tailored for linear programming (LP) problems using row-wise polyhedral uncertainty sets. In addition, the level of conservativeness is flexibly adjusted through changing the value of a pre-determined parameter called *budget-of-uncertainty*. More importantly, according to the principle of strong duality, they showed that, despite the presence of additional variables and constraints, the LP form is preserved in the robust counterpart. They proved the existence of probabilistic bounds on the constraint violations, and thus an acceptable level of performance can be expected with their robust formulation.

Bertsimas and Thiele (2006) first applied the RO to inventory problems under demand uncertainty. The derived robust counterparts not only attain computational tractability, but also show promising results when compared to the classical methods. The methodology is subsequently applied to the sawmill planning problems in Alvarez and Vera (2014) (yield uncertainty) and Varas et al. (2014) (raw material supply and demand uncertainties). Aouam and Brahimi (2013) proposed the RO-based, integrated production planning model with order acceptance, where customer demands are distinguished by marginal revenue and subject to uncertainty. We also see many applications of the RO approach in the production and inventory management literature that involve multi-source uncertainty in their problem settings (see e.g., Wei et al., 2011, Alem and Morabito, 2012, Sanei Bajgiran et al., 2017).

Fewer works simultaneously take supply and demand uncertainties into consideration when the RO is applied. One recent work was Thorsen and Yao (2017), where the authors considered a single-station inventory problem subject to column-wise uncertain lead times and row-wise uncertain demands. Motivated by Bienstock and ÖZbay (2008), they devised Benders' decomposition algorithms to compute the optimal inventory policies.

In this chapter, we build on the multi-period single-station inventory planning problem that was studied in Bertsimas and Thiele (2006) and additionally incorporate supply uncertainty. Specifically, we consider partial supply. Partial supply includes yield loss in the production process, and the loss during the transportation between suppliers and warehouses. Inspired by Gerchak et al. (1988), who used the *stochastic proportional yield*, we introduce an uncertain parameter called *supply ratio* to model partial supply. Moreover, we generalize the theoretical results found in Bertsimas and Thiele (2006) regarding the optimal robust policies, and numerically show the effectiveness of the proposed robust policies.

3.2 Nominal Case

In this section, we consider a single-station inventory problem over a finite discrete planning horizon of T time periods. At the beginning of time period $t \in \mathcal{T} := \{0, 1, \ldots, T-1\}$, after observing current inventory level I_t, the decision maker decides on the ordering quantity. The binary variable δ_t denotes an ordering decision. That is, $\delta_t = 1$ if and only if an order is placed and $\delta_t = 0$, otherwise. Let c_t and K_t be the unit variable cost and setup cost in period t, respectively. The demand d_t occurs during the period. After the demand is realized, the excess stock at the end of period t (i.e., $I_{t+1} \geq 0$) is carried forward to the next time period $t+1$, incurring the holding cost $h_t I_{t+1}$, where h_t is the unit holding cost; while the unsatisfied demand is fully backlogged (i.e., $I_{t+1} < 0$), incurring the shortage cost $-b_t I_{t+1}$, where b_t is the unit shortage cost. Thus, the inventory cost can be expressed as $\max(h_t I_{t+1}, -b_t I_{t+1})$. We assume that $b_t > c_t$ so that it remains a possibility to order until the last period. In representing the supply uncertainty, it is assumed that a certain fraction of the order x_t can be received and we define the supply ratio $\alpha_t \in [0, 1]$. Then we write the received order quantity in a multiplicative form as $\alpha_t x_t$. The initial inventory level is given by I_0.

In the nominal model, both supply ratio and demand in each period are realized in their nominal values with certainty. Therefore, the inventory level at the end of period t with nominal supply ratio $\bar{\alpha}_i$ and nominal demand $\bar{d}_i$ for all $i \leq t$ is expressed as $\bar{I}_{t+1} = I_0 + \sum_{i=0}^{t}(\bar{\alpha}_i x_i - \bar{d}_i)$. The objective aims to minimize the total costs of purchasing (i.e., setup and variable costs), inventory holding and shortage over the entire planning horizon. The nominal single-station problem is:

$$\text{(DS)} \qquad \min \quad \sum_{t=0}^{T-1}(c_t x_t + K_t \delta_t + y_t) \tag{3.1}$$

$$\text{s.t.} \quad y_t \geq h_t\Big(I_0 + \sum_{i=0}^{t}(\bar{\alpha}_i x_i - \bar{d}_i)\Big), \qquad \forall t \in \mathcal{T}, \tag{3.2}$$

$$y_t \geq -b_t\Big(I_0 + \sum_{i=0}^{t}(\bar{\alpha}_i x_i - \bar{d}_i)\Big), \qquad \forall t \in \mathcal{T}, \tag{3.3}$$

$$0 \leq x_t \leq M\delta_t, \quad \delta_t \in \{0,1\}, \qquad \forall t \in \mathcal{T}, \tag{3.4}$$

where y_t is an intermediate variable and is equal to the inventory cost computed at the end of time period t at optimality, namely $y_t^* = \max(h_t \bar{I}_{t+1}^*, -b_t \bar{I}_{t+1}^*)$, and M is a large constant. Constraints (3.2) and (3.3) correspond to the inventory constraints for holding and shortage costs, respectively. Constraint (3.4) ensures that a nonnegative order can be placed if and only if the corresponding ordering decision has been made, namely $\delta_t = 1$.

It is noteworthy that although there exists a difference between the ordered quantity and the received quantity because of the partial supply, the purchasing cost is computed based on the ordered quantity. The problem could be easily adapted to the case where a firm only needs to pay what it actually receives by changing the objective to $\sum_{t=0}^{T-1}(c_t \bar{\alpha}_t x_t + K_t \delta_t + y_t)$. In the following, we apply the RO approach introduced in Bertsimas and Sim (2004) and Bertsimas and Thiele (2006) to DS formulation (3.1)-(3.4).

3.3 Robust Case

Suppose that both supply ratio α_t and demand d_t are subject to data uncertainty. To model the uncertainty, we assume that the supply ratio α_t takes values in the interval $[\bar{\alpha}_t - \hat{\alpha}_t, \bar{\alpha}_t]$ in time period t, which implies that at most the nominal level of ordered quantities can be received. Then, the scaled deviation of the supply ratio from its nominal value is defined as $z_t^\alpha := (\alpha_t - \bar{\alpha}_t)/\hat{\alpha}_t$. Thus, we have $z_t^\alpha \in [-1, 0]$ and $\mathbf{z}_t^\alpha := (z_0^\alpha, z_1^\alpha, \ldots, z_t^\alpha)$. In reality, it is unlikely that all elements in $\mathbf{z}_t^\alpha$ are equal to their worst-case values (i.e., $z_i^\alpha = -1$ for all $i \leq t$); a parameter called budget-of-uncertainty Γ_t^α is imposed to restrict large cumulative deviations in period t as $\sum_{i=0}^t |z_i^\alpha| \leq \Gamma_t^\alpha$, where Γ_t^α satisfies $\Gamma_t^\alpha \in [0, t]$ and $\Gamma_t^\alpha \leq \Gamma_{t+1}^\alpha$ for all t.

Approaching demand uncertainty differently from supply ratio uncertainty, we assume that demand d_t takes values in the *symmetric* interval $[\bar{d}_t - \hat{d}_t, \bar{d}_t + \hat{d}_t]$ in time period t. This is because, while the worst case of supply is always to have less supply than ordered (we will not receive more items than what we paid for), the worst case of demand can be either less or more than expected. The scaled deviation of demand is defined as $z_t^d := (d_t - \bar{d}_t)/\hat{d}_t$. Thus, we have $z_t^d \in [-1, 1]$ and $\mathbf{z}_t^d := (z_0^d, z_1^d, \ldots, z_t^d)$. Given the budget-of-uncertainty Γ_t^d, we have $\sum_{i=0}^t |z_i^d| \leq \Gamma_t^d$, where Γ_t^d satisfies $\Gamma_t^d \in [0, t]$ and $\Gamma_t^d \leq \Gamma_{t+1}^d$ for all t.

According to the above definitions, we write $\alpha_t = \bar{\alpha}_t + \hat{\alpha}_t z_t^\alpha$ and $d_t = \bar{d}_t + \hat{d}_t z_t^d$, respectively. In addition, for $t \in \mathcal{T}$, the budget uncertainty sets are defined as

$$\mathcal{Z}_t^\alpha := \{\mathbf{z}_t^\alpha | -1 \leq z_i^\alpha \leq 0, \sum_{i=0}^t |z_i^\alpha| \leq \Gamma_t^\alpha, \ \forall i \leq t\}, \tag{3.5}$$

and

$$\mathcal{Z}_t^d := \{\mathbf{z}_t^d | -1 \leq z_i^d \leq 1, \sum_{i=0}^t |z_i^d| \leq \Gamma_t^d, \ \forall i \leq t\}. \tag{3.6}$$

The values of budget-of-uncertainty are selected by the decision maker to model the degree of risk aversion. The robust single-station problem can be derived by maximizing the right-hand side (RHS) of Constraints (3.2) and (3.3) with respect to the budget uncertainty sets $\mathcal{Z}_t^\alpha$ and $\mathcal{Z}_t^d$ for all t as follows,

$$\begin{aligned}
\text{(RS)} \qquad \min \quad & \sum_{t=0}^{T-1}(c_t x_t + K_t \delta_t + y_t) \\
\text{s.t.} \quad & y_t \geq h_t\Big(\bar{I}_{t+1} + \max_{\mathcal{Z}_t^\alpha, \mathcal{Z}_t^d} \sum_{i=0}^{t}(-\hat{d}_i z_i^d + \hat{\alpha}_i z_i^\alpha x_i)\Big), && \forall t \in \mathcal{T}, \\
& y_t \geq b_t\Big(- \bar{I}_{t+1} + \max_{\mathcal{Z}_t^\alpha, \mathcal{Z}_t^d} \sum_{i=0}^{t}(\hat{d}_i z_i^d - \hat{\alpha}_i z_i^\alpha x_i)\Big), && \forall t \in \mathcal{T}, \qquad (3.7) \\
& 0 \leq x_t \leq M\delta_t, \quad \delta_t \in \{0,1\}, && \forall t \in \mathcal{T},
\end{aligned}$$

where $\bar{I}_{t+1} = I_0 + \sum_{i=0}^{t}(\bar{\alpha}_i x_i - \bar{d}_i)$.

Noticeably, for the tth pair of inventory holding and backlogging cost constraints in the RS formulation (3.7), we need to solve the auxiliary problems with respect to $\mathbf{z}_t^\alpha$ and $\mathbf{z}_t^d$. Therefore, this formulation is non-convex due to its min-max form, which requires a reformulation to a solvable form. We have the following theorem for RS formulation (3.7).

Theorem 3.1. *The RS formulation (3.7) is equivalent to the following robust counterpart with a mixed integer programming (MIP) structure:*

$$\begin{aligned}
\text{(RSC)} \quad \min \quad & \sum_{t=0}^{T-1}(c_t x_t + K_t \delta_t + y_t) \\
\text{s.t.} \quad & y_t \geq h_t(\bar{I}_{t+1} + o_t \Gamma_t^d + \sum_{i=0}^{t} p_{it}), && \forall t \in \mathcal{T}, \\
& y_t \geq b_t(- \bar{I}_{t+1} + o_t \Gamma_t^d + \sum_{i=0}^{t} p_{it} + q_t \Gamma_t^\alpha + \sum_{i=0}^{t} r_{it}), && \forall t \in \mathcal{T}, \\
& o_t + p_{it} \geq \hat{d}_i, && \forall t \in \mathcal{T}, \forall i \leq t, \quad (3.8) \\
& q_t + r_{it} \geq \hat{\alpha}_i x_i, && \forall t \in \mathcal{T}, \ \forall i \leq t,
\end{aligned}$$

$$o_t \geq 0,\ p_{it} \geq 0,\ q_t \geq 0,\ r_{it} \geq 0, \qquad \forall t \in \mathcal{T},\ \forall i \leq t,$$

$$0 \leq x_t \leq M\delta_t, \quad \delta_t \in \{0,1\}, \qquad \forall t \in \mathcal{T},$$

where $\bar{I}_{t+1} = I_0 + \sum_{i=0}^{t}(\bar{\alpha}_i x_i - \bar{d}_i)$.

Proof. Consider time period t. The solution of the auxiliary problem $\max_{\mathcal{Z}_t^\alpha} \sum_{i=0}^{t} \hat{\alpha}_i x_i z_i^\alpha$ is trivial because z_i^α varies between [-1,0] and thus the problem attains optimality when z_i^α is set to zero for all $i \leq t$. In addition, we next show that the auxiliary problems $\max_{\mathcal{Z}_t^d} \sum_{i=0}^{t} -\hat{d}_i z_i^d$ and $\max_{\mathcal{Z}_t^d} \sum_{i=0}^{t} \hat{d}_i z_i^d$ are essentially equivalent. Because the problems have opposite objectives with symmetric uncertainty set, the former can be rewritten as $\max_{\mathcal{Z}_t^d} \sum_{i=0}^{t} \hat{d}_i z_i^{new}$ where $z_i^{new} = -z_i^d$ for all $i \leq t$, while the uncertainty set remains unchanged with the new variable z_i^{new}.

Bertsimas and Thiele (2006) used a reformulation by invoking strong duality to the robust inventory problems under demand uncertainty with the budget uncertainty sets, which allows them to transform the min-max problem into a solvable form and retain original optimal solutions. Specifically, they considered the following auxiliary problem in period t,

$$\begin{aligned} \max \quad & \sum_{i=0}^{t} \hat{d}_i z_i^d \\ \text{s.t.} \quad & 0 \leq z_i^d \leq 1, \quad \forall i \leq t, \\ & \sum_{i=0}^{t} z_i^d \leq \Gamma_t^d. \end{aligned} \tag{3.9}$$

Notice, this problem is exactly the same auxiliary problem with $\mathbf{z}_t^d$ that comes from the RS formulation (3.7) (because the auxiliary problems with z_i^d embedded in RS formulation (3.7) are equivalent. In the maximization problem $\max_{\mathcal{Z}_t^d} \sum_{i=0}^{t} \hat{d}_i z_i^d$, the scaled deviation z_i^d will take a positive value at optimality for all $i \leq t$. As a result, the budget uncertainty set $\mathcal{Z}_t^d$ becomes $\{0 \le z_i^d \le 1,\ \sum_{i=1}^t z_i^d \le \Gamma_t^d\}$ for all t). Its dual problem is

$$\begin{aligned} \min \quad & o_t \Gamma_t^d + \sum_{i=0}^t p_{it} \\ \text{s.t.} \quad & o_t + p_{it} \ge \hat{d}_i, \quad \forall i \le t, \\ & o_t \ge 0,\ p_{it} \ge 0, \quad \forall i \le t, \end{aligned} \tag{3.10}$$

where o_t and p_{it} are dual variables corresponding to the constrains in auxiliary problem (3.9). By strong duality, they further substitute the dual problem (3.10) instead of auxiliary problem (3.9) into RS formulation (3.7) and obtain the robust counterpart.

Similarly, we consider the following auxiliary problem with $\mathbf{z}_t^\alpha$ that comes from RS formulation (3.7),

$$\begin{aligned} \max \quad & \sum_{i=0}^t \hat{\alpha}_i z_i^\alpha x_i^* \\ \text{s.t.} \quad & 0 \le z_i^\alpha \le 1, \qquad \forall i \le t, \\ & \sum_{i=0}^t z_i^\alpha \le \Gamma_t^\alpha, \end{aligned} \tag{3.11}$$

where x_i^* is an optimal solution of RS formulation (3.7) and considered as given. Then, its dual problem is following,

$$\begin{aligned} \min \quad & q_t \Gamma_t^\alpha + \sum_{i=0}^t r_{it} \\ \text{s.t.} \quad & q_t + r_{it} \ge \hat{\alpha}_i x_i^*, \qquad \forall i \le t, \\ & q_t \ge 0,\ r_{it} \ge 0, \qquad \forall i \le t, \end{aligned} \tag{3.12}$$

where q_t and r_{it} are the dual variables corresponding to the constrains in auxiliary problem (3.11).

After substituting the dual problems (3.10) and (3.12) back into RS formulation (3.7), we obtain the equivalent robust counterpart as shown in RSC formulation (3.8). □

Theorem 3.1 shows that the computational burden of the robust counterpart is not much higher than the corresponding nominal problem. In other words, the robust counterpart remains a LP problem if the nominal problem is a LP problem and a MIP problem if the nominal problem is a MIP problem.

In the following section, we present the theoretical results regarding the optimal robust policy obtained from RSC formulation (3.8).

3.4 Optimal Robust Policy

We now show that RSC formulation (3.8) is equivalent to a (larger) nominal problem and provide theoretical insights into the optimal robust policy. According to Clark and Scarf (1960), the optimal policy of a multi-period inventory problem is called (s, S), if there exists a sequence of parameters (s_t, S_t) for all t such that in time period t we have $x_t^* = S_t - I_t$ when the inventory level is less than a threshold s_t and $x_t^* = 0$ otherwise, with $s_t \leq S_t$. Analogously, the optimal policy is said to be *quasi*-(s, S) if in time period t we have $x_t^* = (S_t - I_t)/\beta_t$ (where $\beta_t > 0$) when $I_t < s_t$ and $x_t^* = 0$ otherwise. In order to present the results, we need the following lemma:

Lemma 3.1 (see Bertsimas and Thiele (2006)). *In the stationary case of DS formulation (3.1)-(3.4). That is, $c_t \equiv c, K_t \equiv K, h_t \equiv h$, $b \equiv b_t$ and $\bar{\alpha}_t \equiv \bar{\alpha}$ for all t, we have:*

(a) If there is no setup cost, the optimal nominal policy is quasi-(s, S) with $s_t = S_t = \bar{d}_t$ for all t. In other words, it is optimal to order $(S_t - I_t)/\bar{\alpha}$ units if $I_t < S_t$ and 0 otherwise in period t.

(b) If there is a setup cost, the optimal nominal policy is quasi-(s, S) with

$$S_j = \sum_{\tau=0}^{U_j} \bar{d}_{t_j+\tau}, \tag{3.13}$$

and

$$s_1 = I_0 - \sum_{\tau=0}^{t_1-1} \bar{d}_\tau, \quad s_j = - \sum_{\tau=U_{j-1}+1}^{L_{j-1}-1} \bar{d}_{t_{j-1}+\tau}, \quad j \geq 2 \tag{3.14}$$

where t_j $(j = 1, \ldots, J)$ denotes the ordering time periods, and $L_j = t_{j+1} - t_j$, $U_j = \lfloor \frac{bL_j}{b+h} \rfloor$ and $U_J = \lfloor \frac{bL_j - c}{b+h} \rfloor$. In other words, it is optimal to order $(S_t - I_t)/\bar{\alpha}$ units if $I_t < s_t$ and 0 otherwise in period t.

Proof. See Bertsimas and Thiele (2006) for the optimality of the *quasi*-(s, S) policy for DS formulation (3.1)-(3.4) when $\bar{\alpha} \equiv 1$ for all t. The results still hold if we rewrite the problem with new variable $x'_t = \bar{\alpha} x_t$ and new unit variable cost $c' \equiv c/\bar{\alpha}$. □

We next present the results regarding the optimal robust policy.

Theorem 3.2 (Optimal robust policy)**.** *Let o^*_t, p^*_{it}, q^*_t and r^*_{it} be an optimal solution of RSC formulation (3.8), then we have:*

(a) The optimal policy in RSC can be obtained by solving the nominal problem with the modified, deterministic demand in period t,

$$d'_t = \bar{d}_t + (\Upsilon_t - \Upsilon_{t-1}) + (\Psi_t - \Psi_{t-1}), \tag{3.15}$$

where $\Upsilon_{-1} = 0$ and $\Upsilon_t := ((b_t - h_t)/(b_t + h_t))A_t$ with $A_t = o^*_t \Gamma^d_t + \sum_{i=0}^{t} p^*_{it}$ being the accumulated deviation of the uncertain demand from its nominal value in period t; $\Psi_{-1} = 0$ and $\Psi_t := (b_t/(b_t + h_t))B_t$ with $B_t = q^*_t \Gamma^\alpha_t + \sum_{i=0}^{t} r^*_{it}$ being the accumulated deviation of the uncertain supply ratio from its nominal value in period t.

(b) In the stationary case, if there is no setup cost, the optimal robust policy is quasi-(s, S) with $s_t = S_t = d'_t$ for all t.

(*c*) *In the stationary case, if there is no setup cost and* $I_0 \leq d'_0$, *the optimal robust policy always results in the same holding and shortage costs for all* t. *Moreover, the inventory cost at optimality can be expressed as*

$$y_t^* = \frac{2bh}{b+h} A_t + \frac{bh}{b+h} B_t. \tag{3.16}$$

(*d*) *In the stationary case, if there is a setup cost, the optimal robust policy is quasi-*(s, S) *with corresponding* s_j *and* S_j, *where* j *indexes the ordering periods, defined in Lemma 3.1, however, applied to the modified demand* d'_t *given in Equation* (3.15).

(*e*) *The optimal cost of RSC is equal to the optimal cost of the nominal problem with the modified demand sequence plus the extra cost* $\sum_{t=0}^{T-1} \frac{2b_t h_t}{b_t+h_t} A_t + \sum_{t=0}^{T-1} \frac{b_t h_t}{b_t+h_t} B_t$.

Proof. We reformulate RSC formulation (3.8) as the nominal problem with the modified demand. Given an optimal solution $(\mathbf{x}^*, \delta^*, \mathbf{o}^*, \mathbf{p}^*, \mathbf{q}^*, \mathbf{r}^*)$ for RSC, note that $\mathbf{x}^*$ and δ^* still solve RSC if the remaining variables are fixed to $\mathbf{o}^*, \mathbf{p}^*, \mathbf{q}^*$ and $\mathbf{r}^*$, which allows us to focus on the following problem,

$$\min_{\mathbf{x} \geq 0} \sum_{t=0}^{T-1} \left[c_t x_t + K_t 1_{\{x_t > 0\}} + \max\left(h_t(\bar{I}_{t+1} + A_t), b_t(-\bar{I}_{t+1} + A_t + B_t) \right) \right], \tag{3.17}$$

where $\bar{I}_{t+1} = I_0 + \sum_{i=0}^{t} (\bar{\alpha}_i x_i - \bar{d}_i)$, $A_t = o_t^* \Gamma_t^d + \sum_{i=0}^{t} p_{it}^*$, $B_t = q_t^* \Gamma_t^\alpha + \sum_{i=0}^{t} r_{it}^*$, and $1_{\{x_t > 0\}}$ function returns 1 if $x_t > 0$ and 0 otherwise.

Next, the modified inventory level I'_t is defined as

$$I'_{t+1} = I'_t + \bar{\alpha}_t x_t - \underbrace{\left(\bar{d}_t + (\Upsilon_t - \Upsilon_{t-1}) + (\Psi_t - \Psi_{t-1}) \right)}_{d'_t}, \tag{3.18}$$

with $I'_0 = I_0$. Obviously, the modified demand d'_t is not subject to uncertainty. Then I'_t can be recursively written as follows:

$$
\begin{aligned}
I'_1 &= \underbrace{I'_0}_{=I_0} + \bar{\alpha}_0 x_0 - \bar{d}_0 - (\Upsilon_0 - \Upsilon_{-1}) - (\Psi_0 - \Psi_{-1}), \\
I'_2 &= I'_1 + \bar{\alpha}_1 x_1 - \bar{d}_1 - (\Upsilon_1 - \Upsilon_0) - (\Psi_1 - \Psi_0) \\
&= I_0 + \sum_{i=0}^{1} (\bar{\alpha}_i x_i - \bar{d}_i) - \Upsilon_1 - \Psi_1, \\
&\vdots \\
I'_{t+1} &= I_0 + \sum_{i=0}^{t} (\bar{\alpha}_i x_i - \bar{d}_i) - \Upsilon_t - \Psi_t \\
&= \bar{I}_{t+1} - \frac{b_t - h_t}{b_t + h_t} A_t - \frac{b_t}{b_t + h_t} B_t.
\end{aligned}
$$

Substituting I'_{t+1} in the following equations, we have

$$
\begin{aligned}
& h_t I'_{t+1} + \frac{2b_t h_t}{b_t + h_t} A_t + \frac{b_t h_t}{b_t + h_t} B_t \\
&= h_t \bar{I}_{t+1} - \frac{h_t(b_t - h_t)}{b_t + h_t} A_t - \frac{b_t h_t}{b_t + h_t} B_t + \frac{2b_t h_t}{b_t + h_t} A_t + \frac{b_t h_t}{b_t + h_t} B_t \\
&= h_t \bar{I}_{t+1} + h_t A_t,
\end{aligned}
$$

and

$$
\begin{aligned}
& - b_t I'_{t+1} + \frac{2b_t h_t}{b_t + h_t} A_t + \frac{b_t h_t}{b_t + h_t} B_t \\
&= -b_t \bar{I}_{t+1} + \frac{b_t(b_t - h_t)}{b_t + h_t} A_t + \frac{b_t^2}{b_t + h_t} B_t + \frac{2b_t h_t}{b_t + h_t} A_t + \frac{b_t h_t}{b_t + h_t} B_t \\
&= -b_t \bar{I}_{t+1} + b_t A_t + b_t B_t.
\end{aligned}
$$

The inventory cost of problem (3.17) in period t can be written as

$$
\max\{h_t(\bar{I}_{t+1} + A_t), b_t(-\bar{I}_{t+1} + A_t + B_t)\}
$$

$$= \max\{h_t I'_{t+1}, -b_t I'_{t+1}\} + \frac{2b_t h_t}{b_t + h_t} A_t + \frac{b_t h_t}{b_t + h_t} B_t. \tag{3.19}$$

We therefore obtain the following nominal problem with the modified demand d'_t given in Equation (3.15) (plus the extra cost $\frac{2b_t h_t}{b_t+h_t} A_t + \frac{b_t h_t}{b_t+h_t} B_t$) in period t,

$$\min_{\mathbf{x}\geq 0} \sum_{t=0}^{T-1} [c_t x_t + K_t 1_{\{x_t>0\}} + \max\{h_t I'_{t+1}, -b_t I'_{t+1}\} + \frac{2b_t h_t}{b_t + h_t} A_t + \frac{b_t h_t}{b_t + h_t} B_t].$$

This proves (a) and (e). Invoking Lemma 3.1, we immediately conclude that the results hold in (b) and (d). According to (b), it is optimal to order

$$x_0^* = \frac{d'_0 - I_0}{\bar{\alpha}} \quad \text{and} \quad x_t^* = \frac{d'_t}{\bar{\alpha}}, \quad t \geq 1$$

Hence, for tth pair of inventory holding and shortage cost constraints in RS formulation (3.7), the inventory holding and shortage costs in the stationary case can be written out as follows,

$$\begin{aligned}
& h(I_0 + (d'_0 - I_0) + \sum_{i=1}^{t} d'_i - \sum_{i=0}^{t} \bar{d}_i + A_t) \\
&= h(\sum_{i=0}^{t} d'_i - \sum_{i=0}^{t} \bar{d}_i + A_t) \\
&= h(\sum_{i=0}^{t} \bar{d}_i + \frac{b-h}{b+h} A_t + \frac{b}{b+h} B_t - \sum_{i=0}^{t} \bar{d}_i + A_t) \\
&= \frac{2bh}{b+h} A_t + \frac{bh}{b+h} B_t,
\end{aligned}$$

and

$$\begin{aligned}
& b(-I_0 - (d'_0 - I_0) - \sum_{i=1}^{t} d'_i + \sum_{i=0}^{t} \bar{d}_i + A_t + B_t) \\
&= b(-\sum_{i=0}^{t} d'_i + \sum_{i=0}^{t} \bar{d}_i + A_t + B_t)
\end{aligned}$$

$$= b(-\sum_{i=0}^{t}\bar{d}_i - \frac{b-h}{b+h}A_t - \frac{b}{b+h}B_t + \sum_{i=0}^{t}\bar{d}_i + A_t + B_t)$$
$$= \frac{2bh}{b+h}A_t + \frac{bh}{b+h}B_t.$$

This completes the proof of (c). □

From Equation (3.16), it is shown that the inventory cost at optimality is a function of both A_t and B_t, where A_t and B_t are defined by strong duality using auxiliary problems (3.9) and (3.11), respectively. It is that clear when the intervals for α_t and d_t get wider (i.e., $\hat{\alpha}_t$ and $\hat{d}_t$ increase), the inventory cost increases accordingly since B_t and A_t increase.

