RISK-AVERSE DECISION MAKING AND BILEVEL STOCHASTIC PROGRAMMING WITH APPLICATIONS

By

DENIZ SEYED ESKANDANI

A dissertation submitted to the

Graduate School—Newark

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Management

Written under the direction of

Jonathan Eckstein

and approved by

____

____

____

____

Newark, New Jersey

January, 2020

ABSTRACT OF THE DISSERTATION

Risk-averse Decision Making and Bilevel Stochastic Programming with Applications

By Deniz Seyed Eskandani

Dissertation Director:

Jonathan Eckstein

We present a novel modeling approach to time-consistently formulate three-stage risk-averse stochastic programming problems, using bilevel programming. For certain classes of applications, we empirically demonstrate that our approach can behave substantially differently from prior formulations of the problem. To obtain these results, we reformulate the $\mathcal{NP}$-hard bilevel model using complementarity constraints and then express it as a disjunctive program. However, this approach does not scale well, even using the best available commercial MIP solvers. To overcome this hurdle, we use a proximal bundle method to efficiently find a lower bound for the optimal solution. We further supplement this procedure with an upper bound by proposing an approach to find a feasible solution. We implement our algorithm in the *gurobipy* module of Python and apply it to various classes of problems and compare our computational results with our earlier disjunctive programming approach. We find that our bounds can provide a better approximation of the optimal solution than the MIP-solver approach and can scale to larger problems.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my advisor Professor Jonathan Eckstein for his continuous support and guidance during my course of research. I have learned tremendously from his vast knowledge. This dissertation would not have been possible without him.

I would also like to very much thank my committee members, Professors Mert Gurbuzbalaban, Welington de Oliveira, and Andrzej Ruszczynski for their insightful and constructive comments and feedback. Professor Ruszczynski is a great mentor who constantly benefited me and this dissertation with his invaluable insights in all stages of this research.

I am grateful to have had incredible professors in the Management Science and Operations Research groups of Rutgers Business School. I have immeasurably learned from their expertise. I especially thank Professor Farid Alizadeh, who provided outstanding guidance throughout my PhD study.

I am also extremely thankful for having wonderful and supportive friends in the PhD program at RBS. Their hard work has always inspired me.

Finally, I owe a deep sense of gratitude to my parents and sister for their indefinite support. They have immensely enthused me in all stages of my life to be the best person I could be. Last but not least, I thank my husband, Arian, who immeasurably supported me with endless love and encouragement. Without his constant support, this dissertation could not be completed.

Table of Contents

Chapter 1

Introduction

In this work, we study a broad class of risk-averse stochastic dynamic programs (RASDPs), in which a risk-averse decision maker (DM) sets an optimal course of action over time under uncertainty. Particularly, we focus on a subset of RASDPs with three stages. In this subset of problems, the DM first sets the values of the first-stage decision variables without full information about the future. In the second stage, the uncertainty partially realizes, and the DM takes a recourse action based on the partial information. Finally, in the third stage, the uncertainty and the DM payoff realize.

There is a wide range of practical problems that could be modeled in this way. For example, in supply chain management, a critical decision is to determine the order quantities of the manufacturing parts required to meet the demand for end products. For instance, it is common among car manufacturers to use various combinations of engines, transmissions, wheel, and bodies to produce different configurations of a car. Given the long manufacturing lead times and a large number of configurations, a manufacturer may need to order parts long in advance of the selling horizon, with little information about the demand of the end products. However, as the selling horizon approaches, manufacturers may learn more about the preferences of the customers and demand for each product variation and configure products accordingly. Finally, the demand of each configuration is realized during the selling horizon, and the manufacturer is rewarded with a profit margin of each unit sold and incurs a penalty for unsold units (such as inventory holding cost and depreciation). In this example, the part ordering phase is the first stage and configuration of the parts is the second stage of the stochastic program, and the demand is the source of the uncertainty in the problem. We return to this example in Chapter 3.1.

Although the literature on theory and applications of this class of problems has mainly focused on risk-neutral DMs, there is evidence suggesting that DMs may be risk averse in practice (Shapiro et al. 2009). For example, Schweitzer & Cachon (2000) finds that inventory managers exhibit risk-averse behavior, particularly for high-value products. Additionally, the *American Express Australia* (2016) survey revealed that CFOs are becoming risk averse as a response to environmental changes: a majority of CFOs surveyed were more risk averse in 2016 than 2015, where 79 percent of them identified the less stable and predictable business environment as the reason for their lower risk appetite. Consequently, a growing body of the literature has turned attention to risk-averse optimization over the past two decades. Artzner et al. (1999) was the first to propose the axiomatic approach of coherent risk measures for modeling risk-averse behavior. A coherent risk measure is a risk function that satisfies the desirable properties of monotonicity, convexity, positive homogeneity, and translation equivariance (see Section 2.1 for a formal definition of these properties). Widely used examples of such risk measures include mean value (resulting in risk-neutral models), mean-semideviation (MSD), and average value at risk (AVaR).

A risk-neutral two-stage problem can be easily converted to its risk-averse equivalent by replacing the expectation operator with a coherent risk measure. However, the use of coherent risk measures for multiperiod problems may be problematic without proper treatment: the optimizer of the first-stage problem may require taking a suboptimal action in the recourse problem—that is, the problem may lack the time consistency property (see Definition 4). To address lack of time consistency, the standard existing modeling approaches force the objective function to have a similar nested structure to the expectation operator. Specifically, instead of optimizing the risk measure of the outcome over all periods at once, the proposed approaches optimize a nested sum of the outcome risks at each stage. This approach guarantees the time consistency property. However, it also creates new challenges: first, the resulting dynamic risk measures lack the law-invariance property, meaning that two different outcomes with the same probability distribution of rewards may be assessed as having different risk levels. Additionally, under this nested structure, the DM's tolerance toward risk changes over time in somewhat an unintuitive fashion. Finally, the complex structure of the objective function might create difficulty for the managers in understanding and assessing the outcome of their actions. These are precisely the challenges addressed in this dissertation.

To overcome the disadvantages of the nested modeling approach, we propose an alternative treatment for modeling risk aversion by adding future recourse problems as constraints to the first-stage optimization problem. For three classes of applications, namely supply chain production planning, portfolio optimization, and hydropower energy planning, we empirically demonstrate that the optimal solution found using our approach can behave dramatically differently from the solutions of prior formulations. For example, in the supply chain problem class with the blended combination of expected value and AVaR risk function, more than 28% of the numerical instances resulted in a difference over 10% in the optimal values, with a maximum difference of 122,774%. However, we also demonstrate that our formulation is $\mathcal{NP}$-hard even for simple risk measures such as MSD and AVaR, and is significantly more time consuming to solve than the nested objective approach (Eckstein et al. 2016).

To address this challenge, we study various methods to efficiently solve our proposed formulation. First, as a benchmark, we convert the most general bilevel formulation of the problem to a linear program with complementarity constraints (LPCC) and solve it using the *gurobipy* module of Python. We find that this method does not scale well in the number of problem parameters and the size of sample space, and cannot solve even relatively small problem instances within a reasonable time. Since our LPCC formulation has a desirable additive structure, we use a proximal bundle method to find a lower bound for the problem. We also provide a technique to find a feasible upper bound. We use the upper and lower bounds provided in our algorithms to approximate the optimal solution of the original formulation and numerically illustrate that our approximation can be tight and may be solved significantly faster than the original LPCC in a general three-stage linear RASDP. Finally, we numerically test our results for the supply chain production planning example and observe that our approach can perform particularly well for such structured problems.

This dissertation first introduces the preliminaries and our formulation of three-stage RASDPs in Chapter 2. Then, in Chapter 3, we present three applications namely, supply chain production planning, portfolio optimization, and hydropower energy planning, and numerically compare the results of our formulation with the traditional approaches. In Chapter 4, we show how we employ a specialized bundle method to find proper lower and feasible upper bounds for a general standard bilevel problem. In Chapter 5, we apply our specialized bundle method algorithm to the supply chain production planning problem.

Chapter 2

Three-stage Risk-averse Stochastic Programming

In this chapter, we introduce the preliminaries of our work and formally define the properties of coherent risk measures. We also formulate the general three-stage RASDP and summarize the drawbacks of its standard time-consistent formulation. We then discuss alternative formulations to address these undesirable properties.

2.1 Preliminaries

Consider a finite probability space $(\Omega, \mathcal{F}, P)$ with a σ-algebra $\mathcal{F}$ (collection of all events) on the sample space Ω and probability measure P on $\mathcal{F}$. Let $\mathfrak{L}$ be the space of all $\mathcal{F}$-measurable random variables and "$\succeq$" be a partial order on $\mathfrak{L}$ such that for $Z, Z' \in \mathfrak{L}$, $Z \succeq Z'$ if and only if $Z(\omega) \geq Z'(\omega)$ for all $\omega \in \Omega$. A coherent risk measure on $\mathfrak{L}$ is defined as follows:

Definition 1 *A* ***coherent risk measure*** *is a function $\rho : \mathfrak{L} \to \bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ such that for random variables $Z, W \in \mathfrak{L}$ we have:*

Monotonicity. If $Z \preceq W$, then $\rho(Z) \leq \rho(W)$.

Convexity. $\rho(\alpha Z + (1-\alpha)W) \leq \alpha\rho(Z) + (1-\alpha)\rho(W)$ for all $\alpha \in [0, 1]$.

Positive homogeneity. $\rho(\alpha Z) = \alpha\rho(Z)$ for all $\alpha \geq 0$.

Translation equivariance. For any $t \in \mathbb{R}$, $\rho(Z + t) = \rho(Z) + t$.

A simple example of a coherent risk measure is the expected value function, which is used in risk-neutral models. Other examples of coherent risk measures frequently used to model risk aversion include mean-upper semideviation (MSD), average value-at-risk (AVaR), and blended risk measures (BRMs).

Example 1 ***(MSD)*** *Consider $Z \in \mathfrak{L}$ and $\gamma \in [0,1]$. The mean-upper semideviation risk is defined as:*

$$\mathrm{MSD}_\gamma(Z) := \mathbb{E}(Z) + \gamma \mathbb{E}[(Z - \mathbb{E}(Z))_+].$$

If Z is used to model cost, the MSD function penalizes the excess of the cost over its mean by a factor γ. A higher value of γ corresponds to a higher degree of risk aversion.

Example 2 ***(AVaR)*** *Consider $Z \in \mathfrak{L}$ and $\alpha \in (0,1]$. The average value at risk is defined as:*

$$\mathrm{AVaR}_\alpha(Z) := \inf_{t \in \mathbb{R}} \{t + \alpha^{-1} \mathbb{E}[Z - t]_+\}.$$

This risk measure averages the loss function over the $\alpha-$quantile of the cost function (Rockafellar et al. 2000) and a higher risk aversion is modeled using lower values of α.

Example 3 ***(BRM)*** *Consider $Z \in \mathfrak{L}$, $\alpha \in (0,1)$, and $\beta \in [0,1]$. We define the blended risk measure as:*

$$BRM^\alpha_\beta(Z) := (1-\beta)\mathbb{E}(Z) + \beta\, \mathrm{AVaR}_\alpha(Z).$$

The blended risk measure is the weighted average of the expected value and AVaR of the cost function. It is a generalization of the AVaR and risk neutral measures, where a higher value of β corresponds to more risk aversion.

Example 4 ***(worst outcome)*** *Consider $Z \in \mathfrak{L}$. The worst-outcome risk is defined as:*

$$\text{ess sup}(Z) := \inf\{b \in \mathbb{R} | P\{Z \le b\} = 1\}.$$

This risk measure is maximally risk averse.

Another desirable characteristic of a risk measure is the *law invariance* property. We have the following definition:

Definition 2 ***(law invariance)*** *A risk measure $\rho(\cdot)$ is law invariant if for $Z,\ W \in \mathfrak{L}$ with identical distributions we have $\rho(Z) = \rho(W)$.*

The law invariance property implies that two random variables with the same probability distribution should naturally induce the same risk for the DM. One can readily check that MSD_γ, AVaR_α, and BRM^α_β are all law-invariant coherent risk measures.

2.2 Three-Stage RASDP Formulations

Consider a risk-averse DM who takes actions over three stages. In the first stage, the DM decides the values of an n_1-dimensional decision variable X_1 with only probabilistic information about the second-and third-stage outcomes. Consequently, we set the $\sigma-$algebra corresponding to this stage to be $\mathcal{F}_1 = \{\emptyset, \Omega\}$, and X_1 is a deterministic, F_1-measurable vector. In the second stage, the first-stage uncertainty is realized and the DM learns the second-stage state, denoted by S. Let $\mathcal{E}_2$ be a partition of Ω with $S \in \mathcal{E}_2$. Also, let $\mathcal{F}_2$ be the σ-algebra corresponding to partition $\mathcal{E}_2$. We then have $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}$. In this stage, the DM sets the values of second-stage n_2-dimensional decision variables X_2, where X_2 is an F_2-measurable random variable. The decision made in scenario S is denoted by X_2^S. In the last stage, given the actions taken in the first two stages, the uncertainty entirely resolves and the DM decides the values of n_3-dimensional decision variables X_3, where X_3 is an $\mathcal{F}$-measurable random variable. We use the notation X_3^ω for the decision variable for atom $\omega \in \Omega$. We also use notation X_3^S for the third-stage decision variable, with its domain restricted to S. Finally, let Z_i be the $\mathcal{F}_i$-measurable cost function for each stage $i \in \{1, 2, 3\}$. The following scenario tree graphically illustrates an example sequence with sample space $\Omega = \{\omega_1, \cdots, \omega_6\}$ and second-stage scenario set $\mathcal{E}_2 = \{S_1, S_2, S_3\}$, where $S_1 = \{\omega_1, \omega_2\},\ S_2 = \{\omega_3, \omega_4\},$ and $S_3 = \{\omega_5, \omega_6\}$.

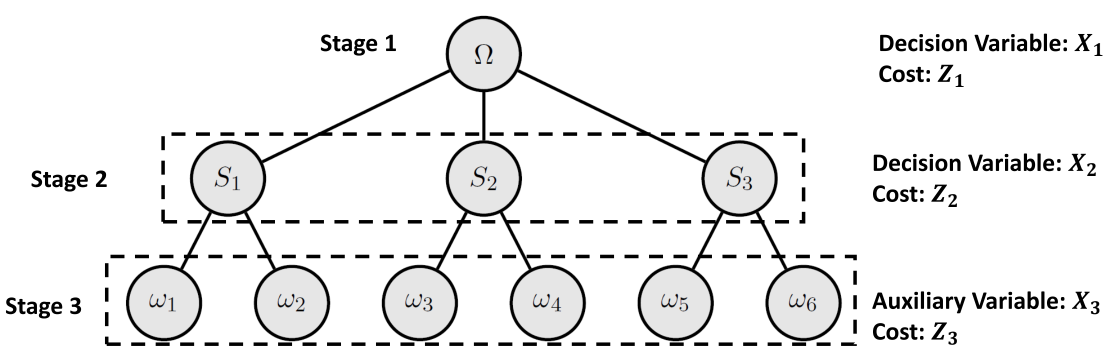

Figure 2.1: Sequence of Events.

In the first stage, the DM's objective is to minimize a coherent risk measure ρ_1 of the total cost. We also consider linear constraints at each time stage. Let A_{11}, A_{2j}^S, and A_{3k}^ω, for $j \in \{1,2\}$, $k \in \{1,2,3\}$, $S \in \mathcal{E}_2$, and $\omega \in \Omega$ be the coefficient matrices of the appropriate dimension and b_1, b_2^S, and b_3^ω be constant vectors of conforming size. The mathematical representation of the first-stage problem is as follows:

$$
\begin{aligned}
\min \quad & \rho_1(Z_1 + Z_2 + Z_3) & \\
\text{s.t.} \quad & A_{11}X_1 = b_1 & \\
& A_{21}^S X_1 + A_{22}^S X_2^S = b_2^S, & \forall S \in \mathcal{E}_2 \\
& A_{31}^\omega X_1 + A_{32}^\omega X_2^S + A_{33}^\omega X_3^\omega = b_3^\omega, & \forall S \in \mathcal{E}_2, \forall \omega \in S \\
& X_1, X_2^S, X_3^\omega \geq 0, & \forall S \in \mathcal{E}_2, \forall \omega \in S.
\end{aligned}
\tag{2.1}
$$

Let $\rho_2(Z_2 + Z_3|S)$ be the second-stage risk measure conditional on scenario $S \in \mathcal{E}_2$. The level-two recourse problem, parametrized by X_1, is formulated as:

$$
\begin{aligned}
\min \quad & \rho_2(Z_2 + Z_3|S) & \\
\text{s.t.} \quad & A_{22}^S X_2^S = b_2^S - A_{21}^S X_1 & \\
& A_{32}^\omega X_2^S + A_{33}^\omega X_3^\omega = b_3^\omega - A_{31}^\omega X_1 & \forall \omega \in S \\
& X_2^S, X_3^\omega \geq 0, & \forall \omega \in S.
\end{aligned}
\tag{2.2}
$$

Note that in the second stage the DM faces $|\mathcal{E}_2|$ recourse problems, one for each possible scenario S. The level-three recourse problem, parametrized by X_1 and X_2^S, with $\omega \in S$, is formulated as:

$$
\begin{aligned}
\min \quad & Z_3(\omega) & \\
\text{s.t.} \quad & A_{33}^\omega X_3^\omega = b_3^\omega - A_{31}^\omega X_1 - A_{32}^\omega X_2^S, & \forall \omega \in S \\
& X_3^\omega \geq 0, & \forall \omega \in S.
\end{aligned}
\tag{2.3}
$$

Next, we establish that these formulations are inconsistent. To see this, we first formally define risk measure time consistency and model time consistency properties.

Definition 3 *(**Risk measure time consistency** (Ruszczyński 2010)) A dynamic risk measure $\{\rho_1, \rho_2\}$ is time consistent if for all second-stage random variables Z_2 and W_2 with $\rho_2(Z_2|Z_1) \leq \rho_2(W_2|Z_1)$, we have $\rho_1(Z_1 + Z_2) \leq \rho_1(Z_1 + W_2)$ for all possible first-stage random variables Z_1.*

Risk measure time consistency implies that if a DM prefers Z_2 over W_2 in the second stage, this preference should stay the same in the first stage when the DM sets identical first-stage decision variable for those two second-stage courses of action.

Let $\mathcal{X}^*_{2,S}(X_1)$ be the set of optimal solutions to recourse problem (2.2). We have the following definition.

Definition 4 *(**Model time consistency**) Let (X_1^*, X_2^*, X_3^*) be a first-stage optimal solution. The system of problems* (2.1)-(2.2) *is **weakly time consistent** if for all $S_i \in \mathcal{E}_2$, there exists $\big((\bar{X}_2^{S_i})^*, (\bar{X}_3^{S_i})^*\big) \in \mathcal{X}^*_{2,S_i}(X_1^*)$ such that $(X_1^*, \bar{X}_2^*, \bar{X}_3^*)$ remains optimal for problem* (2.1)*, where $\bar{X}_j^* = \big((X_j^{S_1})^*, \cdots, (X_j^{S_{i-1}})^*, (\bar{X}_j^{S_i})^*, (X_j^{S_{i+1}})^*, \cdots, (X_j^{S_{\mathcal{E}_2}})^*\big)$ for $j = 2, 3$ (i.e., in the first-stage optimal solution one replaces $(X_2^{S_i})^*$ and $(X_3^{S_i})^*$ by $(\bar{X}_2^{S_i})^*$ and $(\bar{X}_3^{S_i})^*$, respectively).*

*If any optimal recourse solution can be substituted into the stage-one solution without affecting its optimality, then the system is **strongly time consistent**.*

Following Definition 4, weak time consistency requires that at least one of the solutions for the recourse program also appear as the optimal solution in the first-stage problem. The strong time consistency guarantees that all solutions of the recourse problem satisfy this property. In other words, the DM is assured that in the second stage, she does not have any incentive to renege on her optimal first-stage decisions.

From Shapiro (2012), one can show that except for the risk-neutral and worst-outcome risk functions, there always exist examples of the form (2.1)-(2.2) such that the problem lacks weak time consistency. Hence, adjustments to the problem formulation are required to ensure time consistency. A common way to enforce time consistency is to use risk measures of the nested form $\rho_1(Z_1 + \rho_2(Z_2 + Z_3))$, resulting in the formulation:

$$
\begin{array}{lll}
\min_{X_1,X_2,X_3} & \rho_1(Z_1 + \rho_2(Z_2 + Z_3)) & \\
\text{s.t.} & A_{11}X_1 = b_1 & \\
& A_{21}^S X_1 + A_{22}^S X_2^S = b_2^S, & \forall S \in \mathcal{E}_2 \\
& A_{31}^\omega X_1 + A_{32}^\omega X_2^S + A_{33}^\omega X_3^\omega = b_3^\omega, & \forall S \in \mathcal{E}_2, \forall \omega \in S \\
& X_1, X_2^S, X_3^\omega \geq 0, & \forall S \in \mathcal{E}_2, \forall \omega \in S.
\end{array}
\tag{2.4}
$$

Ruszczyński (2010) and Shapiro (2012) show that the system with this nested structure necessarily satisfies the time consistency property. We refer to this modeling approach as the objective time consistent (OTC) formulation. Despite the desirable properties of the OTC formulation, it can be shown that it lacks the law invariance property defined in Definition 2 (Shapiro 2012). In other words, if $Z_1 + Z_2 + Z_3$ and $W_1 + W_2 + W_3$ are the total cost functions with the same distribution, they can have different risk measures; i.e, $\rho_1(Z_1 + \rho_2(Z_2 + Z_3)) \neq \rho_1(W_1 + \rho_2(W_2 + W_3))$. Therefore, such an objective function and its properties might not be intuitive for the DM.

We propose an alternative formulation that is time consistent and law invariant for an arbitrary choice of the coherent risk measure. To achieve this, we add constraints to the first-stage problem to ensure the time consistency of the solution. Specifically, we add constraints $X_2^S, X_3^S \in \mathcal{X}^*_{2,S}(X_1),\ \forall S \in \mathcal{E}_2$ to problem (2.1) to formulate the problem as follows:

$$
\begin{aligned}
\min_{X_1,X_2,X_3} \quad & \rho_1(Z_1 + Z_2 + Z_3) \\
\text{s.t.} \quad & A_{11}X_1 = b_1, \\
& X_1 \geq 0, \\
& X_2^S, X_3^S \in \mathcal{X}^*_{2,S}(X_1),\ \forall S \in \mathcal{E}_2.
\end{aligned}
\tag{2.5}
$$

where $\mathcal{X}^*_{2,S}(X_1)$ is the set of optimal solutions to the following recourse problem:

$$
\begin{aligned}
\min_{X_2^S, X_3^S} \quad & \rho_2(Z_2 + Z_3|S) \\
\text{s.t.} \quad & A_{22}^S X_2^S = b_2^S - A_{21}^S X_1, \\
& A_{32}^\omega X_2^S + A_{33}^\omega X_3^\omega = b_3^\omega - A_{31}^\omega X_1, \quad \forall \omega \in S \\
& X_2^S, X_3^\omega \geq 0, \qquad \forall \omega \in S.
\end{aligned}
\tag{2.6}
$$

Note that the level-three recourse problem (2.3) is subsumed into the second-stage problem (2.6) since there is no possibility of time inconsistency between the second and third stages.

Intuitively, when the first-stage problem is being solved, the DM should keep in mind that the solution should also be subgame perfect (Osborne et al. 2004), i.e., be optimal for the second and third stages. By construction, the solution to the new bilevel program is weakly time consistent (Eckstein et al. 2016, Proposition 2). For a time-consistent risk measure such as expected value, can be shown that these constraints are redundant (Eckstein et al. 2016, Proposition 3). However, for a non-time-consistent risk measure, the constraints are necessary to ensure the time consistency of the problem, so we refer to our modeling approach as the constraint time consistent (CTC) formulation.

In Propositions 6 and 7 of Eckstein et al. (2016), we showed that the CTC formulation is $\mathcal{NP}$-hard, even if we restrict the risk measure to MSD and AVaR functions. If the OTC and CTC models result in close solutions, it will generally be preferable to use the computationally easier OTC formulation. However, in the next chapter, we numerically show that the solution to our formulation can significantly differ from the one to the OTC model. In the rest of this chapter, we present the generic CTC and OTC formulations with the MSD and AVaR risk measures.

2.3 Generic Formulation

In this section, we present the generic formulations of the CTC model (2.5)-(2.6) and OTC model (2.4) with both the MSD and AVaR risk measures. Let N_1 be the number of second-stage scenarios, and suppose that there are N_2 possible third-stage scenarios for each second-stage scenario. Further, assume the cost functions are linear functions and have the following formulation.

$$\begin{aligned} Z_1 &= c_1^T X_1, \\ Z_2^S &= (c_2^S)^T X_2^S, \quad \forall S \in \mathcal{E}_2 \\ Z_3^\omega &= (c_3^\omega)^T X_3^\omega, \quad \forall S \in \mathcal{E}_2, \forall \omega \in S. \end{aligned} \tag{2.7}$$

Define π_i to be the probability of second-stage scenario $i = 1, \ldots, N_1$, and $\pi_{i,j}$ to be the conditional probability of third-stage scenario (i, j), given the occurrence of second-stage scenario i, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$.

2.3.1 Problem Formulation with MSD Risk Measure

Suppose that MSD_{γ_1} and MSD_{γ_2} are the first-and second-stage risk measures, respectively. Table 2.1 gives the variables needed for MSD in both CTC and OTC model.