The extra cost $\sum_{t=0}^{T-1}\frac{2b_th_t}{b_t+h_t}A_t + \sum_{t=0}^{T-1}\frac{b_th_t}{b_t+h_t}B_t$ corresponds to the cost that a firm is willing to pay to guarantee the robustness of the optimal solution against uncertainty. A similar result was first provided in Bertsimas and Thiele (2006) for the robust inventory problem under demand uncertainty, where the extra cost is equal to $\sum_{t=0}^{T-1}\frac{2b_th_t}{b_t+h_t}A_t$. The extra cost $\sum_{t=0}^{T-1}\frac{b_th_t}{b_t+h_t}B_t$ is incurred by incorporating α_t into the problem. Using the same argument above, it gets more expensive to ensure the robustness of the optimal solution as the intervals for α_t and d_t get wider.

One might notice that, in Equation (3.16) as well as the expression of the extra cost, the coefficient for A_t is always twice as much as the coefficient for B_t. This can be explained intuitively because we assumed a symmetric interval for d_t centered around its nominal value, while the interval for α_t is one-sided.

3.5 Extension to Capacitated Cases

In this section, we consider capacitated RSC and investigate how the optimal robust policy is affected.

3.5.1 Capacitated Order

First, we consider the case where there is a time-varying upper bound C_t^{ord} on ordering quantities x_t for all t. The capacitated RSC can be derived by adding the following constraint to RSC formulation (3.8),

$$x_t \leq \mathrm{C}_t^{ord}, \quad \forall t \in \mathcal{T} \tag{3.20}$$

Note that the reformulation does not affect Constraint (3.20), so we have the following theorem.

Theorem 3.3 (Optimal robust policy)**.** *The optimal policy in RSC with capacitated orders can be obtained by solving the nominal problem subject to the modified demand given in Equation* (3.15) *with the capacity* C_t^{ord} *on the orders.*

Proof. The result immediately follows from adding Constraint (3.20) to RSC formulation (3.8) and reformulate the problem as the nominal problem in the same way as in the proof of Theorem 3.2. □

3.5.2 Capacitated Inventory

We now consider the case where there is an upper bound C^{inv} on the storage capacity at the station, namely

$$I_0 + \sum_{i=0}^{t}(\alpha_i x_i - d_i) \leq \mathrm{C}^{inv}, \quad \forall t \in \mathcal{T} \tag{3.21}$$

where $\alpha_i = \bar{\alpha}_i + \hat{\alpha}_i z_i^\alpha$ and $d_i = \bar{d}_i + \hat{d}_i z_i^d$ with $\mathbf{z}_t^\alpha \in \mathcal{Z}_t^\alpha$ and $\mathbf{z}_t^d \in \mathcal{Z}_t^d$. We now write Constraint (3.21) in the robust framework as

$$\bar{I}_{t+1} + \max_{\mathcal{Z}_t^\alpha, \mathcal{Z}_t^d}(-\hat{d}_i z_i^d + \hat{\alpha}_i z_i^\alpha x_i) \leq \mathrm{C}^{inv}, \quad \forall t \in \mathcal{T} \tag{3.22}$$

where $\bar{I}_{t+1} = I_0 + \sum_{i=0}^{t}(\bar{\alpha}_i x_i - \bar{d}_i)$. Using the same auxiliary problem (3.9) as before (the auxiliary problem $\max_{\mathcal{Z}_t^\alpha} \hat{\alpha}_i z_i^\alpha x_i$ is trivially solved), by strong duality, we obtain its robust counterpart as

$$\bar{I}_{t+1} + o_t\Gamma_t^d + \sum_{i=0}^{t} p_{it} \leq \mathrm{C}^{inv}, \quad \forall t \in \mathcal{T} \tag{3.23}$$

The capacitated RSC can be derived by adding Constraint (3.23) to RSC formulation (3.8). Obviously, the reformulation now affects Constraint (3.23). This becomes more clear if we rewrite the constraint with the modified inventory level I'_{t+1} (definition see Proof of Theorem 3.2) as

$$I'_{t+1} \leq \mathrm{C}^{inv} - \frac{2b_t}{b_t + h_t}A_t - \frac{b_t}{b_t + h_t}B_t, \quad \forall t \in \mathcal{T} \tag{3.24}$$

Thus, the optimal robust policy can be characterized as follows.

Theorem 3.4 (Optimal robust policy)**.** *The optimal policy in RSC with capacitated inventory can be obtained by solving the nominal problem subject to the modified demand given in Equation* (3.15) *with inventory capacities* C^{inv} *and* C_{t+1} *on inventory level in periods* 0 *and* $t+1$, $t \geq 0$, *respectively, where* $\mathrm{C}_{t+1} = \mathrm{C}^{inv} - (2b_t/(b_t + h_t))A_t - (b_t/(b_t + h_t))B_t$.

Proof. The result immediately follows from incorporating Constraint (3.24) into the reformulation. □

3.6 Numerical Studies

In this section, we apply RSC formulation (3.8) to a single-station example. For the purpose of comparison, we consider three inventory policies: (1) RO^0, the nominal policy when no uncertainty is considered; (2) RO^1, the robust policy obtained from Bertsimas and Thiele (2006) when only demand uncertainty is considered; (3) RO^2, the proposed robust policy when supply and demand uncertainties are considered.

We estimate the performances by means of simulation with respect to 100 replications of realized supply ratios and demands. In each replication, note that the realized supply ratios and demands are identical across three inventory policies. After a policy is implemented, we obtain the corresponding total cost, denoted by $\mathbf{C}(\text{RO}^i)$, $i = 0, 1, 2$. The effectiveness of RO^2 compared to RO^0 (resp., RO^1) is measured by the relative performance, computed by the ratio $\text{R}^{0-2} = 100 \cdot (\mathbf{C}(\text{RO}^0) - \mathbf{C}(\text{RO}^2))/\mathbf{C}(\text{RO}^0)$ (resp., $\text{R}^{1-2} = 100 \cdot (\mathbf{C}(\text{RO}^1) - \mathbf{C}(\text{RO}^2))/\mathbf{C}(\text{RO}^1)$), in percent. Then the expected relative performance can be computed with respect to those 100 replications, denoted by $\mathbb{E}(\text{R}^{0-2})$ (resp., $\mathbb{E}(\text{R}^{1-2})$), in percent.

Our objective in this study is to find out the added benefit in terms of cost saving by using RO^2. In addition, we investigate the influence of supply variability, demand variability, and uncertainty budgets on the average performance.

3.6.1 Experiment Setting

We consider the planning horizons of $T = 10,\ 20$ and 30 time periods. In a base example, the cost parameters are assumed to be stationary and selected as $c_t = 1$, $h_t = 0.1$ and $b_t = 1.5$ for all t. When the setup cost is explicitly considered, we use $K_t = 35$. There is zero initial inventory at the station.

In the simulation, the stochastic demand in each period is assumed to be i.i.d. and generated from different underlying demand distributions. Specifically, the realized demands are generated in accordance with a lognormal, gamma distribution with the same mean $\mu_d = 100$ and standard deviation $\delta_d = 20$, or a uniform distribution from the interval $[\mu_d - \delta_d, \mu_d + \delta_d]$. Likewise, the stochastic supply ratio in each period is also assumed to be i.i.d. and generated from a distribution. For simplicity, we consider a lognormal distribution with mean $\mu_\alpha = 0.9$ and standard deviation $\delta_\alpha = 0.05$. To avoid oversupply, we set $\alpha_t = 1$ if the realized supply ratio is greater than 1.

In order to capture such variabilities, the following parameters are selected as inputs for the RSC formulation (3.8): Let $\bar{\alpha}_t = 1$ and $\hat{\alpha}_t = (\bar{\alpha}_t - \mu_\alpha + 2\delta_\alpha)$, that is, the supply ratio belongs to the interval $[0.8, 1]$ for all t. Let $\bar{d}_t = \bar{\mu}_d$ and $\hat{d}_t = 2\delta_d$, that is, the demand belongs to the interval $[60, 140]$ for all t. Moreover, the linear budget functions are considered for supply ratio and demand as given by $\Gamma_t^\alpha = \gamma^\alpha + \gamma^\alpha \cdot t$ and $\Gamma_t^d = \gamma^d + \gamma^d \cdot t$, respectively, where γ^α and γ^d are the *budget factors* and we set $\gamma^\alpha = \gamma^d = 0.2$ in the base example. Given the above parameters, we solve the corresponding inventory models and obtain the optimal policies RO^i, $i = 0, 1, 2$.

3.6.2 Computational Effectiveness

The computational effectiveness of obtaining the inventory policies for the base example is reported in Table 3.1. In the table, we report the optimal cost values, the CPU time spent to solve the single-station models in seconds, the percentage gap obtained by each model under each combination of T and K_t. In addition, for the models with setup cost, we also report the number of orders placed in "#Ord." column. Note that the results in "Time (s)" and "Gap%" columns are obtained by solving each model 20 times and taking the average.

The results indicate that the time needed to obtain the robust policy RO^2 scales reasonably well with respect to the length of planning horizon. When the setup cost applies, the number of orders placed under each policy is also reported in the table. We observe that the ordering policy is largely influenced by the supply uncertainty, where in RO^2 both ordering quantity and ordering frequency over the entire planning horizon are increased compared to RO^0 and RO^1.

Table 3.1: Computational results for solving the single-station inventory models using stationary cost parameters $c_t = 1,\ h_t = 0.1,\ b_t = 1.5$ when $T = 10,\ 20$ and 30.

		$K_t = 0$			$K_t = 35$			
Policy	T	Obj.	Time (s)	Gap%	Obj.	Time (s)	Gap%	#Ord.
	10	1000.0	0.00	0.00	1220.0	0.11	0.00	4
RO^0	20	2000.0	0.00	0.00	2435.0	0.45	0.00	7
	30	3000.0	0.01	0.00	3650.0	3.27	0.30	10
	10	1152.5	0.00	0.00	1378.1	0.08	0.00	4
RO^1	20	2455.0	0.01	0.00	2903.3	0.26	0.62	7
	30	3907.5	0.02	0.00	4578.5	2.79	0.07	10
	10	1217.1	0.01	0.00	1519.8	0.13	0.00	5
RO^2	20	2625.9	0.01	0.00	3276.4	14.21	0.89	10
	30	4226.4	0.02	0.00	5265.4	47.06	0.81	20

3.6.3 Comparison of Policies

To evaluate the average performance of using the proposed robust policy RO^2, different demand distributions and cost parameters are considered in this section. Specifically, c_t and b_t are fixed to 1 and 1.5, whereas h_t can take two values $1/15 \cdot b_t = 0.1$ or $1/3 \cdot b_t = 0.5$; K_t can also take two values 35 or 70 when it applies. The results of expected relative performance for $T = 10,\ 20$ and 30 are provided in Table 3.2.

Table 3.2: Expected relative performance, in percent, using stationary cost parameters with the planning horizons of $T = 10,\ 20$ and 30.

		$K_t = 0$		$K_t = 35$		$K_t = 70$	
Demand Dist.	T	$\mathbb{E}(\mathrm{R}^{0-2})$	$\mathbb{E}(\mathrm{R}^{1-2})$	$\mathbb{E}(\mathrm{R}^{0-2})$	$\mathbb{E}(\mathrm{R}^{1-2})$	$\mathbb{E}(\mathrm{R}^{0-2})$	$\mathbb{E}(\mathrm{R}^{1-2})$
$h_t = 0.1,\ b_t = 1.5$							
	10	22.11	7.42	7.97	-0.40	1.48	-6.56
$d_t \sim$ *lognorm*	20	39.50	13.92	21.43	5.13	13.66	-2.62
	30	51.38	20.67	34.16	7.19	28.77	2.69
	10	25.57	9.05	10.89	0.08	1.67	-6.55
$d_t \sim$ *uniform*	20	42.43	17.27	24.08	5.33	14.77	-2.36
	30	50.90	22.39	36.76	8.56	30.82	4.58
	10	28.68	9.52	11.37	0.22	1.95	-6.49
$d_t \sim$ *gamma*	20	46.63	20.28	24.53	5.16	15.40	-1.80
	30	56.32	28.01	40.89	10.47	33.76	5.28
$h_t = 0.5,\ b_t = 1.5$							
	10	16.14	6.74	13.56	5.57	-3.32	-14.52
$d_t \sim$ *lognorm*	20	32.56	11.73	29.06	10.57	6.73	-13.12
	30	36.94	18.82	32.93	16.55	14.23	-5.11
	10	24.32	9.98	17.77	7.79	-1.19	-12.99
$d_t \sim$ *uniform*	20	36.90	20.00	32.77	17.24	9.32	-12.90
	30	40.97	23.11	37.15	20.37	18.29	-2.53
	10	21.02	9.34	17.64	7.66	-2.89	-13.22
$d_t \sim$ *gamma*	20	32.97	16.32	29.21	14.06	6.87	-12.97
	30	36.40	18.92	33.40	16.68	14.69	-4.24

From this table, we see that the robust policy RO^2 outperforms the nominal policy RO^0 in most of the cases. The result also holds true when it compares to the roust policy RO^1 when either no setup cost or the relatively lower setup cost (i.e., $K_t = 35$) is assumed. For $K_t = 70$, we observe that RO^2 does not perform well compared to RO^1. In particular, it performs worse with $h_t = 0.5$ than compared to $h_t = 0.1$. Actually, this phenomenon is understandable because RO^2 accounts for partial supply and demand uncertainties and thus yields larger overall ordering quantities and a higher ordering frequency to prevent potential shortage cost since we assume $h_t < b_t$. For instance, RO^0, RO^1, and RO^2 respectively order 10, 10, and 20 times for the problem with $K_t = 35$ and $T = 20$ (see Table 3.1). However, when the high unit holding and setup costs are assumed, RO^2 is heavily penalized for the unwanted orders and inventories, and therefore it performs poorly compared to the others.

It is also observed that RO^2 offers improved average performance as T increases throughout the table. This is because when it comes to a longer planning horizon, the system involves more cumulative variability, and the unwanted inventories in the previous periods could be used to satisfy large demands in the later periods. In other words, RO^2 yields less chances for stockout and is more likely to lead to a better long-term performance, especially for the system with a negligible setup cost and a relatively lower unit holding cost.

We additionally test the performance of three inventory policies in the non-stationary cost system with $T = 20$. In each time period t, the unit variable cost c_t is generated from a uniform distribution from the interval $[0.6, 1.4]$; the unit holding cost h_t is generated from a uniform distribution from the interval $[0.05, 0.15]$; the unit shortage cost is generated from a uniform distribution from the interval $[1, 2]$. The setup cost K_t remains stationary when it applies, however, with two possible values of 35 and 70. We compare their average performance with respect to 100 replications of realized supply ratios and demands and report the results in Table 3.3. The results show that RO^2 offers consistently better average performance than RO^0 and RO^1.

3.6.4 Impact of Parameters

In this section, we examine the effect of changing the standard deviations and the budgets for uncertainty on the average performance using the base example with $T = 20$. The

Table 3.3: Expected relative performance, in percent, using non-stationary cost parameters with $T = 20$.

	$K_t = 0$		$K_t = 35$		$K_t = 70$	
Demand Dist.	$\mathbb{E}(\mathrm{R}^{0-2})$	$\mathbb{E}(\mathrm{R}^{1-2})$	$\mathbb{E}(\mathrm{R}^{0-2})$	$\mathbb{E}(\mathrm{R}^{1-2})$	$\mathbb{E}(\mathrm{R}^{0-2})$	$\mathbb{E}(\mathrm{R}^{1-2})$
$d_t \sim$ *lognorm*	37.40	8.07	24.38	2.26	17.81	0.67
$d_t \sim$ *uniform*	37.31	7.22	26.84	3.68	18.78	3.17
$d_t \sim$ *gamma*	37.86	7.47	25.09	2.34	16.91	0.93

costs are $c_t = 1$, $h_t = 0.1$ and $b_t = 1.5$. There is no setup cost (i.e., $K_t = 0$). From Table 3.2, it is observed that the demand distribution does not play a significant role in the average performance, so we experiment with the lognormally distributed supply ratios and demands in this section.

Figure 3.1 illustrates how the expected relative performance varies as the ratio $\delta_\alpha/\bar{\alpha}_t$ increases (i.e., as δ_α increases because we fix $\bar{\alpha}_t$ to 1 for all t). It can be seen in the figure that both $\mathbb{E}(\mathrm{R}^{0-2})$ and $\mathbb{E}(\mathrm{R}^{1-2})$ exhibit the similar near-linearly increasing trend by up to 43.13% and 20.04% respectively, as δ_α increases from 0 to 0.1. As expected, when higher supply variability gets involved, the cost benefit of using RO^2 tends to be more visible.

Figure 3.2 shows how the expected relative performance varies as the ratio $\delta_d/\bar{d}_t$ increases (i.e., as δ_d increases because we fix $\bar{d}_t$ to 100 for all t). We see that RO^2 outperforms RO^0 by as much as 43% in terms of cost saving as δ_d increases, but the out-performance starts to get weaker when $\delta_d = 35$. This can be intuitively explained by the fact that there is a pressure to increase the order size in RO^2 to avoid potential shortage cost. However, in the simulation such high variability conversely leads to some lower realized demands, thus RO^2 results in excessive inventory and incurs unexpected inventory holding cost.

In contrast, the out-performance of RO^2 compared to RO^1 gets weaker very early

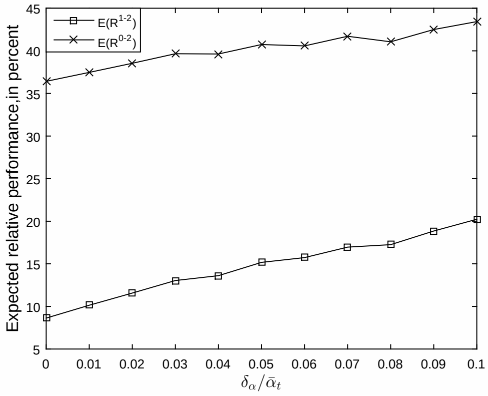

Figure 3.1: Impact of the standard deviation of supply ratio on performance.

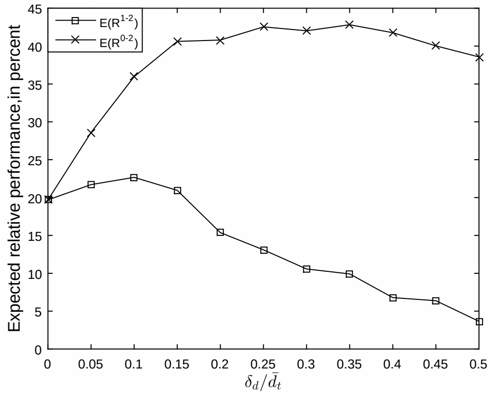

Figure 3.2: Impact of the standard deviation of demand on performance.

as $\mathbb{E}(\mathrm{R}^{1-2})$ starts to decrease when $\delta_d = 10$. This is mainly due to the fact that, as δ_d continues to increase, the demand uncertainty starts to dominate the supply uncertainty. Therefore, we can foresee that $\mathbb{E}(\mathrm{R}^{1-2})$ will eventually approach zero, and by that time RO^2 is reduced to RO^1 since the supply uncertainty is negligible compared to the demand uncertainty. The results in Figures 3.1 and 3.2 suggest that an increase in the demand variability could adversely affect the performance of RO^2.

In Figures 3.3 and 3.4, we study the impact of the budget-of-uncertainty on the average performance through changing the values of γ^α and γ^d. Specifically, Figure 3.3 plots how the expected relative performance reacts to the decision maker's risk-aversion towards the supply uncertainty, i.e., γ^α varies, while γ^d is set to 0.2. From the figure, we see that RO^2 consistently outperforms RO^0 and RO^1, but the highest $\mathbb{E}(\mathrm{R}^{0-2})$ and $\mathbb{E}(\mathrm{R}^{1-2})$ are achieved when $\gamma^\alpha \approx 0.45$. This is because when γ^α exceeds this value, the undesired overly-conservative RO^2 is enforced and hence increases the cost of RO^2 higher relative to RO^0 and RO^1. For instance, the decision maker with $\gamma^\alpha = 0.8$ believes that the worst-case supply ratio α_t=0.8 would be realized almost every two periods. However, in the simulation we observe that this situation rarely occurs. Thus, it is important for the decision maker to perform a similar simulation to find out the appropriate budget values for the system.

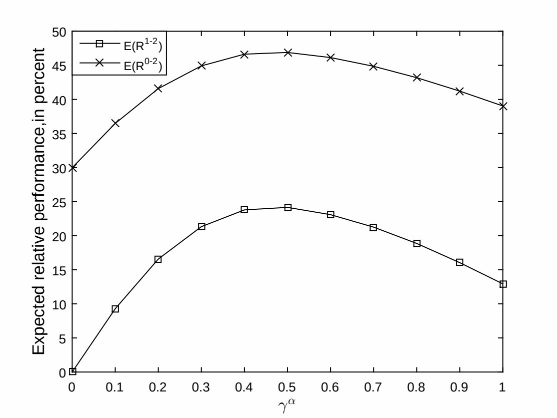

Figure 3.3: Impact of the level of conservatism in supply uncertainty on performance.

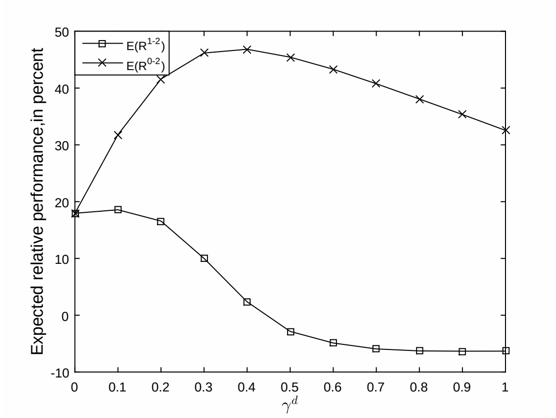

Figure 3.4: Impact of the level of conservatism in demand uncertainty on performance.

Figure 3.4 plots how the relative performance reacts to the decision maker's risk-aversion towards the demand uncertainty, i.e., γ^d varies, while γ^α is set to 0.2. The concavity shape of $\mathbb{E}(\mathrm{R}^{0-2})$ conforms with the results in Figure 3.3. As for $\mathbb{E}(\mathrm{R}^{1-2})$, we see that RO^2 yields a lower average cost by approximately 19% than RO^1, as γ^d increases from 0 to 0.2. By setting γ^d greater than 0.2, it means that the decision maker believes that the demand uncertainty outweighs the supply uncertainty. As γ^d continues to increase, RO^2 is reduced to RO^1 and $\mathbb{E}(\mathrm{R}^{1-2})$ gets closer to zero.

3.7 Concluding Remarks

In this chapter, we have applied the RO with budget uncertainty sets to a classical single-station inventory problem subject to supply and demand uncertainties. We have shown that the resulting robust counterpart formulation is equivalent to the nominal problem with the deterministic demand sequence. In the stationary case, we showed that the optimal robust policy is *quasi*-(s, S), where s and S levels are theoretically computable. In addition, we extended the model to the capacitated cases. The numerical results indicate that the proposed robust policy in general outperforms the nominal policy and the robust policy of Bertsimas and Thiele (2006) in the average performance.

In particular, we found that the robust policy performs exceptionally well when the unit shortage cost is relatively higher than the unit holding cost and no setup cost is present. Moreover, we investigated how deviations of uncertain parameters and budgets for uncertainty affect the performance and provided insights.

Chapter 4

Robust Optimization for a Multi-Echelon System

4.1 Introduction

Chapter 3 focuses on a single-station (or single-echelon) inventory planning problem. However, real-world supply chains are more complex than single stations and could consist of a set of single stations. As a result, the complexities of efficiently and effectively managing inventory within a multi-echelon configuration could be significantly higher, especially in uncertain environments.

The multi-echelon inventory planning problems that deal with demand uncertainty has been studied as early as the 1950's in Whitin (1957), Arrow et al. (1958), and later in Clark and Scarf (1960). In Clark and Scarf (1960), the authors considered systems with a serial structure or a tree structure subject to stochastic demand. In particular, they showed the optimality of base-stock policy for the serial supply chain. The result has been extended and refined in Federgruen and Zipkin (1984), Karmarkar (1981, 1987), and Rosling (1989). As mentioned in Chapter 3, it is imperative in contemporary business environment to incorporate supply uncertainty when designing sound inventory policies because of supply chain globalization. A large proportion of the literature has been conducted under the stochastic framework with a known distribution for the underlying uncertain parameters (i.e., supply and demand). For example, Bollapragada et al. (2004) considered a serial assembly inventory problem subject to stochastic supply and demand. By restricting to the base-stock policy for both component and end-product inventories, the authors showed that the optimal component base-stock level is convex decreasing in the optimal base-stock level of the end-product. However, the distributional information in practice is very difficult to acquire, and under some special circumstances the available historical data does not exhibit any known distributional behavior at all.

As an alternative, robust optimization (RO) has emerged as a promising approach to deal with data uncertainty while it requires very little information on distributions. With a large number of its applications in the single-station inventory problems (see Chapter 2 and 3), we notice that fewer works apply the RO in the context of the multi-echelon systems.

Bertsimas and Thiele (2006) first formulated the Clark and Scarf (1960)'s supply chain network under RO with budget (polyhedral) uncertainty sets. The echelon-specific cost structure allows them to conveniently decompose the network into several single-station problems and thus analyze optimal robust policy by echelons. However, unlike the single-station case, which has also been discussed in the thesis, the optimal robust policy for each echelon is no longer theoretically computable. Rikun (2011) further explored the application of RO framework to network systems with more general topologies and cost structures. Akbari and Karimi (2015) advanced the approach by allowing production capacity requirement (e.g., processing time) to belong to two disjoint polyhedral sets to account for occasional production abnormalities. The authors formulated the problem from the network design point-of-view since the proposed model also determines the location and capacity of distribution centers. Some other works on applying the RO approach to the multi-echelon systems include Ben-Tal et al. (2005, 2009).

In this chapter, we extend the results in Chapter 3 to a multi-echelon supply chain with a tree structure. It is important to note that this extension is built based on an indispensable assumption that the supply uncertainty only affects the orders placed by main storage hubs (see Section 4.2), while orders placed within the network are not subject to uncertainty. In fact, this assumption is commonly valid from a practical perspective, because once the orders have arrived at regional warehouses (i.e., main storage hubs) from overseas suppliers, they will usually be delivered by third-party logistics providers to local warehouses and then to stores in a full quantity and in a timely manner. The numerical results support our intuition that the proposed robust policy outperforms the nominal policy and the robust policy of Bertsimas and Thiele (2006) in average performance as well as the performance stability.

4.2 Nominal Case

As discussed in the introduction, we consider a supply chain network with a tree structure. The network is depicted in Figure 4.1, which contains the set of main storage hubs (MSHs), denoted by $\mathcal{M}$, the set of local storage hubs (LSHs), denoted by $\mathcal{L}$, and the set of stores, denoted by $\mathcal{S}$. The MSHs receive their supplies from an external supplier and then send items throughout the LSHs until they finally arrive at the stores. We let $\mathcal{N}$ be the set of nodes within the network, thus it can be expressed as $\mathcal{N} = \mathcal{M} \cup \mathcal{L} \cup \mathcal{S}$. It is noteworthy that the external supplier is excluded from the network and is numbered as node 0 to distinguish. For notational convenience, we additionally define the set $\mathcal{N}^0 := \{0\} \cup \mathcal{M} \cup \mathcal{L}$. We consider a finite discrete planning horizon of T time periods and use the set $\mathcal{T} = \{0, 1, \ldots, T-1\}$. The initial inventory level at echelon k is given by $X_k(0)$.

It is important to note that the supply uncertainty only affects the orders placed by MSHs, while orders placed within the network are not subject to uncertainty. We use this assumption because we focus on supply uncertainty exogenous to the multi-echelon system. Once the orders have arrived at the storage hubs, they will be delivered without uncertainty. In other words, there is no endogenous uncertainty to the multi-echelon system.

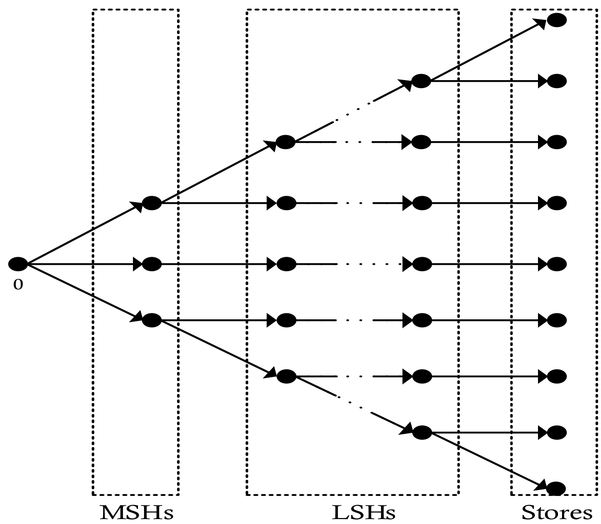

Figure 4.1: Network structure

Our definition of echelon follows Bertsimas and Thiele (2006). For node $k \in \mathcal{N}$, we define the union of all subsequent nodes, including k itself, that directly or indirectly receive supplies from k, and the links in-between as echelon k. The problem is described as follows. At the beginning of time period $t \in \mathcal{T}$, the inventory level at echelon $k \in \mathcal{N}$, $X_k(t)$, can be observed, then an order $D_{i_k k}(t)$ to its supplier i_k can be placed. The binary variable $\delta_{i_k k}(t)$ denotes ordering decisions. The unit variable cost and setup cost to its supplier is given by $c_{i_k k}(t)$ and $K_{i_k k}(t)$, respectively. The demand at store s in period t is denoted by $W_s(t)$. To specify the total demand for echelon k, we denote $\mathcal{S}(k)$ as the set of stores within echelon k. Hence, the total demand within echelon k in period t can be written as $\sum_{s \in \mathcal{S}(k)} W_s(t)$. Moreover, let $\mathcal{N}(k)$ be the set of nodes directly supplied by node k. At the end of the period after demand is realized, the echelon-specific inventory cost is incurred and accounted based on the ending inventory at the echelon as $\max\left(h_k(t)X_k(t+1), -b_k(t)X_k(t+1)\right)$, where $h_k(t)$ and $b_k(t)$ are the unit holding and unit backlogging costs, respectively. We assume that $b_k(t) > c_{i_k k}(t)$ for all k and t.

As mentioned earlier, we assume that the supply uncertainty (i.e., uncertain partial supply) only affects orders placed by MSHs to the external supplier. Hence, the orders made within the network are always received in full. The supply ratio from external supplier 0 to MSH i in period t is denoted by $\alpha_{0i}(t) \in [0, 1]$. Then, the received order quantities at MSHs are written as $\alpha_{0i}(t) \cdot D_{0i}(t)$.

In the nominal model, both supply ratio and demand in each time period are realized as their nominal values with probability one. The objective aims to minimize the total purchasing, inventory holding and shortage costs across all echelons in the network over the entire planning horizon. With the above notations, the nominal network problem is formulated as follows:

$$\text{(DN)} \quad \min \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{N}^0} \sum_{i\in\mathcal{N}(k)} \Big(c_{ki}(t) D_{ki}(t) + K_{ki}(t)\delta_{ki}(t) + Y_i(t) \Big) \tag{4.1}$$

$$\text{s.t. } Y_i(t) \geq h_i(t)\overline{X}_i(t+1), \qquad \forall i \in \mathcal{N},\ \forall t \in \mathcal{T}, \tag{4.2}$$

$$Y_i(t) \geq -b_i(t)\overline{X}_i(t+1), \qquad \forall i \in \mathcal{N},\ \forall t \in \mathcal{T}, \tag{4.3}$$

$$\sum_{i\in\mathcal{N}(k)} D_{ki}(t) \leq \overline{X}_k(t) - \sum_{i\in\mathcal{N}(k)} \overline{X}_i(t), \quad \forall k \in \mathcal{M} \cup \mathcal{L},\ \forall t \in \mathcal{T}, \tag{4.4}$$

$$0 \leq D_{ki}(t) \leq M\delta_{ki}(t), \qquad \forall k \in \mathcal{N}^0,\ \forall i \in \mathcal{N}(k),\ \forall t \in \mathcal{T}, \tag{4.5}$$

$$\delta_{ki}(t) \in \{0, 1\}, \qquad \forall k \in \mathcal{N}^0,\ \forall i \in \mathcal{N}(k),\ \forall t \in \mathcal{T}, \tag{4.6}$$

where $\overline{X}_i(t+1) = X_i(0) + \sum_{\tau=0}^{t} \big(\bar{\alpha}_{0i}(\tau) D_{0i}(\tau) - \sum_{s\in\mathcal{S}(i)} \overline{W}_s(\tau)\big)$ if $i \in \mathcal{M}$; $\overline{X}_i(t+1) = X_i(0) + \sum_{\tau=0}^{t} \big(D_{ki}(\tau) - \sum_{s\in\mathcal{S}(i)} \overline{W}_s(\tau)\big)$ if $i \in \mathcal{L} \cup \mathcal{S}$, and M is a large constant. Constraints (4.2) and (4.3) correspond to the inventory constraints for holding and shortage costs at echelon i, respectively. Constraint (4.4) corresponds to the *coupling constraint*, which guarantees the total orders made to MSHs and LSHs cannot exceed what they have in stock. Equivalently, backlogging is strictly prohibited at MSHs and LSHs, then we note that the derived robust policy could be suboptimal because it may be beneficial to allow for backloggings at these hubs. Constraints (4.5) and (4.6) correspond to the ordering decisions. We next formulate this DN formulation (4.1)-(4.6) under the robust framework.

4.3 Robust Case

In the manner as in Chapter 3, we assume that the supply ratio $\alpha_{0i}(t)$ takes values in the interval $[\bar{\alpha}_{0i}(t)-\widehat{\alpha}_{0i}(t), \bar{\alpha}_{0i}(t)]$ in time period t. The scaled deviation of supply ratio from its nominal value is defined as $Z^{\alpha}_{0i}(t) = (\alpha_{0i}(t) - \bar{\alpha}_{0i}(t))/\widehat{\alpha}_{0i}(t)$. Thus, we have $Z^{\alpha}_{0i}(t) \in [-1, 0]$ and $\mathbf{Z}^{\alpha}_{0i}(t) := (Z^{\alpha}_{0i}(0), Z^{\alpha}_{0i}(1), \ldots, Z^{\alpha}_{0i}(t))$. Moreover, the budget-of-uncertainty $\Gamma^{\alpha}_{0i}(t)$ is imposed to eliminate large deviations in period t as $\sum_{\tau=0}^{t} |Z^{\alpha}_{0i}(\tau)| \leq \Gamma^{\alpha}_{0i}(t)$, where $\Gamma^{\alpha}_{0i}(t) \in [0, t]$ and $\Gamma^{\alpha}_{0i}(t) \leq \Gamma^{\alpha}_{0i}(t+1)$ for all $i \in \mathcal{M}$ and $t \in \mathcal{T}$.

For uncertain demand, we assume that demand $W_s(t)$ takes values in the symmetric interval $[\overline{W}_s(t) - \widehat{W}_s(t), \overline{W}_s(t) + \widehat{W}_s(t)]$ in time period t. The scaled deviation of demand is defined as $Z^W_s(t) = (W_s(t) - \overline{W}_s(t))/\widehat{W}_s(t)$. Thus, $Z^W_s(t) \in [-1, 1]$ and $\mathbf{Z}^W_s(t) := (Z^W_s(0), Z^W_s(1), \ldots, Z^W_s(t))$. Given a budget-of-uncertainty $\Gamma^W_s(t)$, it follows $\sum_{\tau=0}^{t} |Z^W_s(\tau)| \leq \Gamma^W_s(t)$, where $\Gamma^W_s(t) \in [0, t]$ and $\Gamma^W_s(t) \leq \Gamma^W_s(t+1)$ for all $s \in \mathcal{S}$ and $t \in \mathcal{T}$.

According to the above definitions, we have $\alpha_{0i}(t) = \bar{\alpha}_{0i}(t) + \widehat{\alpha}_{0i}(t) Z^{\alpha}_{0i}(t)$ and $W_s(t) = \overline{W}_s(t) + \widehat{W}_s(t) Z^W_s(t)$. In addition, for $t \in \mathcal{T}$, the budget uncertainty sets are defined as

$$\mathcal{Z}^{\alpha}_{0i}(t) = \left\{ \mathbf{Z}^{\alpha}_{0i}(t) | -1 \leq Z^{\alpha}_{0i}(\tau) \leq 0, \sum_{\tau=0}^{t} |Z^{\alpha}_{0i}(\tau)| \leq \Gamma^{\alpha}_{0i}(t), \forall \tau \leq t \right\}, \tag{4.7}$$

and

$$\mathcal{Z}^W_s(t) = \left\{ \mathbf{Z}^W_s(t) | -1 \leq Z^W_s(\tau) \leq 1, \sum_{\tau=0}^{t} |Z^W_s(\tau)| \leq \Gamma^W_s(t), \forall \tau \leq t \right\}. \tag{4.8}$$

The robust network model can be derived by maximizing the right-hand side (RHS) of Constraints (4.2) and (4.3) with respect to the budget uncertainty sets $\mathcal{Z}_{0i}^{\alpha}(t)$ and $\mathcal{Z}_s^W(t)$ for all t as follows,

$$
\begin{aligned}
\text{(RN)} \quad & \min \sum_{t\in\mathcal{T}}\sum_{k\in\mathcal{N}^0}\sum_{i\in\mathcal{N}(k)} \Big(c_{ki}(t)D_{ki}(t) + K_{ki}(t)\delta_{ki}(t) + Y_i(t)\Big) \\
\text{s.t.} \quad & Y_i(t) \geq h_i(t)\Bigg\{\overline{X}_i(t+1) + \sum_{s\in\mathcal{S}(i)}\Bigg(\max_{\mathcal{Z}_s^W(t)}\sum_{\tau=0}^{t} -\widehat{W}_s(\tau)Z_s^W(\tau)\Bigg) \\
& \quad + \max_{\mathcal{Z}_{0i}^{\alpha}(t)}\sum_{\tau=0}^{t}\widehat{\alpha}_{0i}(\tau)Z_{0i}^{\alpha}(\tau)D_{0i}(\tau)\Bigg\}, && \forall i\in\mathcal{M},\ \forall t\in\mathcal{T}, \\
& Y_i(t) \geq b_i(t)\Bigg\{-\overline{X}_i(t+1) + \sum_{s\in\mathcal{S}(i)}\Bigg(\max_{\mathcal{Z}_s^W(t)}\sum_{\tau=0}^{t}\widehat{W}_s(\tau)Z_s^W(\tau)\Bigg) \\
& \quad + \max_{\mathcal{Z}_{0i}^{\alpha}(t)}\sum_{\tau=0}^{t} -\widehat{\alpha}_{0i}(\tau)Z_{0i}^{\alpha}(\tau)D_{0i}(\tau)\Bigg\}, && \forall i\in\mathcal{M},\ \forall t\in\mathcal{T}, \\
& Y_i(t) \geq h_i(t)\Bigg\{\overline{X}_i(t+1) \\
& \quad + \sum_{s\in\mathcal{S}(i)}\Bigg(\max_{\mathcal{Z}_s^W(t)}\sum_{\tau=0}^{t} -\widehat{W}_s(\tau)Z_s^W(\tau)\Bigg)\Bigg\}, && \forall i\in\mathcal{L}\cup\mathcal{S},\ \forall t\in\mathcal{T}, \qquad (4.9) \\
& Y_i(t) \geq b_i(t)\Bigg\{-\overline{X}_i(t+1) \\
& \quad + \sum_{s\in\mathcal{S}(i)}\Bigg(\max_{\mathcal{Z}_s^W(t)}\sum_{\tau=0}^{t}\widehat{W}_s(\tau)Z_s^W(\tau)\Bigg)\Bigg\}, && \forall i\in\mathcal{L}\cup\mathcal{S},\ \forall t\in\mathcal{T}, \\
& \sum_{i\in\mathcal{N}(k)} D_{ki}(t) \leq X_k(t) - \sum_{i\in\mathcal{N}(k)} X_i(t), && \forall k\in\mathcal{M}\cup\mathcal{L},\ \forall t\in\mathcal{T}, \\
& 0 \leq D_{ki}(t) \leq M\delta_{ki}(t), && \forall k\in\mathcal{N}^0,\ \forall i\in\mathcal{N}(k),\ \forall t\in\mathcal{T}, \\
& \delta_{ki}(t)\in\{0,1\}, && \forall k\in\mathcal{N}^0,\ \forall i\in\mathcal{N}(k),\ \forall t\in\mathcal{T},
\end{aligned}
$$

where $\overline{X}_i(t+1) = X_i(0) + \sum_{\tau=0}^{t}\left(\overline{\alpha}_{0i}(\tau)D_{0i}(\tau) - \sum_{s\in\mathcal{S}(i)}\overline{W}_s(\tau)\right)$ and $X_i(t+1) = X_i(0) + \sum_{\tau=0}^{t}\left(\alpha_{0i}(\tau)D_{0i}(\tau) - \sum_{s\in\mathcal{S}(i)} W_s(\tau)\right)$ if $i\in\mathcal{M}$; $\overline{X}_i(t+1) = X_i(0) + \sum_{\tau=0}^{t}\left(D_{ki}(\tau) - \sum_{s\in\mathcal{S}(i)}\overline{W}_s(\tau)\right)$ and $X_i(t+1) = X_i(0) + \sum_{\tau=0}^{t}\left(D_{ki}(\tau) - \sum_{s\in\mathcal{S}(i)} W_s(\tau)\right)$ if $i\in\mathcal{L}\cup\mathcal{S}$.

Nevertheless, the coupling constraint in (4.9) also contains the scaled deviations. Indeed, consider the RHS of coupling constraint for $k \in \mathcal{L}$, we have

$$\begin{aligned}
& X_k(t) - \sum_{i \in \mathcal{N}(k)} X_i(t) \\
&= \left\{ X_k(0) + \sum_{\tau=0}^{t-1} \left(D_{i_k k}(\tau) - \sum_{s \in \mathcal{S}(k)} W_s(\tau)\right) \right\} \\
&\qquad - \sum_{i \in \mathcal{N}(k)} \left\{ X_i(0) + \sum_{\tau=0}^{t-1} \left(D_{ki}(\tau) - \sum_{s \in \mathcal{S}(i)} W_s(\tau)\right) \right\} \\
&= \bar{X}_k(t) - \sum_{i \in \mathcal{N}(k)} \bar{X}_i(t) - \sum_{\tau=0}^{t-1} \sum_{s \in \mathcal{S}(k)} \hat{W}_s(\tau) Z_s^W(\tau) + \sum_{\tau=0}^{t-1} \underbrace{\sum_{i \in \mathcal{N}(k)} \sum_{s \in \mathcal{S}(i)}}_{=\sum_{s \in \mathcal{S}(k)}} \hat{W}_s(\tau) Z_s^W(\tau) \\
&= \bar{X}_k(t) - \sum_{i \in \mathcal{N}(k)} \bar{X}_i(t).
\end{aligned}$$

The scaled deviations cancel each other out and the corresponding coupling constraint is written in its nominal form as

$$\sum_{i \in \mathcal{N}(k)} D_{ki}(t) \leq \bar{X}_k(t) - \sum_{i \in \mathcal{N}(k)} \bar{X}_i(t), \quad \forall\, k \in \mathcal{L}, \forall\, t \in \mathcal{T} \tag{4.10}$$

However, consider the RHS of coupling constraint in (4.9) for $k \in \mathcal{M}$, such result does not hold because

$$\begin{aligned}
& X_k(t) - \sum_{i \in \mathcal{N}(k)} X_i(t) \\
&= \left\{ X_k(0) + \sum_{\tau=0}^{t-1} \left(\alpha_{0k}(\tau) D_{0k}(\tau) - \sum_{s \in \mathcal{S}(k)} W_s(\tau)\right) \right\} \\
&\qquad - \sum_{i \in \mathcal{N}(k)} \left\{ X_i(0) + \sum_{\tau=0}^{t-1} \left(D_{ki}(\tau) - \sum_{s \in \mathcal{S}(i)} W_s(\tau)\right) \right\} \\
&= \bar{X}_k(t) - \sum_{i \in \mathcal{N}(k)} \bar{X}_i(t) + \sum_{\tau=0}^{t-1} \hat{\alpha}_{0k}(\tau) Z_{0k}^{\alpha}(\tau) D_{0k}(\tau).
\end{aligned}$$

Note that the scaled deviations of $Z^{\alpha}_{0k}(\tau)$ cannot be eliminated. Therefore, we rewrite the coupling constraint in its robust form as

$$\sum_{i\in\mathcal{N}(k)} D_{ki}(t) \le \overline{X}_k(t) - \sum_{i\in\mathcal{N}(k)} \overline{X}_i(t) + \min_{\mathcal{Z}^{\alpha}_{0i}(t-1)} \sum_{\tau=0}^{t-1} \widehat{\alpha}_{0k}(\tau) Z^{\alpha}_{0k}(\tau) D_{0k}(\tau), \quad \forall k\in\mathcal{M}, \forall t\in\mathcal{T} \tag{4.11}$$

After substituting Constraints (4.10) and (4.11) back into RN formulation (4.9), we have the following theorem.

Theorem 4.1. *The RN formulation (4.9) is equivalent to the following robust counterpart with a MIP structure:*

$$\begin{aligned}
\text{(RNC)}\quad \min\quad & \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{N}^0} \sum_{i\in\mathcal{N}(k)} \Big(c_{ki}(t) D_{ki}(t) + K_{ki}(t)\delta_{ki}(t) + Y_i(t) \Big) \\
\text{s.t.}\quad & Y_i(t) \ge h_i(t) \Big\{ \overline{X}_i(t+1) \\
& \quad + \sum_{s\in\mathcal{S}(i)} \Big(O_s(t)\Gamma^W_s(t) + \sum_{\tau=0}^{t} P_s(\tau,t) \Big) \Big\}, && \forall i\in\mathcal{N},\ \forall t\in\mathcal{T}, \\
& Y_i(t) \ge b_i(t) \Big\{ -\overline{X}_i(t+1) \\
& \quad + \sum_{s\in\mathcal{S}(i)} \Big(O_s(t)\Gamma^W_s(t) + \sum_{\tau=0}^{t} P_s(\tau,t) \Big) \Big\}, && \forall i\in\mathcal{L}\cup\mathcal{S},\ \forall t\in\mathcal{T}, \\
& Y_i(t) \ge b_i(t) \Big\{ -\overline{X}_i(t+1) + \sum_{s\in\mathcal{S}(i)} \Big(O_s(t)\Gamma^W_s(t) + \sum_{\tau=0}^{t} P_s(\tau,t) \Big) \\
& \quad + Q_{0i}(t)\Gamma^{\alpha}_{0i}(t) + \sum_{\tau=0}^{t} R_{0i}(\tau,t) \Big\}, && \forall i\in\mathcal{M},\ \forall t\in\mathcal{T}, \\
& \sum_{i\in\mathcal{N}(k)} D_{ki}(t) \le \overline{X}_k(t) - \sum_{i\in\mathcal{N}(k)} \overline{X}_i(t), && \forall k\in\mathcal{L},\ \forall t\in\mathcal{T}, \\
& \sum_{i\in\mathcal{N}(k)} D_{ki}(t) \le \overline{X}_k(t) - \sum_{i\in\mathcal{N}(k)} \overline{X}_i(t) \\
& \quad - \Big(Q_{0k}(t-1)\Gamma^{\alpha}_{0k}(t-1) + \sum_{\tau=0}^{t-1} R_{0k}(\tau,t-1) \Big), && \forall k\in\mathcal{M},\ \forall t\in\mathcal{T}, \\
& O_s(t) + P_s(\tau,t) \ge \widehat{W}_s(\tau), && \forall s\in\mathcal{S},\ \forall t\in\mathcal{T},\ \forall \tau\le t,
\end{aligned} \tag{4.12}$$

$$
\begin{aligned}
& Q_{0i}(t) + R_{0i}(\tau,t) \geq \widehat{\alpha}_{0i}(\tau) D_{0i}(\tau), && \forall i \in \mathcal{M},\ \forall t \in \mathcal{T},\ \forall \tau \leq t, \\
& O_s(t) \geq 0,\ \ P_s(\tau,t) \geq 0, && \forall s \in \mathcal{S},\ \forall t \in \mathcal{T},\ \forall \tau \leq t, \\
& Q_{0i}(t) \geq 0,\ \ R_{0i}(t) \geq 0, && \forall i \in \mathcal{M},\ \forall t \in \mathcal{T},\ \forall \tau \leq t, \\
& 0 \leq D_{ki}(t) \leq M\delta_{ki}(t), && \forall k \in \mathcal{N}^0,\ \forall i \in \mathcal{N}(k),\ \forall t \in \mathcal{T}, \\
& \delta_{ki}(t) \in \{0,1\}, && \forall k \in \mathcal{N}^0,\ \forall i \in \mathcal{N}(k),\ \forall t \in \mathcal{T},
\end{aligned}
$$

where $\overline{X}_i(t+1) = X_i(0) + \sum_{\tau=0}^{t} \left(\bar{\alpha}_{0i}(\tau) D_{0i}(\tau) - \sum_{s \in \mathcal{S}(i)} \overline{W}_s(\tau)\right)$ *if* $i \in \mathcal{M}$*;* $\overline{X}_i(t+1) = X_i(0) + \sum_{\tau=0}^{t} \left(D_{ki}(\tau) - \sum_{s \in \mathcal{S}(i)} \overline{W}_s(\tau)\right)$ *if* $i \in \mathcal{L} \cup \mathcal{S}$*.*

Proof. Consider time period t and $i \in \mathcal{M}$. The auxiliary problem with $\mathbf{Z}_{0i}^{\alpha}(t)$ in the first constraint of RN formulation (4.9) attains optimality when $Z_{0i}^{\alpha}(\tau)$ is set to zero for all $\tau \leq t$. The two auxiliary problems with $\mathbf{Z}_s^W(t)$, namely $\max_{\mathcal{Z}_s^W(t)} \sum_{\tau=0}^{t} -\hat{W}_s(\tau) Z_s^W(\tau)$ and $\max_{\mathcal{Z}_s^W(t)} \sum_{\tau=0}^{t} \hat{W}_s(\tau) Z_s^W(\tau)$, are equivalent due to the opposite objectives and symmetry of the uncertainty set. Let us consider the auxiliary problem with $\mathbf{Z}_s^W(t)$ that comes from RN formulation (4.9),

$$
\begin{aligned}
\max \quad & \sum_{\tau=0}^{t} \widehat{W}_s(\tau) Z_s^W(\tau) \\
\text{s.t.} \quad & 0 \leq Z_s^W(\tau) \leq 1, \quad \forall \tau \leq t, \\
& \sum_{\tau=0}^{t} Z_s^W(\tau) \leq \Gamma_s^W(t).
\end{aligned}
\tag{4.13}
$$

The dual problem is

$$
\begin{aligned}
\min \quad & O_s(t)\Gamma_s^W(t) + \sum_{\tau=0}^{t} P_s(\tau,t) \\
\text{s.t.} \quad & O_s(t) + P_s(\tau,t) \geq \widehat{W}_s(\tau), \quad \forall \tau \leq t, \\
& O_s(t) \geq 0,\ \ P_s(\tau,t) \geq 0, \quad \forall \tau \leq t,
\end{aligned}
\tag{4.14}
$$

where $O_s(t)$ and $P_s(\tau,t)$ are the dual variables corresponding to the constraints in auxiliary problem (4.13). The same dual problems (4.14) applies to the auxiliary problems with $\mathbf{Z}_s^W(t)$ embedded in the tth holding/shortage constraints for $i \in \mathcal{L} \cup \mathcal{S}$.