Table 2.1: **Variables needed for** MSD

$K_{i,j}$	Amount by which the combined stage 2 and 3 objective exceeds its overall mean in scenario (i,j), for $i = 1,\ldots,N_1$ and $j = 1,\ldots,N_2$
$Y_{i,j}$	Amount the stage-3 objective exceeds its stage-2 conditional mean in scenario (i,j), for $i = 1,\ldots,N_1$ and $j = 1,\ldots,N_2$
W_i	Second stage cost plus MSD value of third stage cost for OTC model, in second stage scenario i, for $i = 1,\cdots,N_1$
χ_i	Amount W_i exceeds its overall mean in for OTC model second-stage scenario i, for $i = 1,\cdots,N_1$

The first-stage objective function for the CTC model is:

$$\begin{aligned} Z_1 + \mathrm{MSD}_{\gamma_1}(Z_2 + Z_3) &= Z_1 + \mathbb{E}(Z_2 + Z_3) + \gamma_1 \mathbb{E}\big[Z_2 + Z_3 - \mathbb{E}(Z_2 + Z_3)\big]_+ \\ &= Z_1 + \textstyle\sum_{i=1}^{N_1} \pi_i Z_2^i + \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} Z_3^{i,j} \\ &+ \gamma_1 \textstyle\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} K_{i,j} \end{aligned}$$

where

$$\begin{aligned} &K_{i,j} \geq Z_2^i + Z_3^{i,j} - \big(\textstyle\sum_{i'=1}^{N_1} \pi_{i'} Z_2^{i'} + \sum_{i'=1}^{N_1} \sum_{j'=1}^{N_2} \pi_{i'} \pi_{i',j'} Z_3^{i',j'} \big), \ \forall i, \forall j \\ &K_{i,j} \geq 0, \ \forall i, \forall j \end{aligned}$$

and the corresponding second-stage objective function for scenario i is:

$$\begin{aligned} Z_2^i + \mathrm{MSD}_{\gamma_2}(Z_3|i) &= Z_2^i + \mathbb{E}(Z_3|i) + \gamma_2 \mathbb{E}\big(\big[Z_3 - \mathbb{E}(Z_3)\big]_+ |i\big) \\ &= Z_2^i + \textstyle\sum_{j=1}^{N_2} \pi_{i,j} Z_3^{i,j} + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j} \end{aligned}$$

where

$$\begin{aligned} &Y_{i,j} \geq Z_3^{i,j} - \textstyle\sum_{j'=1}^{N_2} \pi_{i,j'} Z_3^{i',j'}, \forall j \\ &Y_{i,j} \geq 0, \forall j. \end{aligned}$$

Above, "$\forall i$" is a shorthand for "$i = 1,\ldots,N_1$", and "$\forall j$" is a shorthand for "$j = 1,\ldots,N_2$." The first-and second-stage optimization problems of CTC model are:

First-stage problem for the CTC model

$$\begin{aligned}
\min_{X_1,X_2,X_3}\quad & c_1^T X_1 + \sum_{i=1}^{N_1}\pi_i (c_2^i)^T X_2^i + \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\pi_i\pi_{i,j}(c_3^{i,j})^T X_3^{i,j} + \gamma_1 \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\pi_i\pi_{i,j}K_{i,j} \\
\text{s.t.}\quad & A_{11}X_1 = b_1 \\
& K_{i,j} \geq (c_2^i)^T X_2^i + (c_3^{i,j})^T X_3^{i,j} - \Big(\sum_{i'=1}^{N_1}\pi_{i'}(c_2^{i'})^T X_2^{i'} \\
& \qquad + \sum_{i'=1}^{N_1}\sum_{j'=1}^{N_2}\pi_{i'}\pi_{i',j'}(c_3^{i',j'})^T X_3^{i',j'}\Big), \qquad \forall i, \forall j \\
& X_1 \geq 0,\ K_{i,j} \geq 0, \qquad \forall i, \forall j \\
& X_2^i, X_3^i \in \mathcal{X}_{2,i}^*(X_1), \qquad \forall i.
\end{aligned} \tag{2.8}$$

Second-stage problem, scenario i in the CTC model

$$\begin{aligned}
\min_{X_2^S,X_3^S}\quad & (c_2^i)^T X_2^i + \sum_{j=1}^{N_2}\pi_{i,j}(c_3^{i,j})^T X_3^{i,j} + \gamma_2 \sum_{j=1}^{N_2}\pi_{i,j}Y_{i,j} \\
\text{s.t.}\quad & A_{22}^i X_2^i = b_2^i - A_{21}^i X_1, \\
& A_{32}^{i,j} X_2^i + A_{33}^{i,j} X_3^{i,j} = b_3^{i,j} - A_{31}^{i,j} X_1, \qquad \forall j \\
& Y_{i,j} \geq (c_3^{i,j})^T X_3^{i,j} - \sum_{j'=1}^{N_2}\pi_{i,j'}(c_3^{i',j'})^T X_3^{i',j'}, \qquad \forall j \\
& Y_{i,j} \geq 0,\ X_2^i \geq 0,\ X_3^{i,j} \geq 0, \qquad \forall j.
\end{aligned} \tag{2.9}$$

If the decision maker uses a nested risk measure, the first-stage objective function is:

$$Z_1 + \text{MSD}_{\gamma_1}\big(Z_2 + \text{MSD}_{\gamma_2}(Z_3)\big) = Z_1 + \sum_{i=1}^{N_1}\pi_i W_i + \gamma_1 \sum_{i=1}^{N_1}\pi_i \chi_i,$$

where

$$\begin{aligned}
& \chi_i \geq W_i - \sum_{i'=1}^{N_1}\pi_{i'}W_{i'}, \qquad \forall i \\
& W_i = Z_2^i + \sum_{j=1}^{N_2}\pi_{i,j}Z_3^{i,j} + \gamma_2 \sum_{j=1}^{N_2}\pi_{i,j}Y_{i,j}, \quad \forall i \\
& Y_{i,j} \geq Z_3^{i,j} - \sum_{j'=1}^{N_2}\pi_{i,j'}Z_3^{i',j'}, \qquad \forall i, \forall j \\
& \chi_i \geq 0,\ Y_{i,j} \geq 0, \qquad \forall i, \forall j.
\end{aligned}$$

OTC model with nested MSD risk measure

$$
\begin{aligned}
\min_{X_1,X_2,X_3} \quad & c_1^T X_1 + \sum_{i=1}^{N_1} \pi_i W_i + \gamma_1 \sum_{i=1}^{N_1} \pi_i \chi_i \\
\text{s.t.} \quad & A_{11} X_1 = b_1 \\
& A_{21}^i X_1 + A_{22}^i X_2^i = b_2^i, && \forall i \\
& A_{31}^{i,j} X_1 + A_{32}^{i,j} X_2^i + A_{33}^{i,j} X_3^{i,j} = b_3^{i,j}, && \forall i, \forall j \\
& \chi_i \geq W_i - \sum_{i'=1}^{N_1} \pi_{i'} W_{i'}, && \forall i \\
& W_i = (c_2^i)^T X_2^i + \sum_{j=1}^{N_2} \pi_{i,j} (c_3^{i,j})^T X_3^{i,j} + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, && \forall i \\
& Y_{i,j} \geq (c_3^{i,j})^T X_3^{i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} (c_3^{i',j'})^T X_3^{i',j'}, && \forall i, \forall j \\
& X_1 \geq 0,\ X_2^i \geq 0,\ X_3^{i,j} \geq 0, \chi_i \geq 0,\ Y_{i,j} \geq 0, && \forall i, \forall j.
\end{aligned}
\tag{2.10}
$$

2.3.2 Problem Formulation with BRM Risk Measure

Let $\mathrm{BRM}_{\beta_1}^{\alpha_1}$ and $\mathrm{BRM}_{\beta_2}^{\alpha_2}$ be the first-and second-stage risk measures, respectively. Table 2.2 gives the variables needed for BRM in both CTC and OTC model.

Table 2.2: **Variables needed for BRM**

g	First-stage AVaR_{α_1} value for CTC model
h_i	AVaR_{α_2} value, in second-stage scenario $i = 1, \ldots, N_1$
t	Helper variable in AVaR_{α_1} definition in stage 1 for CTC model
$K_{i,j}$	Positive part of difference between stage 2 and 3's total cost in scenario (i, j) and t, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
$\bar{t}_i$	Helper variable in AVaR_{α_2} definition in stage 2 scenario i, for $i = 1, \ldots, N_1$
$Y_{i,j}$	Positive part of difference between stage 3's total cost in scenario (i, j) and $\bar{t}_i$, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
ω_i	Second stage cost plus $\mathrm{BRM}_{\beta_2}^{\alpha_2}$ value of third stage cost for OTC model, in second-stage scenario $i = 1, \cdots, N_1$

k	First-stage AVaR_{α_1} value of ω_i for OTC model
τ	Helper variable in AVaR_{α_1} definition in stage 1 for OTC model
χ_i	Positive part of difference between ω_i and τ, for $i = 1, \ldots, N_1$

The first-stage objective function for the CTC model is:

$$\begin{aligned} Z_1 + \text{BRM}_{\beta_1}^{\alpha_1}(Z_2 + Z_3) &= Z_1 + (1-\beta_1)\mathbb{E}(Z_2 + Z_3) + \beta_1 \,\text{AVaR}_{\alpha_1}(Z_2 + Z_3) \\ &= Z_1 + (1-\beta_1)\big(\sum_{i=1}^{N_1} \pi_i Z_2^i + \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \pi_i \pi_{i,j} Z_3^{i,j}\big) + \beta_1 g, \end{aligned}$$

where

$$\begin{aligned} & g \geq t + \tfrac{1}{\alpha_1} \textstyle\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} K_{i,j}, \\ & K_{i,j} \geq Z_2^i + Z_3^{i,j} - t, \ \forall i, \forall j \\ & K_{i,j} \geq 0, \ \forall i, \forall j, \end{aligned}$$

and the second-stage objective function for scenario $i = 1, \ldots, N_1$ is:

$$\begin{aligned} Z_2^i + \text{BRM}_{\beta_2}^{\alpha_2}(Z_3|i) &= Z_2^i + (1-\beta_2)\mathbb{E}(Z_3|i) + \beta_2 \,\text{AVaR}_{\alpha_2}(Z_3|i) \\ &= Z_2^i + (1-\beta_2) \textstyle\sum_{j=1}^{N_2} \pi_{i,j} Z_3^{i,j} + \beta_2 h_i, \end{aligned}$$

where

$$\begin{aligned} & h_i \geq \bar{t}_i + \tfrac{1}{\alpha_2} \textstyle\sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, \\ & Y_{i,j} \geq Z_3^{i,j} - \bar{t}_i, \ \forall j \\ & Y_{i,j} \geq 0, \ \forall j. \end{aligned}$$

Above, "$\forall i$" is a shorthand for "$i = 1, \ldots, N_1$", and "$\forall j$" is a shorthand for "$j = 1, \ldots, N_2$." The first-and second-stage optimization problems of CTC model are:

First-stage problem in CTC model

$$\begin{aligned} \min_{X_1, X_2, X_3} \quad & c_1^T X_1 + (1-\beta_1)\big(\sum_{i=1}^{N_1} \pi_i (c_2^i)^T X_2^i + \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \pi_i \pi_{i,j} (c_3^{i,j})^T X_3^{i,j}\big) + \beta_1 g \\ \text{s.t.} \quad & A_{11} X_1 = b_1 \\ & g \geq t + \frac{1}{\alpha_1} \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} K_{i,j}, \\ & K_{i,j} \geq (c_2^i)^T X_2^i + (c_3^{i,j})^T X_3^{i,j} - t, && \forall i, \forall j \\ & X_1 \geq 0, \ K_{i,j} \geq 0, && \forall i, \forall j \\ & X_2^i, X_3^i \in \mathcal{X}_{2,i}^*(X_1), && \forall i. \end{aligned} \tag{2.11}$$

Second-stage problem, scenario i in CTC model

$$\begin{aligned}
\min_{X_2^S, X_3^S} \quad & (c_2^i)^T X_2^i + (1-\beta_2) \sum_{j=1}^{N_2} \pi_{i,j} (c_3^{i,j})^T X_3^{i,j} + \beta_2 h_i \\
\text{s.t.} \quad & A_{22}^i X_2^i = b_2^i - A_{21}^i X_1, \\
& A_{32}^{i,j} X_2^i + A_{33}^{i,j} X_3^{i,j} = b_3^{i,j} - A_{31}^{i,j} X_1, && \forall j \\
& h_i \geq \bar{t}_i + \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, \\
& Y_{i,j} \geq (c_3^{i,j})^T X_3^{i,j} - \bar{t}_i, && \forall j \\
& Y_{i,j} \geq 0, \ X_2^i \geq 0, \ X_3^{i,j} \geq 0, && \forall j.
\end{aligned} \tag{2.12}$$

If the decision maker uses a nested risk measure, the first-stage objective function is:

$$Z_1 + \text{BRM}_{\beta_1}^{\alpha_1}\big(Z_2 + \text{BRM}_{\beta_2}^{\alpha_2}(Z_3)\big) = Z_1 + (1-\beta_1) \sum_{i=1}^{N_1} \pi_i \omega_i + \beta_1 k,$$

where

$$\begin{aligned}
& k \geq \tau + \frac{1}{\alpha_1} \sum_{i=1}^{N_1} \pi_i \chi_i \\
& \chi_i \geq \omega_i - \tau, && \forall i \\
& \omega_i = Z_2^i + (1-\beta_2) \sum_{j=1}^{N_2} \pi_{i,j} Z_3^{i,j} + \beta_2 h_i, && \forall i \\
& h_i \geq \bar{t}_i + \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, \\
& Y_{i,j} \geq Z_3^{i,j} - \bar{t}_i, && \forall j \\
& \chi_i \geq 0, \ Y_{i,j} \geq 0, && \forall i, \forall j.
\end{aligned}$$

The "OTC" formulation is:

OTC model with nested BRM Risk Measure

$$\begin{aligned}
\min_{X_1, X_2, X_3} \quad & c_1^T X_1 + (1-\beta_1) \sum_{i=1}^{N_1} \pi_i \omega_i + \beta_1 k \\
\text{s.t.} \quad & A_{11} X_1 = b_1 \\
& A_{21}^i X_1 + A_{22}^i X_2^i = b_2^i, && \forall i \\
& A_{31}^{i,j} X_1 + A_{32}^{i,j} X_2^i + A_{33}^{i,j} X_3^{i,j} = b_3^{i,j}, && \forall i, \forall j
\end{aligned} \tag{2.13}$$

$$
\begin{aligned}
& k \geq \tau + \frac{1}{\alpha_1}\sum_{i=1}^{N_1} \pi_i \chi_i \\
& \chi_i \geq \omega_i - \tau, && \forall i \\
& \omega_i = Z_2^i + (1-\beta_2)\sum_{j=1}^{N_2} \pi_{i,j} Z_3^{i,j} + \beta_2 h_i, && \forall i \\
& h_i \geq \bar{t}_i + \frac{1}{\alpha_2}\sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, && \forall i \\
& Y_{i,j} \geq Z_3^{i,j} - \bar{t}_i, && \forall i, \forall j \\
& X_1 \geq,\ X_2^i \geq 0,\ X_3^{i,j} \geq 0, \chi_i \geq 0,\ Y_{i,j} \geq 0, && \forall i, \forall j.
\end{aligned}
$$

In the next chapter, we present CTC and OTC models with MSD and BRM for three practical problems and show that the solutions to these two models can be significantly different in some applications.

Chapter 3

Practical Examples

In this chapter, we model three practical three-stage stochastic programs, namely supply chain production planning, portfolio optimization, and hydropower energy planning. For each problem, we consider two different risk measures, MSD_γ and BRM_β^α, and present their "CTC" and "OTC" formulations.

3.1 Supply Chain Production Planning

Consider the production planning problem described in Section 8 of Collado et al. (2012): a manufacturer produces P_2 different end products by configuring P_1 parts. Before the market demand is known, the manufacturer decides the order size of each part. In the second stage, the market demand realizes, and the manufacturer sets the production quantity of end products using the parts ordered in the first stage. For each unit of unsatisfied demand of product $k \in \{1, \ldots, P_2\}$, the manufacturer incurs "underage" cost l_k. We also assume that each unit of excess inventory can be salvaged at the end of the selling horizon; However, the company incurs inventory holding cost for unsold units. We further allow the holding cost to be stochastic in the first stage and depend on the realized scenario in the second stage.

3.1.1 Problem Formulation with BRM Risk Measure

We consider the case where the manufacturer's objective is to minimize the BRM_β^α of the total cost function. The risk-neutral and AVaR_α measures are the special cases of this model with $\beta = 0$ and $\beta = 1$, respectively. Tables 3.1.1, 3.2, and 3.3 summarize the problem parameters and variables.

Table 3.1: **Given Data**

P_1	Number of Parts
P_2	Number of products
c	Vector of part prices, in $\mathbb{R}^{P_1}$
z_{up}	Vector of maximum quantities of parts that may be ordered, in $\mathbb{R}^{P_1}$
π_i	Probability of second-stage scenario $i = 1, \ldots, N_1$
$\pi_{i,j}$	Conditional probability of third-stage scenario (i, j), given the occurrence of second-stage scenario i, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
r	Vector of revenues from products, in $\mathbb{R}^{P_2}$
l	Vector of unit shortfall costs for products, in $\mathbb{R}^{P_2}$
$H_{i,j}$	Vector of excess-production penalties for each product (in $\mathbb{R}^{P_2}$), in third-stage scenario (i, j), for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
M	$P_1 \times P_2$ matrix indicating how many units of each part are required to build one unit of each product
α_l	Risk aversion parameter for AVaR_{α_l} for stage $l = 1, 2$
β_l	Coefficient of AVaR_{α_l} in the risk measure for stage $l = 1, 2$
$1 - \beta_l$	Coefficient of expected value in the risk measure for stage $i = 1, l$

Table 3.2: **Fundamental model variables**

z	Quantities of parts ordered, in $\mathbb{R}^{P_1}$
g	First-stage AVaR_α value for CTC model
h_i	AVaR_α value, in second-stage scenario $i = 1, \ldots, N_1$
u_i	Vector of amounts of each product made (in $\mathbb{R}^{P_2}$), in second-stage scenario $i = 1, \ldots, N_1$

s_i^-	Vector of product shortfalls below demand (in $\mathbb{R}^{P_2}$), in second-stage scenario $i = 1, \ldots, N_1$
s_i^+	Vector of product units made in excess of demand (in $\mathbb{R}^{P_2}$), in second-stage scenario $i = 1, \ldots, N_1$
w_i	Second stage cost plus $BRM_{\beta_2}^{\alpha_2}$ value for OTC model, in second-stage scenario $i = 1, \ldots, N_1$
k	First-stage AVaR_α value of w for OTC model

Table 3.3: **Auxiliary variables needed for** AVaR

t	Helper variable in AVaR definition in stage 1 for CTC model
$X_{i,j}$	Positive part of difference between stage 2 and 3's total cost in scenario (i, j) and t, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
$\bar{t}_i$	Helper variable in AVaR definition in stage 2 scenario i, for $i = 1, \ldots, N_1$
$Y_{i,j}$	Positive part of difference between stage 3's total cost in scenario (i, j) and t_i, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
τ	Helper variable in AVaR definition in stage 1 for OTC model
χ_i	Positive part of difference between w_i and τ, for $i = 1, \ldots, N_1$

With this notation and the blended risk measure, the first-stage and second-stage problems may be formulated as (3.1) and (3.2).

First-stage problem

$$
\begin{aligned}
\min \quad & c^T z + (1-\beta_1)\sum_{i=1}^{N_1} \pi_i \left(-r^T u_i + l^T s_i^- + (\sum_{j=1}^{N_2} \pi_{i,j} H_{i,j}^T) s_i^+ \right) + \beta_1 g \\
\text{s.t.} \quad & z_{\text{up}} \geq z \geq 0 \\
& g \geq t + \frac{1}{\alpha_1} \sum_{i,j} \pi_i \pi_{i,j} X_{i,j}, \\
& X_{i,j} \geq -r^T u_i + l^T s_i^- + H_{i,j}^T s_i^+ - t,\ X_{i,j} \geq 0, & \forall i, \forall j \\
& (u_i, s_i^-, s_i^+) \in \mathcal{X}_i^*(z), & \forall i.
\end{aligned}
\tag{3.1}
$$

Second-stage problem, scenario i

$$\begin{aligned}
\min \quad & -r^T u_i + l^T s_i^- + (1-\beta_2)(\sum_j \pi_{i,j} H_{i,j}^T) s_i^+ + \beta_2 h_i \\
\text{s.t.} \quad & z \geq M u_i, \\
& s_i^+ \geq u_i - D_i, \ \ s_i^+ \geq 0, \\
& s_i^- \geq -u_i + D_i, \ \ s_i^- \geq 0, \\
& h_i \geq \bar{t}_i + \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, \\
& Y_{i,j} \geq H_{i,j}^T s_i^+ - \bar{t}_i, \ Y_{i,j} \geq 0, \qquad \forall j.
\end{aligned} \tag{3.2}$$

Since the second-stage problems (3.2) are linear, they satisfy standard constraint qualification conditions and one can replace the set of constraints $(u_i, s_i^-, s_i^+) \in \mathcal{X}_i^*(z)$ in (3.1) with KKT optimality conditions of problems (3.2) for each second-stage scenario. With this procedure, we convert the bilevel problem (3.1)-(3.2) into mathematical programming with equilibrium constraints (MPEC) (3.3).

Table 3.4: **Lagrange multiplier variables for second-stage problems**

λ_i	Vector of Lagrange multipliers for constraints $Z \geq M u_i$ below, for $i = 1, \ldots, N_1$
μ_i	Vector of Lagrange multipliers for constraints $s_i^+ \geq u_i - D_i$ below, for $i = 1, \ldots, N_1$
ϵ_i	Vector of Lagrange multipliers for constraints $s_i^- \geq -u_i + D_i$ below, for $i = 1, \ldots, N_1$
ρ_i	Scalar Lagrange multipliers for constraint $h_i \geq \bar{t}_i + \frac{1}{\alpha} \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}$ below, for $i = 1, \ldots, N_1$
$\delta_{i,j}$	Scalar Lagrange multipliers for constraint $Y_{i,j} \geq H_{i,j}^T s_i^+ - \bar{t}_i$ below, for $i = 1, \ldots, N_1, j = 1, \ldots, N_2$

CTC formulation of supply chain problem with BRM risk measure

$$
\begin{aligned}
\min \quad & c^T z + (1-\beta_1)\sum_{i=1}^{N_1}\pi_i\left(-r^T u_i + l^T s_i^- + (\sum_{j=1}^{N_2}\pi_{i,j}H_{i,j}^T)s_i^+\right) + \beta_1 g \\
\text{s.t.} \quad & z_{\text{up}} \geq z \geq 0 \\
& g \geq t + \frac{1}{\alpha_1}\sum_{i,j}\pi_i\pi_{i,j}X_{i,j} \\
& X_{i,j} \geq -r^T u_i + l^T s_i^- + H_{i,j}^T s_i^+ - t, && \forall i, \forall j \\
& z \geq M u_i, && \forall i \\
& s_i^+ \geq u_i - D_i, && \forall i \\
& s_i^- \geq -u_i + D_i, && \forall i \\
& h_i \geq \bar{t}_i + \frac{1}{\alpha_2}\sum_j \pi_{i,j}Y_{i,j}, && \forall i, \forall j \\
& Y_{i,j} \geq H_{i,j}^T s_i^+ - \bar{t}_i, && \forall i, \forall j \\
& -r^T + \lambda_i^T M + \mu_i^T - \epsilon_i^T \geq 0, && \forall i \\
& l^T - \epsilon_i^T \geq 0, && \forall i \\
& -\mu_i^T + \sum_{j=1}^{N_i}\Big((1-\beta_2)\pi_{i,j} + \delta_{i,j}\Big)H_{i,j}^T \geq 0, && \forall i, \forall j \\
& -\delta_{i,j} + \frac{\rho_i}{\alpha_2}\pi_{i,j} \geq 0, && \forall i, \forall j \\
& \beta_2 - \rho_i = 0, && \forall i \\
& -\sum_{j=1}^{N_i}\delta_{i,j} + \rho_i = 0, && \forall i \\
& \lambda_i^T(z - M u_i) = 0, && \forall i \\
& \mu_i^T(s_i^+ - u_i + D_i) = 0, && \forall i \\
& \epsilon_i^T(s_i^- + u_i - D_i) = 0, && \forall i \\
& \delta_{i,j}(Y_{i,j} - H_{i,j}^T s_i^+ + \bar{t}_i) = 0, && \forall i, \forall j \\
& \rho_i(h_i - \bar{t}_i - \frac{1}{\alpha_2}\sum_j \pi_{i,j}Y_{i,j}) = 0, && \forall i, \forall j \\
& (-r^T + \lambda_i^T M + \mu_i^T - \epsilon_i^T)u_i = 0, && \forall i
\end{aligned}
\tag{3.3}
$$

$$
\begin{aligned}
&(l^T - \epsilon_i^T)s_i^- = 0, && \forall i \\
&(-\mu_i^T + \sum_{j=1}^{N_i}((1-\beta_2)\pi_{i,j} + \delta_{i,j})H_{i,j}^T)s_i^+ = 0, && \forall i, \forall j \\
&(-\delta_{i,j} + \frac{\rho_i}{\alpha_2}\pi_{i,j})Y_{i,j} = 0, && \forall i, \forall j \\
&u_i, s_i^-, s_i^+, \lambda_i, \mu_i, \epsilon_i, \rho_i \geq 0, && \forall i \\
&X_{i,j} \geq 0, Y_{i,j} \geq 0, \delta_{i,j} \geq 0, && \forall i, \forall j.
\end{aligned}
$$

Above, "$\forall i$" is a shorthand for "$i = 1, \ldots, N_1$", and "$\forall j$" is a shorthand for "$j = 1, \ldots, N_2$." Using nested risk measure $\rho_1(Z_1 + \rho_2(Z_2 + Z_3))$, we also formulate the "OTC" version of this problem in (3.4).