Similarly, we consider the following auxiliary problem with $\mathbf{Z}_{0i}^{\alpha}(t)$ that comes from RN formulation (4.9),

$$\begin{aligned} \max \quad & \sum_{\tau=0}^{t} \widehat{\alpha}_{0i}(\tau) Z_{0i}^{\alpha}(\tau) D_{0i}^{*}(\tau) \\ \text{s.t.} \quad & 0 \leq Z_{0i}^{\alpha}(\tau) \leq 1, \quad \forall \tau \leq t, \\ & \sum_{\tau=0}^{t} Z_{0i}^{\alpha}(\tau) \leq \Gamma_{0i}^{\alpha}(t), \end{aligned} \tag{4.15}$$

where $D_{0i}^{*}(\tau)$ is an optimal solution of RN formulation (4.9) and considered as given. The dual problem is the following,

$$\begin{aligned} \min \quad & Q_{0i}(t)\Gamma_{0i}^{\alpha}(t) + \sum_{\tau=0}^{t} R_{0i}(\tau, t) \\ \text{s.t.} \quad & Q_{0i}(t) + R_{0i}(\tau, t) \geq \widehat{\alpha}_{0i}(\tau) D_{0i}^{*}(\tau), \quad \forall\ \tau \leq t, \\ & Q_{0i}(t) \geq 0,\ R_{0i}(\tau, t) \geq 0, \quad \forall\ \tau \leq t, \end{aligned} \tag{4.16}$$

where $Q_{0i}(t)$ and $R_{0i}(\tau, t)$ are the dual variables corresponding to the constraints in auxiliary problem (4.15).

To deal with the auxiliary problem in robust coupling constraint (4.11), we rewrite the constraint as

$$\begin{aligned} \sum_{i \in \mathcal{N}(k)} D_{ki}(t) \leq \overline{X}_k(t) - & \sum_{i \in \mathcal{N}(k)} \overline{X}_i(t) \\ & - \max \sum_{\tau=0}^{t-1} \widehat{\alpha}_{0k}(\tau) Z_{0k}^{\alpha}(\tau) D_{0k}(\tau), \quad \forall k \in \mathcal{M}, \forall t \in \mathcal{T} \end{aligned} \tag{4.17}$$

where $0 \leq Z_{0k}^{\alpha}(\tau) \leq 1$ for all $\tau \leq t-1$ and $\sum_{\tau=0}^{t-1} Z_{0k}^{\alpha}(\tau) \leq \Gamma_{0k}^{\alpha}(t-1)$. This allows us to use the auxiliary problem in (4.15), however, for time period $t-1$. Thus, we obtain the robust counterpart of Constraint (4.17) as

$$\sum_{i \in \mathcal{N}(k)} D_{ki}(t) \leq \overline{X}_k(t) - \sum_{i \in \mathcal{N}(k)} \overline{X}_i(t) - \Big(Q_{0k}(t-1)\Gamma_{0k}^{\alpha}(t-1) + \sum_{\tau=0}^{t-1} R_{0k}(\tau, t-1) \Big). \tag{4.18}$$

We substitute the dual problems (4.14) and (4.16), along with Constraint (4.18), back into RN formulation (4.9) and obtain the equivalent robust counterpart as shown in RNC formulation (4.12). □

Although this RNC formulation (4.12) is much complicated than RSC formulation (3.8) proposed in Chapter 3 in terms of problem size, it still maintains the same difficulty as its nominal problem. When there are no setup costs (i.e., $K_{ki}(t) = 0$ for all k, $i \in \mathcal{N}(k)$ and t), it can be solved within polynomial time to an optimal solution.

Next, it is natural to investigate whether the theoretical results regarding the optimal robust policy in RSC exist for RNC formulation.

4.4 Optimal Robust Policy

We present the following results regarding the optimal robust policy for RNC formulation (4.12).

Theorem 4.2 (Optimal robust policy)**.** *Let $O_s^*(t)$, $P_s^*(t)$, $Q_{0k}^*(t)$ and $R_{0k}^*(t)$ be an optimal solution of RNC formulation (4.12), then we have:*

(a) The optimal policy for echelon $k \in \mathcal{M}$ in RNC can be obtained by solving the DS problem (3.1)-(3.4) subject to the modified, deterministic demand $\sum_{s \in \mathcal{S}(k)} W'_{s,k}(t)$ in period t,

$$W'_{s,k}(t) = \overline{W}_s(t) + \Big(\Upsilon_{s,k}(t) - \Upsilon_{s,k}(t-1)\Big) + \Big(\Psi_{0k}(t) - \Psi_{0k}(t-1)\Big), \qquad \forall k \in \mathcal{M} \tag{4.19}$$

where $\Upsilon_{s,k}(-1) = 0$ *and* $\Upsilon_{s,k}(t) := ((b_k(t) - h_k(t))/(b_k(t) + h_k(t)))A_s(t)$ *with* $A_s(t) = O^*_s(t)\Gamma^W_s(t) + \sum_{\tau=0}^{t} P^*_s(\tau, t)$ *being the accumulated deviation of the uncertain demand from its nominal value in period* t *at store* s*;* $\Psi_{0k}(-1) = 0$ *and* $\Psi_{0k}(t) = (b_k(t)/(N_k(b_k(t)+h_k(t)))B_{0k}(t)$ *with* $B_{0k}(t) = Q^*_{0k}(t)\Gamma^\alpha_{0k}(t) + \sum_{\tau=0}^{t} R^*_{0k}(\tau, t)$ *being the accumulated deviation of the uncertain supply ratio from its nominal value in period* t *at MSH* k*, and where we use* N_k *to represent the number of stores within echelon* k *(i.e.,* $N_k = |\mathcal{S}(k)|$*).*

(b) The optimal robust policy for echelon $k \in \mathcal{L} \cup \mathcal{S}$ *in RNC can be obtained by solving the DS problem* (3.1)-(3.4) *with time-varying capacity (given in* (4.23) *and* (4.24)*) on the orders, subject to the modified, deterministic demand* $\sum_{s\in\mathcal{S}(k)} W'_{s,k}(t)$ *in period* t*,*

$$W'_{s,k}(t) = \overline{W}_s(t) + \Big(\Upsilon_{s,k}(t) - \Upsilon_{s,k}(t-1)\Big), \qquad \forall k \in \mathcal{L} \cup \mathcal{S} \tag{4.20}$$

(c) If there is no setup cost, the optimal robust policy for echelon k *also can be obtained by solving the DS problem* (3.1)-(3.4) *with new cost coefficients, subject to the modified, deterministic demand* $\sum_{s\in\mathcal{S}(k)} W'_{s,k}(t)$ *in period* t*.*

(d) The optimal cost of RNC is equal to the total optimal cost of the set of DS problems (3.1)-(3.4) *subject to the modified demand* $\sum_{s\in\mathcal{S}(k)} W'_{s,k}(t)$ *in period* t*, plus the following extra cost incurred by the robust policy,*

$$COST = \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{N}} \frac{2b_k(t)h_k(t)}{b_k(t) + h_k(t)} \sum_{s\in\mathcal{S}(k)} A_s(t) + \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{M}} \frac{b_k(t)h_k(t)}{b_k(t) + h_k(t)} B_{0k}(t).$$

Proof. The proof follows the same way as in the proof of Theorem 3.2 by reformulating the problem for each echelon as the nominal with the modified demand. We first reformulate the problem for echelon $k \in \mathcal{M}$. We assume that $O^*_s(t)$, $P^*_s(t)$, $Q^*_{0k}(t)$ and $R^*_{0k}(t)$ are given, then we need to solve following DS problem,

$$\min_{\mathbf{D}\geq 0} \sum_{t\in\mathcal{T}} \Big\{ c_{0k}(t)D_{0k}(t) + K_{0k}(t)1_{\{D_{0k}(t)>0\}}$$

$$+\max\left(h_k(t)(\overline{X}_k(t+1)+\sum_{s\in\mathcal{S}(k)}A_s(t)),b_k(t)(-\overline{X}_k(t+1)+\sum_{s\in\mathcal{S}(k)}A_s(t)+B_{0k}(t))\right)\Bigg\},\tag{4.21}$$

where $\overline{X}_k(t+1)=X_k(0)+\sum_{\tau=0}^{t}\left(\bar{\alpha}_{0k}(\tau)D_{0k}(\tau)-\sum_{s\in\mathcal{S}(k)}\overline{W}_s(\tau)\right)$, $A_s(t)=O_s^*(t)\Gamma_s^W(t)+\sum_{\tau=0}^{t}P_s^*(\tau,t)$ and $B_{0k}(t)=Q_{0k}^*(t)\Gamma_{0k}^{\alpha}(t)+\sum_{\tau=0}^{t}R_{0k}^*(\tau,t)$. Let us define the modified inventory level variable $X_k'(t+1)$ as

$$\begin{aligned}X_k'(t+1)&=X_k'(t)+\bar{\alpha}_{0k}(t)D_{0k}(t)\\&\quad-\sum_{s\in\mathcal{S}(k)}\underbrace{\left\{\overline{W}_s(t)+\Big(\Upsilon_{s,k}(t)-\Upsilon_{s,k}(t-1)\Big)+\Big(\Psi_{0k}(t)-\Psi_{0k}(t-1)\Big)\right\}}_{W'_{s,k}(t)},\end{aligned}\tag{4.22}$$

with $X_i'(0)=X_i(0)$. Then $X_k'(t)$ can be recursively written as follows:

$$\begin{aligned}X_k'(1)&=X_k'(0)+\bar{\alpha}_{0k}(0)D_{0k}(0)\\&\quad-\sum_{s\in\mathcal{S}(k)}\Big\{\bar{W}_s(0)+(\Upsilon_{s,k}(0)-\Upsilon_{s,k}(-1))+(\Psi_{0k}(0)-\Psi_{0k}(-1))\Big\},\\X_k'(2)&=X_k'(1)+\bar{\alpha}_{0k}(1)D_{0i}(1)\\&\quad-\sum_{s\in\mathcal{S}(k)}\Big\{\bar{W}_s(1)+(\Upsilon_{s,k}(1)-\Upsilon_{s,k}(0))+(\Psi_{0k}(1)-\Psi_{0k}(0))\Big\}\\&=X_k'(0)+\sum_{\tau=0}^{1}\Big(\bar{\alpha}_{0k}(\tau)D_{0k}(\tau)-\sum_{s\in\mathcal{S}(k)}\bar{W}_s(\tau)\Big)-\sum_{s\in\mathcal{S}(k)}\Big(\Upsilon_{s,k}(1)+\Psi_{0k}(1)\Big),\\&\vdots\\X_k'(t+1)&=X_k'(0)+\sum_{\tau=0}^{t}\Big(\bar{\alpha}_{0k}(\tau)D_{0k}(\tau)-\sum_{s\in\mathcal{S}(k)}\bar{W}_s(\tau)\Big)-\sum_{s\in\mathcal{S}(k)}\Big(\Upsilon_{s,k}(t)+\Psi_{0k}(t)\Big)\\&=\bar{X}_k(t+1)-\sum_{s\in\mathcal{S}(k)}\frac{b_k(t)-h_k(t)}{b_k(t)+h_k(t)}A_s(t)-\frac{b_k(t)}{b_k(t)+h_k(t)}B_{0k}(t).\end{aligned}$$

Substituting $X'_k(t+1)$ in the following equations, we have

$$\begin{aligned}
&h_k(t)X'_k(t+1) + \frac{2b_k(t)h_k(t)}{b_k(t)+h_k(t)} \sum_{s\in\mathcal{S}(k)} A_s(t) + \frac{b_k(t)h_k(t)}{b_k(t)+h_k(t)} B_{0k}(t) \\
&\quad = h_k(t)\bar{X}_k(t+1) - \left(\frac{h_k(t)(b_k(t)-h_k(t))}{b_k(t)+h_k(t)} - \frac{2b_k(t)h_k(t)}{b_k(t)+h_k(t)}\right) \sum_{s\in\mathcal{S}(k)} A_s(t) \\
&\qquad\qquad - \left(\frac{b_k(t)h_k(t)}{b_k(t)+h_k(t)} - \frac{b_k(t)h_k(t)}{b_k(t)+h_k(t)}\right) B_{0k}(t) \\
&\quad = h_k(t)\bar{X}_k(t+1) + h_k(t) \sum_{s\in\mathcal{S}(k)} A_s(t),
\end{aligned}$$

and

$$\begin{aligned}
&-b_k(t)X'_k(t+1) + \frac{2b_k(t)h_i(t)}{b_k(t)+h_k(t)} \sum_{s\in\mathcal{S}(k)} A_s(t) + \frac{b_k(t)h_k(t)}{b_k(t)+h_k(t)} B_{0k}(t) \\
&\quad = -b_k(t)\bar{X}_k(t+1) + \left(\frac{b_k(t)(b_k(t)-h_k(t))}{b_k(t)+h_k(t)} + \frac{2b_k(t)h_k(t)}{b_k(t)+h_k(t)}\right) \sum_{s\in\mathcal{S}(k)} A_s(t) \\
&\qquad\qquad + \left(\frac{b_k^2(t)}{b_k(t)+h_k(t)} + \frac{b_k(t)h_k(t)}{b_k(t)+h_k(t)}\right) B_{0k}(t) \\
&\quad = -b_k(t)\bar{X}_k(t+1) + b_k(t) \sum_{s\in\mathcal{S}(k)} A_s(t) + b_k(t)B_{0k}(t).
\end{aligned}$$

The inventory cost of Problem (4.21) in period t can be written as

$$\begin{aligned}
&\max\left(h_k(t)(\overline{X}_k(t+1) + \sum_{s\in\mathcal{S}(k)} A_s(t)), b_k(t)(-\overline{X}_k(t+1) + \sum_{s\in\mathcal{S}(k)} A_s(t) + B_{0k}(t))\right) \\
&= \max\left(h_k(t)X'_k(t+1), -b_k(t)X'_k(t+1)\right) + \frac{2b_k(t)h_k(t)}{b_k(t)+h_k(t)} \sum_{s\in\mathcal{S}(k)} A_s(t) + \frac{b_k(t)h_k(t)}{b_k(t)+h_k(t)} B_{0k}(t),
\end{aligned}$$

therefore, we obtain the following DS problem subject to the modified demand $\sum_{s\in\mathcal{S}(k)} W'_{s,k}(t)$ (plus the fixed cost $\frac{2b_k(t)h_k(t)}{b_k(t)+h_k(t)} \sum_{s\in\mathcal{S}(k)} A_s(t) + \frac{b_k(t)h_k(t)}{b_k(t)+h_k(t)} B_{0k}(t)$) in period t,

$$\min_{\mathbf{D}\geq 0} \sum_{t\in\mathcal{T}} \Big\{ c_{0k}(t)D_{0k}(t) + K_{0k}(t)1_{\{D_{0k}(t)>0\}}$$

$$+ \max\Big(h_k(t)X_k'(t+1), -b_k(t)X_k'(t+1)\Big) + \frac{2b_k(t)h_k(t)}{b_k(t)+h_k(t)} \sum_{s\in\mathcal{S}(k)} A_s(t) + \frac{b_k(t)h_k(t)}{b_k(t)+h_k(t)} B_{0k}(t)\Big\}.$$

This proves (a). Using the same reformulation tactic, we have the following equivalent DS problem for echelon $k \in \mathcal{L} \cup \mathcal{S}$,

$$\min_{\mathbf{D}\geq 0} \sum_{t\in\mathcal{T}} \Big\{ c_{i_k k}(t) D_{i_k k}(t) + K_{i_k k}(t) 1_{\{D_{i_k k}(t)>0\}} \\ + \max\Big(h_k(t)X_k'(t+1), -b_k(t)X_k'(t+1)\Big) + \frac{2b_k(t)h_k(t)}{b_k(t)+h_k(t)} \sum_{s\in\mathcal{S}(k)} A_s(t)\Big\}.$$

Hence, the result in (d) follows from combining the extra costs incurred at all the echelons.

It is noteworthy that for echelon $k \in \mathcal{L} \cup \mathcal{S}$ and $i_k \in \mathcal{M}$, the orders placed in period t have to satisfy

$$D_{i_k k}(t) \leq \overline{X}_{i_k}(t) - \sum_{i\in\mathcal{N}(i_k)} \overline{X}_i(t) - \sum_{j\in\mathcal{N}(i_k), j\neq k} D_{i_k j}(t) - B_{0 i_k}(t-1), \tag{4.23}$$

while the orders placed by echelon $k \in \mathcal{L} \cup \mathcal{S}$ and $i_k \in \mathcal{L}$ have to satisfy

$$D_{i_k k}(t) \leq \overline{X}_{i_k}(t) - \sum_{i\in\mathcal{N}(i_k)} \overline{X}_i(t) - \sum_{j\in\mathcal{N}(i_k), j\neq k} D_{i_k j}(t). \tag{4.24}$$

The result in (b) immediately follows from setting the RHS of Constraints (4.23) and (4.24) to optimality.

To prove (c), since there is no setup cost, we can effectively decouple the echelons by dualizing the coupling constraints in RNC formulation (4.12) according to the Lagrangian multiplier method. This transforms RNC into a LP problem where the feasible set is separable in the echelons and the resulting objective contains the Lagrangian multipliers. After combining the terms in the objective, we obtain a new objective which is the sum of several uncapacitated single-station problems. The result in (c) follows from invoking the strong duality for the Lagrange dual objective and applying the reformulation tactic. □

Similar to that in RSC formulation (3.8), when intervals for $\alpha_{0k}(t)$ and $W_s(t)$ get wider (i.e., $\widehat{\alpha}_{0k}(t)$ and $\widehat{W}_s(t)$ increase), it gets more expensive to protect the robustness of the optimal solution against uncertainty since the extra cost increases accordingly.

Note that the optimal robust policies are distinguished by echelons. That is, the equivalent problem for echelon k is the uncapacitated DS problem with the modified demand involving both $A_s(t)$ and $B_{0k}(t)$ if $k \in \mathcal{M}$; otherwise, it is the capacitated DS problem with the modified demand only involving $A_s(t)$. The difference in the expressions of the modified demand is caused by the fact that the supply uncertainty only influences the orders placed by MSHs. Note that the optimal robust policy for echelon $k \in \mathcal{L} \cup \mathcal{S}$ is tightly connected to the single-station problem with a time-varying capacity C_t^{ord} on the maximal order (see Section 3.5.1) in the way that we can view $\mathrm{C}_t^{ord} = \overline{X}_{i_k}(t) - \sum_{i \in \mathcal{N}(i_k)} \overline{X}_i(t) - \sum_{j \in \mathcal{N}(i_k), j \neq k} D_{i_k j}(t) - B_{0 i_k}(t-1)$ for echelon $k \in \mathcal{L} \cup \mathcal{S}$ and $i_k \in \mathcal{M}$; $\mathrm{C}_t^{ord} = \overline{X}_{i_k}(t) - \sum_{i \in \mathcal{N}(i_k)} \overline{X}_i(t) - \sum_{j \in \mathcal{N}(i_k), j \neq k} D_{i_k j}(t)$ for echelon $k \in \mathcal{L} \cup \mathcal{S}$ and $i_k \in \mathcal{L}$.

The reason why the expression of $\Psi_{0k}(t)$ contains N_k can be interpreted in the way that the effect of incorporating the supply uncertainty into an echelon has been evenly divided in terms of the modified demand to the stores within the echelon.

4.5 Numerical Studies

In this section, we compare the performance of three inventory policies: RO^0 (nominal policy), RO^1(robust policy of Bertsimas and Thiele, 2006) and RO^2 (proposed robust policy) in an network example through simulations. We focus on the average performance as well as the performance stability of using RO^2. To achieve that, we measure the average performance by ratio $\mathbb{E}(\mathrm{R}^{i-2})$, $i = 0, 1$ (definition see Section 3.6), that is computed with respect to 100 replications of realized supply ratios and demands, while the performance stability is estimated by observing standard deviation in the sample distribution of R^{i-2}, $i = 0, 1$. We study how the setup costs and the length of planning horizon affect the performance stability.

4.5.1 Experiment Setting

The network is depicted in Figure 4.2 and described as follows. There are one central warehouse (i.e., node 1) and two stores (i.e., nodes 2 and 3) in the network. The warehouse receives supplies from an external supplier (i.e., node 0) and ships the items directly to the stores. Recall the definition of echelon given in Section 4.2, the network therefore consists of totally three echelons, where: Echelon 1 consists of nodes 1, 2 and 3, and the links in-between; Echelons 2 and 3 only consist of nodes 2 and 3, respectively.

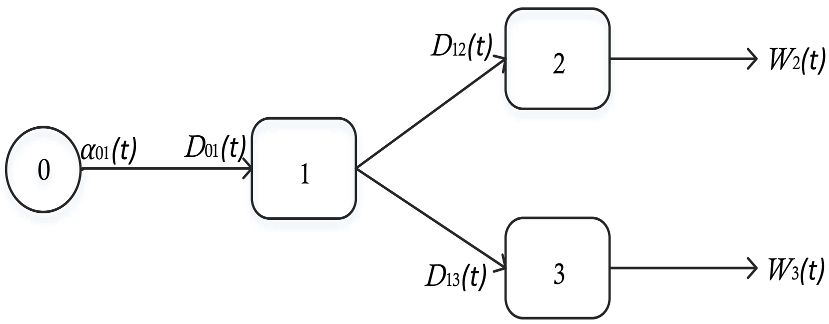

Figure 4.2: A three-echelon network

We consider the network with the planning horizons of $T = 10$, 20 and 30. The cost parameters are assumed to be stationary and selected as follows. Let $c_{01}(t) = c_{12}(t) = c_{13}(t) = 1$, $h_1(t) = 0.1$, $b_1(t) = 4$, $h_2(t) = h_3(t) = 0.2$ and $b_2(t) = b_3(t) = 5$ for all t. When the setup costs are explicitly considered, we use $K_{01}(t) = 35$ and

$K_{12}(t) = K_{13}(t) = 10$. The initial inventories at Echelon 1, 2 and 3 are given by 80, 20 and 10, respectively.

In the simulation, demands $W_2(t)$ and $W_3(t)$ are assumed to be i.i.d. and generated from two lognormal, gamma distributions with the same mean $\mu_W = 100$ and standard deviation $\delta_W = 20$, or two uniform distributions from the same interval $[\mu_W - \delta_W, \mu_W + \delta_W]$; the supply ratio $\alpha_{01}(t)$ is also assumed to be i.i.d. and generated from a lognormal distribution with mean $\mu_\alpha = 0.9$ and standard deviation $\delta_\alpha = 0.05$. If the generated α_t is greater than 1, we set it to 1 to avoid oversupply.

To obtain the inventory policies, we let $\alpha_{01}(t)$ take values from the interval $[\mu_\alpha - 2\delta_\alpha, 1]$, i.e., $\bar{\alpha}_{01}(t) = 1$ and $\widehat{\alpha}_{01}(t) = \bar{\alpha}_{01}(t) - \mu_\alpha + 2\delta_\alpha = 0.2$. We let both $W_2(t)$ and $W_3(t)$ take values from the interval $[\mu_W - 2\delta_W, \mu_W + 2\delta_W]$, i.e., $\overline{W}_i = \mu_W$ and $\widehat{W}_i = 2\delta_W$, $i = 2, 3$. In addition, the linear budget functions are given by $\Gamma_2^W(t) = \gamma_2^W + \gamma_2^W \cdot t$, $\Gamma_3^W(t) = \gamma_3^W + \gamma_3^W \cdot t$, and $\Gamma_{01}^\alpha(t) = \gamma_{01}^\alpha + \gamma_{01}^\alpha \cdot t$ for $W_2(t)$, $W_3(t)$ and $\alpha_{01}(t)$, respectively. In the experiment, we set $\gamma_2^W = \gamma_3^W = \gamma_{01}^\alpha = 0.2$. With the above parameters, the policies RO^0, RO^1 and RO^2 can be obtained by solving the corresponding network models.

The computational effectiveness of the network models are reported in Table 4.1. In the table, we report the optimal costs of the network, the time spent to solve the network models in seconds, the percentage gap obtained by each model under each combination of the planning horizon and the setup costs. In addition, for the models with setup costs, we also report the number of orders placed by the echelons in the sequence of Echelons 1, 2 and 3 in the last column. Again, the results in "Time (s)" and "Gap%" columns are the average values. The results indicate that the computational time of the proposed robust model scales reasonably well with respect to the length of the planning periods. It takes 0.26 seconds, 23.07 seconds, and 83.23 seconds on average to obtain the robust policy RO^2 for the problem with $T = 10$, 20 and 30, respectively.

We find that the proposed robust network model shows its potential for practical implementation. It typically takes less than 90 seconds with the average of 83.23 seconds to obtain RO^2 for the network with $T = 30$ when the setup costs are present. Note that we also report the number of orders placed by each echelon for the setup cost case. An interesting observation is that incorporating supply uncertainty heavily affects the ordering frequency at Echelon 1, whereas the ordering frequencies at Echelons 2 and 3 are not affected at all, as they remain the same across three policies. Intuitively, this makes sense because the deviation of uncertain partial supply is not very high in the model when $\widehat{\alpha}_t = 0.2$. The effect of deviation is further divided at Echelons 2 and 3 and thus does not cause too much impact on the ordering policy at Echelons 2 and 3. However, the nominal policy RO^0 at Echelons 2 and 3 can be easily distinguished from the robust policies RO^1 and RO^2 in terms of the order size because of the demand uncertainty considered in the robust models.

Table 4.1: Computational results for solving the network models using stationary cost parameters $c_{01}(t) = c_{12}(t) = c_{13}(t) = 1,\ h_1(t) = 0.1,\ h_2(t) = h_3(t) = 0.2,\ b_1(t) = 4,\ b_2(t) = b_3(t) = 5$ when $T = 10,\ 20$ and 30.

		$K_{01}(t) = K_{12}(t) = K_{13}(t) = 0$			$K_{01}(t) = 35,\ K_{12}(t) = K_{13}(t) = 10$			
Policy	T	Obj.	Time (s)	Gap%	Obj.	Time (s)	Gap%	#Ord.^{1}/Ord.^{2}/Ord.^{3}
	10	4670.0	0.01	0.00	5115.0	0.15	0.61	5 / 9 /10
RO^0	20	8870.0	0.01	0.00	9790.0	0.44	0.88	10 / 20 / 19
	30	13070.0	0.02	0.00	14465.0	2.97	0.31	15 / 30 / 29
	10	5565.2	0.02	0.00	6017.0	0.14	0.46	5 / 9 / 10
RO^1	20	11511.0	0.02	0.00	12455.2	0.54	0.52	10 / 20 / 19
	30	18380.0	0.03	0.00	19797.0	3.55	0.78	15 / 30 / 29
	10	5730.4	0.02	0.00	6243.6	0.26	0.88	9 / 9 / 10
RO^2	20	11902.1	0.02	0.00	12972.7	23.07	0.67	19 / 20 / 19
	30	19086.3	0.05	0.00	20707.8	83.23	0.87	29 / 30 / 29

4.5.2 Comparison of Policies

The expected relative performances are summarized in Table 4.2. It also provides the lowest relative performances for R^{0-2} (resp., R^{1-2}) among 100 replications in $\mathrm{R}^{0-2}_{\text{low.}}$ (resp., $\mathrm{R}^{1-2}_{\text{low.}}$) columns. From this table, it can be seen that RO^2 performs exceptionally well compared to RO^0 with all the lowest R^{0-2} staying positive (at least 1.85% higher in terms of cost saving), which suggests the importance of incorporating uncertainty into decision-making. While the lowest R^{1-2} has negative values in the table, RO^2 consistently outperforms RO^1 in the average performance. Moreover, we see that RO^2 provides improved performance as T increases, which coincides with our conclusion in Section 3.6.3 that RO^2 is well-suited for the system in which the decision maker is more concerned with the long-term performance. It is also important to note that the expected relative performance is negatively affected by the setup costs.

Table 4.2: Expected relative performance, in percent, using stationary cost parameters with the planning horizons of $T = 10,\ 20$ and 30.

		$K_{01}(t) = K_{12}(t) = K_{13}(t) = 0$				$K_{01}(t) = 35,\ K_{12}(t) = K_{13}(t) = 10$			
Demand Dist.	T	$\mathbb{E}(\mathrm{R}^{0-2})$	$\mathbb{E}(\mathrm{R}^{1-2})$	$\mathrm{R}^{0-2}_{\text{low.}}$	$\mathrm{R}^{1-2}_{\text{low.}}$	$\mathbb{E}(\mathrm{R}^{0-2})$	$\mathbb{E}(\mathrm{R}^{1-2})$	$\mathrm{R}^{0-2}_{\text{low.}}$	$\mathrm{R}^{1-2}_{\text{low.}}$
$W_2(t)$~*lognorm* $W_3(t)$~*lognorm*	10	41.23	6.52	14.12	1.39	35.96	5.16	10.72	-2.29
	20	61.52	22.90	42.63	3.73	55.81	11.82	32.12	-0.99
	30	65.90	29.33	43.82	14.77	60.71	17.39	41.48	4.12
$W_2(t)$~*uniform* $W_3(t)$~*uniform*	10	33.64	5.40	5.05	-3.89	22.99	3.14	1.85	-7.32
	20	50.32	10.05	21.41	-1.09	41.09	8.64	10.15	-6.54
	30	56.35	15.09	28.87	3.46	47.66	11.27	18.05	-3.88
$W_2(t)$~*gamma* $W_3(t)$~*gamma*	10	44.47	8.33	13.64	-1.32	36.39	6.07	4.85	-3.71
	20	63.72	27.76	30.46	2.63	55.20	13.32	19.67	-1.56
	30	66.94	31.44	33.40	12.46	59.27	18.80	20.17	2.02

While we have simulated the system with different demand distributions, the demand distribution does not seem to have a significant impact on the performance except the degree of out-performance. In particular, we find that the out-performance of RO^2 is stronger under the gamma distribution of demand.

To supplement the analysis, we show the performance stability between two robust policies RO^1 and RO^2 through the view of sample distribution. In the following, we show the results of the lognormally distributed supply ratios and demands only, because similar results are observed for other demand distributions.

The variability of R^{1-2} with (resp., without) setup costs is visualized in Figure 4.3 (resp., Figure 4.4). We see that RO^2 offers the better average performance relative to RO^1 as T increases, however, at the expense of the slightly higher standard deviation of the sample distribution, i.e., it ranges from 3.35% (resp., 3.80%) as $T = 10$ to 5.39% (resp., 6.34%) as $T = 30$ in the case of zero (resp., non-zero) setup costs. The results also reveal that although the presence of setup costs negatively affects the average performance, it does not much hurt the performance stability of the robust policy RO^2.

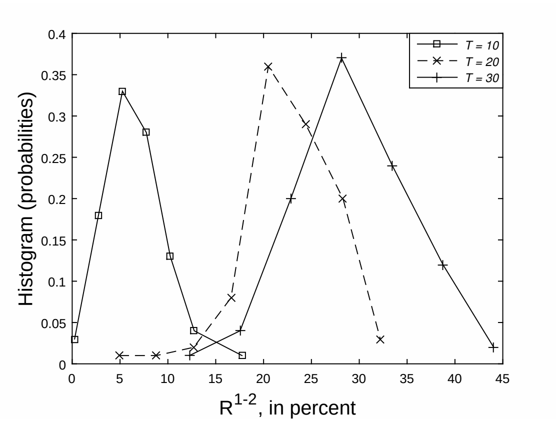

Figure 4.3: Sample distribution of the relative performances without setup costs when T=10, 20 and 30.

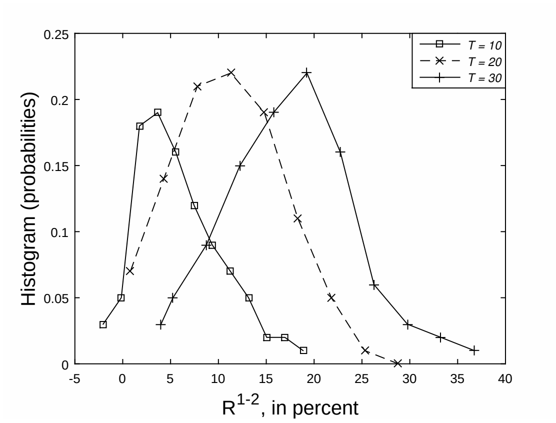

Figure 4.4: Sample distribution of the relative performances with setup costs when T=10, 20 and 30.

Figures 4.5 to 4.10 show the sample distributions of the costs with the lognormal demands at both stores. The figures clearly demonstrate that RO^2 performs better as T increases, since it yields lower average costs and standard deviation than RO^0 and

RO^1. On the other hand, when T is small, we observe in Figures 4.5 and 4.6 that the difference between two robust policies RO^1 and RO^2 is limited. Overall, we conclude that RO^2 outperforms RO^0 and RO^1, in terms of both average cost saving and stability, when applied to the network example, especially in the situations of the long planning horizon.

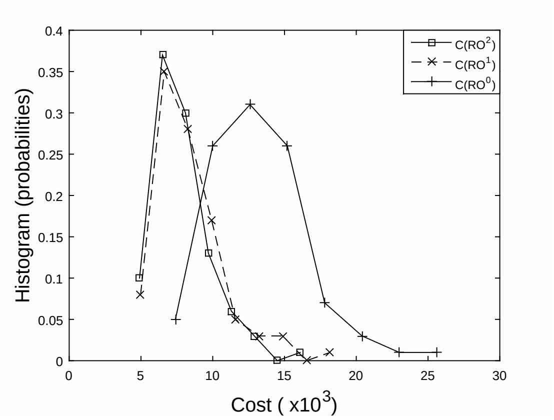

Figure 4.5: Sample distribution of costs with setup costs when $T = 10$.

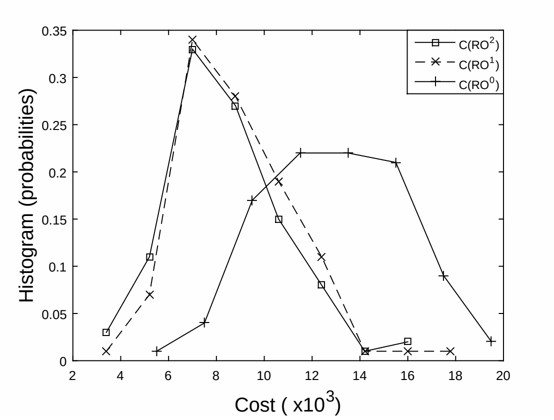

Figure 4.6: Sample distribution of costs without setup costs when $T = 10$.

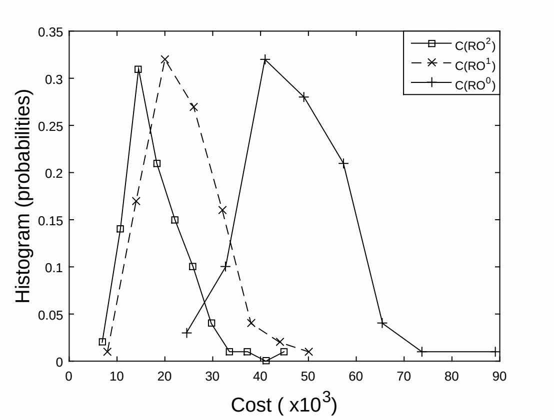

Figure 4.7: Sample distribution of the costs with setup costs when $T = 20$.

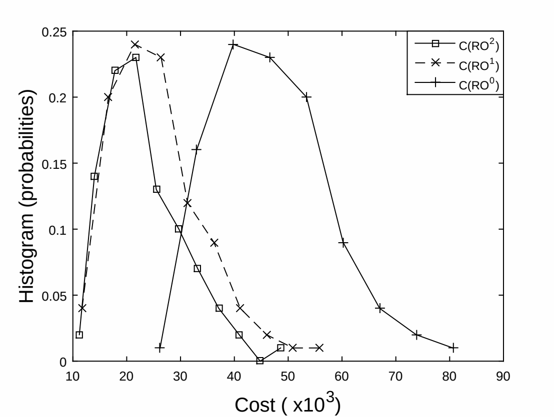

Figure 4.8: Sample distribution of costs without setup costs when $T = 20$.

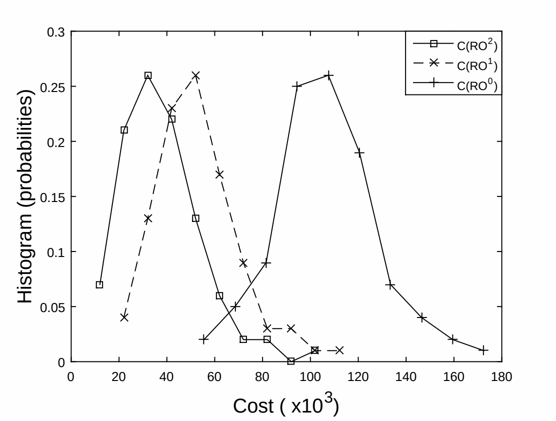

Figure 4.9: Sample distribution of the costs with setup costs when $T = 30$.

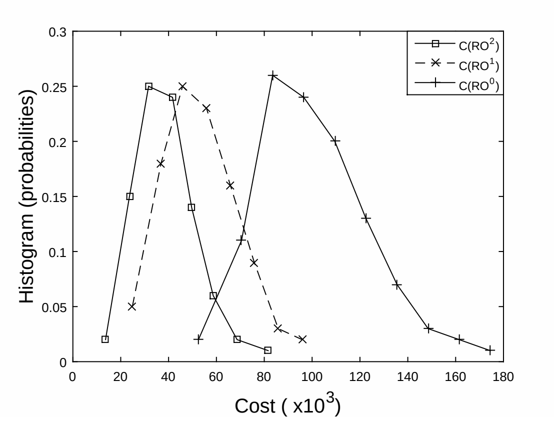

Figure 4.10: Sample distribution of costs without setup costs when $T = 30$.

We also test the performance of the inventory policies in the non-stationary cost system with $T = 20$ and report the results in Table 4.3. In each time period t, the unit variable costs $c_{01}(t), c_{12}(t)$ and $c_{13}(t)$ are generated from a uniform distributions from the interval $[0.6, 1.4]$; the unit holding cost $h_1(t)$ is generated from a uniform distribution from the interval $[0.05, 0.15]$, while the unit holding costs $h_2(t)$ and $h_3(t)$ are generated from a uniform distributions from the interval $[0.15, 0.25]$; the unit shortage cost $b_1(t)$ is generated from a uniform distribution from the interval $[3, 5]$, while the unit shortage costs $b_2(t)$ and $b_3(t)$ are generated from a uniform distribution from the interval $[4, 6]$. The setup costs remain stationary when they apply, however, with two possible combinations of $K_{01}(t) = 35, K_{12}(t) = K_{13}(t) = 10$ and $K_{01}(t) = 70, K_{12}(t) = K_{13}(t) = 15$. With the exception of the case with relatively higher setup costs and uniform demands, the proposed robust policy RO^2 performs well compared to the nominal policy RO^0 and the robust policy RO^1, which further demonstrates the high potential of the proposed robust network model for use in more realistic supply chain settings.

Table 4.3: Expected relative performance, in percent, using non-stationary cost parameters with $T = 20$.

	$K_{01}(t) = 0$ $K_{12}(t) = K_{13}(t) = 0$		$K_{01}(t) = 35$ $K_{12}(t) = K_{13}(t) = 10$		$K_{01}(t) = 70$ $K_{12}(t) = K_{13}(t) = 15$	
Demand Dist.	$\mathbb{E}(\mathrm{R}^{0-2})$	$\mathbb{E}(\mathrm{R}^{1-2})$	$\mathbb{E}(\mathrm{R}^{0-2})$	$\mathbb{E}(\mathrm{R}^{1-2})$	$\mathbb{E}(\mathrm{R}^{0-2})$	$\mathbb{E}(\mathrm{R}^{1-2})$
$W_2(t) \sim$ *lognorm* $W_3(t) \sim$ *lognorm*	48.16	13.58	26.80	5.71	20.18	2.61
$W_2(t) \sim$ *uniform* $W_3(t) \sim$ *uniform*	27.24	6.12	23.78	2.73	16.99	-3.64
$W_2(t) \sim$ *gamma* $W_3(t) \sim$ *gamma*	33.41	8.28	25.84	3.42	17.54	1.76

4.6 Concluding Remarks

In this chapter, we have presented a RO-based model for a tree-structured supply chain subject to supply and demand uncertainties. The proposed RNC formulation (4.12) maintain computational tractability and thus hold potential for use in practice. Moreover, the echelon-specific cost structure allows us to analyze optimal robust policy by echelons. Therefore, we have shown that the robust network counterpart formulation is made up of several interconnected (capacitated and uncapacitated) single-station problems subject to the deterministic demands. Furthermore, if there are no setup costs in the network, then the problem can be decomposed into several uncapacitated single-station problems with new cost parameters subject to the same deterministic demands. The numerical results indicate that the proposed robust policy could yield significantly better performance than the nominal policy and the robust policy of Bertsimas and Thiele (2006) while achieving good performance stability.

Chapter 5

Stochastic Programming for a Periodic-Review Assemble-To-Order System

5.1 Introduction

Increasing high-tech manufacturing firms (e.g., IBM and Dell Computer) are adopting Assemble-To-Order (ATO) systems for their production and inventory control since they enable the firms to provide responsive service to a variety of customer orders and effectively eliminate the inventory of final products (i.e., inventory consists only of components).

In an ATO system (see Figure 5.1), the components are ordered from outside suppliers and stocked in advance. Once a customer order for a final product has arrived, the needed components are obtained from inventory and assembled into the product in response to the demand. In this way, the firms benefit from the ATO system if the replenishment lead times of the components are substantial compared with the assembly times of the final products. In the context of multi-component, multi-product ATO systems, different final products may share common components. Through the postponement strategy, the ATO systems exploit component commonality and provide product variety at low cost (Song and Zhao, 2009). In case that one or more needed components are out of stock, the demand will be backlogged until the components are replenished.

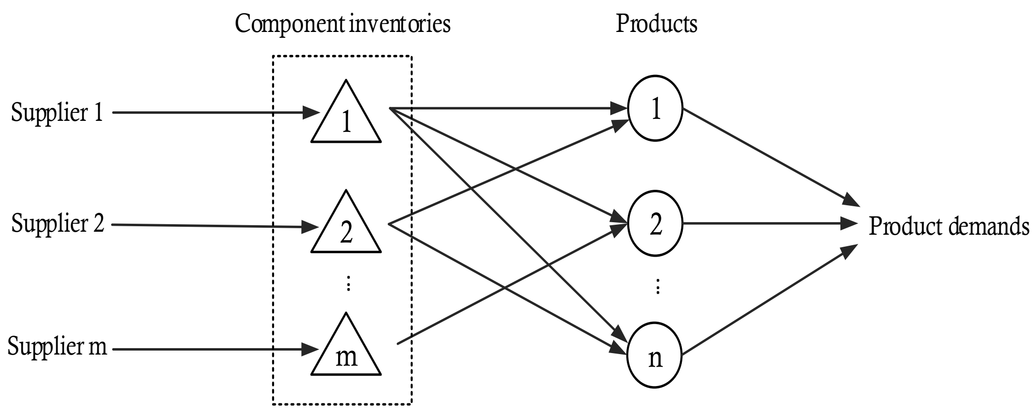

Figure 5.1: ATO system

While ATO systems are popular in practice, in general they are very difficult to analyze and solve to optimality. One difficulty comes from joint optimization, which involves *inventory replenishment* and *component allocation* problems. These two problems need to be addressed simultaneously due to their strong correlations. For example, consider a simple system consisting of only one final product that requires multiple components. Apparently, allocation decisions for a particular component cannot be determined independently because it depends on inventory availability of not only itself but other components. If a *Fist-Come-First-Served* (FCFS) allocation rule is applied (i.e., the product demand of a product in an earlier periods must be satisfied before the demand of that product in later periods), those available inventories remain unused due to lack of other components, called *remnant stock* (Huang and de Kok, 2015)

The optimization of ATO literature can be broadly classified according to three aspects, namely review periods, types of objective of the optimization problem, and the decisions involved in the optimization problems concerning either inventory replenishment optimization or joint optimization. Regarding the review periods, the ATO models can be further classified into single-period models (e.g., Baker et al., 1986, Hsu et al., 2006), periodic-review models (e.g., Zhang, 1997, Agrawal and Cohen, 2001, Akçay and Xu, 2004, Huang and de Kok, 2015) and continuous-review models (e.g., Song et al., 1999, Song and Yao, 2002, Song and Zhao, 2009). In terms of the objectives of the optimization problems, the ATO models are commonly concerned with minimizing the inventory cost (or level) (e.g., Zhang, 1997, Hsu et al., 2006, Huang and de Kok, 2015), or maximizing the service level (e.g., Hausman et al., 1998, Akçay and Xu, 2004). In particular, Huang and de Kok (2015) explicitly investigated the impact of the remnant stock holding cost in a periodic-review model. Depending on the decisions, several works focus on the inventory replenishment optimization (e.g., Zhang, 1997, Agrawal and Cohen, 2001, Lu and Song, 2005). However, Akçay and Xu (2004) have numerically demonstrated that the inventory replenishment and component allocation decisions have a similar impact on the aggregated service performance (i.e., fill rate), which suggests the importance of the joint optimization. Other joint optimization models have been proposed in Hsu et al. (2006), Huang and de Kok (2015), van Jaarsveld and Scheller-Wolf (2015). To alleviate the complexity of the joint optimization, some simple component allocation heuristics have been proposed in the literature; for example, the product-based priority rule (Zhang, 1997), fair-share rule (Agrawal and Cohen, 2001), and the order-based component allocation heuristic (Akçay and Xu, 2004).

With a few exceptions, the above ATO literature assumes the uncertain product demands, while the component replenishment lead times are assumed to be deterministic but possibly differentiated by components. In contrast, Kumar (1989) considered a single-period assembly system with deterministic demands and stochastic procurement lead times. They characterized the expected holding time for each component until it is finally assembled, and showed that the optimal ordering policies under a linear cost structure are independent of the lead time variability. Similarly, Fujiwara and Sedarage (1997) studied an EOQ-based ATO system under (Q, r) inventory control policy, where known demands but stochastic procurement lead times were assumed. Relevant works in the ATO context with deterministic demands and random lead times can also be found in Yano (1987), Chu et al. (1993), Mauroy and Wardi (1995), Shore (1995), and Proth et al. (1997).

Efforts have also been made to simultaneously consider lead time and demand uncertainties in ATO environments. Song and Yao (2002) analyzed a single-product ATO system under the assumptions of stochastic lead times and Poisson demands. They modeled the component supply process as a set of $M/G/\infty$ queues. Later, Lu et al. (2003) extended the model of Song and Yao (2002) to the multi-product setting. Gallien and Wein (2001) derived an approximate-optimal procurement policy for the same ATO system as in Song and Yao (2002). In Zhao and Simchi-Levi (2006), the authors studied a base-stock ATO system with stochastic, sequential (i.e., no order crossing) lead times and stochastic customer demands, and provided exact analysis for the system properties. Besides, Song et al. (2000) proposed a single-period ATO model to determine the optimal component procurement policy, where stochastic lead times and a single demand with stochastic demand timing and quantity were assumed. Moreover, they presented some structural results regarding the optimal policies. However, none of these papers specifically address the joint optimization of inventory replenishment and component allocation problems in the context of periodic-review ATO systems.

It is worth noting that the optimal policies of the inventory replenishment for the general ATO systems are unknown. In the literature, most of the ATO research focuses on the base-stock policy (i.e., order-up-to level) due to its simplicity and effectiveness (e.g., Akçay and Xu, 2004, Lu and Song, 2005, Zhao and Simchi-Levi, 2006, Huang and de Kok, 2015). That is, there is a target inventory level for each component, once the inventory position of a component drops below the corresponding target level, a replenishment order will be placed to raise the inventory position to the target level again. In addition, we acknowledge that certain component allocation rules are assumed to ensure the analysis in a tractable manner. In particular, the FCFS allocation rule is widely used, which, however, is not in general the optimal allocation rule. Huang (2014) considered two classes of non-FCFS allocation rules, and showed that these rules may significantly outperform the FCFS rule in terms of the fill-rate in a periodic-review ATO system with differentiated demands. In addition, Doğru et al. (2010) proposed a two-stage stochastic program for a continuous-review ATO system with identical lead times. They demonstrated that the superior performance can be achieved in violation of the FCFS rule. Nevertheless, the FCFS rule is still commonly used in practice because of its ease of implementation (e.g., Amazon, see Xu et al., 2009).

In this chapter, we study a periodic-review, multi-component and multi-product ATO system in the presence of stochastic component replenishment lead times and stochastic product demands. We jointly consider the inventory replenishment and component allocation problems in the system, and formulate the problems as two-stage or multi-stage stochastic programs, depending on the assumptions on the lead times. The *sample average approximation* (SAA) algorithms are used to solve the proposed stochastic models. We test the long-term performance of the derived base-stock levels by simulation.

5.2 Model Development

We consider a periodic-review ATO system involving m components and n products. We define $\mathcal{M} = \{1, \cdots, m\}$ and $\mathcal{N} = \{1, \cdots, n\}$ as the sets of components and products, respectively. The replenishment of each component follows an *independent* base-stock policy, where the base-stock level for component i is denoted by S_i for all $i \in \mathcal{M}$. We determine the base-stock levels with a *budget constraint* for the inventory investment among the different components. That is, a budget of B is used to constrain the total investment for the base-stock levels, where the value of the budget can be interpreted as the maximum dollars invested or storage space for the total inventory. We use c_i to denote the unit inventory investment for component i, and it must satisfy $\sum_{i=1}^{m} c_i S_i \leq B$. Let the matrix $(a_{i,j})$ be the Bill of Materials (BOM) where the element $a_{i,j}$ represents the number of component i used in each unit of product j for all $i \in \mathcal{M}$ and $j \in \mathcal{N}$. The customer demand for product j in period t is denoted by $P_{j,t}$, which is assumed to be an integer-valued, stochastically distributed random variable and independent across periods. When demand $P_{j,t}$ arrives, if one or more needed components have insufficient availability, then it will be fully backlogged until the components are replenished. Note that the assembly time is assumed to be negligible. The system enforces the FCFS rule for product demand fulfillment. It is noteworthy that we allow partial fulfillment of demands over subsequent periods, while they must be satisfied on the FCFS basis. The total demand for component i in period t is driven by the product demands in that period. With the BOM, $D_{i,t}$ can be expressed as $D_{i,t} = \sum_{j=1}^{n} a_{i,j} P_{j,t}$. Let w_j denote the response time window of product j, and we assume that a reward $r_{j,k}$ is collected if one unit of product demand $P_{j,t}$ is satisfied in period $t + k$ where $k \leq w_j$.

The same system was studied in Akçay and Xu (2004); Huang and de Kok (2015). In this thesis, we add an important assumption on this system, namely, the replenishment orders of the components have stochastic lead times. Specifically, we assume that lead time l_i^t of the replenishment order for component i placed in period t is a discrete random variable with finite support $\{\underline{L}_i, \cdots, \overline{L}_i\}$ whose probabilities are given by $\{p_{\underline{L}_i}, \cdots, p_{\overline{L}_i}\}$. The notations of the system are further summarized in Table 5.1.

The sequence of events within each time period is described as follows: At the beginning of each period, the present inventory position of each component is reviewed and replenishment orders are placed according to the corresponding base-stock levels. The lead times of the replenishment orders are immediately realized. Then the earlier replenishment orders arrive and the inventory positions are updated. After the product demands arrive, the available components are allocated and assembled into the products

Table 5.1: ATO system notation

t:	index of periods; period t is defined as the interval $[t, t+1)$,
i:	index of components, where $i \in \mathcal{M} = \{1, \cdots, m\}$,
j:	index of products, where $j \in \mathcal{N} = \{1, \cdots, n\}$,
S_i:	base-stock level of component i,
w_j:	response time window of product j,
$r_{j,k}$:	reward of assembling one unit product j in period $t+k$ to satisfy demand in period t
$a_{i,j}$:	the number of component i used in each unit of product j,
c_i:	the investment of each unit of base-stock of component i,
B:	the total budget given for overall base-stock investment; i.e., $\sum_{i=1}^{m} c_i S_i \leq B$,
l_i^t:	lead time of the replenishment order of component i placed in period t,
$P_{j,t}$:	demand for product j in period t,
$D_{i,t}$:	demand for component i in period t; that is, $D_{i,t} = \sum_{j=1}^{n} a_{i,j} P_{j,t}$,
$\underline{L}_i$:	minimum lead time of component i, where $\underline{L}_i \geq 0$,
$\overline{L}_i$:	maximum lead time of component i, where $\overline{L}_i \geq \underline{L}_i$,
L:	maximum lead time among all components; that is, $\max_{i \in \mathcal{M}} \overline{L}_i$; $\mathcal{L} := \{0, \cdots, L\}$.

in response to these demands. The rewards are collected at the end of the period if the product demands are fulfilled within the desired response time windows.

For the purpose of analysis, we assume that there are no order crossings for one type of component; however, the orders of different types of components may cross. Therefore, we assume,

Condition 5.1. *The realized lead times for each component must satisfy*

$$t + l_i^t \leq t + 1 + l_i^{t+1}, \quad \forall i \in \mathcal{M}, \ \forall t$$

The above condition implies that the lead times of a component are *correlated* across different periods. That is, the possible lead time values of each component in the present period is conditioned on the realized lead times of that component in the earlier periods. For example, suppose that component i initially has a lead time support $\{0, 1, 2\}$ with probabilities $\{0.2, 0.3, 0.5\}$. Let t be the present period, then the lead time support of l_i^t becomes $\{1, 2\}$ if $l_i^{t-1} = 2$, and the corresponding probabilities are updated as $\{\frac{0.3}{0.3+0.5}, \frac{0.5}{0.3+0.5}\} = \{0.375, 0.625\}$.