OTC formulation of supply chain problem with nested BRM risk measure

$$
\begin{aligned}
\min \quad & c^T z + (1-\beta_1)\sum_{i=1} \pi_i w_i + \beta_1 k \\
\text{s.t.} \quad & z_{\text{up}} \geq z \geq 0 \\
& z \geq M u_i, && \forall i \\
& s_i^+ \geq u_i - D_i, && \forall i \\
& s_i^- \geq -u_i + D_i, && \forall i \\
& w_i = l^T s_i^- - r^T u_i + (1-\beta_2)(\sum_{j=1}^{N_i} \pi_{i,j} H_{i,j}^T s_i^+) + \beta_2 h_i, && \forall i \\
& k \geq \tau + \frac{1}{\alpha_1}\sum_i \pi_i \chi_i, && \forall i \\
& \chi_i \geq w_i - t, && \forall i \\
& Y_{i,j} \geq H_{i,j}^T s_i^+ - \bar{t}_i, && \forall i \\
& h_i \geq \bar{t}_i + \frac{1}{\alpha_2}\sum_{j=1}^{N_i} \pi_{i,j} Y_{i,j}, && \forall i, \forall j \\
& \chi_i, Y_{i,j}, u_i, s_i^+, s_i^- \geq 0, && \forall i, \forall j.
\end{aligned}
\tag{3.4}
$$

3.1.2 Problem Formulation with MSD Risk Measure

We follow the same procedure to formulate the CTC and OTC models of this problem with risk measure MSD_γ. New parameters and variables required to formulate these models are presented in Tables 3.5 and 3.6. Also, Tables 3.7 and 3.8 summarize the auxiliary variables and Lagrange multiplier variables needed for MSD risk measure, respectively.

Table 3.5: **Given Data**

γ_i	Risk aversion parameter for MSD_{γ_l} risk measure for stage $l = 1, 2$

Table 3.6: **Fundamental model variables**

V_i	Second stage cost plus MSD_{γ_2} value for OTC model, in second-stage scenario $i = 1, \ldots, N_1$

Table 3.7: **Auxiliary variables needed for** MSD

$X_{i,j}$	Amount combined stage 2 and 3 objective exceeds its overall mean in scenario (i, j), for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
$Y_{i,j}$	Amount stage 3 objective exceeds its stage 2 conditional mean in scenario (i, j), for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
χ_i	Amount V_i exceeds its overall mean for OTC model in second stage scenario i, for $i = 1, \ldots, N_1$

Table 3.8: **Lagrange multiplier variables** MSD

λ_i	Vector of Lagrange multipliers for constraints $z \geq Mu_i$ below, for $i = 1, \ldots, N_1$
ξ_i	Vector of Lagrange multipliers for constraints $s_i^- \geq -u_i + D_i$ below, for $i = 1, \ldots, N_1$
μ_i	Vector of Lagrange multipliers for constraints $s_i^+ \geq u_i - D_i$ below, for $i = 1, \ldots, N_1$

$\delta_{i,j}$ Scalar Lagrange multiplier for each constraint $Y_{i,j} \geq \left(H_{i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} H_{i,j'}\right)^T s_i^+$ below, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$

CTC formulation of supply chain problem with MSD risk measure

$$
\begin{aligned}
\min \quad & c^T z + \sum_{i=1}^{N_1} \pi_i \left(-r^T u_i + l^T s_i^- + \left(\sum_{j=1}^{N_2} \pi_{i,j} H_{i,j}^T \right) s_i^+ \right) + \gamma_1 \sum_{i=1}^{N_1} \pi_i \left(\sum_{j=1}^{N_2} \pi_{i,j} X_{i,j} \right) \\
\text{s.t.} \quad & 0 \leq z \leq z_{\text{up}} \\
& X_{i,j} \geq -r^T u_i + l^T s_i^- + H_{i,j}^T s_i^+ \\
& \qquad + \sum_{i'=1}^{N_1} \pi_i' \left(-r^T u_{i'} + l^T s_{i'}^- + \left(\sum_{j'=1}^{N_1} \pi_{i',j'} H_{i',j'}^T \right) s_{i'}^+ \right), && \forall i, \forall j \\
& z \geq M u_i, && \forall i \\
& s_i^+ \geq u_i - D_i, && \forall i \\
& s_i^- \geq -u_i + D_i, && \forall i \\
& Y_{i,j} \geq \left(H_{i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} H_{i,j'} \right)^T s_i^+, && \forall i, \forall j \\
& M^T \lambda_i + \mu_i - \xi_i \geq r, && \forall i \\
& \xi_i^T \leq l, && \forall i \\
& -\mu_i + \sum_{j=1}^{N_2} \left(\pi_{i,j} \left(1 - \sum_{j'=1}^{N_2} \delta_{i,j'} \right) + \delta_{i,j} \right) H_{i,j} \geq 0, && \forall i \\
& \delta_{i,j} \leq \gamma_2 \pi_{i,j}, && \forall i, \forall j \\
& \lambda_i^T (z - M u_i) = 0, && \forall i \\
& \mu_i^T (s_i^+ - u_i + D_i) = 0, && \forall i \\
& \xi_i^T (s_i^- + u_i - D_i) = 0, && \forall i \\
& \delta_{i,j} \left(Y_{i,j} - H_{i,j}^T s_i^+ + \sum_{j'=1}^{N_2} \pi_{i,j'} H_{i,j'}^T s_i^+ \right) = 0, && \forall i, \forall j \\
& (-r + M^T \lambda_i + \mu_i - \xi_i)^T u_i = 0, && \forall i
\end{aligned}
\tag{3.5}
$$

$$(l-\xi_i)^T s_i^- = 0, \qquad \forall i$$

$$\left(-\mu_i + \sum_{j=1}^{N_2}\left(\pi_{i,j}\big(1-\sum_{j'=1}^{N_2}\delta_{i,j'}\big)+\delta_{i,j}\right)H_{i,j}\right)^T s_i^+ = 0, \qquad \forall i$$

$$(\gamma_2\pi_{i,j}-\delta_{i,j})Y_{i,j} = 0, \qquad \forall i$$

$$u_i, s_i^-, s_i^+, \lambda_i, \mu_i, \xi_i \geq 0, \qquad \forall i$$

$$X_{i,j}, Y_{i,j}, \lambda_i, \delta_{i,j} \geq 0, \qquad \forall i, \forall j$$

OTC formulation of supply chain problem with nested MSD risk measure

$$\begin{aligned}
\min \quad & c^T z + \sum_{i=1}^{N_1}\pi_i V_i + \gamma_1\Big(\sum_{i=1}^{N_1}\pi_i\chi_i\Big) \\
\text{s.t.} \quad & z_{\text{up}} \geq z \geq 0 \\
& z \geq M u_i, && \forall i \\
& s_i^+ \geq u_i - D_i, && \forall i \\
& s_i^- \geq -u_i + D_i, && \forall i \\
& \chi_i \geq V_i - \sum_{i=1}^{N_1}\pi_i V_i, && \forall i \\
& V_i = l^T s_i^- - r^T u_i + \sum_{j=1}^{N_2}\pi_{i,j}(H_{i,j}^T s_i^+) + \gamma_2\Big(\sum_{'=1}^{N_2}\pi_{i,j}Y_{i,j}\Big), && \forall i \\
& Y_{i,j} \geq H_{i,j}^T s_i^+ - \sum_{j'=1}^{N_2}\pi_{i,j'}H_{i,j'}^T s_i^+, && \forall i, \forall j \\
& \chi_i, Y_{i,j}, u_i, s_i^+, s_i^- \geq 0, && \forall i, \forall j.
\end{aligned} \tag{3.6}$$

3.1.3 Numerical Study

The numerical study is comprised of $5,000$ randomly generated instances with 5 parts, 5 end products, and 5 possible second-stage demand scenarios. In each second-stage scenario, we also consider 5 possible outcomes for the holding cost in the third stage.We repeated the numerical study for each risk measure. To study various degrees of risk aversion, we consider parameter values $\gamma = 0.3$, 0.7, and 0.9 for MSD_γ and $(\alpha,\beta) = (0.01, 0.5)$, $(0.05, 0.5)$, $(0.25, 0.5)$, $(0.05, 0.25)$, and $(0.25, 0.75)$ for BRM_β^α.

We formulated each optimization problem in AMPL and solved it using Gurobi. SOS1 constraints[1] were used to model the CTC's complementarity constraints. Note that all optimal solutions of "OTC" are feasible for "CTC." Therefore, we can plug the solution of OTC into the CTC's objective function to obtain an upper bound for CTC. Let Obj^i and Sol^i, be the objective function and optimal solution of model $i \in \{\mathrm{OTC}, \mathrm{CTC}\}$, respectively. To measure the deviation of the CTC model from the OTC formulation, define the relative difference of the two models as

$$\mathrm{Diff} = \frac{\mathrm{Obj}^{\mathrm{CTC}}(\mathrm{Sol}^{\mathrm{OTC}}) - \mathrm{Obj}^{\mathrm{CTC}}(\mathrm{Sol}^{\mathrm{CTC}})}{\mathrm{Obj}^{\mathrm{CTC}}(\mathrm{Sol}^{\mathrm{OTC}})}.$$

A larger difference translates to a worse solution suggested by the OTC approach and justifies the computational costs of the CTC model. Also, let Δ_1 (Δ_2) denote the maximum percentage difference between the first-stage (second-stage) optimal solutions of the OTC and CTC problems. Consequently, a higher value of Δ_i corresponds to a higher deviation between the two models.

Tables 3.9 and 3.10 summarize the results for risk measures MSD_γ and BRM_β^α, respectively. The first five rows report the percentage of instances with a Diff value in the specified range. For example, 22.76% of the instances resulted in a Diff value between 0.1% and 1% for risk measure $\mathrm{MSD}_{0.3}$ as illustrated in Table 3.9. Row 6 reports the maximum Diff value observed among all the numerical instances.

We observe that although the solutions of OTC and CTC can potentially match in some cases (the first row), in a significant fraction of instances, they could lead to dramatically different solutions: in 7.12% of the instances with $\gamma = .9$, we observed a gap higher than 10%. We also encountered gaps as large as 1,094,395% as a result of $\mathrm{Obj}^{\mathrm{CTC}}(\mathrm{Sol}^{\mathrm{OTC}}) = -0.0016$, and $\mathrm{Obj}^{\mathrm{CTC}} = -17.4291$ for an instance with $\gamma = 0.7$, suggesting that the two formulations can differ from one another by arbitrarily large amounts. We also see that the fraction of the instances with relatively large gaps (Diff $\geq 10\%$) increases as a function of γ, i.e., for higher degrees of risk aversion.

Interestingly, Table 3.10 shows that the solution to the OTC and CTC formulations

[1] SOS1 (special ordered set type 1) constraints enforce that at most one of a set of decision variables can be nonzero. In this case, the SOS1 sets each consisted of two variables.

	γ =0.3	γ =0.7	γ =0.9
Diff < 0.01%	54.42%	20.90%	13.56%
0.01% ≤ Diff < 0.1%	16.76%	12.28%	8.56%
0.1% ≤ Diff < 1%	22.76%	37.32%	35.40%
1% ≤ Diff < 10%	5.50%	25.18%	35.36%
10% ≤ Diff	0.56%	4.32%	7.12%
Max Diff	3,333.66%	1,094,395%	2,596.47%
Δ_1	8.84%	4.20%	6.33%
Δ_2	19.18%	13.31%	13.92%

Table 3.9: Computational results of the supply chain problem with the MSD risk measure

	$\beta = 0.5$ $\alpha = 0.01$	$\beta = 0.5$ $\alpha = 0.05$	$\beta = 0.75$ $\alpha = 0.05$	$\beta = 0.5$ $\alpha = 0.25$	$\beta = 0.75$ $\alpha = 0.25$
Diff < 0.01%	5.32%	5.30%	4.16%	6.84%	16.10%
0.01% ≤ Diff < 0.1%	2%	2.66%	1.76%	5.00%	2.62%
0.1% ≤ Diff < 1%	14.34%	16.98%	15.44%	26.42%	15.10%
1% ≤ Diff < 10%	50.00%	51.62%	52.80%	46.42%	42.86%
10% ≤ Diff	28.36%	23.44%	25.84%	15.32%	23.32%
Max Diff	122,774%	51,609.07%	5,627.64%	21,819.79%	4,619,215%
Δ_1	1.66%	2.99%	3.04%	11.59%	8.72%
Δ_2	3.74%	58.51%	10.44%	14.48%	24.48%

Table 3.10: Computational results of the supply chain problem with the BRM risk measure.

significantly differ in an even larger fraction of the instances under the BRM than the MSD risk measure. Also, similar to MSD, we again observe that the gaps between the OTC and CTC models are larger for higher degrees of risk aversion.

To see if the CTC model can properly be scaled, we randomly generated 300 instances with 10 parts, 10 end products, and 10 possible second-stage demand scenarios, each leading to 10 third-stage scenarios. In this experiment, we considered the $\text{BRM}_{0.75}^{0.05}$ risk measure. OTC model was able to solve all instances within 1s with an average time of 0.05s. However, CTC model could not be solved in 174 of the instances within 900s. Furthermore, the average solution time for the rest of the instances was 108.5s, which is more than 2000 times greater than the OTC's average solving time.

In sum, it appears that the solution to the OTC model can be poor approximation of the CTC formulation for the supply chain problem, particularly for more risk-averse DMs. Next, we compare the performance of the two formulations for another class of problems.

3.2 Portfolio Optimization Problem

In this section, we study the portfolio optimization example studied in Gülten & Ruszczyński (2015), where a risk-averse investor allocates a fixed budget to a portfolio of available assets to maximize the value of the portfolio over three stages. The return rates of the assets are unknown in each stage and realize after the investment decisions are made. In the first period, the investor buys a portfolio of assets. In the second stage, in addition to the option of investing in new assets, the investor can also sell or buy more of the owned assets. At the end of the third stage, uncertainty resolves and the investor earns the value of the portfolio.

3.2.1 Problem Formulation with MSD Risk Measure

We consider the case where the investor's objective is to maximize the MSD_γ of the portfolio value. Tables 3.11, 3.12, and 3.13 summarize the problem parameters and variables for portfolio problem with MSD risk measure.

Table 3.11: **Given Data**

r	Number of assets

R_i^1	Vector of return rates in second-stage scenario i in $\mathbb{R}^r$, for $i = 1, \ldots, N_1$
$R_{i,j}^2$	Vector of return rates in third-stage scenario (i, j) in $\mathbb{R}^r$, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
π_i	Probability of second-stage scenario $i = 1, \ldots, N_1$
$\pi_{i,j}$	Conditional probability of third-stage scenario (i, j), given the occurrence of second-stage scenario i, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
ϵ	Vector of relative transaction cost of securities in $\mathbb{R}^r$
γ_i	Risk aversion parameter for MSD_{γ_l} risk measure for stage $l = 1, 2$

Table 3.12: **Fundamental model variables**

z	Vector of investment assets money value in first stage, in $\mathbb{R}^r$
y_i	Vector of investment assets money value in second-stage scenario i, in $\mathbb{R}^r$, for $i = 1, \ldots, N_1$
s_i	Money value vector of sold assets in second-stage scenario i, in $\mathbb{R}^r$, for $i = 1, \ldots, N_1$
b_i	Money value vector of bought assets in second-stage scenario i, in $\mathbb{R}^r$, for $i = 1, \ldots, N_1$
$W_{i,j}$	Scalar total value of portfolio in third-stage scenario (i, j), for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
V_i	MSD value of $W_{i,j}$ for OTC model, in second-stage scenario i, for $i = 1, \ldots, N_1$

Table 3.13: **Auxiliary variables needed for** MSD

$X_{i,j}$	Amount $W_{i,j}$ is under its overall mean in scenario (i, j), for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
$Y_{i,j}$	Amount $W_{i,j}$ is under its conditional mean in scenario (i, j), for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$

χ_i Amount V_i exceeds its overall mean for OTC model in second-stage scenario i, for $i = 1, \ldots, N_1$

If we consider MSD_γ as our risk measure, the first-stage objective function is:

$$\begin{aligned}\text{MSD}_{\gamma_1}(Z_3) &= \mathbb{E}(Z_3) + \gamma_1 \mathbb{E}\big[\big(Z_3 - \mathbb{E}(Z_3)\big)_+\big] = -\mathbb{E}(W) + \gamma_1 \mathbb{E}\big[\big(\mathbb{E}(W) - W\big)_+\big] \\ &= -\textstyle\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \pi_i \pi_{i,j} W_{i,j} + \gamma_1 \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j}\end{aligned}$$

where

$$\begin{aligned}&X_{i,j} \geq \textstyle\sum_{i'=1}^{N_1}\sum_{j'=1}^{N_2} \pi_{i'} \pi_{i',j'} W_{i',j'} - W_{i,j}, i = 1, \ldots, N_1, \quad j = 1, \ldots, N_2 \\ &X_{i,j} \geq 0, \qquad i = 1, \ldots, N_1, \quad j = 1, \ldots, N_2\end{aligned}$$

and the second-stage scenario i objective function is:

$$\begin{aligned}\text{MSD}_{\gamma_2}(Z_3|i) &= \mathbb{E}(Z_3|i) + \gamma_2 \mathbb{E}\big[\big(Z_3 - \mathbb{E}(Z_3)\big)_+|i\big] = -\mathbb{E}(W|i) + \gamma_2 \mathbb{E}\big[\big(\mathbb{E}(W) - W\big)_+|i\big] \\ &= -\textstyle\sum_{j=1}^{N_2} \pi_{i,j} W_{i,j} + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}\end{aligned}$$

where

$$\begin{aligned}&Y_{i,j} \geq \textstyle\sum_{j'=1}^{N_2} \pi_{i,j'} W_{i,j'} - W_{i,j}, \quad j = 1, \ldots, N_2 \\ &Y_{i,j} \geq 0, \qquad j = 1, \ldots, N_2.\end{aligned}$$

Suppose, DM has one dollar in the first stage and wants to invest it in r assets. The first and second-stage optimization problems are as follow:

First-stage problem

$$\begin{aligned}\min \quad & -\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \pi_i \pi_{i,j} W_{i,j} + \gamma_1 \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j} \\ \text{s.t.} \quad & z^T \mathbf{1} = 1 \\ & X_{i,j} \geq \sum_{i'=1}^{N_1}\sum_{j'=1}^{N_2} \pi_{i'} \pi_{i',j'} W_{i',j'} - W_{i,j}, \qquad \forall i, \forall j \\ & z \geq \mathbf{0}, \ X_{i,j} \geq 0, \qquad \forall i, \forall j \\ & W_i \in \mathcal{X}_i^*(z), \qquad \forall i.\end{aligned} \tag{3.7}$$

Second-stage problem, scenario i:

$$
\begin{aligned}
\min \quad & -\sum_{j=1}^{N_2} \pi_{i,j} W_{i,j} + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j} \\
\text{s.t.} \quad & y_i = (\mathbf{1} + R_i^1) \circ z + (\mathbf{1} - \epsilon) \circ b_i - (\mathbf{1} + \epsilon) \circ s_i, \\
& b_i^T \mathbf{1} - s_i^T \mathbf{1} = 0, \\
& W_{i,j} = (\mathbf{1} + R_{i,j}^2)^T y_i, && \forall j \\
& Y_{i,j} \geq \sum_{j'=1}^{N_2} \pi_{i,j'} W_{i,j'} - W_{i,j} && \forall j \\
& y_i \geq \mathbf{0},\ b_i \geq \mathbf{0},\ s_i \geq \mathbf{0},\ Y_{i,j} \geq 0, && \forall j.
\end{aligned}
\tag{3.8}
$$

In the above problem, "$\circ$" is Hadamard product, element-wise product of two matrices with the same dimension. Also, $\mathbf{1}$ is a vector of all ones with appropriate dimension. Now, define the Lagrange multiplier for the second-stage problems to use them in converting the bilevel linear problem to a MPEC formulation.

Table 3.14: **Lagrange multiplier variables for MPEC formulation**

λ_i	Vector of Lagrange multipliers for constraints $y_i - (\mathbf{1} + R_i^1) \circ z - (\mathbf{1} - \epsilon) \circ b_i + (\mathbf{1} + \epsilon) \circ s_i = 0$, for $i = 1, \ldots, N_1$
μ_i	Scalar Lagrange multiplier for constraints $b_i^T \mathbf{1} - s_i^T \mathbf{1} = 0$, for $i = 1, \ldots, N_1$
$\rho_{i,j}$	Scalar Lagrange multiplier for constraints $W_{i,j} - (\mathbf{1} + R_{i,j}^2)^T y_i = 0$, for $i = 1, \ldots, N_1,\ j = 1, \ldots, N_2$
$\sigma_{i,j}$	Scalar Lagrange multiplier for constraints $Y_{i,j} \geq \sum_{j'=1}^{N_2} \pi_{i,j'} W_{i,j'} - W_{i,j}$ below, for $i = 1, \ldots, N_1,\ j = 1, \ldots, N_2$

The second-stage Lagrange function $L_i(y_i, b_i, s_i, W_{i,j}, Y_{i,j}, \lambda_i, \mu_i, \rho_{i,j}, \sigma_{i,j})$ is:

$$
\begin{aligned}
L_i = \quad & -\sum_{j=1}^{N_2} \pi_{i,j} W_{i,j} + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j} - \lambda_i^T \big(y_i - (\mathbf{1} + R_i^1) \circ z - (\mathbf{1} - \epsilon) \circ b_i \\
& + (\mathbf{1} + \epsilon) \circ s_i\big) - \mu_i (b_i^T \mathbf{1} - s_i^T \mathbf{1}) - \sum_{j=1}^{N_2} \rho_{i,j} (W_{i,j} - (\mathbf{1} + R_{i,j}^2)^T y_i) \\
& - \sum_{j=1}^{N_2} \sigma_{i,j} (Y_{i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} W_{i,j'} + W_{i,j}) = y_i^T (-\lambda_i + \sum_{j=1}^{N_2} \rho_{i,j} (\mathbf{1} + R_{i,j}^2)) \\
& + b_i^T \big((1 - \epsilon) \circ \lambda_i - \mu_i \mathbf{1}\big) + s_i^T \big(- (1 + \epsilon) \circ \lambda_i + \mu_i \mathbf{1}\big) \\
& + \sum_{j=1}^{N_2} W_{i,j} (-\pi_{i,j} - \rho_{i,j} - \sigma_{i,j} + \pi_{i,j} \sum_{j'=1}^{N_2} \sigma_{i,j'}) + \sum_{j=1}^{N_2} Y_{i,j} (\gamma_2 \pi_{i,j} - \sigma_{i,j}) \\
& + \lambda_i^T (\mathbf{1} + R_i^1) \circ z
\end{aligned}
$$

By replacing constraints $W_i \in \mathcal{X}_i^*(z)$ in (3.7) with KKT optimality conditions of (3.8) for each second-stage scenario $i = 1, \ldots, N_1$, the "CTC" primal problem is:

CTC formulation of portfolio problem with MSD_γ risk measure

$$
\begin{aligned}
\min \quad & -\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \pi_i \pi_{i,j} W_{i,j} + \gamma_1 \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j} && \\
\text{s.t.} \quad & z^T \mathbf{1} = 1,\ z \geq \mathbf{0}, && \forall i \\
& X_{i,j} \geq \sum_{i'=1}^{N_1}\sum_{j'=1}^{N_2} \pi_{i'} \pi_{i',j'} W_{i',j'} - W_{i,j},\ X_{i,j} \geq 0, && \forall i, \forall j \\
& y_i = (\mathbf{1} + R_i^1) \circ z + (\mathbf{1} - \epsilon) \circ b_i - (\mathbf{1} + \epsilon) \circ s_i, && \forall i \\
& b_i^T \mathbf{1} - s_i^T \mathbf{1} = 0, && \forall i \\
& W_{i,j} = (\mathbf{1} + R_{i,j}^2)^T y_i, && \forall i, \forall j \\
& Y_{i,j} \geq \sum_{j'=1}^{N_2} \pi_{i,j'} W_{i,j'} - W_{i,j}, && \forall i, \forall j \\
& -\lambda_i + \sum_{j=1}^{N_2} \rho_{i,j} (\mathbf{1} + R_{i,j}^2) \geq \mathbf{0}, && \forall i \\
& (1 - \epsilon) \circ \lambda_i - \mu_i \mathbf{1} \geq \mathbf{0}, && \forall i \qquad (3.9) \\
& -(1 + \epsilon) \circ \lambda_i + \mu_i \mathbf{1} \geq \mathbf{0}, && \forall i \\
& -\pi_{i,j} - \rho_{i,j} - \sigma_{i,j} + \pi_{i,j} \sum_{j'=1}^{N_2} \sigma_{i,j'} = 0, && \forall i, \forall j \\
& \gamma_2 \pi_{i,j} - \sigma_{i,j} \geq 0, && \forall i, \forall j \\
& \sigma_{i,j} (Y_{i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} W_{i,j'} + W_{i,j}) = 0, && \forall i, \forall j \\
& y_i^T (-\lambda_i + \sum_{j=1}^{N_2} \rho_{i,j} (\mathbf{1} + R_{i,j}^2)) = 0, && \forall i \\
& b_i^T \big((1 - \epsilon) \circ \lambda_i - \mu_i \mathbf{1} \big) = \mathbf{0}, && \forall i \\
& s_i^T \big(-(1 + \epsilon) \circ \lambda_i + \mu_i \mathbf{1} \big) = \mathbf{0}, && \forall i \\
& Y_{i,j} (\gamma_2 \pi_{i,j} - \sigma_{i,j}) = 0, && \forall i, \forall j \\
& y_i \geq \mathbf{0},\ b_i \geq \mathbf{0},\ s_i \geq \mathbf{0},\ Y_{i,j} \geq 0,\ \sigma_{i,j} \geq 0, && \forall i, \forall j.
\end{aligned}
$$

Above, "$\forall i$" is a shorthand for "$i = 1, \ldots, N_1$", and "$\forall j$" is a shorthand for "$j = 1, \ldots, N_2$." If the decision maker uses a nested risk measure, $\rho_1(\rho_2(Z_3))$, for the first-stage objective function, the "OTC" formulation is:

OTC formulation of portfolio problem with nested MSD risk measure

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{N_1} \pi_i V_i + \gamma_1 (\sum_{i=1}^{N_1} \pi_i \chi_i) \\
\text{s.t.} \quad & z^T \mathbf{1} = 1 \\
& y_i = (\mathbf{1} + R_i^1) \circ z + (\mathbf{1} - \epsilon) \circ b_i - (\mathbf{1} + \epsilon) \circ s_i, && \forall i \\
& b_i^T \mathbf{1} - s_i^T \mathbf{1} = 0, && \forall i \\
& W_{i,j} = (\mathbf{1} + R_{i,j}^2)^T y_i, && \forall i, \forall j \\
& Y_{i,j} \geq \sum_{j'=1}^{N_2} \pi_{i,j'} W_{i,j'} - W_{i,j}, && \forall i, \forall j \\
& V_i = -(\sum_{j=1}^{N_2} \pi_{i,j} W_{i,j}) + \gamma_2 (\sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}), && \forall i \\
& \chi_i \geq V_i - \sum_{i=1}^{N_1} \pi_i V_i, && \forall i \\
& z \geq \mathbf{0},\ b_i \geq \mathbf{0},\ d_i \geq \mathbf{0},\ \chi_i \geq 0,\ Y_{i,j} \geq 0, && \forall i, \forall j.
\end{aligned}
\tag{3.10}
$$

3.2.2 Problem Formulation with BRM Risk Measure

We also follow the same procedure to formulate the CTC and OTC models of this problem with risk measure BRM_β^α. Some of the problem parameters and fundamental variables are presented in the Tables 3.11 and 3.12. Tables 3.15 and 3.16 present the rest of required parameters and fundamental variables for portfolio problem with BRM. Table 3.17 summarizes the auxiliary variables needed for BRM risk measure.

Table 3.15: **Given Data**

α_l	Risk aversion parameter for AVaR_{α_l} for stage $l = 1, 2$
β_l	Coefficient of AVaR_{α_l} in the risk measure for stage $l = 1, 2$
$1 - \beta_l$	Coefficient of expected value in the risk measure for stage $l = 1, 2$

Table 3.16: **Fundamental model variables**

g	First-stage AVaR_α value for CTC model
h_i	AVaR_α value, in second-stage scenario $i = 1, \ldots, N_1$
V_i	Second stage cost plus $\text{BRM}_{\beta_2}^{\alpha_2}$ value for OTC model, in second-stage scenario $i = 1, \cdots, N_1$
k	First-stage AVaR_{α_1} value of V_i for OTC model

Table 3.17: **Auxiliary variables needed for BRM**

t	Helper variable in AVaR definition in stage 1 for CTC model
$X_{i,j}$	Positive part of difference between stage 2 and 3's total cost in scenario (i, j) and t, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
$\bar{t}_i$	Helper variable in AVaR definition in stage 2 scenario i, for $i = 1, \ldots, N_1$
$K_{i,j}$	Positive part of difference between stage 3's total cost in scenario (i, j) and $\bar{t}_i$, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
τ	Helper variable in AVaR_{α_1} definition in stage 1 for OTC model
χ_i	Positive part of difference between V_i and τ, for $i = 1, \ldots, N_1$

The first stage objective function for CTC formulation is:

$$\rho_1(Z_3) = (1 - \beta_1)\mathbb{E}(Z_3) + \beta_1 * \text{AVaR}(Z_3) = (1 - \beta_1)(\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j}(-W_{i,j})) + \beta_1 g$$

where

$$g \geq t + \frac{1}{\alpha_1}(\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j}),$$

$$X_{i,j} \geq -W_{i,j} - t, \qquad i = 1, \ldots, N_1, \quad j = 1, \ldots, N_2$$

$$X_{i,j} \geq 0, \qquad i = 1, \ldots, N_1, \quad j = 1, \ldots, N_2$$

and the second stage objective function is:

$$\rho_2(Z_3|i) = (1 - \beta_2)\mathbb{E}(Z_3|i) + \beta_2 * \text{AVaR}(Z_3|i) = (1 - \beta_2) \sum_{j=1}^{N_2} \pi_{i,j}(-W_{il}) + \beta_2 h_i$$

where,

$$h_i \geq \bar{t}_i + \frac{1}{\alpha_2}\sum_{j=1}^{N_2} \pi_{i,j} K_{i,j}, \qquad j = 1, \ldots, N_2,$$

$$K_{i,j} \geq -W_{i,j} - \bar{t}_i, \qquad j = 1, \ldots, N_2,$$

$$K_{i,j} \geq 0, \qquad j = 1, \ldots, N_2.$$

First-stage Problem

$$\begin{aligned}
\min \quad & (1-\beta_1)(\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \pi_i \pi_{i,j}(-W_{i,j})) + \beta_1 g \\
\text{s.t.} \quad & z^T \mathbf{1} = 1 \\
& g \geq t + \frac{1}{\alpha_1}(\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j}), \\
& X_{i,j} \geq -W_{i,j} - t, && \forall i, \forall j \\
& z \geq \mathbf{0}, \; X_{i,j} \geq 0, && \forall i, \forall j \\
& W_i \in \mathcal{X}_i^*(z), && \forall i.
\end{aligned} \tag{3.11}$$

Second-stage problem, scenario i:

$$\begin{aligned}
\min \quad & (1-\beta_2)\sum_{j=1}^{N_2} \pi_{i,j}(-W_{il}) + \beta_2 h_i \\
\text{s.t.} \quad & y_i = (\mathbf{1} + R_i^1) \circ z + (\mathbf{1} - \epsilon) \circ b_i - (\mathbf{1} + \epsilon) \circ s_i, \\
& b_i^T \mathbf{1} - s_i^T \mathbf{1} = 0, \\
& W_{i,j} = (\mathbf{1} + R_{i,j}^2)^T y_i, && \forall j \\
& h_i \geq \bar{t}_i + \frac{1}{\alpha_2}\sum_{j=1}^{N_2} \pi_{i,j} K_{i,j}, \\
& K_{i,j} \geq -W_{i,j} - \bar{t}_i, && \forall j \\
& K_{i,j} \geq 0, && \forall j \\
& y_i \geq \mathbf{0}, \; b_i \geq \mathbf{0}, \; s_i \geq \mathbf{0}, \; Y_{i,j} \geq 0, && \forall j.
\end{aligned} \tag{3.12}$$

In the above problem, "$\circ$" is Hadamard product, element-wise product of two matrices with the same dimension. Also, $\mathbf{1}$ is a vector of all ones with appropriate dimension.

Now, define the Lagrange multiplier for the second-stage problems to use them in converting the bilevel linear problem to a MPEC formulation.

Table 3.18: **Lagrange multiplier variables for MPEC formulation**

λ_i	Vector of Lagrange multipliers for constraints $y_i - (\mathbf{1} + R_i^1) \circ z - (\mathbf{1} - \epsilon) \circ b_i + (\mathbf{1} + \epsilon) \circ s_i = 0$, for $i = 1, \ldots, N_1$
μ_i	Scalar Lagrange multiplier for constraints $b_i^T \mathbf{1} - s_i^T \mathbf{1} = 0$, for $i = 1, \ldots, N_1$
$\rho_{i,j}$	Scalar Lagrange multiplier for constraints $W_{i,j} - (\mathbf{1} + R_{i,j}^2)^T y_i = 0$, for $i = 1, \ldots, N_1,\ j = 1, \ldots, N_2$
θ_i	Scalar Lagrange multiplier for constraints $h_i \geq \bar{t}_i + \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j} K_{i,j}$ for $i = 1, \ldots, N_1$
$\sigma_{i,j}$	Scalar Lagrange multiplier for constraints $K_{i,j} \geq -W_{i,j} - \bar{t}_i$ below, for $i = 1, \ldots, N_1,\ j = 1, \ldots, N_2$

The second-stage Lagrange function $L_i(y_i, s_i, b_i, W_{i,j}, h_i, \bar{t}_i, K_{i,j}, \lambda_i, \mu_i, \rho_{i,j}, \theta_i, \sigma_{i,j})$ is:

$$\begin{aligned}
L_i = \quad & (1 - \beta_2) \textstyle\sum_{j=1}^{N_2} \pi_{i,j}(-W_{il}) + \beta_2 h_i - \lambda_i \big(y_i - (\mathbf{1} + R_i^1) \circ z - (\mathbf{1} - \epsilon) \circ b_i \\
& +(\mathbf{1} + \epsilon) \circ s_i\big) - \mu_i (b_i^T \mathbf{1} - s_i^T \mathbf{1}) - \textstyle\sum_{j=1}^{N_2} \rho_{i,j}(W_{i,j} - (\mathbf{1} + R_{i,j}^2)^T y_i) \\
& -\theta_i (h_i - \bar{t}_i - \tfrac{1}{\alpha_2} \textstyle\sum_{j=1}^{N_2} \pi_{i,j} K_{i,j}) - \textstyle\sum_{j=1}^{N_2} \sigma_{i,j}(K_{i,j} + W_{i,j} + \bar{t}_i) \\
& = y_i^T \big(- \lambda_i + \textstyle\sum_{j=1}^{N_2} \rho_{i,j}(\mathbf{1} + R_{i,j}^2)\big) + b_i^T \big((1 - \epsilon) \circ \lambda_i - \mu_i \mathbf{1}\big) \\
& + s_i^T \big(- (1 + \epsilon) \circ \lambda_i + \mu_i \mathbf{1}\big) + \textstyle\sum_{j=1}^{N_2} W_{i,j} \big(- (1 - \beta_2)\pi_{i,j} - \rho_{i,j} - \sigma_{i,j}\big) \\
& + h_i(\beta_2 - \theta_i) + \bar{t}_i(\theta_i - \textstyle\sum_{j=1}^{N_2} \sigma_{i,j}) + \textstyle\sum_{j=1}^{N_2} K_{i,j}(\tfrac{1}{\alpha_2}\pi_{i,j}\theta_i - \sigma_{i,j}) \\
& + \lambda_i^T (\mathbf{1} + R_i^1) \circ z
\end{aligned}$$

CTC formulation of portfolio problem with BRM risk measure

$$\begin{aligned}
\min \quad & (1 - \beta_1)(\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j}(-W_{i,j})) + \beta_1 g \\
\text{s.t.} \quad & z^T \mathbf{1} = 1 \\
& g \geq t + \frac{1}{\alpha_1}(\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j}), \\
& X_{i,j} \geq -W_{i,j} - t, \qquad \forall i, \forall j \\
& y_i = (\mathbf{1} + R_i^1) \circ z + (\mathbf{1} - \epsilon) \circ b_i - (\mathbf{1} + \epsilon) \circ s_i, \quad \forall i
\end{aligned} \tag{3.13}$$

$$b_i^T \mathbf{1} - s_i^T \mathbf{1} = 0, \qquad \forall i$$

$$W_{i,j} = (\mathbf{1} + R_{i,j}^2)^T y_i, \qquad \forall i, j$$

$$h_i \geq \bar{t}_i + \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j} K_{i,j}, \qquad \forall i$$

$$K_{i,j} \geq -W_{i,j} - \bar{t}_i, \qquad \forall i, j$$

$$-\lambda_i + \sum_{j=1}^{N_2} \rho_{i,j}(1 + R_{i,j}^2) \geq 0, \qquad \forall i$$

$$(1 - \epsilon) \circ \lambda_i - \mu_i \mathbf{1} \geq 0, \qquad \forall i$$

$$-(1 + \epsilon) \circ \lambda_i + \mu_i \mathbf{1} \geq 0, \qquad \forall i$$

$$-(1 - \beta_2)\pi_{i,j} - \rho_{i,j} - \sigma_{i,j} \geq 0, \qquad \forall i, j$$

$$\beta_2 - \theta_i = 0, \qquad \forall i$$

$$\theta_i - \sum_{j=1}^{N_2} \sigma_{i,j} = 0, \qquad \forall i$$

$$\frac{1}{\alpha_2} \pi_{i,j} \theta_i - \sigma_{i,j} \geq 0, \qquad \forall i$$

$$\theta_i (h_i - \bar{t}_i - \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j} K_{i,j}) = 0, \qquad \forall i$$

$$\sigma_{i,j}(K_{i,j} + W_{i,j} + \bar{t}_i) = 0, \qquad \forall i, j$$

$$y_i^T(-\lambda_i + \sum_{j=1}^{N_2} \rho_{i,j}(\mathbf{1} + R_{i,j}^2)) = 0, \qquad \forall i$$

$$b_i^T \big((1 - \epsilon) \circ \lambda_i - \mu_i \mathbf{1}\big) = 0, \qquad \forall i,$$

$$s_i^T \big(-(1 + \epsilon) \circ \lambda_i + \mu_i \mathbf{1}\big) = 0, \qquad \forall i$$

$$\sum_{j=1}^{N_2} K_{i,j} (\frac{1}{\alpha_2} \pi_{i,j} \theta_i - \sigma_{i,j}) = 0, \qquad \forall i$$

$$z \geq \mathbf{0}, X_{i,j} \geq 0, \qquad \forall i, j$$

$$y_i \geq \mathbf{0}, b_i, s_i \geq \mathbf{0}, K_{i,j} \geq 0, \qquad \forall i, j$$

$$\theta_i, \sigma_{i,j} \geq 0, \qquad \forall i, j$$

Using nested risk measure $\rho_1(Z_1 + \rho_2(Z_2 + Z_3))$, we also formulate the "OTC" version of this problem.

OTC formulation of portfolio problem with nested BRM risk measure

$$
\begin{aligned}
\min \quad & (1-\beta_1)(\sum_{i=1}^{N_1} \pi_i V_i) + \beta_1 k \\
\text{s.t.} \quad & z^T \mathbf{1} = 1 \\
& V_i = (1-\beta_2)(\sum_{j=1}^{N_2} -\pi_{i,j} W_{i,j}) + \beta_2 h_i, & \forall i \\
& k \geq \tau + \frac{1}{\alpha_1}(\sum_{i=1}^{N_1} \pi_i \chi_i), \\
& \chi_i \geq V_i - \tau, & \forall i \\
& y_i = (\mathbf{1} + R_i^1) \circ z + (\mathbf{1} - \epsilon) \circ b_i - (\mathbf{1} + \epsilon) \circ s_i, & \forall i \\
& b_i^T \mathbf{1} - s_i^T \mathbf{1} = 0, & \forall i \\
& W_{i,j} = (\mathbf{1} + R_{i,j}^2)^T y_i, & \forall i, j \\
& h_i \geq \bar{t}_i + \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j} K_{i,j}, & \forall i \\
& K_{i,j} \geq -W_{i,j} - \bar{t}_i, & \forall i, j \\
& z \geq \mathbf{0}, \chi_i \geq 0, & \forall i \\
& y_i \geq \mathbf{0}, b_i, s_i \geq \mathbf{0}, K_{i,j} \geq 0, & \forall i, j
\end{aligned}
\tag{3.14}
$$

3.2.3 Numerical Study

The numerical study is comprised of 5,000 randomly generated instances with 5 different assets and 5 possible second-stage return rate scenarios. In each second-stage scenario, we also consider 5 possible outcomes for the return rates in the third stage. Similar to the supply chain problem, we consider two risk measures MSD_γ and BRM_β^α with parameter values $\gamma = 0.3, 0.7, \text{and } 0.9$ and $(\alpha, \beta) = (0.01, 0.5), (0.05, 0.5), (0.25, 0.5),$ $(0.05, 0.25)$, and $(0.25, 0.75)$. Tables 3.19 and 3.20 summarize the results for MSD_γ and BRM_β^α risk measures, respectively.

With the MSD risk measure, in approximately 81%, 58%, and 49% of the instances with γ values 0.3, 0.7, and 0.9, respectively, we do not observe a significant difference between the two models (Diff<0.01%), suggesting that the OTC formulation can serve

	γ =0.3	γ =0.7	γ =0.9
Diff < 0.01%	80.98%	58.04%	48.76 %
0.01% $\leq$ Diff < 0.1%	6.04%	9.56 %	9.02%
0.1% $\leq$ Diff < 1%	11.62%	22.74 %	27.3%
1% $\leq$ Diff < 10%	1.36%	9.66%	14.92%
10% $\leq$ Diff	0%	0%	0%
Max Diff	2.14%	6.59%	9.03%
Δ_1	121.33%	99.67%	102.45%
Δ_2	40.70%	62.04%	58.20%

Table 3.19: Computational results of the portfolio optimization problem with the MSD risk measure.

	$\beta = 0.5$ $\alpha = 0.01$	$\beta = 0.5$ $\alpha = 0.05$	$\beta = 0.75$ $\alpha = 0.05$	$\beta = 0.5$ $\alpha = 0.25$	$\beta = 0.75$ $\alpha = 0.25$
Diff < 0.01%	17.86%	19.88%	10.3%	30.1%	16.88%
0.01% $\leq$ Diff < 0.1%	3.46%	4.5%	4.8%	6.9%	5.54%
0.1% $\leq$ Diff < 1%	23.12%	26.14%	32.52%	28.48%	30.86%
1% $\leq$ Diff < 10%	52.3%	47.78%	51.24%	34.24%	45.22%
10% $\leq$ Diff	3.26%	1.7%	1.14%	0.28%	1.5%
Max Diff	34.18%	29.85%	37.05%	16.49%	43.50%
Δ_1	162.13%	102.31%	50.51%	112.47%	102.86%
Δ_2	84.066%	66.87%	56.53%	61.41%	105.52%

Table 3.20: Computational results for portfolio problems with the BRM risk measure.

as a good approximator of the CTC model. This is especially the case for lower degrees of risk aversion (i.e., cases with small γ). However, we still find instances with relatively different solutions for the OTC and CTC models. For example, for $\gamma = 0.9$, in 15% of the cases, we observed a gap of more than 1% with a maximum of 9.03%. We also confirm our finding that for more risk-averse problems, the OTC and CTC formulations are more likely to result in different solutions.

Table 3.20 illustrates that when using the BRM risk measure, the OTC and CTC models differ more significantly: in 55.56% of the numerical instances with $\beta = 0.5$ and $\alpha = 0.01$, we observe a gap higher than 1%. Additionally, the gap difference could be as high as 43.5% for the instances with $\beta = 0.75$ and $\alpha = 0.25$. However, similarly to the MSD_γ risk model, the OTC and CTC formulations yield closer solutions in portfolio optimization than for the supply chain problem.

3.3 Hydropower Energy Planning Problem

In this section, we present the optimal scheduling of a hydropower system problem considered in Pereira & Pinto (1991), where an electric utility "the company" simultaneously generates electricity from hydro and thermal plants. The company may own multiple plants for each source of energy and decides the electricity generation targets for each plant over three stages to meet the market demand while minimizing cost. The cost functions is a sum of two elements: the cost of electricity generation and penalties for load-shedding, which occurs when the market demand cannot be met.

Hydropower plants have relatively low marginal cost of energy production. However, they have limited and uncertain production capacity, because they depend on the *hydrologic trend* which consists of such factors as the amount of rainfall, evaporation, and streamflow discharge. On the other hand, thermal plants have a predictable production capacity, while the variable cost of electricity generation depends on factors such as the fuel price, which is uncertain. Notably, the variable cost can vary over time and be different for plants in different geographical locations.

To capture these comparative properties, without loss of generality, we assume a zero variable production cost for the hydroelectric plants, and no uncertainty in production capacity of the thermal plants. We also assume that at the beginning of the second and third stages, the cost of generating electricity in each thermal plant, the hydrologic trend at each hydro plant, and the market demand are unknown.

Given that the planner aims to minimize cost, it desires to use the cheapest mode of production at each stage, i.e., the hydropower plant. However, if it depletes the reservoir storage early on in the planning horizon, it may not have the hydropower capacity in the future when the fuel cost may be high. On the other hand, if it does not use the capacity of the hydropower plant at the right time, then there is a chance of spillage in the case of high inflow, and capacity may be wasted.

3.3.1 Problem Formulation with MSD Risk Measure

We consider the case where the company's objective function is to minimize the MSD_γ of the total cost function. Tables 3.21, 3.22, and 3.23 summarize the problem parameters and decision variables.

Table 3.21: **Given Data**

T	Set of thermal plants
H	Set of hydro plants
c_1	Vector of unit cost of generating electricity in thermal plants, in $\mathbb{R}^{\vert T\vert}$ in first stage ($\$/GJ$)
$c_{2,i}$	Vector of unit cost of generating electricity in thermal plants in second-stage scenario i, for $i = 1, \cdots, N_1$, in $\mathbb{R}^{\vert T\vert}$ ($\$/GJ$)
$c_{3,i,j}$	Vector of unit cost of generating electricity in thermal plants in third-stage scenario (i, j), for $i = 1, \cdots, N_1$ and $j = 1, \cdots, N_2$, in $\mathbb{R}^{\vert T\vert}$ ($\$/GJ$)
c_l	Unit cost of load-shedding ($\$/GJ$)
ρ	Production coefficient vector for hydro plants, in $\mathbb{R}^{\vert H\vert}$
$\bar{g}$	Thermal generation capacity vector
a_1	Vector of hydrologic trend of plants in the first stage, in $\mathbb{R}^{\vert H\vert}$
$a_{2,i}$	Vector of hydrologic trend of plants in second-stage scenario i, for $i = 1, \dots, N_1$, in $\mathbb{R}^{\vert H\vert}$
$a_{3,i,j}$	Vector of hydrologic trend of plants in third-stage scenario (i, j), for $i = 1, \cdots, N_1$ and $j = 1, \cdots, N_2$, in $\mathbb{R}^{\vert H\vert}$

v_1	Reservoir storage level of hydro plants in the beginning of first stage, in $\mathbb{R}^{\|H\|}$
π_i	Probability of second-stage scenario i, for $i = 1, \ldots, N_1$
$\pi_{i,j}$	Conditional probability of third-stage scenario (i, j), given the occurrence of second-stage scenario i, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
d_1	Energy demand in first stage
$d_{2,i}$	Energy demand in second-stage scenario i, for $i = 1, \ldots, N_1$
$d_{3,i,j}$	Energy demand in third-stage scenario (i, j), for $i = 1, \cdots, N_1$ and $j = 1, \cdots, N_2$
M	Incidence matrix of hydro plants
$\bar{v}$	Vector of reservoir capacities, in $\mathbb{R}^{\|H\|}$
$\bar{q}$	Limits on turbine outflow, in $\mathbb{R}^{\|H\|}$
$\underline{q}$	Lower bounds on total outflow, in $\mathbb{R}^{\|H\|}$
γ_l	Risk aversion parameter for $\mathrm{MSD}^1_{\gamma_l}$ risk measure for stage $l = 1, 2$

Table 3.22: **Fundamental model variables**

G_1	Total generation of thermal plants in first stage, in $\mathbb{R}^{\|T\|}$
$G_{2,j}$	Total generation of thermal plants in second-stage scenario i, for $i = 1, \ldots, N_1$, in $\mathbb{R}^{\|T\|}$
$G_{3,i,j}$	Total generation of thermal plants in third-stage scenario (i, j), for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$, in $\mathbb{R}^{\|T\|}$
L_1	Load-shedding in first stage
$L_{2,i}$	Load-shedding in second-stage scenario i, for $i = 1, \cdots, N_1$
$L_{3,i,j}$	Load-shedding in third-stage scenario (i, j), for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
$Q_1, (S_1)$	Turbine outflow (spilled volumes) of hydro plants in first-stage, in $\mathbb{R}^{\|H\|}$

$Q_{2,i}, (S_{2,i})$	Turbine outflow (spilled volumes) of hydro plants in second-stage scenario i, for $i = 1, \ldots, N_1$, in $\mathbb{R}^{\|H\|}$
$Q_{3,i,j}, (S_{3,i,j})$	Turbine outflow, (spilled volumes) of hydro plants in third-stage scenario (i, j), for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$, in $\mathbb{R}^{\|T\|}$
V_2	Reservoir storage level of hydro plants in the beginning of second-stage scenario in $\mathbb{R}^{\|H\|}$
$V_{3,i}$	Reservoir storage level of hydro plants in the beginning of third-stage scenario i, for $i = 1, \ldots, N_1$, in $\mathbb{R}^{\|H\|}$

Table 3.23: **Auxiliary variables needed for** MSD

$X_{i,j}$	Amount combined stage 2 and 3 objective exceeds its overall mean in scenario (i, j), for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
$Y_{i,j}$	Amount stage-3 objective exceeds its stage-2 conditional mean in
W_i	Second stage cost plus MSD value of third stage cost for OTC model, in second stage scenario i, for $i = 1, \cdots, N_1$ scenario (i, j), for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
χ_i	Amount W_i exceeds its overall mean in for OTC model second-stage scenario i, for $i = 1, \cdots, N_1$

With this notation and the MSD risk measure, the first-stage and second-stage problems can be formulated as problems (3.15) and (3.16).