Consider product demand $P_{j,t}$ for product j arriving in period t, it is important to specify the maximum waiting time until this demand can be fully satisfied. Since each product might be assembled from multiple components, we need to specify when the associated component demands can be satisfied. For our ATO system, we have the following theorem regarding the maximum waiting time until the product demand $P_{j,t}$ has been fully satisfied.

Theorem 5.1. *Under Condition 5.1, the component demand $D_{i,t}$ can be fully satisfied before or in period $t + \overline{L}_i + 1$ for all $i \in \mathcal{M}$ and t. The product demand $P_{j,t}$ can be fully satisfied before or in period $t + L + 1$ for all $j \in \mathcal{N}$ and t, where $L = \max_{i \in \mathcal{M}} \overline{L}_i$.*

Proof. Under the base-stock policy, the replenishment order for component i placed in period $t + 1$ is always triggered by the component demand in period t (i.e., $D_{i,t}$) and exactly equals it. In other words, the size of the replenishment order for component i that will be placed in period $t + 1$ is known at the end of period t. When the FCFS rule is applied, the worst-case scenario is that the component demand $D_{i,t}$ will be satisfied by the order triggered by itself and this order will arrive by period $t+\overline{L}_i+1$ since $l_i^{t+1} \leq \overline{L}_i$. On the other hand, it is still possible that the product demand $P_{j,t}$ won't be satisfied in period $t + \overline{L}_i + 1$ due to lack of other components. In the worst-case scenario, we know that all the component demands associated with the product demand $P_{j,t}$ will be fully satisfied by period $t + L + 1$. Therefore, product demand $P_{j,t}$ will be fully satisfied by period $t + L + 1$. □

A similar result can be found in Huang and de Kok (2015); however, the authors assumed deterministic lead times for all replenishment orders (i.e., $l_i^t \equiv L_i$ for all $i \in \mathcal{M}$

and t, where L_i corresponds to the expected lead time of component i and $\underline{L}_i \leq L_i \leq \overline{L}_i$). The difference between $t + \overline{L}_i + 1$ and $t + L + 1$ in the theorem is caused by multiple components contained in one product, a phenomenon called *"multi-matching"* also illustrated in Huang and de Kok (2015).

With Theorem 5.1, it allows us to formulate the ATO system as a two-stage or multi-stage stochastic integer program with recourse, depending on the knowledge of the random lead times.

5.2.1 Knowledge of the Realized Lead Times

In this section, we assume that in period t, all the lead times of replenishment orders corresponding to demands prior to period $t+2$ are known. In other words, these random lead times are realized. As noted above, the demand $D_{i,t}$ arriving in period t will trigger a replenishment order of the same size in period $t + 1$. So the assumption requires that l_i^{t+1} is *known* when the component allocation decisions for $D_{i,t}$ are made in period t. This assumption is reasonable if an ATO manufacturer engages in close relationships with its suppliers, and therefore it has complete information of the suppliers, including the random lead times that will be realized in the next period. In such a circumstance, although the lead times are stochastic, the ATO manufacturer can make component allocation decisions with the realized lead times.

Consider component i. In period t, note that the earliest order that has not arrived in period t yet could only be triggered by the demand in period $t - \overline{L}_i$. Define $\mathcal{A}_i^0$ as the set of all orders triggered by demands within $[t - \overline{L}_i, t - 1]$ that arrive before or in period t. That is,

$$\mathcal{A}_i^0 = \{t' \in [t - \overline{L}_i, t - 1] : t' + 1 + l_i^{t'+1} \leq t\}. \tag{5.1}$$

Similarly, in period $t+u$, define $\mathcal{A}_i^u$ as the set of all orders triggered by demands within $[t-\overline{L}_i, t]$ that arrive *exactly* in period $t+u$. That is,

$$\mathcal{A}_i^u = \{t' \in [t-\overline{L}_i, t] : t'+1+l_i^{t'+1} = t+u\}, \tag{5.2}$$

where $1 \le u \le \overline{L}_i + 1$. Clearly, if $0 \le k, u \le \overline{L}_i + 1$ and $k \ne u$, then $\mathcal{A}_i^k \cap \mathcal{A}_i^u = \emptyset$. On the other hand, we have $\cup_{k=0}^{\overline{L}_i+1} \mathcal{A}_i^k = \{t-\overline{L}_i, \cdots, t\}$.

With the $\mathcal{A}_i^u$ notation, we can define, by period $t+k$, the total available on-hand inventory that will be used to satisfy the demand $D_{i,t}$ for component i in period t as follows:

$$\widetilde{O}_i^k = \min\{(S_i - D_i^{\overline{L}_i} + \sum_{l \in \cup_{u=0}^k \mathcal{A}_i^u} D_{i,l})^+, D_{i,t}\}, \tag{5.3}$$

where $0 \le k \le \overline{L}_i + 1$, and $D_i^{\overline{L}_i} = \sum_{l=t-\overline{L}_i}^{t-1} D_{i,l}$. Note that $\widetilde{O}_i^{\overline{L}_i+1} \equiv D_{i,t}$. Clearly, $\widetilde{O}_i^k$ is a piece-wise linear, non-convex function of S_i.

Component demand $D_{i,t}$ can be satisfied in multiple periods, namely $t+0, t+1, \ldots, t+\overline{L}_i+1$, while product demand $P_{j,t}$ can be satisfied in $t+0, t+1, \ldots, t+L+1$ (see Theorem 5.1). We denote the decision variable $x_{j,k}$ as the amount of product j assembled in period $t+k$ for all $j \in \mathcal{N}$, $0 \le k \le L+1$ and t. Thus, by period $t+k$, the total amount $\sum_{u=0}^{k} \sum_{j=1}^{n} a_{i,j} x_{j,u}$ of component i is obtained from the on-hand inventories and assembled (with other components) into different products.

According to our assumption on the lead times, when the base-stock levels are given, the available on-hand inventories $\widetilde{O}_i^k$ are deterministic, so we have the following component allocation problem,

$$Q(\mathbf{S}, \xi(\omega)) = \max \quad \sum_{j=1}^{n} \sum_{k=0}^{w_j} r_{j,k} x_{j,k} \tag{5.4}$$

$$\text{s.t.} \quad \sum_{k=0}^{L+1} x_{j,k} = P_{j,t}, \qquad \forall j \in \mathcal{N}, \tag{5.5}$$

$$\sum_{u=0}^{k}\sum_{j=1}^{n} a_{i,j}x_{j,u} \leq \widetilde{O}_i^k, \qquad \forall i \in \mathcal{M}, \forall k \in \mathcal{L}, \tag{5.6}$$

$$x_{j,k} \in \mathbb{Z}_+, \qquad \forall j \in \mathcal{N}, \forall k \in \mathcal{L}, \tag{5.7}$$

where the vector $\mathbf{S} = (S_i)_{i\in\mathcal{M}}$ denotes the base-stock levels, and the vector $\xi(\omega)$ denotes the (realized) random demands $(P_{j,k})_{j\in\mathcal{N},k\in[t-\overline{L}_i,t]}$ and (realized) random lead times $(l_i^k)_{i\in\mathcal{M},k\in[t-\overline{L}_i+1,t+1]}$. The objective (5.4) aims to maximize the total reward from satisfying the product demands $(P_{j,t})_{j\in\mathcal{N}}$ within the given response time windows. Constraint (5.5) guarantees that the demands $(P_{j,t})_{j\in\mathcal{N}}$ will be satisfied no later than period $t+L+1$. Constraint (5.6) is the component availability constraint, which requires that the allocation of component i in each period could only take place when there are enough inventories of component i for all $i \in \mathcal{M}$ and $0 \leq k \leq L+1$. Constraint (5.7) requires that the allocation decisions can only take nonnegative integer values.

With a given base-stock vector $\mathbf{S}$ and realized random vector $\xi(\omega)$, we can optimally solve the formulation (5.4)-(5.7). To determine the optimal base-stock levels with a budget of B allocated to the total inventory investment, we need to solve the following stochastic program,

$$\max \quad \mathbb{E}_\xi[Q(\mathbf{S},\xi)] \tag{5.8}$$

$$\text{s.t.} \quad \sum_{i=1}^{m} c_i S_i \leq B, \tag{5.9}$$

$$S_i \in \mathbb{Z}_+, \qquad \forall i \in \mathcal{M}. \tag{5.10}$$

The objective function (5.8) is to maximize the expected value of $Q(\mathbf{S},\xi)$, where $Q(\mathbf{S},\xi)$ is the recourse function and it equals the optimal objective value of formulation (5.4)-(5.7). Constraint (5.9) ensures that the total investment for the base-stock levels cannot exceed the budget B. Constraint (5.10) requires that the base-stock levels can only take nonnegative integer values. Therefore, formulation (5.8)-(5.10) is a two-stage stochastic integer program. The decisions in the first-stage are the optimal base-stock levels, namely $\mathbf{S} = (S_i)_{i\in\mathcal{M}}$. In the second-stage, the component allocation decisions are made repeatedly over time under the optimal base-stock levels.

As mentioned earlier, the same joint optimization of base-stock levels and component allocation, however, with the expected (deterministic) lead time for each component (i.e., $l_i^t \equiv L_i$ for all $i \in \mathcal{M}$) has been addressed in Akçay and Xu (2004). In fact, it is readily to show that the formulation (5.8)-(5.10) is equivalent to the two-stage stochastic program proposed in Akçay and Xu (2004) when the lower and upper bounds of the lead time for each component are set to the corresponding expected lead time (i.e., $\underline{L}_i = \overline{L}_i = L_i$ for all $i \in \mathcal{M}$).

5.2.2 Knowledge of the Random Lead Time Distributions

In this section, we assume that when facing the component allocation decisions, the decision maker of the ATO system does not have the full knowledge of the realized lead times. Instead, the decision maker only has the knowledge of random lead time distributions in time period $t + 1$ when he makes component allocation decisions in period t.

In this circumstance, the decision maker of the ATO system can only construct a scenario tree from all previous samples and use this scenario tree for component allocation optimization. Let $\mathcal{T}$ be the scenario tree. For each node g of the scenario tree, let $\mathcal{P}(g)$ be the path from node g to the root node; p_g be the probability of node g; $t(g)$ be the period of node g. Note that $t(g) = t$ if g is the root node. The set $\mathcal{L}_f$ denotes the leaf nodes. If g is a leaf node (i.e., $g \in \mathcal{L}_f$), then $\mathcal{P}(g)$ corresponds to a *scenario*. There are $L+2$ layers in the scenario tree, corresponding to periods $t, \cdots, t+L+1$. An illustration of the scenario tree is provided in Figure 5.2.

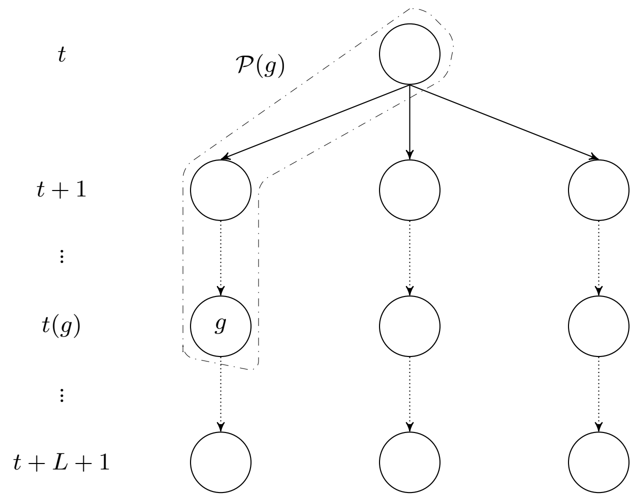

Figure 5.2: The scenario tree $\mathcal{T}$

In order to make an optimal component allocation decision for product demands arriving in period t, we need to know the realized $P_{j,t}$ for all $j \in \mathcal{N}$ and l_i^u for all $i \in \mathcal{M}$ and $u \leq t+1$. The root node in layer t of the tree corresponds to the realized demands $(P_{j,t})_{j\in\mathcal{N}}$ and realized lead times $(l_i^k)_{i\in\mathcal{M},k\in[t-\overline{L}_i,t]}$. According to the sequence of events, this information is known when we solve the component allocation problem in period t, so the probability of the root node is equal to 1. In layer $t+1$, there is a finite number of nodes since the lead time l_i^{t+1} for all $i \in \mathcal{M}$ has a finite discrete support. Note that the nodes in this layer are distinguishable since each node corresponds to a specific realization of the random lead times for all the components in period $t+1$. Hence, the probability of each node is the joint probability associated with the corresponding lead time realization at that node. For instance, if the lead time for component i, $\forall i \in \mathcal{M}$, has a realization $\widetilde{L}_i$ in period $t+1$ with the probability $p_{\widetilde{L}_i}$, then a node g in layer $t+1$ that corresponds to the realization $(\widetilde{L}_i)_{i\in\mathcal{M}}$ has the probability $p_g = \Pi_{i\in\mathcal{M}}\, p_{\widetilde{L}_i}$.

It is important to note that the number of nodes in layer $t+1$ is determined by the realized lead times in earlier periods $t-\overline{L}_i, \cdots, t$ for all $i \in \mathcal{M}$ due to the correlated lead times under Condition 5.1. Although there are $L+2$ layers in the scenario tree, it essentially consists of two stages since the nodes in the layers after $t+1$ reveals no new information in terms of the realization of stochastic variables. This explains why the scenarios no longer branch after layer $t+1$. Hence, the probability of a node after layer $t+1$ in a scenario is equal to the probability of its ancestor node in layer $t+1$ of that scenario. For more details of scenario tree notations, refer to Huang and Ahmed (2009).

With the scenario tree notations, the component allocation problem can be modeled as a two-stage stochastic integer program, where in the first-stage the component allocation decisions corresponding to the root node in layer t of the scenario tree are made; then based on the realization of the lead times in period $t+1$, the component allocation decisions corresponding to the nodes from layers $t+1$ to $t+L+1$ are made in the second-stage problem. The component allocation problem is formulated as follows,

$$Q(\mathbf{S},\xi(\omega)) = \max \sum_{j\in\mathcal{N}} \sum_{g\in\mathcal{T},t(g)\leq w_j} p_g r_{j,t(g)} x_{j,g} \tag{5.11}$$

$$\text{s.t.} \sum_{g\in\mathcal{P}(q)} x_{j,g} = P_{j,t}, \qquad \forall j\in\mathcal{N}, \forall q\in\mathcal{L}_f, \tag{5.12}$$

$$\sum_{g\in\mathcal{P}(q)} \sum_{j\in\mathcal{N}} a_{i,j} x_{j,g} \leq \widetilde{O}_i^q, \qquad \forall i\in\mathcal{M}, \forall q\in\mathcal{T}, \tag{5.13}$$

$$x_{j,g}\in\mathbb{Z}_+, \qquad \forall j\in\mathcal{N}, \forall g\in\mathcal{T}, \tag{5.14}$$

where $\widetilde{O}_i^q$ is the available on-hand inventory at node q of the scenario tree. Note that the nonanticipativity constraints are satisfied in the scenario tree notations (Birge and Louveaux, 2011). The objective (5.11) is to maximize the expected total reward over the entire scenario tree. Constraint (5.12) guarantees that the demands $(P_{j,t})_{j\in\mathcal{N}}$ will be satisfied before or in period $t+L+1$ for each scenario. Constraint (5.13) ensures that for each node in the scenario tree, the component allocation will only happen when sufficient number of component inventories are available. Constraint (5.14) requires that the allocation decisions for each node of the scenario tree be nonnegative integer values. The above formulation is a two-stage stochastic integer program with a finite number of different scenarios, and the number of the scenarios equals the number of nodes in layer

$t + 1$ of the corresponding scenario tree.

To determine the optimal base-stock levels, we need to solve the formulation (5.8)-(5.10) where $Q(\mathbf{S}, \xi)$ is defined as the optimal objective value of formulation (5.11)-(5.14). Importantly, the formulation (5.8)-(5.10) now has become a multi-stage stochastic integer program (the number of stages is three).

5.3 Sample Average Approximation Algorithms

Solving the proposed stochastic programs exactly is computationally prohibitive due to the uncountable number of possible scenarios for the random vector ξ. In this section, we use a Monte Carlo simulation-based approach, the SAA algorithm, to obtain the near-optimal base-stock levels for the proposed stochastic programs.

The main idea of the SAA algorithm is to estimate the objective function $Q(\mathbf{S}, \xi)$ by sampling a set of scenarios of the random vector ξ. Specifically, M samples are generated independently, with N realizations of the random vector ξ in each sample. The different optimal second-stage decisions $x_{j,k}$ for all $j \in \mathcal{N}$ and $0 \leq k \leq L + 1$ are made for each realization, while the optimal first-stage decisions S_i for all $i \in \mathcal{M}$ are made over all the realizations in a sample. The SAA algorithm yields M sets of base-stock levels. Next, the performance of the M candidate solutions is tested with a new sample of N' realizations of the random vector ξ, where $N' \gg N$. At last, the one out of the M candidate solutions that yields the largest expected reward is considered as the optimal base-stock levels for our proposed stochastic programs. It has been shown that the probability of the SAA algorithm producing an exact optimal solution for a two-stage stochastic program with integer recourse approaches one exponentially fast in the sample size M (Ahmed et al., 2002, Kleywegt et al., 2002).

The SAA algorithm has been studied in the literature to find the near-optimal solutions (e.g., Verweij et al., 2003, Schütz et al., 2009). In particular, it has been successfully used to solve the joint optimization problem in the context of ATO systems (see e.g., Akçay and Xu, 2004, Huang and de Kok, 2015). In the following, we describe the SAA algorithms for the proposed two-stage and multi-stage models, respectively.

5.3.1 The Two-Stage Model

Let $\xi(\omega_l^1), \cdots, \xi(\omega_l^N)$ be the l-th sample, $l = 1, \cdots, M$, where $\xi(\omega_l^h)$ denotes the realized lead times $(l_i^k(\omega_l^h))_{i\in\mathcal{M},k\in[t-\overline{L}_i+1,t+1]}$ and realized demands $(P_{j,k}(\omega_l^h))_{j\in\mathcal{N},k\in[t-\overline{L}_i,t]}$ of the h-th realization, $h = 1, \cdots, N$, for the sample. For each sample, we solve the following approximation of the proposed two-stage model, which is referred to as the SAA problem.

$$
\begin{aligned}
\widehat{Q}_N(\widehat{\mathbf{S}}_l) = \max \quad & \frac{1}{N}\sum_{h=1}^{N}\sum_{j=1}^{n}\sum_{k=0}^{w_j} r_{j,k} x_{j,k}^{\omega_l^h} \\
\text{s.t.} \quad & \sum_{k=0}^{L+1} x_{j,k}^{\omega_l^h} = P_{j,t}(\omega_l^h), && \forall j \in \mathcal{N}, h = 1, \dots, N, \\
& \sum_{u=0}^{k}\sum_{j=1}^{n} a_{i,j} x_{j,u}^{\omega_l^h} \le \widetilde{O}_i^k(\omega_l^h), && \forall i \in \mathcal{M}, \forall k \in \mathcal{L}, h = 1, \dots, N, \\
& \sum_{i=1}^{m} c_i S_i^l \le B, \\
& S_i^l \in \mathbb{Z}_+, && \forall i \in \mathcal{M}, \\
& x_{j,k}^{\omega_l^h} \in \mathbb{Z}_+, && \forall j \in \mathcal{N}, \forall k \in \mathcal{L}, h = 1, \dots, N,
\end{aligned}
\tag{5.15}
$$

where $\widehat{Q}_N(\widehat{\mathbf{S}}_l)$ denotes the optimal objective value of the formulation (5.15) and $\widehat{\mathbf{S}}_l = (S_i^l)_{i\in\mathcal{M}}$ corresponds to the vector of optimal base-stock levels for the l-th sample. Note that we add the superscript ω_l^h for the allocation decisions $x_{j,k}$ since they depend on $\xi(\omega_l^h)$. The right-hand side $\widetilde{O}_i^k(\omega_l^h) = \min\{(S_i - D_i^{\overline{L}_i}(\omega_l^h) + \sum_{l\in\cup_{u=0}^{k}\mathcal{A}_i^u} D_{i,l}(\omega_l^h))^+, D_{i,t}(\omega_l^h)\}$ is a piece-wise, non-convex function of S_i, so we rewrite the second constraint of the formulation (5.15) as

$$
\sum_{u=0}^{k}\sum_{j=1}^{n} a_{ij} x_{j,u}^{\omega_l^h} \le D_{i,t}(\omega_l^h), \tag{5.16}
$$

$$\sum_{u=0}^{k}\sum_{j=1}^{n} a_{ij} x_{j,u}^{\omega_l^h} \leq (S_i - D_i^{\overline{L}_i}(\omega_l^h) + \sum_{l \in \cup_{u=0}^{k} \mathcal{A}_i^u} D_{i,l}(\omega_l^h))^+. \tag{5.17}$$

Then, we use the "Big-M" method to linearize Constraint (5.17) by

$$\sum_{u=0}^{k}\sum_{j=1}^{n} a_{ij} x_{j,u}^{\omega_l^h} \leq M z_{i,k}^{\omega_l^h}, \tag{5.18}$$

$$\sum_{u=0}^{k}\sum_{j=1}^{n} a_{ij} x_{j,u}^{\omega_l^h} \leq (S_i - D_i^{\overline{L}_i}(\omega_l^h) + \sum_{l \in \cup_{u=0}^{k} \mathcal{A}_i^u} D_{i,l}(\omega_l^h)) + M(1 - z_{i,k}^{\omega_l^h}), \tag{5.19}$$

$$S_i - D_i^{\overline{L}_i}(\omega_l^h) + \sum_{l \in \cup_{u=0}^{k} \mathcal{A}_i^u} D_{i,l}(\omega_l^h) \leq M z_{i,k}^{\omega_l^h}, \tag{5.20}$$

$$z_{i,k}^{\omega_l^h} \in \{0, 1\}, \tag{5.21}$$

where M is a large constant.

In Akçay and Xu (2004), the plus sign "+" in the right-hand side of the component availability constraint is dropped to facilitate the computation of their SAA algorithm. However, Deza et al. (2018) analyzed the impact of such the relaxation and pointed out that it may cause the infeasibility issue at the low budget levels.

Although the SAA problem (5.15) above is still not easy to solve, the number of scenarios involved is considerably decreased than the original two-stage model. In addition, it can be used to derive the bounds on the optimal objective value of the original model. Let $Q^*(\mathbf{S}^*)$ be the optimal objective value of the original two-stage model, and $\mathbf{S}^*$ be the corresponding optimal base-stock level vector. Moreover, we denote the average of these M optimal objective values of the SAA problem by $\overline{Q}_N^M = \frac{1}{M}\sum_{l=1}^{M} \widehat{Q}_N(\widehat{\mathbf{S}}_l)$. It is well-known that

$$Q^*(\mathbf{S}^*) \leq \mathbb{E}(\overline{Q}_N^M),$$

hence, $\overline{Q}_N^M$ provides the estimated upper bound of $Q^*(\mathbf{S}^*)$.

Clearly, $\mathbb{E}_\xi[Q(\widehat{\mathbf{S}}_l, \xi)]$ is a lower bound of $Q^*(\mathbf{S}^*)$ because $\widehat{\mathbf{S}}_l$ is a feasible solution of the original two-stage model. To obtain an unbiased estimator of $\mathbb{E}_\xi[Q(\widehat{\mathbf{S}}_l,\xi)]$, an extra sample of N' realizations is independently generated. Let $\xi(\omega_0^1),\ldots,\xi(\omega_0^{N'})$ be the sample of N' realizations, where $\xi(\omega_0^h)$ denotes the realized lead times $(l_i^k(\omega_0^h))_{i\in\mathcal{M},k\in[t-\overline{L}_i+1,t+1]}$ and realized demands $(P_{j,k}(\omega_0^h))_{j\in\mathcal{N},k\in[t-\overline{L}_i,t]}$ of the h-th realization, $h=1,\ldots,N'$. Therefore, the lower bound $\mathbb{E}_\xi[Q(\widehat{\mathbf{S}}_l,\xi)]$ can be estimated by the following for $l=1,\ldots,M$,

$$\begin{aligned} Q_{N'}(\widehat{\mathbf{S}}_l) = \max\ & \frac{1}{N'}\sum_{h=1}^{N'}\sum_{j=1}^{n}\sum_{k=0}^{w_j} r_{j,k}x_{j,k}^{\omega_0^h} \\ \text{s.t.}\ & \sum_{k=0}^{L+1} x_{j,k}^{\omega_0^h} = P_{j,t}(\omega_0^h), && \forall j\in\mathcal{N}, h=1,\ldots,N', \\ & \sum_{u=0}^{k}\sum_{j=1}^{n} a_{i,j}x_{j,u}^{\omega_0^h} \le \widetilde{O}_i^k(\omega_0^h), && \forall i\in\mathcal{M},\forall k\in\mathcal{L}, h=1,\ldots,N', \\ & x_{j,k}^{\omega_0^h}\in\mathbb{Z}_+, && \forall j\in\mathcal{N},\forall k\in\mathcal{L}, h=1,\ldots,N'. \end{aligned} \tag{5.22}$$

Note that given $\widehat{\mathbf{S}}_l$ and $\xi(\omega_0^1),\ldots,\xi(\omega_0^{N'})$ for the formulation (5.22), the right-hand side of the component availability constraint (i.e., $\widetilde{O}_i^k(\omega_0^h)$) can be computed in advance and thus is completely deterministic. Besides, the formulation is decomposable with respect to the N' realizations. Therefore, we can solve the component allocation problem for each $\xi(\omega_0^h)$, $h=1,\ldots,N'$, and compute the optimal objective value $Q_{N'}(\widehat{\mathbf{S}}_l)$ by taking the average over the N' realizations.

After solving formulation (5.22) repeatedly with M candidate solutions, it is natural to take the base-stock levels that yield the largest $Q_{N'}(\cdot)$, i.e., $\widehat{\mathbf{S}}^*\in\operatorname{argmax}\{Q_{N'}(\widehat{\mathbf{S}}_l): l=1,\ldots,M\}$, as the optimal base-stock levels. Clearly, $Q_{N'}(\widehat{\mathbf{S}}^*)$ serves as a lower bound estimation of $Q^*(\mathbf{S}^*)$. That is,

$$Q_{N'}(\widehat{\mathbf{S}}^*) \le Q^*(\mathbf{S}^*).$$

The SAA algorithm for the two-stage model is further summarized in Algorithm 1.

Algorithm 1 The SAA algorithm for the two-stage model

```
Initialize: Select values for M, N, and N';
for l = 1, ..., M do
    Generate an independent sample ξ(ω_l^1), ..., ξ(ω_l^N);
    Solve the SAA problem (5.15), and record Ŝ_l and Q̂_N(Ŝ_l);
end for
Calculate the estimated upper bound of Q*(S*) using Q̄_N^M = (1/M) Σ_{l=1}^M Q̂_N(Ŝ_l);
Generate an independent sample ξ(ω_0^1), ..., ξ(ω_0^{N'});
for l = 1, ..., M do
    Solve the optimization problem (5.22) using Ŝ_l, and record Q_{N'}(Ŝ_l);
end for
Select Ŝ* ∈ argmax{Q_{N'}(Ŝ_l) : l = 1, ..., M}, and a lower bound of Q*(S*) is given
    by Q_{N'}(Ŝ*).
```

5.3.2 The Multi-Stage Model

In this section, we modify the SAA algorithm described in Section 5.3.1 to solve the proposed multi-stage model. In the same manner, we generate M independent samples, with N realizations of random vector in each sample. However, it is important to note that in the multi-stage setting, when we make component allocation decisions for demands arriving in period t, the random lead times $(l_i^{t+1})_{i\in\mathcal{M}}$ are not yet known and thus must be excluded from the realizations of the SAA algorithm. Instead, we must take into account all possible realizations of the lead times $(l_i^{t+1})_{i\in\mathcal{M}}$ for component allocation optimization.

To distinguish from the SAA algorithm for the two-stage model, we use $\xi'(\omega_l^1),\ldots,\xi'(\omega_l^N)$ to denote the l-th sample, $l=1,\ldots,M$, where $\xi'(\omega_l^h)$ denotes the realized lead times $(l_i^k(\omega_l^h))_{i\in\mathcal{M},k\in[t-\overline{L}_i+1,t]}$ and realized demands $(P_{j,k}(\omega_l^h))_{j\in\mathcal{N},k\in[t-\overline{L}_i,t]}$ of the h-th realization, $h=1,\ldots,N$, for the sample. For each $\xi'(\omega_l^h)$, we specify the possible realizations of the lead times $(l_i^{t+1})_{i\in\mathcal{M}}$ under Condition 5.1 and construct the corresponding scenario tree, denoted by $\mathcal{T}_{\xi'(\omega_l^h)}$. For each sample, the proposed multi-stage model is approximated by the following SAA problem.

$$\begin{aligned}
\widehat{Q}_N(\widehat{\mathbf{S}}_l) = \max\ & \frac{1}{N}\sum_{h=1}^{N}\Big(\sum_{j\in\mathcal{N}}\sum_{g\in\mathcal{T}_{\xi'(\omega_l^h)},t(g)\le w_j} p_g r_{j,t(g)} x_{j,g}^{\omega_l^h}\Big) \\
\text{s.t.}\ & \sum_{g\in\mathcal{P}(q)} x_{j,g}^{\omega_l^h} = P_{j,t}(\omega_l^h), && \forall j\in\mathcal{N},\forall q\in\mathcal{L}_f\subset\mathcal{T}_{\xi'(\omega_l^h)}, h=1,\dots,N, \\
& \sum_{g\in\mathcal{P}(q)}\sum_{j\in\mathcal{N}} a_{i,j} x_{j,g}^{\omega_l^h} \le \widetilde{O}_i^q(\omega_l^h), && \forall i\in\mathcal{M},\forall q\in\mathcal{T}_{\xi'(\omega_l^h)}, h=1,\dots,N, \\
& \sum_{i\in\mathcal{M}} c_i S_i^l \le B, && \qquad (5.23) \\
& S_i^l \in \mathbb{Z}_+, && \forall i\in\mathcal{M}, \\
& x_{j,g}^{\omega_l^h} \in \mathbb{Z}_+, && \forall j\in\mathcal{N},\forall g\in\mathcal{T}_{\xi'(\omega_l^h)}, h=1,\dots,N.
\end{aligned}$$

It is worthing noting that the SAA problem (5.23) involves solving multiple number of formulation (5.15) simultaneously, which may require extensive computational effort depending on the size of the corresponding scenario tree $\mathcal{T}_{\xi'(\omega_l^h)}$.

Once the SAA problem has been solved repeatedly with different samples, we obtain M candidate solutions of the base-stock levels. The estimated upper bound of $Q^*(\mathbf{S}^*)$ is again computed by $\overline{Q}_N^M = \frac{1}{M}\sum_{l=1}^{M}\widehat{Q}_N(\widehat{\mathbf{S}}_l)$, where $Q^*(\mathbf{S}^*)$ corresponds to the optimal objective value of the proposed multi-stage model. Then, we generate an independent sample of N' realizations $\xi'(\omega_0^1),\dots,\xi'(\omega_0^{N'})$ and solve the following deterministic problem with the base-stock vector $\widehat{\mathbf{S}}_l$ for $l = 1,\dots,M$,

$$\begin{aligned}
Q_{N'}(\widehat{\mathbf{S}}_l) = \max\ & \frac{1}{N'}\sum_{h=1}^{N'}\Big(\sum_{j\in\mathcal{N}}\sum_{g\in\mathcal{T}_{\xi'(\omega_l^h)},t(g)\le w_j} p_g r_{j,t(g)} x_{j,g}^{\omega_l^h}\Big) \\
\text{s.t.}\ & \sum_{g\in\mathcal{P}(q)} x_{j,g}^{\omega_l^h} = P_{j,t}(\omega_l^h), && \forall j\in\mathcal{N},\forall q\in\mathcal{L}_f\subset\mathcal{T}_{\xi'(\omega_l^h)}, h=1,\dots,N', \\
& \sum_{g\in\mathcal{P}(q)}\sum_{j\in\mathcal{N}} a_{i,j} x_{j,g}^{\omega_l^h} \le \widetilde{O}_i^q(\omega_l^h), && \forall i\in\mathcal{M},\forall q\in\mathcal{T}_{\xi'(\omega_l^h)}, h=1,\dots,N', \quad (5.24) \\
& x_{j,g}^{\omega_l^h} \in \mathbb{Z}_+, && \forall j\in\mathcal{N},\forall g\in\mathcal{T}_{\xi'(\omega_l^h)}, h=1,\dots,N'.
\end{aligned}$$

Finally, the base-stock level vector $\widehat{\mathbf{S}}^*$ that yields the largest $Q_{N'}(\cdot)$ is selected as the optimal base-stock levels for the multi-stage model and the estimated lower bound of $Q^*(\mathbf{S}^*)$ is provided by $Q_{N'}(\widehat{\mathbf{S}}^*)$.

We summarize the SAA algorithm for the multi-stage model in Algorithm 2.

Algorithm 2 The SAA algorithm for the multi-stage model

Initialize: Select values for M, N, and N';
for $l = 1, \ldots, M$ **do**
  Generate an independent sample $\xi'(\omega_l^1), \ldots, \xi'(\omega_l^N)$;
  Construct the scenario three $\mathcal{T}_{\xi'(\omega_l^1)}, \ldots, \mathcal{T}_{\xi'(\omega_l^N)}$;
  Solve the SAA problem (5.23), and record $\widehat{\mathbf{S}}_l$ and $\widehat{Q}_N(\widehat{\mathbf{S}}_l)$;
end for
Calculate the estimated upper bound of $Q^*(\mathbf{S}^*)$ using $\overline{Q}_N^M = \frac{1}{M}\sum_{l=1}^M \widehat{Q}_N(\widehat{\mathbf{S}}_l)$;
Generate an independent sample $\xi'(\omega_0^1), \ldots, \xi'(\omega_0^{N'})$;
Construct the scenario three $\mathcal{T}_{\xi'(\omega_0^1)}, \ldots, \mathcal{T}_{\xi'(\omega_0^{N'})}$;
for $l = 1, \ldots, M$ **do**
  Solve the optimization problem (5.24) using $\widehat{\mathbf{S}}_l$, and record $Q_{N'}(\widehat{\mathbf{S}}_l)$;
end for
Select $\widehat{\mathbf{S}}^* \in \operatorname{argmax}\{Q_{N'}(\widehat{\mathbf{S}}_l) : l = 1, \ldots, M\}$, and a lower bound of $Q^*(\mathbf{S}^*)$ is given by $Q_{N'}(\widehat{\mathbf{S}}^*)$.

5.3.3 SAA Performance

The *optimality gap* (i.e., $\overline{Q}_N^M - Q_{N'}(\widehat{\mathbf{S}}^*)$) is primarily used to evaluate the quality of the SAA solutions. Kleywegt et al. (2002) showed that the tighter optimality gap tends to be achieved by using larger values of N and N'. On the other hand, the computational complexity for solving the optimization problems in the SAA algorithms increases significantly. In our case, we select a relatively smaller N while selecting a larger N' since formulations (5.22) and (5.24) are decomposable by realizations.

The variance of the optimality gap is also commonly used to evaluate the solution quality. However, we only report the optimality gap in the following numerical experiments. For completeness, we provide the way based on Ahmed et al. (2002) to compute the estimated variance of the optimality gap yielded by Algorithms 1 and 2.

The variances of $\overline{Q}_N^M$ and $Q_{N'}(\widehat{\mathbf{S}}^*)$ can be estimated by

$$\sigma_{\overline{Q}_N^M} = \frac{1}{M(M-1)} \sum_{l=1}^{M} \left[\widehat{Q}_N(\widehat{\mathbf{S}}_l) - \overline{Q}_N^M \right]^2,$$

and

$$\sigma_{Q_{N'}(\widehat{\mathbf{S}}^*)} = \frac{1}{N'(N'-1)} \sum_{h=1}^{N'} \left[Q(\widehat{\mathbf{S}}^*, \xi(\omega_0^h)) - Q_{N'}(\widehat{\mathbf{S}}^*) \right]^2,$$

respectively. The variance of the optimality gap is estimated by

$$\sigma^2_{\overline{Q}_N^M - Q_{N'}(\widehat{\mathbf{S}}^*)} = \sigma^2_{\overline{Q}_N^M} + \sigma^2_{Q_{N'}(\widehat{\mathbf{S}}^*)}. \tag{5.25}$$

In the case that the values of the optimality gap and (or) the variance are too large, one might consider to increase the values of N and N', and re-perform the SAA algorithms. The detailed SAA statistics have been discussed in Ahmed et al. (2002); Kleywegt et al. (2002).

5.4 Numerical Studies

In this section, we report numerical experiment results for two data sets from the ATO literature (Zhang, 1997, Agrawal and Cohen, 2001). Our purposes of this experiment are twofold. Firstly, we test the effectiveness of the SAA algorithms. Secondly, we carry out simulations to evaluate the long-term performance of the derived base-stock levels with respect to randomly generated lead times and demands. We focus on the average performance as well as the performance stability.

For each data set, we first determine the base-stock levels for the proposed stochastic programs using their corresponding SAA algorithms under different budget levels. We denote by SAA-TS and SAA-MS the SAA results of the proposed two-stage and multi-stage models yielded by Algorithm 1 and Algorithm 2, respectively. For the purpose of comparison, we additionally consider the SAA results of Akçay and Xu (2004), which is denoted by SAA-AX. The SAA results reported include the optimal base-stock levels, the lower bound (LB), the upper bound (UB), and the optimality gap (Gap = UB − LB). Since rewards are identical across all products in the two data sets under consideration, we present the results of LB, UB, and Gap in the form of the aggregated *type-II* service level (in percent), i.e., $100\% \times \sum_{j=1}^{n} \sum_{k=0}^{w_j} x_{j,k} / \sum_{j=1}^{n} \mathbb{E}[P_{j,t}]$, where $\mathbb{E}[P_{j,t}]$ represents the mean of random demand $P_{j,t}$.

In the simulation, given the base-stock levels, we repeatedly solve the component allocation problem for each period using the *optimal component allocation* (OA) policy under the FCFS rule. Then, the long-term performance is estimated with respect to 1000 periods of realized lead time and demand. The simulated results are also represented by the type-II service level (in percent).

5.4.1 Agrawal and Cohen System

The ATO system of Agrawal and Cohen (2001) involves four products ($j = 1, \ldots, 4$) and two components ($i = 1, 2$), as described in Table 5.2. Specifically, it is assumed that in each period the demand for product j is normally distributed with the mean μ_j and standard deviation σ_j, while the lead time of component i in each period follows a discrete uniform distribution between $\underline{L}_i$ and $\overline{L}_i$ denoted by $Uniform[\underline{L}_i, \overline{L}_i]$. We assume that component 1 suffers a higher lead time variability and thus let $l_1^t \in Uniform[4, 6]$, whereas the lead time of component 2 has two possible values with $l_2^t \in Uniform[6, 7]$. Note that the orders of component 1 may cross with the given distribution. To avoid that, when the realized lead time of component 1 in any period is 6, the lead time distribution in the following period becomes $Uniform[5, 6]$ in the experiment. To obtain SAA-AX, we consider the deterministic lead times of 4 and 6 for components 1 and 2, respectively. The reward of 1 is collected to each product assembled within its response time window and the response time windows for all products are set to 0. Therefore, the service performance corresponds to the *off-shelf* type-II service level (in percent) in this section.

Table 5.2: Problem setting of Agrawal and Cohen (2001) system

			Products				
			j	1	2	3	4
			μ_j	15	18	18	15
			σ_j	3	3	3	3
			$r_{j,t}$	1	1	1	1
	Components		w_j	0	0	0	0
i	c_i	*Uniform*$[\underline{L}_i, \overline{L}_i]$		BOM ($a_{i,j}$)			
1	10	[4, 6]		1	3	3	1
2	10	[6, 7]		2	1	1	2

We obtain SAA-AX, SAA-TS, and SAA-MS using parameters $M = 500$, $N = 30$, and $N' = 500$ under different budget levels. The results are reported in Table 5.3. The CPU-time for running Algorithm 1 varies between 24 min and 28 min. For Algorithm 2, the CPU-time required is significantly increased. This is not surprising as each SAA problem (formulation (5.23)) involves a multiple number of scenarios, which is not decomposable. The CPU-time for running Algorithm 2 varies between 106 min and 114 min.

In Table 5.3, we note that the gaps between the LBs and UBs range from 0.04% to 0.63% throughout the table, which indicates that the SAA algorithms produce provably high quality solutions. We also note that the base-stock levels under the three SAA solutions are monotonely increasing as the budget level increases. However, it is observed that the increment of base-stock levels under SAA-AX slows down when the budget level exceeds 15,000. This is because the corresponding type-II service level almost achieves

100% with a budget of 15,000, the inventory budget constraint will no longer be active in the optimal base-stock levels with a larger budget level. In other words, it means that a budget of $15,000$ would be sufficient if only demand uncertainty is present in the system.

Table 5.3: Computational results for solving Agrawal and Cohen (2001) system using the SAA algorithms

	SAA-AX					SAA-TS					SAA-MS				
B	S_1^*	S_2^*	LB%	UB%	Gap%	S_1^*	S_2^*	LB%	UB%	Gap%	S_1^*	S_2^*	LB%	UB%	Gap%
13,000	647	653	70.61	70.92	0.31	662	638	11.68	11.89	0.21	684	616	10.80	10.97	0.17
14,000	708	692	94.86	95.03	0.17	725	675	29.23	29.48	0.25	741	659	28.06	28.56	0.50
15,000	757	743	99.14	99.20	0.06	768	732	57.26	57.80	0.54	780	720	54.83	55.32	0.49
16,000	798	762	99.79	99.89	0.10	820	780	78.45	78.53	0.08	842	758	79.42	79.89	0.47
17,000	822	780	100.83	100.87	0.04	900	800	94.00	94.56	0.56	920	780	94.06	94.67	0.61
18,000	834	792	100.71	100.78	0.07	968	832	99.56	100.02	0.46	982	818	98.55	99.18	0.63

In addition, we notice that the base-stock levels for the two components are comparatively closer to each other under SAA-AX in contrast to that they become more unequal under SAA-TS (also SAA-MS) at each budget level. It appears that the deviations of the base-stock levels get larger under SAA-TS (also SAA-MS) as the budget level increases. Under SAA-TS and SAA-MS, we observe that the base-stock level of component 1 is always larger than that of component 2, which is very likely due to the higher lead time variability of component 1. As a result, a higher base-stock level for component 1 is required to ensure a more reliable performance. Moreover, it can be seen that the LB and UB under SAA-AX are significantly higher than those under SAA-TS and SAA-MS for a budget between 13,000 and 16,000, which indicates that the lead time uncertainty could significantly degrade the service performance, especially with the restrictive inventory budget.

Next, we simulate the derived base-stock levels with 1000 periods of realized lead time and demand. Note that the lead times and demands are generated using the same distributions and parameters in Table 5.2. Given the base-stock levels, we optimally make component allocation decisions (i.e., OA policy) for each period. The average service performance is estimated based on these 1000 periods. Table 5.4 compares the average type-II service level (Avg. OA), the maximum service level (Max OA), and the standard deviation of the service level (SD) under the three sets of base-stock levels.

Table 5.4: Comparison of simulated results (in percent) based on 1000 simulation runs for Agrawal and Cohen (2001) system under different SAA base-stock levels

	SAA-AX			SAA-TS			SAA-MS		
B	Avg. OA	Max OA	SD	Avg. OA	Max OA	SD	Avg. OA	Max OA	SD
13,000	5.61	52.58	26.15	8.23	55.52	22.17	10.83	56.97	20.18
14,000	24.29	63.67	24.21	26.53	69.64	21.64	28.71	71.11	19.02
15,000	51.83	89.79	21.38	53.00	94.73	20.54	54.88	97.21	18.21
16,000	66.50	93.18	19.70	77.75	99.79	18.51	79.10	99.73	17.73
17,000	71.74	94.73	19.49	93.33	105.24	16.50	94.64	107.79	15.21
18,000	72.31	94.21	19.34	99.04	113.79	12.06	99.14	110.27	10.45

From the table above, it is clearly that the performance measures under SAA-TS and SAA-MS is consistently better than SAA-AX. For tight budget levels, it is worth pointing out that the simulated average service level under SAA-AX is significantly lower than the corresponding SAA results. For example, we see that at a budget of $13,000$, the LB and UB under SAA-AX are 70.61% and 70.92%, respectively, in Table 5.4, while the simulated result is merely 5.61%. This observation highlights that the lead time uncertainty could severely impact the system performance.

We also note that the performance measures improve as the budget level increases, while the improvement under SAA-AX is rather insignificant once the budget level exceeds 15,000. Under SAA-AX, we see that the average service level is only increased by

5.81% when the budget level is increased from 16,000 to 18,000. In contrast, the average service level is increased by 42.21% when the budget level is increased from 14,000 to 16,000. This is because, as noted earlier, the base-stock levels under SAA-AX only marginally increase when the budget gets larger than 15,000 (see Table 5.3). The result with regard to the average service level is further illustrated in Figure 5.3.

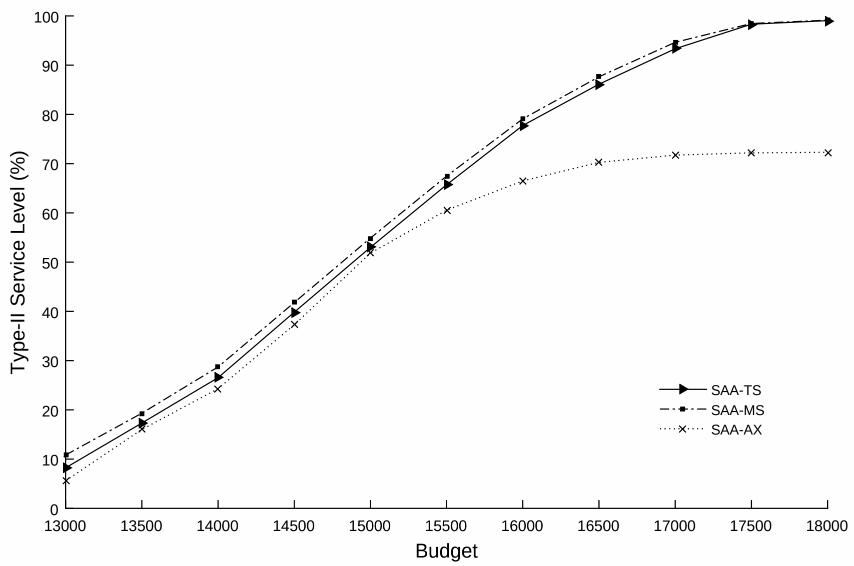

Figure 5.3: Average service levels of Agrawal and Cohen (2001) system using different SAA base-stock levels

The figure clearly shows the average service performance under SAA-TS compares well to SAA-MS at all budget levels under consideration. The percentage difference between SAA-TS and SAA-MS becomes almost negligible as the budget increases. The implication is that one can potentially use the proposed two-stage model instead of the multi-stage model for the base-stock optimization for large ATO systems, especially when a sufficient budget is allocated. After all, solving the SAA algorithm of the two-stage model requires less computational effort.

5.4.2 Zhang System

In this section, we consider an ATO system studied in Zhang (1997) that consists of four products ($j = 1, \ldots, 4$) and five components ($i = 1, \ldots, 5$) as summarized in Table 5.5. The product demands in each period are assumed to be normally distributed with the mean μ_j and standard deviation σ_j. The components are ordered from external suppliers with uncertain lead times. For the sake of simplicity, we assume that the lead times for *some* components are discretely distributed while the others are deterministic.

Table 5.5: Problem setting of Zhang (1997) system

			Products				
			j	1	2	3	4
			μ_j	100	150	50	30
			σ_j	25	30	15	11
			$r_{j,t}$	1	1	1	1
Components			w_j	1	1	1	1
i	c_i	$\{\underline{L}_i, \ldots, \overline{L}_i\}$		BOM ($a_{i,j}$)			
1	2	{3}		1	1	–	–
2	3	{2, 3}		2	1	1	–
3	6	{2, 3}		1	1	1	–
4	4	{4, 5}		–	–	1	1
5	1	{4}		–	–	–	1

Specifically, the lead times of components 2, 3 and 4 have lead time uncertainty with $l_2^t \in [2,3], l_3^t \in [2,3]$ and $l_4^t \in [4,5]$. For these components, note that the lower bound lead time values (i.e., 2, 2, and 4) are used to obtain SAA-AX. To demonstrate the lead time variability on the system performance, we consider two lead time distributions, they are: *Uniform*$[\underline{L}_i,\ \overline{L}_i]$ and *Two-point*$[\underline{L}_i,\ \overline{L}_i]$ (where $\Pr(l_i^t = \underline{L}_i) = 0.75$ and $\Pr(l_i^t = \overline{L}_i) = 0.25$). The response time window of 1 is set for all the products, and the reward of 1 is collected as long as the products are satisfied within their response time windows.

Thus, the type-II service level (in percent) reported in this section corresponds to the fraction of demands satisfied within one period of their arrival.

We select the parameters of $M = 500$, $N = 25$ and $N' = 500$ for the SAA algorithms. Tables 5.6, 5.7, and 5.8 show the SAA results when the lead times are deterministic, two-point distributed, and uniformly distributed, respectively. The CPU-time for running Algorithm 1 varies between 1 hour and 1.25 hours. For the Algorithm 2, it typically takes around 2.5 hours to obtain the solution.

Table 5.6 presents the SAA results under the deterministic lead times. The results indicate that it would be better off to allocate a majority of the inventory budget to the base-stock levels for components 1, 2, and 3 to satisfy demands of products 1 and 2. In particular, at a budget of 8,000, we observe that the base-stock levels of components 4 and 5 are set to zero. This is intuitive since the lead times of components 1, 2, and 3 are

Table 5.6: Computational results for solving Zhang (1997) system using the SAA algorithms where lead times are deterministic

	SAA-AX							
B	S_1^*	S_2^*	S_3^*	S_4^*	S_5^*	LB%	UB%	Gap%
8,000	874	838	623	0	0	74.53	75.27	0.56
10,000	867	873	664	373	171	97.56	98.17	0.61
12,000	930	1156	796	430	174	99.74	100.15	0.41
14,000	954	1183	801	449	235	99.77	100.05	0.29

lower than components 4 and 5. In fact, a similar trend is observed for the base-stock levels under SAA-TS and SAA-MS; however, there is a lag until the base-stock levels of components 4 and 5 become non-zeros (see Tables 5.7 and 5.8), which is definitely caused by the lead time uncertainty. Moreover, we observe that the increment of the base-stock levels under SAA-AX slows down once the budget reaches 12,000, which implies that such the budget level would handle demand uncertainty well. This is verified through the fact that the type-II service level almost achieves 100% with a budget level $B = 12,000$ in the table.

In Tables 5.7 and 5.8, we see that the quality of the derived base-stock levels under SAA-TS and SAA-MS is guaranteed by the tightness of the gaps between the LBs and UBs, which range from 0.02% to 0.78% across the tables. For a low budget level, we note that the values of LB and UB are largely decreased compared to those in Table 5.6. This coincides with the results for Agrawal and Cohen (2001) that the lead time uncertainty significantly degenerates the system performance, especially under a restrictive inventory budget.

Different from the results in Table 5.4, we find that not all the base-stock levels monotonely increase in this system. Indeed, when the budget level is increased from $10,000$ to $12,000$ in Tables 5.7 and 5.8, the base-stock level of component 1 under SAA-TS and SAA-MS decreases. Meanwhile, the base-stock levels of components 4 and 5 become non-zeros. One reason for this may be that product 1 requires 2 units of component 2 while product 3 requires only 1 unit of the component, the system benefits from the component commonality through satisfying less demand for product 1, and shifting a part of the inventory budget for the base-stock levels of components 4 and 5 so as to satisfy demands of products 3 and 4.

Compare Tables 5.7 and 5.8, another interesting observation is that the lead time distribution seemingly has more impact on the base-stock levels when a tight budget is assumed. For example, we can see that at a budget between $8,000$ and $10,000$, a noticeable difference is that the base-stock level of component 2 under the uniform lead times is higher than that under the two-point lead times. The main reason is that the uniform distribution has the higher probability of producing larger lead times for the component than the two-point distribution; since component 2 is required for assembling of several products, it would be more beneficial to have a relatively higher base-stock level for component 2 against its lead time uncertainty under the uniform distribution.