First-stage problem

$$\begin{aligned}
\min \quad & c_l^T L_1 + \textstyle\sum_{i=1}^{N_1} \pi_i c_l^T L_{2,i} + \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} c_l^T L_{3,i,j} + c_1^T G_1 + \sum_{i=1}^{N_1} \pi_i c_{2,i}^T G_{2,i} \\
& + \textstyle\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} c_{3,i,j}^T G_{3,i,j} + \gamma_1 \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j} \\
\text{s.t.} \quad & X_{i,j} \geq c_l^T (L_{2,i} + L_{3,i,j}) + c_{2,i}^T G_{2,i} + c_{3,i,j}^T G_{3,i,j} - \textstyle\sum_{i'=1}^{N_1} \pi_{i'} \big[c_l^T L_{2,i'} \\
& \quad + c_{2,i'}^T G_{2,i'} + \textstyle\sum_{j'=1}^{N_2} \pi_{i',j'} (c_l^T L_{3,i',j'} + c_{3,i',j'}^T G_{3,i',j'}) \big]
\end{aligned} \tag{3.15}$$

$$
\begin{aligned}
&\rho^T Q_1 + \mathbf{1}^T G_1 + L_1 = d_1 \\
&V_2 = v_1 + a_1 + M(Q_1 + S_1), \\
&V_2 \le \bar{v}, \\
&Q_1 \le \bar{q}, \\
&Q_1 + S_1 \ge \underline{q}, \\
&G_1 \le \bar{g} \\
&L_1, G_1, X_{i,j}, Q_1, V_2, S_1 \ge 0, && \forall i, j \\
&(G_{2,i}, G_{3,i,j}, L_{2,i}, L_{3,i,j}) \in \mathcal{X}_i^*(V_2), && \forall i
\end{aligned}
$$

Second-stage problem, scenario i

$$
\begin{aligned}
\min \quad & c_l^T L_{2,i} + c_{2,i}^T G_{2,i} + \sum_{j=1}^{N_2} \pi_{i,j}(c_l^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j}) + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j} \\
\text{s.t.} \quad & Y_{i,j} \ge c_l^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'}(c_l^T L_{3,i,j'} + c_{3,i,j'}^T G_{3,i,j'}) \\
& \rho^T Q_{2,i} + \mathbf{1}^T G_{2,i} + L_{2,i} = d_{2,i} \\
& V_{3,i} = V_2 + a_{2,i} + M(Q_{2,i} + S_{2,i}), \quad V_{3,i} \le \bar{v}, \\
& Q_{2,i} \le \bar{q}, \\
& Q_{2,i} + S_{2,i} \ge \underline{q}, \\
& G_{2,i} \le \bar{g} \\
& \rho^T Q_{3,i,j} + G_{3,i,j} + L_{3,i,j} = d_{3,i,j}, && \forall j \\
& V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \ge 0, && \forall j \\
& V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \le \bar{v}, && \forall j \\
& Q_{3,i,j} \le \bar{q}, && \forall j \\
& Q_{3,i,j} + S_{3,i,j} \ge \underline{q}, && \forall j \\
& G_{3,i,j} \le \bar{g}, && \forall j \\
& L_{2,i}, L_{3,i,j}, G_{2,i}, G_{3,i,j}, Y_{i,j} \ge 0, && \forall j \\
& Q_{2,i}, S_{2,i}, Q_{3,i,j}, S_{3,i,j}, V_{3,i} \ge 0, && \forall j
\end{aligned}
\tag{3.16}
$$

Since the second-stage problems (3.16) are linear, they satisfy standard constraint qualification conditions and one can replace the set of constraints $(G_{2,i}, G_{3,i,j}, L_{2,i}, L_{3,i,j}) \in \mathcal{X}_i^*(V_2)$ in (3.15) with KKT optimality conditions for each scenario. This procedure converts the bilevel problem (3.15)-(3.16) into mathematical programming with equilibrium constraints (MPEC) (3.17).

Table 3.24: **Lagrange multiplier variables for second-stage problems**

$\lambda_{i,j}$	Vector of Lagrange multipliers for constraints $Y_{i,j} \geq c_l^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'}(c_l^T L_{3,i,j'} + c_{3,i,j'}^T G_{3,i,j'})$ below, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
$\alpha_{2,i}$	Vector of Lagrange multipliers for constraints $\rho^T Q_{2,i} + \mathbf{1}^T G_{2,i} + L_{2,i} = d_{2,i}$, for $i = 1, \ldots, N_1$
$\beta_{2,i}$	Vector of Lagrange multipliers for constraints $V_{3,i} = V_2 + a_{2,i} + M(Q_{2,i} + S_{2,i})$, for $i = 1, \ldots, N_1$
$\sigma_{2,i}$	Vector of Lagrange multipliers for constraint $V_{3,i} \leq \bar{v}$ below, for $i = 1, \ldots, N_1$
$\theta_{2,i}$	Vector of Lagrange multipliers for constraint $Q_{2,i} \leq \bar{q}$ below, for $i = 1, \ldots, N_1$
$\delta_{2,i}$	Vector of Lagrange multipliers for constraint $Q_{2,i} + S_{2,i} \geq \underline{q}$ below, for $i = 1, \ldots, N_1$
$\mu_{2,i}$	Vector of Lagrange multipliers for constraint $G_{2,i} \leq \bar{g}$ below, for $i = 1, \ldots, N_1$
$\alpha_{3,i,j}$	Vector of Lagrange multipliers for constraint $\rho^T Q_{3,i,j} + G_{3,i,j} + L_{3,i,j} = d_{3,i,j}$, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
$\beta_{3,i,j}$	Vector of Lagrange multipliers for constraint $V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \geq 0$ below, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
$\sigma_{3,i,j}$	Vector of Lagrange multipliers for constraint $V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \leq \bar{v}$ below, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$

$\theta_{3,i,j}$ Vector of Lagrange multipliers for constraint $Q_{3,i,j} \leq \bar{q}$ below, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$

$\delta_{3,i,j}$ Vector of Lagrange multipliers for constraint $Q_{3,i,j}+S_{3,i,j} \geq \underline{q}$ below, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$

$\mu_{3,i,j}$ Vector of Lagrange multipliers for constraint $G_{3,i,j} \leq \bar{g}$ below, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$

The second-stage Lagrangian function $L_i(G_{2,i}, G_{3,i,j}, L_{2,i}, L_{3,i,j}, Q_{2,i}, S_{2,i}, Q_{3,i,j}, S_{3,i,j}, V_{3,i}, Y_{i,j}, \lambda_{i,j}, \alpha_{2,i}, \beta_{2,i}, \sigma_{2,i}, \theta_{2,i}, \delta_{2,i}, \mu_{2,i}, \alpha_{3,i,j}, \beta_{3,i,j}, \sigma_{3,i,j}, \theta_{3,i,j}, \delta_{3,i,j}, \mu_{3,i,j})$ is:

$$
\begin{aligned}
L_i = \; & c_l^T L_{2,i} + c_{2,i}^T G_{2,i} + \textstyle\sum_{j=1}^{N_2} \pi_{i,j}(c_l^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j}) + \gamma_2 \textstyle\sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j} \\
& - \textstyle\sum_{j=1}^{N_2} \lambda_{i,j}^T (Y_{i,j} - c_l^T L_{3,i,j} - c_{3,i,j}^T G_{3,i,j} + \textstyle\sum_{j'=1}^{N_2} \pi_{i,j'}(c_l^T L_{3,i,j'} + c_{3,i,j'}^T G_{3,i,j'}) \\
& - \alpha_{2,i}^T(\rho^T Q_{2,i} + \mathbf{1}^T G_{2,i} + L_{2,i} - d_{2,i}) - \beta_{2,i}^T(V_2 + a_{2,i} + M(Q_{2,i} + S_{2,i}) - V_{3,i}) \\
& - \sigma_{2,i}^T(\bar{v} - V_{3,i}) - \theta_{2,i}^T(\bar{q} - Q_{2,i}) - \delta_{2,i}^T(Q_{2,i} + S_{2,i} - \underline{q}) - \mu_{2,i}^T(\bar{g} - G_{2,i}) \\
& - \textstyle\sum_{j=1}^{N_2} \alpha_{3,i,j}^T(\rho^T Q_{3,i,j} + G_{3,i,j} + L_{3,i,j} - d_{3,i,j}) \\
& - \textstyle\sum_{j=1}^{N_2} \beta_{3,i,j}^T(V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j})) \\
& - \textstyle\sum_{j=1}^{N_2} \sigma_{3,i,j}^T(\bar{v} - V_{3,i} - a_{3,i,j} - M(Q_{3,i,j} + S_{3,i,j})) - \textstyle\sum_{j=1}^{N_2} \theta_{3,i,j}^T(\bar{q} - Q_{3,i,j}) \\
& - \textstyle\sum_{j=1}^{N_2} \delta_{3,i,j}^T(Q_{3,i,j} + S_{3,i,j} - \underline{q}) - \textstyle\sum_{j=1}^{N_2} \mu_{3,i,j}^T(\bar{g} - G_{3,i,j})
\end{aligned}
$$

CTC formulation of hydropower energy planning problem with MSD

$$
\begin{aligned}
\min \quad & c_l^T L_1 + c_1^T G_1 + \sum_{i=1}^{N_1} \pi_i (c_l^T L_{2,i} + c_{2,i}^T G_{2,i}) \\
& + \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} (c_{3,i,j}^T G_{3,i,j} + c_l^T L_{3,i,j}) + \gamma_1 \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j} \\
\text{s.t.} \quad & X_{i,j} \geq c_l^T (L_{2,i} + L_{3,i,j}) + c_{2,i}^T G_{2,i} + c_{3,i,j}^T G_{3,i,j} - \sum_{i'=1}^{N_1} \pi_{i'} \big[c_l^T L_{2,i'} \\
& \qquad + c_{2,i'}^T G_{2,i'} + \sum_{j'=1}^{N_2} \pi_{i',j'} (c_l^T L_{3,i',j'} + c_{3,i',j'}^T G_{3,i',j'}) \big], \qquad \forall i, \forall j \\
& \rho^T Q_1 + \mathbf{1}^T G_1 + L_1 = d_1, \\
& V_2 = v_1 + a_1 + M(Q_1 + S_1), \\
& V_2 \leq \bar{v}, \\
& Q_1 \leq \bar{q},
\end{aligned} \tag{3.17}
$$

$$Q_1 + S_1 \geq \underline{q},$$

$$G_1 \leq \bar{g},$$

$$Y_{i,j} \geq c_l^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} (c_l^T L_{3,i,j'} + c_{3,i,j'}^T G_{3,i,j'}), \quad \forall i, \forall j$$

$$\rho^T Q_{2,i} + \mathbf{1}^T G_{2,i} + L_{2,i} = d_{2,i}, \quad \forall i$$

$$V_{3,i} = V_2 + a_{2,i} + M(Q_{2,i} + S_{2,i}), \quad \forall i$$

$$V_{3,i} \leq \bar{v}, \quad \forall i$$

$$Q_{2,i} \leq \bar{q}, \quad \forall i$$

$$Q_{2,i} + S_{2,i} \geq \underline{q}, \quad \forall i$$

$$G_{2,i} \leq \bar{g}, \quad \forall i$$

$$\rho^T Q_{3,i,j} + G_{3,i,j} + L_{3,i,j} = d_{3,i,j}, \quad \forall i, \forall j$$

$$V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \geq 0, \quad \forall i, \forall j$$

$$V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \leq \bar{v}, \quad \forall i, j$$

$$Q_{3,i,j} \leq \bar{q}, \quad \forall i, \forall j$$

$$Q_{3,i,j} + S_{3,i,j} \geq \underline{q}, \quad \forall i, \forall j$$

$$G_{3,i,j} \leq \bar{g}, \quad \forall i, \forall j$$

$$c_{2,i} - \mathbf{1}\alpha_{2,i} + \mu_{2,i} \geq 0, \quad \forall i$$

$$\pi_{i,j} c_{3,i,j} + c_{3,i,j} \lambda_{i,j} - \pi_{i,j} c_{3,i,j} (\sum_{j'=1}^{N_2} \lambda_{i,j'}) - \alpha_{3,i,j} + \mu_{3,i,j} \geq 0, \quad \forall i, \forall j$$

$$c_l - \alpha_{2,i} \geq 0, \quad \forall i$$

$$\pi_{i,j} c_l + c_l \lambda_{i,j} - \pi_{i,j} c_l (\sum_{j'=1}^{N_2} \lambda_{i,j'}) - \alpha_{3,i,j} \geq 0, \quad \forall i, j$$

$$-\rho\alpha_{2,i} - M^T \beta_{2,i} + \theta_{2,i} - \delta_{2,i} \geq 0, \quad \forall i$$

$$-M^T \beta_{2,i} - \delta_{2,i} \geq 0, \quad \forall i$$

$$-\rho\alpha_{3,i,j} - M^T \beta_{3,i,j} + M^T \sigma_{3,i,j} + \theta_{3,i,j} - \delta_{3,i,j} \geq 0, \quad \forall i, \forall j$$

$$-M^T \beta_{3,i,j} + M^T \sigma_{3,i,j} - \delta_{3,i,j} \geq 0, \quad \forall i, \forall j$$

$$\beta_{2,i} + \sigma_{2,i} - \sum_{j=1}^{N_2} \beta_{3,i,j} + \sum_{j=1}^{N_2} \sigma_{3,i,j} \geq 0, \quad \forall i, \forall j$$

$$\gamma_2\pi_{i,j} - \lambda \geq 0, \qquad \forall i, \forall j$$

$$\lambda_{i,j}^T(Y_{i,j} - c_l^T L_{3,i,j} - c_{3,i,j}^T G_{3,i,j} + \sum_{j'=1}^{N_2} \pi_{i,j'}(c_l^T L_{3,i,j'} + c_{3,i,j'}^T G_{3,i,j'})) = 0, \quad \forall i, \forall j$$

$$\sigma_{2,i}^T(\bar{v} - V_{3,i}) = 0, \qquad \forall i, \forall j$$

$$\theta_{2,i}^T(\bar{q} - Q_{2,i}) = 0, \qquad \forall i$$

$$\delta_{2,i}^T(Q_{2,i} + S_{2,i} - \underline{q}) = 0, \qquad \forall i$$

$$\mu_{2,i}^T(\bar{g} - G_{2,i}) = 0, \qquad \forall i$$

$$\beta_{3,i,j}^T(V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j})) = 0, \qquad \forall i, \forall j$$

$$\sigma_{3,i,j}^T(\bar{v} - V_{3,i} - a_{3,i,j} - M(Q_{3,i,j} - S_{3,i,j})) = 0, \qquad \forall i, \forall j$$

$$\theta_{3,i,j}^T(\bar{q} - Q_{3,i,j}) = 0, \qquad \forall i, \forall j$$

$$\delta_{3,i,j}^T(Q_{3,i,j} + S_{3,i,j} - \underline{q}) = 0, \qquad \forall i, \forall j$$

$$\mu_{3,i,j}^T(\bar{g} - G_{3,i,j}) = 0, \qquad \forall i$$

$$G_{2,i}^T(c_{2,i} - \mathbf{1}\alpha_{2,i} + \mu_{2,i}) = 0, \qquad \forall i$$

$$G_{3,i,j}^T(\pi_{i,j}c_{3,i,j} + c_{3,i,j}\lambda_{i,j} - \pi_{i,j}c_{3,i,j}(\sum_{j'=1}^{N_2}\lambda_{i,j'}) - \alpha_{3,i,j} + \mu_{3,i,j}) = 0, \qquad \forall i, \forall j$$

$$L_{2,i}^T(c_l - \alpha_{2,i}) = 0, \qquad \forall i$$

$$L_{3,i,j}^T(\pi_{i,j}c_l + c_l\lambda_{i,j} - \pi_{i,j}c_l(\sum_{j'=1}^{N_2}\lambda_{i,j'}) - \alpha_{3,i,j}) = 0, \qquad \forall i, \forall j$$

$$Q_{2,i}^T(-\rho\alpha_{2,i} - M^T\beta_{2,i} + \theta_{2,i} - \delta_{2,i}) = 0, \qquad \forall i$$

$$S_{2,i}^T(-M^T\beta_{2,i} - \delta_{2,i}) = 0, \qquad \forall i$$

$$Q_{3,i,j}^T(-\rho\alpha_{3,i,j} - M^T\beta_{3,i,j} + M^T\sigma_{3,i,j} + \theta_{3,i,j} - \delta_{3,i,j}) = 0, \qquad \forall i, \forall j$$

$$S_{3,i,j}^T(-M^T\beta_{3,i,j} + M^T\sigma_{3,i,j} - \delta_{3,i,j}) = 0, \qquad \forall i, \forall j$$

$$V_{3,i}^T(\beta_{2,i} + \sigma_{2,i} - \sum_{j=1}^{N_2}\beta_{3,i,j} + \sum_{j=1}^{N_2}\sigma_{3,i,j}) = 0, \qquad \forall i, \forall j$$

$$Y_{i,j}^T(\gamma_2\pi_{i,j} - \lambda_{i,j}) = 0, \qquad \forall i, \forall j$$

$$\lambda_{i,j}, \alpha_{2,i}, \beta_{2,i}, \sigma_{2,i}, \theta_{2,i}, \delta_{2,i}, \mu_{2,i} \geq 0 \qquad \forall i$$

$$\alpha_{3,i,j}, \beta_{3,i,j}, \sigma_{3,i,j}, \theta_{3,i,j}, \delta_{3,i,j}, \mu_{3,i,j} \geq 0, \qquad \forall i, \forall j$$

$$L_1, G_1, X_{i,j}, Q_1, V_2, S_1 \geq 0, \qquad \forall i$$
$$L_{2,i}, L_{3,i,j}, G_{2,i}, G_{3,i,j}, Y_{i,j}, Q_{2,i}, S_{2,i}, Q_{3,i,j}, S_{3,i,j}, V_{3,i} \geq 0, \qquad \forall i, \forall j.$$

Above, "$\forall i$" is a shorthand for "$i = 1, \ldots, N_1$", and "$\forall j$" is a shorthand for "$j = 1, \ldots, N_2$." Using nested risk measure $\rho_1(Z_1 + \rho_2(Z_2 + Z_3))$, we also formulate the "OTC" version of this problem in (3.18).

OTC formulation of hydropower energy planning problem with MSD

$$
\begin{aligned}
\min \quad & c_l^T L_1 + c_1^T G_1 + \sum_{i=1}^{N_1} \pi_i (W_i + \gamma_1 \chi_i) \\
\text{s.t.} \quad & \chi_i \geq W_i - \sum_{i'=1}^{N_1} \pi_{i'} W_{i'}, && \forall i \\
& W_i = c_l^T L_{2,i} + c_{2,i}^T G_{2,i} + \sum_{j=1}^{N_2} \pi_{i,j} (c_l^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j}) + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, && \forall i \\
& Y_{i,j} \geq c_l^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} (c_l^T L_{3,i,j'} + c_{3,i,j'}^T G_{3,i,j'}), && \forall i, \forall j \\
& \rho^T Q_1 + \mathbf{1}^T G_1 + L_1 = d_1, \\
& V_2 = v_1 + a_1 + M(Q_1 + S_1), \\
& V_2 \leq \bar{v}, \\
& Q_1 \leq \bar{q}, \\
& Q_1 + S_1 \geq \underline{q}, \\
& G_1 \leq \bar{g}, \\
& \rho^T Q_{2,i} + \mathbf{1}^T G_{2,i} + L_{2,i} = d_{2,i}, && \forall i \\
& V_{3,i} = V_2 + a_{2,i} + M(Q_{2,i} + S_{2,i}), && \forall i \\
& V_{3,i} \leq \bar{v}, && \forall i \\
& Q_{2,i} \leq \bar{q}, && \forall i \\
& Q_{2,i} + S_{2,i} \geq \underline{q}, && \forall i \\
& G_{2,i} \leq \bar{g}, && \forall i
\end{aligned}
\tag{3.18}
$$

$$\rho^T Q_{3,i,j} + G_{3,i,j} + L_{3,i,j} = d_{3,i,j} \qquad \forall i, \forall j$$

$$V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \geq 0, \qquad \forall i, \forall j$$

$$V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \leq \bar{v}, \qquad \forall i, \forall j$$

$$Q_{3,i,j} \leq \bar{q}, \qquad \forall i, \forall j$$

$$Q_{3,i,j} + S_{3,i,j} \geq \underline{q}, \qquad \forall i, \forall j$$

$$G_{3,i,j} \leq \bar{g}, \qquad \forall i, \forall j$$

$$L_1, G_1, \chi_i, Q_1, V_2, S_1 \geq 0, \qquad \forall i$$

$$L_{2,i}, L_{3,i,j}, G_{2,i}, G_{3,i,j}, Y_{i,j}, Q_{2,i}, S_{2,i}, Q_{3,i,j}, S_{3,i,j}, V_{3,i} \geq 0, \qquad \forall i, \forall j.$$

3.3.2 Problem Formulation with BRM Risk Measure

We also follow the same steps to formulate the CTC and OTC models of this problem with risk measure BRM_β^α. Tables 3.25, 3.26, 3.27, and 3.28 present the rest of the required parameters, fundamental variables, auxiliary variables, and Lagrange multiplier variables, respectively, for hydropower energy planning problem with BRM.

Table 3.25: **Given Data**

α_l	Risk aversion parameter for AVaR_{α_l} for stage $l = 1, 2$
β_l	Coefficient of AVaR_{α_l} in the risk measure for stage $l = 1, 2$
$1 - \beta_l$	Coefficient of expected value in the risk measure for stage $l = 1, 2$

Table 3.26: **Fundamental model variables**

g	First-stage AVaR_α value for CTC model
h_i	AVaR_α value, in second-stage scenario $i = 1, \ldots, N_1$
ω_i	Second stage cost plus $\text{BRM}_{\beta_2}^{\alpha_2}$ value for OTC model, in second-stage scenario $i = 1, \cdots, N_1$
k	First-stage AVaR_{α_1} value of ω_i for OTC model

Table 3.27: **Auxiliary variables needed for BRM**

t	Helper variable in AVaR definition in stage 1 for CTC model

$X_{i,j}$	Positive part of difference between stage 2 and 3's total cost in scenario (i,j) and t, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
$\bar{t}_i$	Helper variable in AVaR definition in stage 2 scenario i, for $i = 1, \ldots, N_1$
$Y_{i,j}$	Positive part of difference between stage 3's total cost in scenario (i,j) and $\bar{t}_i$, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$
τ	Helper variable in AVaR_{α_1} definition in stage 1 for OTC model
χ_i	Positive part of difference between ω_i and τ, for $i = 1, \ldots, N_1$

Table 3.28: **Lagrange multiplier variables for second-stage problems**

ζ_i	Vector of Lagrange multipliers for constraints $h_i \geq \bar{t}_i + \frac{1}{\alpha_2}\sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}$, for $i = 1, \cdots, N_1$

CTC formulation of hydropower energy planning problem with BRM_β^α

$$\begin{aligned}
\min \quad & c_l^T L_1 + c_1^T G_1 + (1-\beta_1) \sum_{i=1}^{N_1} \Big[\pi_i (c_l^T L_{2,i} + c_{2,i}^T G_{2,i}) \\
& + \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} (c_{3,i,j}^T G_{3,i,j} + c_l^T L_{3,i,j}) \Big] + \beta_1 g \\
\text{s.t.} \quad & g \geq t + \frac{1}{\alpha_1} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j} \\
& X_{i,j} \geq c_l^T (L_{2,i} + L_{3,i,j}) + c_{2,i}^T G_{2,i} + c_{3,i,j}^T G_{3,i,j} - t, \quad \forall i, \forall j \\
& \rho^T Q_1 + \mathbf{1}^T G_1 + L_1 = d_1, \\
& V_2 = v_1 + a_1 + M(Q_1 + S_1), \\
& V_2 \leq \bar{v}, \\
& Q_1 \leq \bar{q}, \\
& Q_1 + S_1 \geq \underline{q}, \\
& G_1 \leq \bar{g}, \\
& h_i \geq \bar{t}_i + \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, \qquad \forall i
\end{aligned} \tag{3.19}$$

$$Y_{i,j} \geq c_l^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j} - \bar{t}_i, \qquad \forall i, \forall j$$

$$\rho^T Q_{2,i} + \mathbf{1}^T G_{2,i} + L_{2,i} = d_{2,i}, \qquad \forall i$$

$$V_{3,i} = V_2 + a_{2,i} + M(Q_{2,i} + S_{2,i}), \qquad \forall i$$

$$V_{3,i} \leq \bar{v}, \qquad \forall i$$

$$Q_{2,i} \leq \bar{q}, \qquad \forall i$$

$$Q_{2,i} + S_{2,i} \geq \underline{q}, \qquad \forall i$$

$$G_{2,i} \leq \bar{g}, \qquad \forall i$$

$$\rho^T Q_{3,i,j} + G_{3,i,j} + L_{3,i,j} = d_{3,i,j}, \qquad \forall i, \forall j$$

$$V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \geq 0, \qquad \forall i, j$$

$$V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \leq \bar{v}, \qquad \forall i, j$$

$$Q_{3,i,j} \leq \bar{q}, \qquad \forall i, \forall j$$

$$Q_{3,i,j} + S_{3,i,j} \geq \underline{q}, \qquad \forall i, \forall j$$

$$G_{3,i,j} \leq \bar{g}, \qquad \forall i, \forall j$$

$$c_{2,i} - \mathbf{1}\alpha_{2,i} + \mu_{2,i} \geq 0, \qquad \forall i$$

$$(1 - \beta_2)\pi_{i,j} c_{3,i,j} + c_{3,i,j}\lambda_{i,j} - \alpha_{3,i,j} + \mu_{3,i,j} \geq 0, \qquad \forall i, \forall j$$

$$c_l - \alpha_{2,i} \geq 0, \qquad \forall i$$

$$(1 - \beta_2)\pi_{i,j} c_l + c_l \lambda_{i,j} - \alpha_{3,i,j} \geq 0, \qquad \forall i, j$$

$$-\rho\alpha_{2,i} - M^T \beta_{2,i} + \theta_{2,i} - \delta_{2,i} \geq 0, \qquad \forall i$$

$$-M^T \beta_{2,i} - \delta_{2,i} \geq 0, \qquad \forall i$$

$$-\rho\alpha_{3,i,j} - M^T \beta_{3,i,j} + M^T \sigma_{3,i,j} + \theta_{3,i,j} - \delta_{3,i,j} \geq 0, \qquad \forall i, \forall j$$

$$-M^T \beta_{3,i,j} + M^T \sigma_{3,i,j} - \delta_{3,i,j} \geq 0, \qquad \forall i, \forall j$$

$$\beta_{2,i} + \sigma_{2,i} - \sum_{j=1}^{N_2} \beta_{3,i,j} + \sum_{j=1}^{N_2} \sigma_{3,i,j} \geq 0, \qquad \forall i, \forall j$$

$$\frac{\zeta_i}{\alpha_2}\pi_{i,j} - \lambda \geq 0, \qquad \forall i, \forall j$$

$$\beta_2 - \zeta_i = 0, \qquad \forall i$$

$$\zeta_i - \sum_{j=1}^{N_2} \lambda_{i,j} = 0, \qquad \forall i$$

$$\zeta_i^T(h_i - \bar{t}_i - \frac{1}{\alpha}\sum_{j=1}^{N_2}\pi_{i,j}Y_{i,j}) = 0, \qquad \forall i$$

$$\lambda_{i,j}^T(Y_{i,j} - c_l^T L_{3,i,j} - c_{3,i,j}^T G_{3,i,j} + \bar{t}_i) = 0, \qquad \forall i, \forall j$$

$$\sigma_{2,i}^T(\bar{v} - V_{3,i}) = 0, \qquad \forall i, \forall j$$

$$\theta_{2,i}^T(\bar{q} - Q_{2,i}) = 0, \qquad \forall i$$

$$\delta_{2,i}^T(Q_{2,i} + S_{2,i} - \underline{q}) = 0, \qquad \forall i$$

$$\mu_{2,i}^T(\bar{g} - G_{2,i}) = 0, \qquad \forall i$$

$$\beta_{3,i,j}^T(V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j})) = 0, \qquad \forall i, \forall j$$

$$\sigma_{3,i,j}^T(\bar{v} - V_{3,i} - a_{3,i,j} - M(Q_{3,i,j} - S_{3,i,j})) = 0, \qquad \forall i, \forall j$$

$$\theta_{3,i,j}^T(\bar{q} - Q_{3,i,j}) = 0, \qquad \forall i, \forall j$$

$$\delta_{3,i,j}^T(Q_{3,i,j} + S_{3,i,j} - \underline{q}) = 0, \qquad \forall i, \forall j$$

$$\mu_{3,i,j}^T(\bar{g} - G_{3,i,j}) = 0, \qquad \forall i$$

$$G_{2,i}^T(c_{2,i} - \mathbf{1}\alpha_{2,i} + \mu_{2,i}) = 0, \qquad \forall i$$

$$G_{3,i,j}^T((1-\beta_2)\pi_{i,j}c_{3,i,j} + c_{3,i,j}\lambda_{i,j} - \alpha_{3,i,j} + \mu_{3,i,j}) = 0, \qquad \forall i, \forall j$$

$$L_{2,i}^T(c_l - \alpha_{2,i}) = 0, \qquad \forall i$$

$$L_{3,i,j}^T((1-\beta_2)\pi_{i,j}c_l + c_l\lambda_{i,j} - \alpha_{3,i,j}) = 0, \qquad \forall i, \forall j$$

$$Q_{2,i}^T(-\rho\alpha_{2,i} - M^T\beta_{2,i} + \theta_{2,i} - \delta_{2,i}) = 0, \qquad \forall i$$

$$S_{2,i}^T(-M^T\beta_{2,i} - \delta_{2,i}) = 0, \qquad \forall i$$

$$Q_{3,i,j}^T(-\rho\alpha_{3,i,j} - M^T\beta_{3,i,j} + M^T\sigma_{3,i,j} + \theta_{3,i,j} - \delta_{3,i,j}) = 0, \qquad \forall i, \forall j$$

$$S_{3,i,j}^T(-M^T\beta_{3,i,j} + M^T\sigma_{3,i,j} - \delta_{3,i,j}) = 0, \qquad \forall i, \forall j$$

$$V_{3,i}^T(\beta_{2,i} + \sigma_{2,i} - \sum_{j=1}^{N_2}\beta_{3,i,j} + \sum_{j=1}^{N_2}\sigma_{3,i,j}) = 0, \qquad \forall i, \forall j$$

$$Y_{i,j}^T(\frac{\zeta_i}{\alpha_2}\pi_{i,j} - \lambda) = 0, \qquad \forall i, \forall j$$

$$\lambda_{i,j}, \alpha_{2,i}, \beta_{2,i}, \sigma_{2,i}, \theta_{2,i}, \delta_{2,i}, \mu_{2,i} \geq 0 \qquad \forall i$$

$$\alpha_{3,i,j}, \beta_{3,i,j}, \sigma_{3,i,j}, \theta_{3,i,j}, \delta_{3,i,j}, \mu_{3,i,j} \geq 0, \qquad \forall i, \forall j$$

$$L_1, G_1, X_{i,j}, Q_1, V_2, S_1 \geq 0, \qquad \forall i$$

$$L_{2,i}, L_{3,i,j}, G_{2,i}, G_{3,i,j}, Y_{i,j}, Q_{2,i}, S_{2,i}, Q_{3,i,j}, S_{3,i,j}, V_{3,i} \geq 0, \qquad \forall i, \forall j.$$

Using nested risk measure $\rho_1(Z_1+\rho_2(Z_2+Z_3))$, we also formulate the "OTC" version of this problem in (3.20).

OTC formulation of hydropower energy planning problem with $\text{BRM}^{\alpha}_{\beta}$

$$
\begin{aligned}
\min \quad & c_l^T L_1 + c_1^T G_1 + (1-\beta_1)\sum_{i=1}^{N_1} \pi_i \omega_i + \beta_1 k \\
\text{s.t.} \quad & k \geq \tau + \frac{1}{\alpha}\sum_{i=1}^{N_1} \pi_i \chi_i \\
& \chi_i \geq \omega_i - \tau, && \forall i \\
& \omega_i = c_l^T L_{2,i} + c_{2,i}^T G_{2,i} + (1-\beta_2)\big(\sum_{j=1}^{N_2} \pi_{i,j}(c_l^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j})\big) + \beta_2 h_i, && \forall i \\
& h_i \geq \bar{t}_i + \frac{1}{\alpha_2}\sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, && \forall i \\
& Y_{i,j} \geq c_l^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j} - \bar{t}_i, && \forall i, \forall j \\
& \rho^T Q_1 + \mathbf{1}^T G_1 + L_1 = d_1, \\
& V_2 = v_1 + a_1 + M(Q_1 + S_1), \\
& V_2 \leq \bar{v}, \\
& Q_1 \leq \bar{q}, \\
& Q_1 + S_1 \geq \underline{q}, \\
& G_1 \leq \bar{g}, \\
& \rho^T Q_{2,i} + \mathbf{1}^T G_{2,i} + L_{2,i} = d_{2,i}, && \forall i \\
& V_{3,i} = V_2 + a_{2,i} + M(Q_{2,i} + S_{2,i}), && \forall i \\
& V_{3,i} \leq \bar{v}, && \forall i \\
& Q_{2,i} \leq \bar{q}, && \forall i \\
& Q_{2,i} + S_{2,i} \geq \underline{q}, && \forall i \\
& G_{2,i} \leq \bar{g}, && \forall i \\
& \rho^T Q_{3,i,j} + G_{3,i,j} + L_{3,i,j} = d_{3,i,j} && \forall i, \forall j
\end{aligned}
$$

$$V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \geq 0, \qquad \forall i, \forall j$$

$$V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \leq \bar{v}, \qquad \forall i, \forall j$$

$$Q_{3,i,j} \leq \bar{q}, \qquad \forall i, \forall j$$

$$Q_{3,i,j} + S_{3,i,j} \geq \underline{q}, \qquad \forall i, \forall j$$

$$G_{3,i,j} \leq \bar{g}, \qquad \forall i, \forall j$$

$$Y_{i,j} \geq c_l^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'}(c_l^T L_{3,i,j'} + c_{3,i,j'}^T G_{3,i,j'}), \qquad \forall i, \forall j$$

$$W_i = c_l^T L_{2,i} + c_{2,i}^T G_{2,i} + \sum_{j=1}^{N_2} \pi_{i,j}(c_l^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j}) + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, \quad \forall i$$

$$\chi_i \geq W_i - \sum_{i=1}^{N_1} \pi_i W_i, \qquad \forall i$$

$$L_1, G_1, \chi_i, Q_1, V_2, S_1 \geq 0, \qquad \forall i$$

$$L_{2,i}, L_{3,i,j}, G_{2,i}, G_{3,i,j}, Y_{i,j}, Q_{2,i}, S_{2,i}, Q_{3,i,j}, S_{3,i,j}, V_{3,i} \geq 0, \qquad \forall i, \forall j.$$

3.3.3 Numerical Study

We randomly generate 1000 instances with 6 thermal plants, 3 hydro plants, 5 second-stage scenarios, and in each second-stage scenario, 5 third-stage scenarios. For each numerical instance, we consider parameter values $\gamma = 0.3,\ 0.7$, and 0.9 for MSD_γ and $(\alpha, \beta) = (0.01, 0.5)$, $(0.05, 0.5)$, $(0.25, 0.5)$, $(0.05, 0.25)$, and $(0.25, 0.75)$ for BRM_β^α, giving a total of 8 risk measures.

Tables 3.29 and 3.30 summarize the results for instances with risk measure MSD_γ and BRM_β^α, respectively. With MSD risk measure, the average solution time were 0.48s and 0.02s for CTC and OTC models, respectively. For $\text{MSD}_{0.9}$, we observed an average Diff value of more than 10% in 0.6% of instances with a maximum value of 25.88%. The average solution time for model with BRM risk measure were 4.22s and 0.02s for CTC and OTC models, respectively, and the largest gap was 68.22% observed in the model with $\text{BRM}_{0.75}^{0.25}$. Overall, for all risk measures considered in this section, more than 90% of numerical instances resulted in Diff value of less than 0.01%. This observation suggests that for hydropower planning problem, OTC formulation can provide relatively a good heuristic for the CTC model. However, one should still keep in mind that the difference between the two models can be significant, particularly for high degrees of risk aversion.

	γ =0.3	γ =0.7	γ =0.9
Diff < 0.01%	93.90%	90.70%	90.40%
0.01% ≤ Diff < 0.1%	1.10%	1.10%	1.00%
0.1% ≤ Diff < 1%	3.50%	3.20%	3.60%
1% ≤ Diff < 10%	1.50%	4.80%	4.40%
10% ≤ Diff	0%	0.20%	0.60%
Max Diff	9.06%	14.19%	25.88 %

Table 3.29: Computational results of the hydropower planning problem with MSD_γ.

	$\beta = 0.5$ $\alpha = 0.01$	$\beta = 0.5$ $\alpha = 0.05$	$\beta = 0.75$ $\alpha = 0.05$	$\beta = 0.5$ $\alpha = 0.25$	$\beta = 0.75$ $\alpha = 0.25$
Diff < 0.01%	94.90%	94.90%	95.60%	91.40%	90.40%
0.01% ≤ Diff < 0.1%	0.10%	0.10%	0.20%	0.10%	0%
0.1% ≤ Diff < 1%	0.70%	1.00%	1.30%	1.40%	1.2%
1% ≤ Diff < 10%	3.00%	3.20%	1.40%	3.60%	4.00 %
10% ≤ Diff	1.30%	0.80%	1.5%	3.50%	4.4%
Max Diff	34.09%	26.34%	57.47%	46.89%	68.22%

Table 3.30: Computational results of the hydropower planning problem with BRM^α_β.

To study the impact of the choice of the risk measure on the deviation of the CTC model from the OTC formulation, for each risk measure, we selected the problem instances with the highest Diff value. Table 3.31 presents the Diff values of those 8 instances for different risk measures. For example, among all the numerical instances, the largest Diff value of BRM model with $(\alpha, \beta) = (0.05, 0.5)$ is 26.34%, which occurred at the problem instance **E**. At the same time, the Diff values of problem instance **E** are equal to zero for the rest of the risk measures. Table 3.31 suggests that even if a problem has a large Diff value for a particular risk measure, it can result in zero Diff value for another risk measure. As such, deviation of the CTC model from the OTC formulation strongly depends on the risk measure.

	γ =0.3	γ =0.7	γ =0.9	$\alpha = 0.01$ $\beta = 0.5$	$\alpha = 0.05$ $\beta = 0.5$	$\alpha = 0.05$ $\beta = 0.75$	$\alpha = 0.25$ $\beta = 0.5$	$\alpha = 0.25$ $\beta = 0.75$
A	9.06%	0%	0%	0%	0%	0%	0%	0%
B	0%	14.19%	8.43%	0%	0%	0%	0%	0%
C	0%	0%	25.88%	0%	0%	0%	4.45%	0%
D	0%	0%	0%	34.09%	0%	47.52%	0%	0%
E	0%	0%	0%	0%	26.34%	0%	0%	0%
F	0%	0%	0%	1.38%	0%	57.47%	0%	0%
G	0%	0%	0%	12.77%	0.41%	0%	46.89%	0%
H	0%	0%	0%	22.58%	0%	0%	0%	68.22%

Table 3.31: The Diff value for 5 hydropower planning instances with MSD_γ and $\mathrm{BRM}_\beta^\alpha$.

We conclude this chapter by noting that for three general examples, namely supply chain production planning, portfolio optimization, and hydropower energy planning, OTC and CTC formulations can yield significantly different solutions. However, this difference is problem specific and depends on the degree of risk aversion. Consequently, the OTC formulation cannot provide a good approximation for the CTC model. However, as was observed in our numerical study, solving the CTC model is time consuming. As such, proper algorithms should be utilized to solve the CTC model within a reasonable time. In the next chapter, we propose a bundle method approach to achieve this goal.

Chapter 4

Bundle Method for Bilevel Stochastic Programming

For the numerical examples considered in Chapter 3, both the OTC and CTC formulations were solved quickly with average solution time less than 1 Second. However, the OTC formulation was solved, on average, 10 times faster than the CTC model. Additionally, for slightly larger problems, the OTC model was still able to find the solution within a few seconds while the CTC model was unable to provide a feasible solution within 900 seconds. Therefore, using the CTC formulation for larger problems requires a specialized solution algorithm that exploits problem structure. As our CTC formulation posits bilevel structure, we focus on finding a fast method to solve a general bilevel stochastic program. After trying different methods, such as using semidefinite programming, we observed that a specialized proximal bundle method is a promising approach to find a quick and high quality lower bound for bilevel stochastic programs. Additionally, we introduce a procedure to find a feasible solution, which provides an upper bound for the problem. The combination of the lower and upper bounds enables us to quantify the optimality gap.

4.1 Bilevel Stochastic Programming Formulation

Consider the following bilevel stochastic program:

$$\begin{aligned} \min_{X_1, X_2} \quad & c_1^T X_1 + \textstyle\sum_{S \in \mathcal{E}_2} (c_2^S)^T X_2^S \\ \text{s.t.} \quad & A_{11} X_1 = b_1 \\ & X_1 \geq 0 \\ & X_2^S \in \mathcal{X}^*_{2,S}(X_1), \qquad \forall\, S \in \mathcal{E}_2 \end{aligned} \tag{4.1}$$

in which $\mathcal{E}_2$ is the set of second-stage scenarios, $c_1 \in \mathbb{R}^{n_1}$, $c_2 = (c_2^{S_1}, \cdots, c_2^{S_{|\mathcal{E}_2|}}) \in \mathbb{R}^{|\mathcal{E}_2|n_2}$, and $b_1 \in \mathbb{R}^{m_1}$ are given, $X_1 \in \mathbb{R}^{n_1}$, and $X_2 = (X_2^{S_1}, \cdots, X_2^{S_{|\mathcal{E}_2|}}) \in \mathbb{R}^{|\mathcal{E}_2|n_2}$ are the decision variables, and given X_1 and $S \in \mathcal{E}_2$, $\mathcal{X}^*_{2,S}(X_1)$ is the optimal solution set of the second-stage problem:

$$\begin{aligned} \min_{X_2^S} \quad & (d_2^S)^T X_2^S \\ \text{s.t.} \quad & A_{22}^S X_2^S = b_2^S - A_{21}^S X_1 \\ & X_2^S \geq 0, \end{aligned} \tag{4.2}$$

with $d_2^S \in \mathbb{R}^{n_2}$ and $b_2^S \in \mathbb{R}^{m_2}$. Note that all third-stage decision variables X_3, can be modeled as second-stage decision variables within (4.2) by defining $X_{2,\text{new}}^S = \begin{pmatrix} X_2^S \\ X_3^S \end{pmatrix}$ and $A_{22,\text{new}}^S = \begin{bmatrix} A_{22}^S & 0 \\ A_{32}^S & A_{33}^S \end{bmatrix}, \forall S \in \mathcal{E}_2$.

4.2 LPCC Formulation

Now, we follow the same procedure that we used for CTC formulation of each practical problem with MSD and BRM risk measures to convert them into the LPCC formulation. Replace $X_2^S \in \mathcal{X}^*_{2,S}(X_1)$ in (4.1) with the KKT optimality conditions of second-stage problems to convert the bilevel formulation (4.1)-(4.2) into a single level problem. Define $\lambda^S \in \mathbb{R}^{m_2}$ as the Lagrange multiplier of equality constraint in (4.2); the Lagrange function is:

$$L(X_2^S, \lambda^S) = (d_2^S)^T X_2^S - (\lambda^S)^T (A_{21}^S X_1^S + A_{22}^S X_2^S - b_2^S).$$

So the KKT optimality conditions are:

$$\begin{aligned} & d_2^S - (A_{22}^S)^T \lambda^S \geq 0 \\ & (X_2^S)^T (d_2^S - (A_{22}^S)^T \lambda^S) = 0. \end{aligned}$$

Now, we obtain the following equivalent LPCC formulation:

$$\begin{aligned} F = \min_{X_1, X_2, \lambda} \quad & c_1^T X_1 + \sum_{S \in \mathcal{E}_2} (c_2^S)^T X_2^S \\ \text{s.t.} \quad & A_{11} X_1 = b_1 \\ & A_{21}^S X_1 + A_{22}^S X_2^S = b_2^S, \quad \forall S \in \mathcal{E}_2 \end{aligned} \tag{4.3}$$

$$d_2^S - (A_{22}^S)^T \lambda^S \geq 0, \qquad \forall\, S \in \mathcal{E}_2$$

$$(X_2^S)^T (d_2^S - (A_{22}^S)^T \lambda^S) = 0, \quad \forall\, S \in \mathcal{E}_2$$

$$X_1 \geq 0$$

$$X_2^S \geq 0, \qquad \forall\, S \in \mathcal{E}_2$$

Next, we discuss our method to find a lower bound for the LPCC formulation.

4.3 Lower Bound for LPCC Formulation

First, we aim to break problem (4.3) into $|\mathcal{E}_2|$ smaller subproblems that can be solved efficiently and provide us with a lower bound on the optimal objective. Let $Y = (Y^{S_1}, \ldots, Y^{S_{|\mathcal{E}_2|}}) \in \mathbb{R}^{n_1 \times |\mathcal{E}_2|}$. Consider the following optimization problem for each $S \in \mathcal{E}_2$:

$$\begin{aligned} f^S(Y) = \quad & \min_{X_1^S, X_2^S, \lambda^S} && \tfrac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y^S \rangle \\ & \text{s.t.} && A_{11} X_1^S = b_1 \\ &&& A_{21}^S X_1^S + A_{22}^S X_2^S = b_2^S \\ &&& d_2^S - (A_{22}^S)^T \lambda^S \geq 0 \\ &&& (X_2^S)^T (d_2^S - (A_{22}^S)^T \lambda^S) = 0 \\ &&& X_1^S,\ X_2^S \geq 0. \end{aligned} \tag{4.4}$$

Lemma 5 *Function $f^S(Y)$ is concave on $\mathbb{R}^{n_1 \times |\mathcal{E}_2|}$.*

Proof: First, note that the feasible set of optimization problem (4.4) is the same for any $Y \in \mathbb{R}^{n_1 \times |\mathcal{E}_2|}$ and call it $\mathbb{X}^S$. For any $Y_1, Y_2 \in \mathbb{R}^{n_1 \times |\mathcal{E}_2|}$, and $\alpha \in [0, 1]$:

$$\begin{aligned} f^S(\alpha Y_1 + (1-\alpha) Y_2) &= \min_{(X_1^S, X_2^S, \lambda^S) \in \mathbb{X}^S} \quad \frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, \alpha Y_1^S + (1-\alpha) Y_2^S \rangle \\ &= \min_{(X_1^S, X_2^S, \lambda^S) \in \mathbb{X}^S} \quad \{\alpha(\frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y_1^S \rangle) \\ &\qquad + (1-\alpha)(\frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y_2^S \rangle)\} \\ &\geq \min_{(X_1^S, X_2^S, \lambda^S) \in \mathbb{X}^S} \quad \alpha(\frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y_1^S \rangle) \\ &\quad + \min_{(X_1^S, X_2^S, \lambda^S) \in \mathbb{X}^S} \quad (1-\alpha)(\frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y_2^S \rangle) \\ &= \alpha \min_{(X_1^S, X_2^S, \lambda^S) \in \mathbb{X}^S} \quad (\frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y_1^S \rangle) \end{aligned}$$

$$+ (1-\alpha) \min_{(X_1^S, X_2^S, \lambda^S) \in \mathbb{X}^S} \left(\frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y_2^S \rangle\right)$$

$$= \alpha f^S(Y_1) + (1-\alpha) f^S(Y_2),$$

so f^S is concave. ■

Let K be the linear subspace $\{(z^1, \cdots, z^{|\mathcal{E}_2|}) \mid z^1, \cdots, z^{|\mathcal{E}_2|} \in \mathbb{R}^{n_1}, z^i = z^j \ \forall i, j = 1, \cdots, |\mathcal{E}_2|\}$ and $K^* := \{(y^1, \cdots, y^{|\mathcal{E}_2|}) \mid y^1, \cdots, y^{|\mathcal{E}_2|} \in \mathbb{R}^{n_1}, y^1 + \cdots + y^{|\mathcal{E}_2|} = 0\}$ be its orthogonal complement within $\mathbb{R}^{n_1 \times |\mathcal{E}_2|}$.