Table 5.7: Computational results for solving Zhang (1997) system using the SAA algorithms where lead times follow two-point distributions

	SAA-TS							
B	$\widehat{S}_1^*$	$\widehat{S}_2^*$	$\widehat{S}_3^*$	$\widehat{S}_4^*$	$\widehat{S}_5^*$	LB	UB	Gap
8,000	799	848	643	0	0	38.71	39.33	0.62
10,000	839	1096	839	0	0	68.27	68.91	0.64
12,000	760	1108	879	430	162	91.31	91.97	0.66
14,000	883	1400	1015	446	160	98.73	99.08	0.35
	SAA-MS							
B	$\widehat{S}_1^*$	$\widehat{S}_2^*$	$\widehat{S}_3^*$	$\widehat{S}_4^*$	$\widehat{S}_5^*$	LB	UB	Gap
8,000	661	874	676	0	0	40.55	41.33	0.78
10,000	794	1106	849	0	0	70.07	70.77	0.70
12,000	763	1122	869	432	166	93.19	93.68	0.49
14,000	878	1313	1054	453	167	99.63	100.15	0.52

Table 5.8: Computational results for solving Zhang (1997) system using the SAA algorithms where lead times follow uniform distributions

	SAA-TS							
B	$\widehat{S}_1^*$	$\widehat{S}_2^*$	$\widehat{S}_3^*$	$\widehat{S}_4^*$	$\widehat{S}_5^*$	LB	UB	Gap
8,000	754	1054	555	0	0	35.09	35.76	0.66
10,000	815	1114	838	0	0	68.14	68.25	0.10
12,000	769	1126	879	415	150	88.01	88.79	0.78
14,000	925	1360	1010	460	170	97.48	97.92	0.44
	SAA-MS							
B	$\widehat{S}_1^*$	$\widehat{S}_2^*$	$\widehat{S}_3^*$	$\widehat{S}_4^*$	$\widehat{S}_5^*$	LB	UB	Gap
8,000	769	1068	543	0	0	36.81	37.19	0.38
10,000	821	1140	823	0	0	69.04	69.53	0.49
12,000	778	1153	857	422	154	88.15	88.47	0.32
14,000	914	1362	1018	454	162	99.80	99.82	0.02

In the simulation, for each set of the base-stock levels, we optimally solve the component allocation problem for 1000 periods of realized lead time and demand. The results including the average type-II service level (Avg. OA), the maximum service level (Max OA), and the standard deviation of the service level (SD) are reported in Tables 5.9 and 5.10 when the two-point distributed and uniformly distributed lead times are assumed, respectively.

Table 5.9: Comparison of simulated results (in percent) based on 1000 simulation runs for Zhang (1997) system under different base-stock levels where lead times follow two-point distributions

	SAA-AX			SAA-TS			SAA-MS		
B	Avg OA	Max OA	SD	Avg OA	Max OA	SD	Avg OA	Max OA	SD
8,000	30.32	76.36	29.48	34.34	81.55	27.05	38.87	84.21	24.68
10,000	45.10	88.61	30.58	69.09	93.03	15.04	71.88	98.70	14.74
12,000	76.47	99.18	19.30	92.88	105.55	12.56	93.82	107.36	12.68
14,000	82.35	103.91	18.35	99.73	112.45	11.83	99.85	113.91	11.19

Table 5.10: Comparison of simulated results (in percent) based on 1000 simulation runs for Zhang (1997) system under different base-stock levels where lead times follow uniform distributions

	SAA-AX			SAA-TS			SAA-MS		
B	Avg OA	Max OA	SD	Avg OA	Max OA	SD	Avg OA	Max OA	SD
8,000	28.88	72.02	30.23	36.53	78.88	28.42	36.72	79.33	28.80
10,000	42.66	86.91	30.77	68.06	91.03	19.06	68.62	93.67	15.98
12,000	74.40	96.03	20.60	91.18	107.64	14.86	91.27	103.64	13.48
14,000	80.44	103.85	15.78	97.14	111.09	12.04	97.17	110.73	11.62

From the tables, although all three sets of the base-stock levels offer improved performance measures as the budget level increases, the benefit of incorporating lead time uncertainty is clearly demonstrated. We observe that the base-stock levels under SAA-TS (respectively, SAA-MS) outperform SAA-AX with at least 4.02% (respectively, 7.65%) higher average service level, and 1.18% (respectively, 1.43%) lower standard deviation. When it comes to compare the system performance between SAA-TS and SAA-MS, we notice that the base-stock levels under SAA-TS performs similarly to SAA-MS, especially with the uniformly distributed lead times. This finding is well illustrated in Figures 5.4 and 5.5.

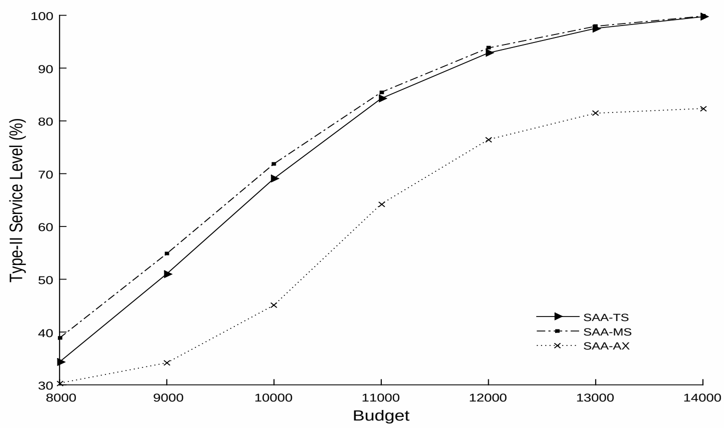

Figure 5.4: Average service levels of Zhang (1997) system using different SAA base-stock levels where lead times follow two-point distributions

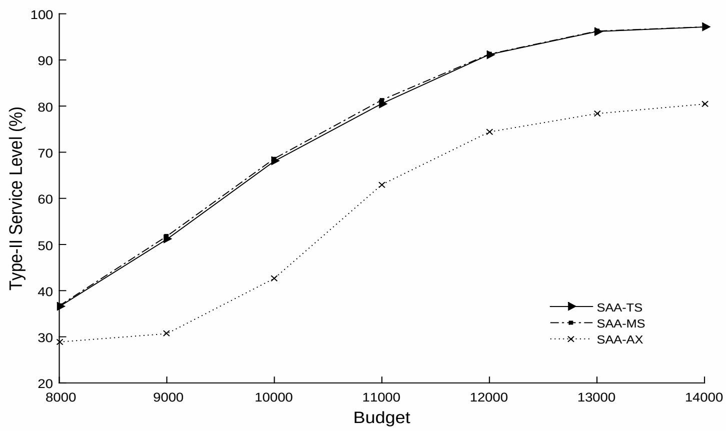

Figure 5.5: Average service levels of Zhang (1997) system using different SAA base-stock levels where lead times follow uniform distributions

In the figures, it is also interesting to observe that the improvement of the average service level under SAA-AX is very insignificant relative to SAA-TS and SAA-MS, when the budget level is increased from 80,000 to 90,000. At a budget of $9,000$, we find that the base-stock levels under SAA-AX are $(S_1^*, S_2^*, S_3^*, S_4^*, S_5^*) = (769, 780, 597, 348, 148)$, in which the values for components 1, 2, and 3 are essentially decreased compared to the optimal base-stock levels at a budget of $8,000$ (see Table 5.6). Implementing such base-stock levels will certainly enhance the service levels of products 3 and 4, while the service improvement needs to compensate for the service reduction of products 1 and 2. That explains why the aggregated service level only insignificantly increases in the figures.

5.5 Concluding Remarks

In this chapter, we have proposed stochastic integer programs for a periodic-review ATO system with component base-stock policy and FCFS allocation rule. The proposed stochastic models jointly determine the optimal base-stock levels and the optimal component allocation decisions in the presence of stochastic lead times and demands. The proposed multi-stage model can handle realistic situations where the decision maker only knows the distributions of lead times, however, at the cost of higher degree of computational complexity. We used the SAA algorithms to obtain the near-optimal base-stock levels, as well as estimated upper and lower bounds, for the proposed stochastic models. The effectiveness of the SAA algorithms has been tested with two examples from the ATO literature. In addition, we carried out the simulation studies with the derived base-stock levels as well as the base-stock levels of Akçay and Xu (2004). The simulated results show that better system performance can be achieved with the base-stock policies that incorporate the lead time uncertainty. Moreover, we found that the two-stage model compares well to the multi-stage model in terms of the base-stock level optimization because the similar performance measures have been observed in our simulated results, especially when the lead times follow uniform distributions and the higher inventory budget is allocated.

Chapter 6

Conclusions and Future Research Directions

In this chapter, the major contributions of this thesis are summarized. In addition, we suggest some future research directions.

6.1 Conclusions

This thesis presented mathematical models based on robust optimization (RO) and stochastic programming (SP) to address the inventory management decisions for three periodic-review, finite-horizon inventory systems in the presence of supply and demand uncertainties.

First, we considered a multi-period single-station inventory problem, where we modeled the uncertain partial supply and uncertain demand with budget polyhedral uncertainty sets. We formulated the problem under the RO framework and showed that the derived robust counterpart is equivalent to a nominal problem with a modified deterministic demand sequence. Furthermore, we theoretically characterized the optimal robust policy for the stationary case of the problem. We showed that the optimal robust policy is of *quasi*-(s, S) form, where s and S levels are theoretical computable. The simulation-based numerical studies suggest that the proposed robust policy could significantly outperform the nominal policy as well as the robust policy of Bertsimas and Thiele (2006) in terms of average performance when the relatively low (or ideally no) setup cost is present and the unit holding cost is much smaller than the unit shortage cost, and it offers improved average performance as the length of planning horizon increases. We also examined how the supply variability, demand variability, and the budgets for uncertainty affect the average performance.

Next, we extended the RO framework to a multi-echelon supply chain with a tree structure. By assuming that the partial supply only affects the orders placed by the main storage hubs, we showed that the robust counterpart of the supply chain problem is decomposable by echelons; however, the optimal robust policies for the echelons are not necessarily identical. Specifically, we have shown that the optimal robust policy for echelons of the main storage hubs can be obtained by solving a single-station problem with modified deterministic demands, whereas the optimal robust policy for echelons of the local storage hubs or stores can be obtained by solving a capacitated single-station problem with modified deterministic demands. The extensive numerical studies indicate that the proposed robust policy performs well compared to the nominal policy and the robust policy of Bertsimas and Thiele (2006) in terms of the average performance and the performance stability. In addition, we found that an increase in the length of planning horizon yields the better performance stability while the presence of setup costs seemingly does not much affect the performance stability.

In the context of a periodic-review Assemble-To-Order (ATO) system, we addressed the joint optimization of inventory replenishment and component allocation decisions in the face of stochastic component replenishment lead times and stochastic product demands. We analyzed the problems with independent base-stock policy for inventory replenishment and First-Come-First-Served (FCFS) rule for component allocation. In the case that the decision maker has full knowledge of the realized lead times, we formulated the problems under a two-stage SP framework where in the first-stage we decide on the base-stock levels with an inventory budget constraint, and the second-stage decisions are concerned with the component allocation. Subsequently, we assumed that the decision maker only has knowledge of random lead time distributions, which leads to a scenario tree for optimization. In this case, we formulated the problems under a multi-stage SP framework. We determined the base-stock levels using the sample average approximation (SAA) algorithms, and tested the longer-term performance of the derived base-stock levels by simulation. The simulated results indicate that our proposed base-stock levels provide better and more stable system performance compared to the base-stock levels of Akçay and Xu (2004), which highlights the importance of incorporating the lead time uncertainty into decision-making. Moreover, we found that the performance measures of the two-stage SP compare well with the multi-stage SP, which demonstrates that the two-stage SP could be potentially used as an approximation for the multi-stage SP for the base-stock level optimization. This is particularly appealing in the application of the proposed SP models for large ATO systems since solving the two-stage SP model requires significantly less computational effort than that of the multi-stage SP model.

6.2 Future Research Directions

The studies in this thesis can be extended into several directions. Firstly, for the static RO approach used in Chapters 3 and 4, it is well-known that this RO formulation usually results in a overly conservative solution. To avoid over-conservativeness, future work includes developing a Benders' algorithm based on Bienstock and ÖZbay (2008) to solve the "true" min-max version of the proposed robust models. Alternatively, adjustable RO (ARO) approach proposed by Ben-Tal et al. (2004) can produce less conservative solutions since it allows the decision maker to dynamically incorporate the information of the recent realized uncertain parameters and make a set of recourse actions accordingly. Therefore, it would be highly interesting to formulate the inventory problems under the ARO framework and compare the derived adjustable policies with our proposed *quasi*-(s, S) policy. Secondly, it would be useful and interesting to integrate a decomposition-based algorithm into the SAA algorithm for the multi-stage SP model proposed in Chapter 5 so as to facilitate the computation. Also, it is of practical interest to extend the existing component allocation heuristics to our multi-stage problem setting for large-scale ATO systems.

Bibliography

Agrawal, N. and Cohen, M. A. (2001). Optimal material control in an assembly system with component commonality. *Naval Research Logistics*, 48(5):409–429.

Ahmed, S., Shapiro, A., and Shapiro, E. (2002). The sample average approximation method for stochastic programs with integer recourse. *SIAM Journal of Optimization*, 24:479–502.

Akbari, A. A. and Karimi, B. (2015). A new robust optimization approach for integrated multi-echelon, multi-product, multi-period supply chain network design under process uncertainty. *The International Journal of Advanced Manufacturing Technology*, 79(1-4):229–244.

Akçay, Y. and Xu, S. H. (2004). Joint inventory replenishment and component allocation optimization in an assemble-to-order system. *Management Science*, 50(1):99–116.

Alem, D. J. and Morabito, R. (2012). Production planning in furniture settings via robust optimization. *Computers and Operations Research*, 39(2):139–150.

Alvarez, P. P. and Vera, J. R. (2014). Application of robust optimization to the sawmill planning problem. *Annals of Operations Research*, 219(1):457–475.

Aouam, T. and Brahimi, N. (2013). Integrated production planning and order acceptance under uncertainty: A robust optimization approach. *European Journal of Operational Research*, 228(3):504–515.

Arrow, K. J., Karlin, S., Scarf, H. E., et al. (1958). Studies in the mathematical theory of inventory and production. Stanford University Press, CA.

Baker, K. R., Magazine, M. J., and Nuttle, H. L. (1986). The effect of commonality on safety stock in a simple inventory model. *Management Science*, 32(8):982–988.

Bakir, M. A. and Byrne, M. D. (1998). Stochastic linear optimisation of an MPMP production planning model. *International Journal of Production Economics*, 55(1):87–96.

Barbarosoğlu, G. and Arda, Y. (2004). A two-stage stochastic programming framework for transportation planning in disaster response. *Journal of the Operational Research Society*, 55(1):43–53.

Ben-Tal, A., Boaz, G., and Shimrit, S. (2009). Robust multi-echelon multi-period inventory control. *European Journal of Operational Research*, 199(3):922–935.

Ben-Tal, A., Golany, B., Nemirovski, A., and Vial, J.-P. (2005). Retailer-supplier flexible commitments contracts: A robust optimization approach. *Manufacturing & Service Operations Management*, 7(3):248–271.

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

Ben-Tal, A. and Nemirovski, A. (1999). Robust solutions of uncertain linear programs. *Operations Research Letters*, 25(1):1–13.

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

Bertsimas, D., Iancu, D. A., and Parrilo, P. A. (2011). A hierarchy of near-optimal policies for multistage adaptive optimization. *IEEE Transactions on Automatic Control*, 56(12):2809–2824.

Bertsimas, D. and Pachamanova, D. (2008). Robust multiperiod portfolio management in the presence of transaction costs. *Computers & Operations Research*, 35(1):3–17.

Bertsimas, D. and Sim, M. (2003). Robust discrete optimization and network flows. *Mathematical Programming*, 98(1):49–71.

Bertsimas, D. and Sim, M. (2004). The price of robustness. *Operations Research*, 52(1):35–53.

Bertsimas, D. and Thiele, A. (2006). A robust optimization approach to inventory theory. *Operations Research*, 54(1):150–168.

Bienstock, D. and ÖZbay, N. (2008). Computing robust basestock levels. *Discrete Optimization*, 5(2):389–414.

Birge, J. R. and Louveaux, F. (2011). Introduction to stochastic programming. Springer Science & Business Media.

Bohle, C., Maturana, S., and Vera, J. (2010). A robust optimization approach to wine grape harvesting scheduling. *European Journal of Operational Research*, 200(1):245–252.

Bollapragada, R., Rao, U. S., and Zhang, J. (2004). Managing two-stage serial inventory systems under demand and supply uncertainty and customer service level requirements. *IIE Transactions*, 36(1):73–85.

Brandimarte, P. (2006). Multi-item capacitated lot-sizing with demand uncertainty. *International Journal of Production Research*, 44(15):2997–3022.

Chu, C., Proth, J. M., and Xie, X. (1993). Supply management in assembly systems. *Naval Research Logistics*, 40(7):933–949.

Ciarallo, F. W., Akella, R., and Morton, T. E. (1994). A periodic review, production planning model with uncertain capacity and uncertain demand—optimality of extended myopic policies. *Management Science*, 40(3):320–332.

Clark, A. J. and Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4):475–490.

Dantzig, G. B. (1955). Linear programming under uncertainty. *Management Science*, 1(3-4):197–206.

Dantzig, G. B. and Glynn, P. W. (1990). Parallel processors for planning under uncertainty. *Annals of Operations Research*, 22(1):1–21.

Deza, A., Huang, K., Liang, H., and Wang, X. J. (2018). On component commonality for periodic review assemble-to-order systems. *Annals of Operations Research*, 265:29–46.

Dillon, M., Oliveira, F., and Abbasi, B. (2017). A two-stage stochastic programming model for inventory management in the blood supply chain. *International Journal of Production Economics*, 187:27–41.

Doğru, M. K., Reiman, M. I., and Wang, Q. (2010). A stochastic programming based inventory policy for assemble-to-order systems with application to the W model. *Operations Research*, 58(4-part-1):849–864.

El Agizy, M. (1969). Dynamic inventory models and stochastic programming. *IBM Journal of Research and Development*, 13(4):351–356.

Escudero, L. F., Kamesam, P. V., King, A. J., and Wets, R. J. (1993). Production planning via scenario modelling. *Annals of Operations research*, 43(6):309–335.

Federgruen, A. and Zipkin, P. (1984). Computational issues in an infinite-horizon, multiechelon inventory model. *Operations Research*, 32(4):818–836.

Fragniere, E., Gondzio, J., and Vial, J.-P. (2000). Building and solving large-scale stochastic programs on an affordable distributed computing system. *Annals of Operations Research*, 99(1):167–187.

Fujiwara, O. and Sedarage, D. (1997). An optimal (Q, r) policy for a multipart assembly system under stochastic part procurement lead times. *European Journal of Operational Research*, 100(3):550–556.

Gabrel, V., Murat, C., and Thiele, A. (2014). Recent advances in robust optimization: An overview. *European Journal of Operational Rresearch*, 235(3):471–483.

Gallien, J. and Wein, L. M. (2001). A simple and effective component procurement policy for stochastic assembly systems. *Queueing Systems*, 38(2):221–248.

Gerchak, Y., Vickson, R. G., and Parlar, M. (1988). Periodic review production models with variable yield and uncertain demand. *IIE Transactions*, 20(2):144–150.

Hausman, W. H., Lee, H. L., and Zhang, A. X. (1998). Joint demand fulfillment probability in a multi-item inventory system with independent order-up-to policies. *European Journal of Operational Research*, 109(3):646–659.

He, X. X., Xu, S. H., Ord, J. K., and Hayya, J. C. (1998). An inventory model with order crossover. *Operations Research*, 46(3-supplement-3):S112–S119.

Hood, S. J., Bermon, S., and Barahona, F. (2003). Capacity planning under demand uncertainty for semiconductor manufacturing. *IEEE Transactions on Semiconductor Manufacturing*, 16(2):273–280.

Hsu, V. N., Lee, C. Y., and So, K. C. (2006). Optimal component stocking policy for assemble-to-order systems with lead-time-dependent component and product pricing. *Management Science*, 52(3):337–351.

Huang, K. (2014). Benchmarking non-first-come-first-served component allocation in an assemble-to-order system. *Annals of Operations Research*, 223(1):217–237.

Huang, K. and Ahmed, S. (2009). The value of multistage stochastic programming in capacity planning under uncertainty. *Operations Research*, 57(4):893–904.

Huang, K. and de Kok, T. (2015). Optimal FCFS allocation rules for periodic-review assemble-to-order systems. *Naval Research Logistics*, 62(2):158–169.

Kall, P. and Wallace, S. W. (1994). Stochastic programming. Wiley, Chichester.

Karmarkar, U. S. (1981). The multiperiod multilocation inventory problem. *Operations Research*, 29(2):215–228.

Karmarkar, U. S. (1987). The multilocation multiperiod inventory problem: Bounds and approximations. *Management Science*, 33(1):86–94.

Kazemi Zanjani, M., Nourelfath, M., and Ait-Kadi, D. (2010). A multi-stage stochastic programming approach for production planning with uncertainty in the quality of raw materials and demand. *International Journal of Production Research*, 48(16):4701–4723.

Kleywegt, A. J., Shapiro, A., and Homem-de Mello, T. (2002). The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502.

Kumar, A. (1989). Component inventory costs in an assembly problem with uncertain supplier lead-times. *IIE Transactions*, 21(2):112–121.

Lu, Y. and Song, J. (2005). Order-based cost minimization in assemble-to-order systems. *Operations Research*, 53(1):151–169.

Lu, Y., Song, J., and Yao, D. (2003). Order fill rate, lead time variability, and advance demand information in an assemble-to-order system. *Operations Research*, 51(2):292–308.

Mauroy, G. and Wardi, Y. (1995). Cost optimization in supply management policies for assembly systems with random component yield times. In *Emerging Technologies and Factory Automation, 1995. ETFA'95, Proceedings., 1995 INRIA/IEEE Symposium on*, volume 3, pages 445–453. IEEE.

MirHassani, S. A., Lucas, C., Mitra, G., Messina, E., and Poojari, C. A. (2000). Computational solution of capacity planning models under uncertainty. *Parallel Computing*, 26(5):511–538.

Mulvey, J. M., Vanderbei, R. J., and Zenios, S. A. (1995). Robust optimization of large-scale systems. *Operations Research*, 43(2):264–281.

Ordóñez, F. and Zhao, J. (2007). Robust capacity expansion of network flows. *Networks*, 50(2):136–145.

Ozbay, K. and Ozguven, E. (2007). Stochastic humanitarian inventory control model for disaster planning. *Transportation Research Record: Journal of the Transportation Research Board*, (2022):63–75.

Parlar, M. and Berkin, D. (1991). Future supply uncertainty in EOQ models. *Naval Research Logistics*, 38(1):107–121.

Paul, J. A. and Wang, X. J. (2015). Robust optimization for United States Department of Agriculture food aid bid allocations. *Transportation Research Part E: Logistics and Transportation Review*, 82:129–146.

Proth, J. M., Mauroy, G., Wardi, Y., Chu, C., and Xie, X. (1997). Supply management for cost minimization in assembly systems with random component yield times. *Journal of Intelligent Manufacturing*, 8(5):385–403.

Rikun, A. A. (2011). Applications of robust optimization to queueing and inventory management. PhD thesis, Massachusetts Institute of Technology.

Rosling, K. (1989). Optimal inventory policies for assembly systems under random demands. *Operations Research*, 37(4):565–579.

Ruszczynski, A. (1999). Some advances in decomposition methods for stochastic linear programming. *Annals of Operations Research*, 85:153–172.

Sanei Bajgiran, O., Kazemi Zanjani, M., and Nourelfath, M. (2017). Forest harvesting planning under uncertainty: a cardinality-constrained approach. *International Journal of Production Research*, 55(7):1914–1929.

Schütz, P., Tomasgard, A., and Ahmed, S. (2009). Supply chain design under uncertainty using sample average approximation and dual decomposition. *European Journal of Operational Research*, 199(2):409–419.

See, C. T. and Sim, M. (2010). Robust approximation to multiperiod inventory management. *Operations Research*, 58(3):583–594.

Shapiro, A. and Nemirovski, A. (2005). On complexity of stochastic programming problems. In Jeyakumar, V. and Rubinov, A. M., editors, *Continuous Optimization: Current Trends and Applications*, pages 111–144. Springer.

Shore, H. (1995). Setting safety lead-times for purchased components in assembly systems: A general solution procedure. *IIE Transactions*, 27(5):638–645.

Solyalı, O., Cordeau, J. F., and Laporte, G. (2015). The impact of modeling on robust inventory management under demand uncertainty. *Management Science*, 62(4):1188–1201.

Song, J. S., Xu, S. H., and Liu, B. (1999). Order-fulfillment performance measures in an assemble-to-order system with stochastic leadtimes. *Operations Research*, 47(1):131–149.

Song, J. S., Yano, C. A., and Lerssrisuriya, P. (2000). Contract assembly: Dealing with combined supply lead time and demand quantity uncertainty. *Manufacturing & Service Operations Management*, 2(3):287–296.

Song, J. S. and Yao, D. D. (2002). Performance analysis and optimization of assemble-to-order systems with random lead times. *Operations Research*, 50:889–903.

Song, J. S. and Zhao, Y. (2009). The value of component commonality in a dynamic inventory system with lead times. *Manufacturing & Service Operations Management*, 11(3):493–508.

Soyster, A. L. (1973). Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5):1154–1157.

Srinivasan, M., Novack, R., and Thomas, D. (2011). Optimal and approximate policies for inventory systems with order crossover. *Journal of Business Logistics*, 32(2):180–193.

Thiele, A. (2006). Single-product pricing via robust optimization. Working Paper, Lehigh University.

Thiele, A. (2010). A note on issues of over-conservatism in robust optimization with cost uncertainty. *Optimization*, 59(7):1033–1040.

Thorsen, A. and Yao, T. (2017). Robust inventory control under demand and lead time uncertainty. *Annals of Operations Research*, 257:207–236.

Tsiakis, P., Shah, N., and Pantelides, C. C. (2001). Design of multi-echelon supply chain networks under demand uncertainty. *Industrial and Engineering Chemistry Research*, 40(16):3585–3604.

van Jaarsveld, W. and Scheller-Wolf, A. (2015). Optimization of industrial-scale assemble-to-order systems. *INFORMS Journal on Computing*, 27(3):544–560.

Varas, M., Maturana, S., Pascual, R., Vargas, I., and Vera, J. (2014). Scheduling production for a sawmill: A robust optimization approach. *International Journal of Production Economics*, 150:37–51.

Verweij, B., Ahmed, S., Kleywegt, A. J., Nemhauser, G., and Shapiro, A. (2003). The sample average approximation method applied to stochastic routing problems: a computational study. *Computational Optimization and Applications*, 24(2-3):289–333.

Wei, C., Li, Y., and Cai, X. (2011). Robust optimal policies of production and inventory with uncertain returns and demand. *International Journal of Production Economics*, 134(2):357–367.

Whitin, T. M. (1957). *Theory of inventory management.* Princeton University Press.

Xu, P. J., Allgor, R., and Graves, S. C. (2009). Benefits of reevaluating real-time order fulfillment decisions. *Manufacturing & Service Operations Management*, 11(2):340–355.

Yano, C. A. (1987). Stochastic leadtimes in two-level assembly systems. *IIE Transactions*, 19(4):371–378.

Yano, C. A. and Lee, H. L. (1995). Lot sizing with random yields: A review. *Operations Research*, 43(2):311–334.

Zhang, A. (1997). Demand fulfillment rates in an assemble-to-order system with multiple products and dependent demands. *Production and Operations Management*, 6(3):309–324.

Zhao, Y. and Simchi-Levi, D. (2006). Performance analysis and evaluation of assemble-to-order systems with stochastic sequential lead times. *Operations Research*, 54(4):706–724.

Zipkin, P. H. (2000). Foundations of Inventory Management. McGraw-Hill, New York.