Lemma 6 *For all $Y \in K^*$, we have $\sum_{S \in \mathcal{E}_2} f^S(Y) \equiv f(Y) \leq F$.*

Proof: First, obtain the scenario formulation of F in (4.3) by replicating $|\mathcal{E}_2|$ copies of X_1, indexed $X_1^{S_1}, X_1^{S_2}, \ldots, X_1^{S_{|\mathcal{E}_2|}}$, and add the nonanticipativity constraint $\mathcal{X}_1 := (X_1^{S_1}, \ldots, X_1^{S_{|\mathcal{E}_2|}}) \in K$. This transformation results in the following equivalent reformulation of (4.3):

$$\begin{array}{lll}
\min_{\mathcal{X}_1, X_2, \lambda} & \sum_{S \in \mathcal{E}_2} \frac{c_1^T}{|\mathcal{E}_2|} X_1^S + \sum_{S \in \mathcal{E}_2} (c_2^S)^T X_2^S & \\
\text{s.t.} & A_{11} X_1^S = b_1, & \forall\, S \in \mathcal{E}_2 \\
& A_{21}^S X_1^S + A_{22}^S X_2^S = b_2^S, & \forall\, S \in \mathcal{E}_2 \\
& d_2^S - (A_{22}^S)^T \lambda^S \geq 0, & \forall\, S \in \mathcal{E}_2 \\
& (X_2^S)^T (d_2^S - (A_{22}^S)^T \lambda^S) = 0, & \forall\, S \in \mathcal{E}_2 \\
& \mathcal{X}_1 = (X_1^{S_1}, \ldots, X_1^{S_{|\mathcal{E}_2|}}) \in K & \\
& X_1^S \geq 0 & \\
& X_2^S \geq 0, & \forall\, S \in \mathcal{E}_2.
\end{array} \tag{4.5}$$

This problem is clearly equivalent to (4.3). Suppose $(\mathcal{X}_1^*, X_2^*)$ is an optimal solution for problem (4.5). Consequently, the tuple (X_1^*, X_2^*) with $X_1^* := \sum_{S \in \mathcal{E}_2} \frac{1}{|\mathcal{E}_2|} X_1^{S^*}$ and $X_2^* := (X_2^{S_1^*}, \ldots, X_2^{S^*_{|\mathcal{E}_2|}})$ is an optimal solution to problem (4.3). We have

$$\begin{aligned}
F &= c_1^T X_1^* + \textstyle\sum_{S \in \mathcal{E}_2} (c_2^S)^T X_2^{S^*} \\
&= \textstyle\sum_{S \in \mathcal{E}_2} \left(\frac{1}{|\mathcal{E}_2|} c_1^T X_1^{S^*} + (c_2^S)^T X_2^{S^*}\right) \\
&= \textstyle\sum_{S \in \mathcal{E}_2} \left(\frac{1}{|\mathcal{E}_2|} c_1^T X_1^{S^*} + (c_2^S)^T X_2^{S^*}\right) + \langle \mathcal{X}_1^*, Y \rangle \\
&= \textstyle\sum_{S \in \mathcal{E}_2} \left(\frac{1}{|\mathcal{E}_2|} c_1^T X_1^{S^*} + (c_2^S)^T X_2^{S^*} + \langle X_1^{S^*}, Y^S \rangle\right) \\
&\geq \textstyle\sum_{S \in \mathcal{E}_2} f^S(Y) = f(Y),
\end{aligned}$$

where the third equality follows since K and K^* are orthogonal complements, with $\mathcal{X}_1^* = (X_1^{S_1^*}, \ldots, X_1^{S^*_{|\mathcal{E}_2|}}) \in K$ and $Y \in K^*$. ■

This lemma equips us with a relatively fast method to find a lower bound for the original CTC formulation: first, choose any $Y \in K^*$ (e.g., $Y = \mathbf{0}$). Then, solve subproblems (4.4) for all $S \in \mathcal{E}_2$. Finally, add the optimal value of the objective functions to find the lower bound.

Also, observe that $f(Y)$ is the Lagrangian relaxation of problem (4.5) obtained by dualizing the nonanticipativity constraint $\mathcal{X}_1 \in K$. To see this, let $\mathbb{X}^S$ be the feasible set of problem (4.4) and $\mathbb{X} = \{(\mathcal{X}_1, X_2, \lambda) | (X_1^S, X_2^S, \lambda^S) \in \mathbb{X}^S, \ \forall S \in \mathcal{E}_2\}$. Note that $\mathbb{X}$ is the feasible set of all (4.5) constraints except for the nonanticipativity constraint. Then, Lagrangian relaxation of problem (4.5) obtained by dualizing $\mathcal{X}_1 \in K$ is:

$$
\begin{aligned}
&\min_{(\mathcal{X}_1, X_2, \lambda) \in \mathbb{X}} \textstyle\sum_{S \in \mathcal{E}_2} \left(\frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S \right) + \langle \mathcal{X}_1, Y \rangle \\
&= \min_{(\mathcal{X}_1, X_2, \lambda) \in \mathbb{X}} \textstyle\sum_{S \in \mathcal{E}_2} \left(\frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y^S \rangle \right) \\
&= \textstyle\sum_{S \in \mathcal{E}_2} \min_{(X_1^S, X_2^S, \lambda^S) \in \mathbb{X}^S} \frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y^S \rangle \\
&= \textstyle\sum_{S \in \mathcal{E}_2} f^S(Y) = f(Y).
\end{aligned}
$$

Therefore, $f(Y)$ is the Lagrangian function and $\max_{Y \in K^*} f(Y)$ is the Lagrangian dual problem of (4.5). Accordingly, to find a lower bound, it is sufficient to solve this dual problem. In the next section, we use a form of the bundle method discussed in Ruszczyński (2006) and de Oliveira & Eckstein (2015) to solve this maximization problem. Ruszczyński (2006) gives a general overview of bundle methods, while de Oliveira & Eckstein (2015) describes a bundle method for problems with additive structure. Van Ackooij & Frangioni (2018) describe an algorithm subsuming the one the technical report of de Oliveira & Eckstein (2015). Also, Dempe (2000) and Dempe & Bard (2001) apply bundle methods to bilevel programming, but without exploiting additive scenario structure in the follower problem, as proposed here. For other variations and extensions of the bundle method also see Kiwiel (1991), Frangioni (2002), and Belloni & Sagastizábal (2009). We refer the interested reader to Mäkelä (2002) for a survey on the bundle method.

4.4 The Proximal Bundle Method

In this section, we present a bundle method to find a lower and upper bound for the optimization problem $\max_{Y\in K^*} f(Y)$. To achieve this, first, we construct a piecewise linear approximation of $f(Y)$ by using subgradient inequalities.

Lemma 7 *Suppose* $(X_1^{S^*}, X_2^{S^*})$ *is the optimal solution for problem* (4.4) *at* $Z \in K^*$. *Then* $X_1^{S^*}$ *is a subgradient of* f^S.

Proof: Choose an arbitrary $Y \in K^*$.

$$\begin{aligned} f^S(Y) &\le \tfrac{c_1^T}{|\mathcal{E}_2|} X_1^{S^*} + (c_2^S)^T X_2^{S^*} + \langle X_1^{S^*}, Y^S\rangle \\ &= \tfrac{c_1^T}{|\mathcal{E}_2|} X_1^{S^*} + (c_2^S)^T X_2^{S^*} + \langle X_1^{S^*}, Z^S\rangle + \langle X_1^{S^*}, Y^S - Z^S\rangle \\ &= f^S(Z) + \langle X_1^{S^*}, Y^S - Z^S\rangle. \end{aligned}$$

where the first inequality follows from the definition of $f^S(Y)$ and the last equality follows from the lemma's assumption. ■

For each $S \in \mathcal{E}_2$ and iteration $k \ge 1$, define a set of indices $\mathcal{B}_k^S \subseteq \{1, \cdots, k\}$ and call it the *bundle*. From Lemma 7,

$$\hat{f}_k^S(Y) \equiv \min_{j\in\mathcal{B}_k^S} \{f^S(Y_j) + \langle X_{1j}^{S^*}, Y^S - Y_j^S\rangle\} \tag{4.6}$$

is a piecewise linear approximation of $f^S(Y)$ with $(X_{1j}^{S^*}, X_{2j}^{S^*})$ the optimal solution of $f^S(Y_j)$. Let $\hat{Y}_k$, also called a *center*, to be the solution of the master problem at iteration k. In order to obtain the next point $Y_{k+1} \in K^*$, we solve the following ***master*** problem.

$$\begin{aligned} \max_{r,Y} \quad & \sum_{S\in\mathcal{E}_2} r^S - \frac{1}{2t_k} \left\| Y - \hat{Y}_k \right\|^2 \\ \text{s.t.} \quad & r^S \le f^S(Y_j) + \langle X_{1j}^{S^*}, Y^S - Y_j^S\rangle, \ \forall S \in \mathcal{E}_2, \ j \in \mathcal{B}_k^S \\ & Y = (Y^{S_1}, \cdots, Y^{S_{|\mathcal{E}_2|}}) \in K^*. \end{aligned} \tag{4.7}$$

The quadratic term and the parameter t_k in the objective function of (4.7) are responsible for keeping Y_{k+1} close enough to the current center. Let (r^*, Y_{k+1}) be the optimal solution of (4.7). For each $S \in \mathcal{E}_2$, define:

$$\begin{aligned}
&I_k^S = \{j \mid r^{S^*} = f^S(Y_j) + \langle X_{1j}^{S^*}, Y_{k+1}^S - Y_j^S \rangle,\ j \in \mathcal{B}_k^S\},\\
&f_k^m(Y_{k+1}) = \sum_{S \in \mathcal{E}_2} r^{S^*},\\
&\delta_k = f_k^m(Y_{k+1}) - f(\hat{Y}_k),\\
&\gamma_k = f(\hat{Y}_k) + \kappa \delta_k,\ \kappa \in (0,1).
\end{aligned} \tag{4.8}$$

Here I_k^S is the active set of first constraint in (4.7) for scenario S, $f_k^m(Y_{k+1})$ is the current estimate of $f(Y_{k+1})$, δ_k is the maximum improvement of the objective function value from the current best estimation by choosing Y_{k+1} as the new center, and γ_k is the target value for the next best estimation. Note that $\gamma_k = (1-\kappa) f(\hat{Y}_k) + \kappa f_k^m(Y_{k+1})$.

Next, we calculate $f(Y_{k+1})$. There might exist $Y \in K^*$ such that some of subproblems (4.4) are unbounded. To determine which subproblems cannot be solved to the optimality at Y_{k+1}, for each scenario S define ψ_{k+1}^S equal to 1 when $f^S(Y_{k+1})$ is unbounded and equal to 0 otherwise. Consequently, $\max_{S \in \mathcal{E}_2}\{\psi_{k+1}^S\} = 1$ means that we cannot calculate $f(Y_{k+1})$ and must keep the center the same as the previous step, $\hat{Y}_{k+1} = \hat{Y}_k$. When $\max_{S \in \mathcal{E}_2}\{\psi_{k+1}^S\} = 0$, we have the value of $f(Y_{k+1})$ and we are able to compare $f(Y_{k+1})$ with γ_k. If $f(Y_{k+1}) \geq \gamma_k$, set the next center point $\hat{Y}_{k+1} = Y_{k+1}$ (**serious step**); otherwise $\hat{Y}_{k+1} = \hat{Y}_k$ (**null step**).

It is obvious that as we add more cuts to the master problem, we have a better approximation for our problem. However, as the number of cuts get larger, it takes more time to solve (4.7). From de Oliveira & Eckstein (2015), we consider the bundle for iteration $k+1$ for scenario $S \in \mathcal{E}_2$ as follow:

$$\mathcal{B}_{k+1}^S = \begin{cases} I_k^S \cup \{k+1\}, & \text{if } \psi_{k+1}^S = 0 \\ I_k^S, & \text{otherwise.} \end{cases} \tag{4.9}$$

We present our bundle method procedure in Algorithm 1, but before that, we present our method to find a feasible solution for problem (4.3) at each iteration of our proximal bundle method.

4.4.1 Feasible Upper Bound

Before finding a feasible solution, first, we reformulate (4.7) and then determine the corresponding dual problem of the master problem. For $S \in \mathcal{E}_2$ and $j \in \mathcal{B}_k^S$, define the linearization error e_j^S at the center point $\hat{Y}_k$ as:

$$e_j^S = [f^S(Y_j) + \langle X_{1j}^{S^*}, \hat{Y}_k^S - Y_j^S \rangle] - f^S(\hat{Y}_k), \quad \forall S \in \mathcal{E}_2,\ j \in \mathcal{B}_k^S. \tag{4.10}$$

Next, we have:

$$\begin{aligned} f^S(Y_j) + \langle X_{1j}^{S^*}, Y^S - Y_j^S \rangle \quad &= f^S(Y_j) + \langle X_{1j}^{S^*}, Y^S - Y_j^S \rangle + e_j^S - f^S(Y_j) \\ &\quad - \langle X_{1j}^{S^*}, \hat{Y}_k^S - Y_j^S \rangle + f^S(\hat{Y}_k) \\ &= f^S(\hat{Y}_k) + e_j^S + \langle X_{1j}^{S^*}, Y^S - Y_j^S - \hat{Y}_k^S + Y_j^S \rangle \\ &= f^S(\hat{Y}_k) + e_j^S + \langle X_{1j}^{S^*}, Y^S - \hat{Y}_k \rangle \end{aligned} \tag{4.11}$$

Now, we can reformulate the master problem (4.7) as:

$$\begin{aligned} \max_{r,Y} \quad & \sum_{S \in \mathcal{E}_2} r^S - \frac{1}{2t_k} \left\| Y - \hat{Y}_k \right\|^2 \\ \text{s.t.} \quad & r^S \le f^S(\hat{Y}_k) + e_j^S + \langle X_{1j}^{S^*}, Y^S - \hat{Y}_k^S \rangle, \quad \forall S \in \mathcal{E}_2,\ j \in \mathcal{B}_k^S \\ & Y = (Y^{S_1}, \cdots, Y^{S_{|\mathcal{E}_2|}}) \in K^* \end{aligned} \tag{4.12}$$

in which the first constraints follow from (4.11). Introducing α_j^S and β as the Lagrange multipliers of first and second constraints in problem (4.12), respectively; the Lagrangian function is:

$$\begin{aligned} L(r, Y, \alpha, \beta) = & \sum_{S \in \mathcal{E}_2} r^S - \frac{1}{2t_k} \left\| Y - \hat{Y}_k \right\|^2 - \sum_{S \in \mathcal{E}_2} \sum_{j \in \mathcal{B}_k^S} \alpha_j^S \big(r^S - f^S(\hat{Y}_k) \\ & - e_j^S - \langle X_{1j}^{S^*}, Y^S - \hat{Y}_k^S \rangle \big) - \sum_{S \in \mathcal{E}_2} \beta Y^S \\ = & \sum_{S \in \mathcal{E}_2} r^S (1 - \sum_{j \in \mathcal{B}_k^S} \alpha_j^S) + \sum_{S \in \mathcal{E}_2} \Big[- \frac{1}{2t_k} \left\| Y^S - \hat{Y}_k^S \right\|^2 \\ & + \sum_{j \in \mathcal{B}_k^S} \alpha_j^S \langle X_{1j}^{S^*}, Y^S - \hat{Y}_k^S \rangle - \beta Y^S \Big] + \sum_{S \in \mathcal{E}_2} \sum_{j \in \mathcal{B}_k^S} \alpha_j^S \big(f^S(\hat{Y}_k) + e_j^S \big). \end{aligned}$$

At the optimal point (r^*, Y^*), we have:

$$\begin{aligned} \nabla_{r^{S*}} L &= (1 - \sum_{j \in \mathcal{B}_k^S} \alpha_j^S) = 0 \\ \nabla_{Y^{S*}} L &= -\frac{1}{t_k} (Y^{S^*} - \hat{Y}_k^S) + \sum_{j \in \mathcal{B}_k^S} \alpha_j^S X_{1j}^{S^*} - \beta = 0, \end{aligned}$$

for all $S \in \mathcal{E}_2$. From the last equation, we have:

$$Y^{S^*} = \hat{Y}_k^S + \sum_{j \in \mathcal{B}_k^S} t_k \alpha_j^S X_{1j}^{S^*} - t_k \beta.$$

Therefore,

$$\begin{aligned}
\max_{r,Y} L(r, Y, \alpha, \beta) \quad &= \sum_{S \in \mathcal{E}_2} -\frac{1}{2t_k} \left\| \sum_{j \in \mathcal{B}_k^S} t_k \alpha_j^S X_{1j}^{S^*} - t_k \beta \right\|^2 \\
&+ \sum_{S \in \mathcal{E}_2} \sum_{j \in \mathcal{B}_k^S} \alpha_j^S \langle X_{1j}^{S^*}, \sum_{j \in \mathcal{B}_k^S} t_k \alpha_j^S X_{1j}^{S^*} - t_k \beta \rangle \\
&- \sum_{S \in \mathcal{E}_2} \beta(\hat{Y}_k^S + \sum_{j \in \mathcal{B}_k^S} t_k \alpha_j^S X_{1j}^{S^*} - t_k \beta) + \sum_{S \in \mathcal{E}_2} \sum_{j \in \mathcal{B}_k^S} \alpha_j^S \big(f^S(\hat{Y}_k) + e_j^S\big) \\
&= \sum_{S \in \mathcal{E}_2} \frac{t_k}{2} \langle \sum_{j \in \mathcal{B}_k^S} \alpha_j^S X_{1j}^{S^*} + \beta, \sum_{j \in \mathcal{B}_k^S} \alpha_j^S X_{1j}^{S^*} - \beta \rangle \\
&- \sum_{S \in \mathcal{E}_2} \beta(\hat{Y}_k^S + \sum_{j \in \mathcal{B}_k^S} t_k \alpha_j^S X_{1j}^{S^*} - t_k \beta) + \sum_{S \in \mathcal{E}_2} \sum_{j \in \mathcal{B}_k^S} \alpha_j^S \big(f^S(\hat{Y}_k) + e_j^S\big) \\
&= \sum_{S \in \mathcal{E}_2} \frac{t_k}{2} (\sum_{j \in \mathcal{B}_k^S} X_{1j}^{S^*} \alpha_j^S - \beta)^2 - \sum_{S \in \mathcal{E}_2} \hat{Y}_k^S \beta \\
&+ \sum_{S \in \mathcal{E}_2} \sum_{j \in \mathcal{B}_k^S} \big(f^S(\hat{Y}_k) + e_j^S\big) \alpha_j^S
\end{aligned}$$

The dual of master problem is:

$$\begin{aligned}
\min_{\alpha,\beta} \quad & \sum_{S \in \mathcal{E}_2} \frac{t_k}{2} (\nu^S)^2 - \sum_{S \in \mathcal{E}_2} \hat{Y}_k^S \beta + \sum_{S \in \mathcal{E}_2} \sum_{j \in \mathcal{B}_k^S} \big(f^S(\hat{Y}_k) + e_j^S\big) \alpha_j^S \\
\text{s.t.} \quad & \sum_{j \in \mathcal{B}_k^S} \alpha_j^S = 1, \ \forall S \in \mathcal{E}_2 \\
& \nu^S = \sum_{j \in \mathcal{B}_k^S} X_{1j}^{S^*} \alpha_j^S - \beta, \ \forall S \in \mathcal{E}_2 \\
& \alpha_j^S \geq 0, \ \forall S \in \mathcal{E}_2, j \in \mathcal{B}_k^S.
\end{aligned} \tag{4.13}$$

In the master problem (4.12), constraint $Y \in K^*$ is the dual of constraint $\mathcal{X}_1 \in K$ in problem (4.5). Also, note that in the master dual problem (4.13), β is the Lagrange multiplier of constraint $Y \in K^*$. Accordingly, we can choose β as a candidate point for X_1. Additionally, we add the first and second stage linear constraints (4.3) to the dual master problem (4.13), to guarantee that β satisfies these constraints. This procedure, gives us the following optimization problem:

$$\begin{aligned}
\min_{\alpha,\beta,X_2} \quad & \sum_{S\in\mathcal{E}_2} \frac{t_k}{2}(\nu^S)^2 - \sum_{S\in\mathcal{E}_2} \hat{Y}_k^S \beta + \sum_{S\in\mathcal{E}_2} \sum_{j\in\mathcal{B}_k^S} \left(f^S(\hat{Y}_k) + e_j^S\right)\alpha_j^S \\
\text{s.t.} \quad & \sum_{j\in\mathcal{B}_k^S} \alpha_j^S = 1,\ \forall S \in \mathcal{E}_2 \\
& \nu^S = \sum_{j\in\mathcal{B}_k^S} X_{1j}^{S^*} \alpha_j^S - \beta,\ \forall S \in \mathcal{E}_2 \\
& \alpha_j^S \geq 0,\ \forall S \in \mathcal{E}_2,\ j \in \mathcal{B}_k^S \\
& A_{11}\beta = b_1 \\
& A_{21}^S \beta + A_{22}^S X_2^S = b_2^S,\ \forall S \in \mathcal{E}_2.
\end{aligned} \tag{4.14}$$

Let β^* be the optimal value of β in the above problem. Set $X_1 = \beta^*$ and solve all the second-stage problems (4.2) to find an optimal point $X_2^{S^*} \in \mathcal{X}_{2,S}^*(\beta^*)$. Note that (β^*, X_2^*) is a feasible point for problem (4.3) and by plugging it in the objective function, we obtain an upper bound for problem (4.3). Denote the value of this upper bound in iteration k by U_k. We use this upper bound in the algorithm proposed in the next section to compute an optimality gap. This upper bound is particularly valuable when the LPCC formulation fails to provide a feasible solution within a reasonable time limit. In this case, our upper bound, can provide a reasonable course of action to the DM.

4.4.2 Specialized Bundle Method Algorithm

Let $f(\hat{Y}_k)$ and U_k be the lower and upper bounds discussed in the previous sub-sections, respectively, in each iteration k. Also, let $L = \min_{l\in\{1,\ldots,k\}} f(\hat{Y}_l)$ be the best lower bound found by iteration k. Similarly, define $U = \max_{l\in\{1,\ldots,k\}} U_l$ to be the best upper bound found by iteration k.

Let $\tau, \epsilon \geq 0$ be small thresholds. The algorithm stops if any of the following four criteria is met:

(i) $(f_k^m(Y_{k+1}) - L)/L \leq \tau$

(ii) $(U - L)/L \leq \tau$

(iii) $||Y_{k+1} - Y_k||_\infty \leq \epsilon$

(iv) $||Y_{k+1} - Y_{k-1}||_\infty \leq \epsilon$, for $k \geq 1$.

Therefore, the algorithm stops if the estimated value of the next iteration is sufficiently close to the lower bound (in which case it approximates the optimal value to the desired level); or if the upper bound is sufficiently close to the lower bound (in which case the feasible solution is near optimal); or if the steps are not sufficiently large (potential improvements are very slow in each step); or if it is "stuck" in a two-iteration loop. Denote the set that is parameterized by these stopping criteria by **STOP**.

The following algorithm summarizes our proximal bundle method:

Algorithm 1 Proximal Bundle Method

Step 0. Let $t_0,\ \tau > 0,\ k = 0,\ \kappa \in (0,1),\ Y_0 \in K^*$, and $\hat{Y}_0 = Y_0$. Compute $f^S(\hat{Y}_0)$ and e_0^S. Set $L = \sum_{S \in E_2} f^S(\hat{Y}_0),\ U = +\infty$, and $\mathcal{B}_0^S = \{0\},\ \forall S \in \mathcal{E}_2$.

Step 1. Solve the master problem (4.12) and set (r^*, Y_{k+1}) as its optimal solution.

Step.2 Solve problem (4.14). Compute U_k. If $U_k < U,\ U = U_k$.

Step 3. Compute $f_k^m(Y_{k+1}),\ \delta_k$, and γ_k based on (4.8).

Step 4. If **STOP**, Return L and U.

Step 5. Calculate $f(Y_{k+1})$ and $\psi_{k+1}^S, \forall S \in \mathcal{E}_2$.

Step 6. If $\max_{S \in \mathcal{E}_2} \psi_{k+1}^S = 0$ and $f(Y_{k+1}) \geq \gamma_k$, set $L = \sum_{S \in E_2} f^S(\hat{Y}_{k+1})$, $\hat{Y}_{k+1} = Y_{k+1}$ and choose $t_{k+1} \geq t_k$, otherwise $\hat{Y}_{k+1} = \hat{Y}_k$ and $t_{k+1} = t_k$.

Step 7. Update $\mathcal{B}_{k+1}^S,\ \forall S \in E_2$ based on (4.9).

Step 8. Set $k = k + 1$, and go to **Step 1**.

4.5 Bundle Method Computational Results

In this section, we numerically validate the performance of Algorithm 1 for general stochastic bilevel problem (4.1)-(4.2). Consider the notation in Table 4.1. We studied six problem categories of different sizes, as summarized in Table 4.2. For each class, we randomly generated 100 numerical instances, resulting in a total of 600 instances. Among these instances, there were several infeasible cases, which were removed from the numerical sets. This left us with 395 feasible instances to study.

Table 4.1: **Notations**

n_1	Dimension of first-stage decision variable X_1
m_1	Number of first-stage equality constraints in problem (4.1)
$\lvert\mathcal{E}_2\rvert$	Number of second-stage problems
n_2	Dimension of second-stage decision variable $X_2^S,\ S \in \mathcal{E}_2$
m_2	Number of second-stage equality constraints in problem (4.2)
$\mathcal{L}$	Lower bound obtained by Gurobi solver for the LPCC model
L	Lower bound obtained by the bundle method for f function
U	Upper bound obtained by the bundle method for the LPCC model
AD_1	Average of $\frac{L-\mathcal{L}}{\mathcal{L}}$
AD_2	Average of $\frac{U-L}{L}$

We solved each numerical instance by implementing Algorithm 1 in Python on a workstation with Intel Xeon E5-2683 v4 2.10GHz CPUs and 256 GiB memory. To benchmark the performance of our algorithm, we also used gurobipy to solve the LPCC formulation (4.3) for each numerical instance. For both our algorithm and the benchmark, we limited the run time to 300s for each instance. If an instance could not be solved to the optimality in the LPCC formulation, then the best lower bound found by Gurobi was chosen as the benchmark. Table 4.2 reports the computational results.

Gurobi was not able to solve any of the instances to the optimality for the LPCC formulation. However, for a majority of the cases, the bundle method stopped within the time threshold with an average of 152s over all the instances, and a much smaller average for smaller problems (Classes I, II, and IV).

Also, we observe that our bundle method provides better lower bounds than gurobipy: overall, on average, our algorithm provided at least 5.13% better lower bound than gurobipy, measured by $\frac{L-\mathcal{L}}{\mathcal{L}}$. Notably, this gap was on average 81.39% in Category III.

We also note that the performance of our algorithm is sensitive to the number of first-stage variables (n_1): when $n_1 = 2$, the average gap between the upper and lower bounds (AD_2) is less than 7%, while for $n_2 = 5$ or 25, AD_2 is larger than 55%. However, Table 4.2 suggests that the algorithm is not as sensitive to the number of second-stage variables (n_2), second-stage scenarios ($|\mathcal{E}_2|$), and second-stage constraints (m_2).

Finally, the last row in Table 4.2 shows that for several instances, AD_2 was less than 1%. For example, in Category III, this happened for 20.73% of the cases. We also observed $AD_2 \leq 1\%$ for 2.27% of Category IV problems.

Category	I	II	III	IV	V	VI
n_1	2	2	2	5	5	25
m_1	1	1	1	2	2	10
$\|\mathcal{E}_2\|$	50	100	250	100	250	250
n_2	20	20	50	25	25	50
m_2	5	5	10	5	5	10
Feasible instances	48	42	82	72	63	88
LPCC average time	300s	300s	300s	300s	300s	300s
Bundle average time	22s	45s	242s	52s	117s	298s
AD_1	7.75%	15.03%	81.39%	25.95%	24.25%	5.13%
AD_2	6.84%	6.43%	6.64%	55.81%	76.62%	111.10%
Fraction $AD_2 \leq 1\%$	16.67%	9.52%	20.73%	8.33%	4.76%	2.27%

Table 4.2: Computational results of bundle method.

Chapter 5

Using Bundle Method to Solve Supply Chain Problems with BRM Risk Measure

In this chapter, we use the proposed bundle method to solve the CTC formulation of the supply chain production planning problem discussed in Section 3.1 with the BRM risk measure. The first step is to reformulate (3.1)-(3.2) to the standard bilevel stochastic formulation (4.1)-(4.2) to be processed by Algorithm 1.

5.1 Reformulation to the Standard Form

First, we remove constraint $g \geq t + \frac{1}{\alpha_1}\sum_{i,j}\pi_i\pi_{i,j}X_{i,j}$ from (3.1) and replace g by $t + \frac{1}{\alpha_1}\sum_{i,j}\pi_i\pi_{i,j}X_{i,j}$ in the objective function. Also, we perform the same procedure to remove h_i from problem (3.2). Next, let $t = t^+ - t^-$ and $\bar{t}_i = \bar{t}_i^+ - \bar{t}_i^-$ for some nonnegative decision variables $t^+, t^-, \bar{t}_i^+, \bar{t}_i^-$. After these changes, formulations (5.1)-(5.2) below are equivalent to (3.1)-(3.2):

First-stage problem

$$\begin{aligned}
\min \quad & c^T z + (1-\beta_1)\sum_{i=1}^{N_1}\pi_i\left(-r^T u_i + l^T s_i^- + (\sum_{j=1}^{N_2}\pi_{i,j}H_{i,j}^T)s_i^+\right) \\
& + \beta_1(t^+ - t^- + \frac{1}{\alpha_1}\sum_{i,j}\pi_i\pi_{i,j}X_{i,j}) \qquad (5.1) \\
\text{s.t.} \quad & z_{\text{up}} \geq z \geq 0, \\
& (u_i, s_i^-, s_i^+, X_{i,j}) \in \mathcal{X}_i^*(z), \qquad \forall i.
\end{aligned}$$

Second-stage problem, scenario i

$$\begin{aligned}
\min \quad & -r^T u_i + l^T s_i^- + (1-\beta_2)(\sum_j \pi_{i,j} H_{i,j}^T) s_i^+ \\
& + \beta_2(\bar{t}_i^+ - \bar{t}_i^- + \frac{1}{\alpha_2}\sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}) \\
\text{s.t.} \quad & z \geq M u_i, \\
& s_i^+ \geq u_i - D_i, \;\; s_i^+ \geq 0, \\
& s_i^- \geq -u_i + D_i, \;\; s_i^- \geq 0, \\
& X_{i,j} \geq -r^T u_i + l^T s_i^- + H_{i,j}^T s_i^+ - t, \;\; X_{i,j} \geq 0, \quad \forall j \\
& Y_{i,j} \geq H_{i,j}^T s_i^+ - \bar{t}_i, \;\; Y_{i,j} \geq 0, \qquad \forall j.
\end{aligned} \tag{5.2}$$

By introducing proper slack variables, we convert the inequality constraints to equality to obtain standard form (4.1)-(4.2).

5.2 Master Problem Adjustments

By applying the proposed bundle method to the standard form of the supply chain problem, we observe that the master problem may return a dual point Y such that all sub-problems become unbounded. When this phenomenon occurs the algorithm cannot progress to the next iteration to improve the lower bound. To overcome this hurdle, we exploit the problem structure to adjust the master problem.

Suppose $\hat{z}$ is the nonnegative slack variable for the constraint $z_{\text{up}} \geq z$, so that $z + \hat{z} = z_{\text{up}}$. We replicate the first-stage decision variables $z, \hat{z}, t^+$, and t^- for each second-stage scenario, denoting them by $z_i, \hat{z}_i, t_i^+$, and t_i^-, for $i = 1, \cdots, N_1$. Let $Y_i = (Y_i^z, Y_i^{\hat{z}}, Y_i^{t^+}, Y_i^{t^-})^T \in \mathbb{R}^{n_1}$ for $i = 1, \cdots, N_1$. The objective function of (4.4) at $Y = (Y_1, \cdots, Y_{N_1})$ for scenario i in the supply chain problem is:

$$\begin{aligned}
\text{O.F.} =& \frac{c^T}{N_1} z_i + \frac{\beta_1}{N_1}(t_i^+ - t_i^-) + (1-\beta_1)\pi_i(-r^T u_i + l^T s_i^- + (\sum_{j=1}^{N_2} \pi_{i,j} H_{i,j}^T) s_i^+) \\
&+ \frac{\beta_1}{\alpha_1}\sum_j \pi_i \pi_{i,j} X_{i,j} + \langle z_i, Y_i^z \rangle + \langle \hat{z}_i, Y_i^{\hat{z}} \rangle + t_i^+ Y_i^{t^+} + t_i^- Y_i^{t^-}.
\end{aligned}$$

Note that decision variables t_i^+ and t_i^- only appear in constraints $X_{i,j} \geq -r^T u_i + l^T s_i^- + H_{i,j}^T s_i^+ - t_i^+ + t_i^-$, for $j = 1, \cdots, N_2$, and their coefficients in O.F. are $(\frac{\beta_1}{N_1} +$

$Y_i^{t^+}$) and $(-\frac{\beta_1}{N_1} + Y_i^{t^-})$, respectively. Consider a feasible point for subproblem $f^i(Y)$. The objective function O.F. may become unbounded in Algorithm 1 for the following reasons:

(i) If $(\frac{\beta_1}{N_1} + Y_i^{t^+}) + (-\frac{\beta_1}{N_1} + Y_i^{t^-}) = Y_i^{t^+} + Y_i^{t^-} < 0$, then by increasing both t_i^+ and t_i^-, the subproblem value $f^i(Y)$ may be driven to $-\infty$. To prevent this case from occurring, we add cuts $Y_i^{t^+} + Y_i^{t^-} \geq 0$, for all i, to the master problem.

(ii) If $(\frac{\beta_1}{N_1} + Y_i^{t^+}) < 0$, then increasing t_i^+ can drive the subproblem value to $-\infty$. To address this case, we add constraints $Y_i^{t^+} \geq -\frac{\beta_1}{N_1}$, for all i, to the master problem.

(iii) If $(-\frac{\beta_1}{N_1} + Y_i^{t^-}) + \frac{\beta_1}{\alpha_1}\sum_j \pi_i\pi_{i,j} < 0$, then by increasing t_i^- and $X_{i,j}$, $j = 1, \cdots, N_2$, simultaneously, one can drive the subproblem objective value to $-\infty$. To avoid this difficulty, we add constraints $Y_i^{t^-} \geq \frac{\beta_1}{N_1} - \frac{\beta_1}{\alpha_1}\sum_j \pi_i\pi_{i,j}$, for all i, to the master problem.

Therefore, we consider the following adjusted master problem for the supply chain problem:

$$\begin{aligned}
\max_{r,Y} \quad & \sum_{i=1}^{N_1} r_i - \frac{1}{2t_k}\left\|Y - \hat{Y}_k\right\|^2 \\
\text{s.t.} \quad & r^i \leq f^i(\hat{Y}_k) + e_j^i + \langle X_{1j}^{i^*}, Y^i - \hat{Y}_{k,i}\rangle, \quad \forall i \in \{1, \cdots, N_1\},\ j \in \mathcal{B}_k^i \\
& Y = (Y_1, \cdots, Y_{N_1}) \in K^* \\
& Y_i^{t^+} + Y_i^{t^-} \geq 0, \qquad \forall i \in \{1, \cdots, N_1\} \\
& Y_i^{t^+} \geq -\frac{\beta_1}{N_1}, \qquad \forall i \in \{1, \cdots, N_1\} \\
& Y_i^{t^-} \geq \frac{\beta_1}{N_1} - \frac{\beta_1}{\alpha_1}\sum_j \pi_i\pi_{i,j}, \qquad \forall i \in \{1, \cdots, N_1\}
\end{aligned} \tag{5.3}$$

where $(X_{1,j}^{i^*}, X_{2,j}^{i^*})$ is the optimal solution of $f^i(Y_j)$ and e_j^i is the linearization error at the center point $\hat{Y}_k$.

We also add an additional cut to the master problem to improve the convergence of the algorithm: note that by adding all constraints of second-stage problem (5.2) to the first-stage problem (5.1), one can obtain an LP which can be solved efficiently to find a lower bound for the original bilevel problem. Let L^{LP} be the value of this lower bound. We numerically observed that for the supply chain planning problem, L^{LP} is often quite close to the optimal solution value of the bilevel problem, but the corresponding solution is infeasible for that problem since it violates most of the complementarity constraints. To exploit this observation, we add the cut $\sum_{i=1}^{N_1} r_i \geq L^{LP}$ to the master problem (5.3). As a result, the master problem does not explore dual points that offers a lower bound worse than L^{LP}. This cut usually resulted in a faster convergence in our numerical study.[1] Additionally, after the algorithm stops, we take the maximum of the lower bound found in our algorithm (L) and L^{LP} to find a tighter lower bound. In other words, the lower bound reported in our algorithm is $L^{Final} = \max(L, L^{LP})$.

Next, we formulate the dual problem of (5.3) to find a feasible point for the supply chain problem.

5.3 Dual of Master Problem

We follow the same procedure as in Section 4.4.1 to find a feasible solution to (5.1)-(5.2). First, we formulate the dual of the master problem. Table 5.1 summarizes the Lagrange multipliers for the constraints in problem (5.3).

Table 5.1: **Lagrange multiplier**

$\alpha_{j,i}$	Scalar Lagrange multipliers for constraints $r^i \leq f^i(\hat{Y}_k) + e^i_j + \langle X^{i*}_{1j}, Y_i - \hat{Y}_{k,i} \rangle$ above, for $i = 1, \ldots, N_1$ and $j \in \mathcal{B}^i_k$
β	Vector of Lagrange multipliers for constraint $Y_1 + \cdots + Y_{N_1} = 0$
λ_i	Scalar Lagrange multipliers for constraints $Y_i^{t^+} + Y_i^{t^-} \geq 0$ above, for $i = 1, \ldots, N_1$
μ_i	Scalar Lagrange multipliers for constraint $Y_i^{t^+} \geq -\frac{\beta_1}{N_1}$ above, for $i = 1, \ldots, N_1$
δ_i	Scalar Lagrange multipliers for constraint $Y_i^{t^-} \geq \frac{\beta_1}{N_1} - \frac{\beta_1}{\alpha_1} \sum_j \pi_i \pi_{i,j}$ above, for $i = 1, \ldots, N_1$

[1] Note that addition of this cut to the master problem does not always guarantee a faster convergence in a general problem. This is because it can limit the search space around a center in an iteration, which can possibly slow down convergence. Therefore, the extent of the improvement in the convergence strongly depends on the quality of L^{LP}, which was good for the supply chain planning problem in our numerical study.

Note that $X_{1j}^{i*} = (z_{j,i}^*, \hat{z}_{j,i}^*, t_{j,i}^{+*}, t_{j,i}^{-*})$. The Lagrangian function $L(r, Y, \alpha, \beta, \lambda, \mu, \delta)$ is:

$$\sum_{i=1}^{N_1} r^i - \frac{1}{2t_k}\left\|Y - \hat{Y}_k\right\|^2 - \sum_{i=1}^{N_1}\sum_{j\in\mathcal{B}_k^i} \alpha_{j,i}\big(r^i - f^i(\hat{Y}_k) - e_j^i - \langle X_{1j}^{i*}, Y_i - \hat{Y}_{k,i}\rangle\big)$$
$$-\sum_{i=1}^{N_1}\beta Y_i - \sum_{i=1}^{N_1}\lambda_i(-Y_i^{t^+} - Y_i^{t^-}) - \sum_{i=1}^{N_1}\mu_i(-Y_i^{t^+} - \frac{\beta_1}{N_1})$$
$$-\sum_{i=1}^{N_1}\delta_i(-Y_i^{t^-} + \frac{\beta_1}{N_1} - \frac{\beta_1}{\alpha_1}\sum_j \pi_i\pi_{i,j})$$
$$= \sum_{i=1}^{N_1} r^i(1 - \sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}) + \sum_{i=1}^{N_1}\big[-\frac{1}{2t_k}\left\|Y_i^z - \hat{Y}_{k,i}^z\right\|^2 + \sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}\langle z_{j,i}^*, Y_i^z - \hat{Y}_{k,i}^z\rangle - \beta^z Y_i^z\big]$$
$$+\sum_{i=1}^{N_1}\big[-\frac{1}{2t_k}\left\|Y_i^{\hat{z}} - \hat{Y}_{k,i}^{\hat{z}}\right\|^2 + \sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}\langle \hat{z}_{j,i}^*, Y_i^{\hat{z}} - \hat{Y}_{k,i}^{\hat{z}}\rangle - \beta^{\hat{z}} Y_i^{\hat{z}}\big]$$
$$+\sum_{i=1}^{N_1}\big[-\frac{1}{2t_k}(Y_i^{t^+} - \hat{Y}_{k,i}^{t^+})^2 + \sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}t_{j,i}^{+*}(Y_i^{t^+} - \hat{Y}_{k,i}^{t^+}) - \beta^{t^+}Y_i^{t^+} + \lambda_i Y_i^{t^+} + \mu_i Y_i^{t^+}\big]$$
$$+\sum_{i=1}^{N_1}\big[-\frac{1}{2t_k}(Y_i^{t^-} - \hat{Y}_{k,i}^{t^-})^2 + \sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}t_{j,i}^{-*}(Y_i^{t^-} - \hat{Y}_{k,i}^{t^-}) - \beta^{t^-}Y_i^{t^-} + \lambda_i Y_i^{t^-} + \delta_i Y_i^{t^-}\big]$$
$$+\sum_{i=1}^{N_1}\sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}\big(f^i(\hat{Y}_k) + e_j^i\big) + \sum_{i=1}^{N_1}\mu_i\frac{\beta_1}{N_1} + \sum_{i=1}^{N_1}\delta_i(-\frac{\beta_1}{N_1} + \frac{\beta_1}{\alpha_1}\sum_j \pi_i\pi_{i,j}).$$

At the optimal point (r^*, Y^*), we have:

$$\nabla_{r^{i*}} L = (1 - \sum_{j\in\mathcal{B}_k^S}\alpha_{j,i}) = 0,$$
$$\nabla_{Y_i^{z*}} L = -\frac{1}{t_k}(Y_i^{z^*} - \hat{Y}_{k,i}^z) + \sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}z_{j,i}^* - \beta^z = 0,$$
$$\nabla_{Y_i^{\hat{z}*}} L = -\frac{1}{t_k}(Y_i^{\hat{z}^*} - \hat{Y}_{k,i}^{\hat{z}}) + \sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}\hat{z}_{j,i}^* - \beta^{\hat{z}} = 0,$$
$$\nabla_{(Y_i^{t^+})^*} L = -\frac{1}{t_k}((Y_i^{t^+})^* - \hat{Y}_{k,i}^{t^+}) + \sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}t_{j,i}^{+*} - \beta^{t^+} + \lambda_i + \mu_i = 0,$$
$$\nabla_{(Y_i^{t^-})^*} L = -\frac{1}{t_k}((Y_i^{t^-})^* - \hat{Y}_{k,i}^{t^-}) + \sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}t_{j,i}^{-*} - \beta^{t^-} + \lambda_i + \delta_i = 0,$$

for all $i \in \{1, \cdots, N_1\}$. From the last four equations, we have:

$$Y_i^{z^*} = \hat{Y}_{k,i}^z + \sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}z_{j,i}^* - t_k\beta^z,$$

$$Y_i^{\hat{z}^*} = \hat{Y}_{k,i}^{\hat{z}} + \sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}\hat{z}_{j,i}^* - t_k\beta^{\hat{z}},$$

$$Y_i^{t^+} = \hat{Y}_{k,i}^{t^+} + \sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}t_{j,i}^{+*} - t_k\beta^{t^+} + t_k\lambda_i + t_k\mu_i,$$

$$Y_i^{t^-} = \hat{Y}_{k,i}^{t^-} + \sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}t_{j,i}^{-*} - t_k\beta^{t^-} + t_k\lambda_i + t_k\delta_i,$$

By plugging these values in the Lagrangian function, $\max_{r,Y} L(r,Y,\alpha,\beta,\lambda,\mu,\delta)$ is given by:

$$\sum_{i=1}^{N_1}\big[-\frac{1}{2t_k}\left\|\sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}z_{j,i}^* - t_k\beta^z\right\|^2 + \sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}\langle z_{j,i}^*, \sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}z_{j,i}^* - t_k\beta^z\rangle$$

$$-\beta^z(\hat{Y}_{k,i}^z + \sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}z_{j,i}^* - t_k\beta^z)\big] + \sum_{i=1}^{N_1}\big[-\frac{1}{2t_k}\left\|\sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}\hat{z}_{j,i}^* - t_k\beta^{\hat{z}}\right\|^2$$

$$+\sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}\langle\hat{z}_{j,i}^*, \sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}\hat{z}_{j,i}^* - t_k\beta^{\hat{z}}\rangle - \beta^{\hat{z}}(\hat{Y}_{k,i}^{\hat{z}} + \sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}\hat{z}_{j,i}^* - t_k\beta^{\hat{z}})\big]$$

$$+\sum_{i=1}^{N_1}\big[-\frac{1}{2t_k}(\sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}t_{j,i}^{+*} - t_k\beta^{t^+} + t_k\lambda_i + t_k\mu_i)^2$$

$$+\sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}t_{j,i}^{+*}(\sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}t_{j,i}^{+*} - t_k\beta^{t^+} + t_k\lambda_i + t_k\mu_i)$$

$$+(-\beta^{t^+} + \lambda_i + \mu_i)(\hat{Y}_{k,i}^{t^+} + \sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}t_{j,i}^{+*} - t_k\beta^{t^+} + t_k\lambda_i + t_k\mu_i)\big]$$

$$+\sum_{i=1}^{N_1}\big[-\frac{1}{2t_k}(\sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}t_{j,i}^{-*} - t_k\beta^{t^-} + t_k\lambda_i + t_k\delta_i)^2$$

$$+\sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}t_{j,i}^{-*}(\sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}t_{j,i}^{-*} - t_k\beta^{t^-} + t_k\lambda_i + t_k\delta_i)$$

$$+(-\beta^{t^-} + \lambda_i + \delta_i)(\hat{Y}_{k,i}^{t^-} + \sum_{j\in\mathcal{B}_k^i} t_k\alpha_{j,i}t_{j,i}^{-*} - t_k\beta^{t^-} + t_k\lambda_i + t_k\delta_i)\big]$$

$$+\sum_{i=1}^{N_1}\sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}\big(f^i(\hat{Y}_k) + e_j^i\big) + \sum_{i=1}^{N_1}\mu_i\frac{\beta_1}{N_1} + \sum_{i=1}^{N_1}\delta_i(-\frac{\beta_1}{N_1} + \frac{\beta_1}{\alpha_1}\sum_j\pi_i\pi_{i,j})$$

$$=\sum_{i=1}^{N_1}\frac{t_k}{2}(\sum_{j\in\mathcal{B}_k^i} z_{j,i}^*\alpha_{j,i} - \beta^z)^2 - \sum_{i=1}^{N_1}\beta^z\hat{Y}_{k,i}^z + \sum_{i=1}^{N_1}\frac{t_k}{2}(\sum_{j\in\mathcal{B}_k^i}\hat{z}_{j,i}^*\alpha_{j,i} - \beta^{\hat{z}})^2 - \sum_{i=1}^{N_1}\beta^z\hat{Y}_{k,i}^z$$

$$+\sum_{i=1}^{N_1}\frac{t_k}{2}(\alpha_{j,i}t_{j,i}^{+*} - \beta^{t^+} + \lambda_i + \mu_i)^2 + \sum_{i=1}^{N_1}(-\beta^{t^+} + \lambda_i + \mu_i)\hat{Y}_{k,i}^{t^+}$$

$$+\sum_{i=1}^{N_1}\frac{t_k}{2}(\alpha_{j,i}t_{j,i}^{-*}-\beta^{t^-}+\lambda_i+\delta_i)^2+\sum_{i=1}^{N_1}(-\beta^{t^-}+\lambda_i+\delta_i)\hat{Y}_{k,i}^{t^-}$$
$$+\sum_{i=1}^{N_1}\sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}\big(f^i(\hat{Y}_k)+e_j^i\big)+\sum_{i=1}^{N_1}\mu_i\frac{\beta_1}{N_1}+\sum_{i=1}^{N_1}\delta_i(-\frac{\beta_1}{N_1}+\frac{\beta_1}{\alpha_1}\sum_j\pi_i\pi_{i,j})$$

Therefore, the dual of the master problem is:

$$\begin{aligned}
\min_{\alpha,\beta,\lambda,\mu,\delta}\quad & \sum_{i=1}^{N_1}\frac{t_k}{2}[(\nu_i^z)^2+(\nu_i^{\hat{z}})^2+(\nu_i^{t^+})^2+(\nu_i^{t^-})^2] \\
& +\sum_{i=1}^{N_1}[-\beta^z\hat{Y}_{k,i}^z-\beta^z\hat{Y}_{k,i}^z+(-\beta^{t^+}+\lambda_i+\mu_i)\hat{Y}_{k,i}^{t^+}+(-\beta^{t^-}+\lambda_i+\delta_i)\hat{Y}_{k,i}^{t^-}] \\
& +\sum_{i=1}^{N_1}\sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}\big(f^i(\hat{Y}_k)+e_j^i\big)+\sum_{i=1}^{N_1}\mu_i\frac{\beta_1}{N_1}+\sum_{i=1}^{N_1}\delta_i(-\frac{\beta_1}{N_1}+\frac{\beta_1}{\alpha_1}\sum_j\pi_i\pi_{i,j}) \\
\text{s.t.}\quad & \sum_{j\in\mathcal{B}_k^i}\alpha_{j,i}=1, && \forall i \\
& \nu_i^z=\sum_{j\in\mathcal{B}_k^i}z_{j,i}^*\alpha_{j,i}-\beta^z, && \forall i \\
& \nu_i^{\hat{z}}=\sum_{j\in\mathcal{B}_k^i}\hat{z}_{j,i}^*\alpha_{j,i}-\beta^{\hat{z}}, && \forall i \\
& \nu_i^{t^+}=\alpha_{j,i}t_{j,i}^{+*}-\beta^{t^+}+\lambda_i+\mu_i, && \forall i \\
& \nu_i^{t^-}=\alpha_{j,i}t_{j,i}^{-*}-\beta^{t^-}+\lambda_i+\delta_i, && \forall i \\
& \alpha_{i,j},\beta,\lambda_i,\mu_i,\delta_i\geq 0, && \forall i,j\in\mathcal{B}_k^i.
\end{aligned}\tag{5.4}$$

Let β^* be the optimal value of β in the above problem. Set $X_1=\beta^*$ and solve all the second-stage problems (5.2) to find an optimal point $X_2^{S^*}\in\mathcal{X}_{2,S}^*(\beta^*)$. By plugging (β^*,X_2^*) in the objective function (5.1), we obtain a feasible upper bound for the problem.

Next, we solve the adjusted master and dual problems using Algorithm 1.

5.4 Numerical Study

Similar to Section 4.5, we generated six problem categories of different sizes, as summarized in Table 5.2. P_1 is the number of parts in the first stage, P_2 is the number of products in the second stage, N_1 is the number of second-stage scenarios, and $N2$ is the number of third-stage scenarios for each second-stage scenario. Importantly, note that P_1 (P_2), is not the number of first-stage (second-stage) decision variables. For example, a problem with 4 parts in the first stage has 10 first-stage decision variables (4 variables for parts (z), 4 helper variables for inequality constraints ($\hat{z}$), and 2 variables for risk measure (t^+, t^-)). Also, the total number of possible scenarios in each problem is $N_1 \times N_2$.

Each problem category contains 50 randomly generated instances, resulting in a total of 300 numerical problems. For brevity, we focused on $BRM_{0.75}^{0.05}$ risk measure, because the CTC and OTC formulations with this risk measure resulted in dramatically different outcomes (see Table 3.10). We solved each problem instance using *gurobipy* for the LPCC formulation and bundle method for the bilevel formulation (5.1)-(5.2). For both algorithms, we limited the run time to 300s for each instance. Table 5.2 reports the computational results.

Category	I	II	III	IV	V	VI
P_1	2	2	4	4	4	4
P_2	5	5	10	10	25	25
N_1	10	20	10	20	10	25
N_2	10	5	5	5	5	5
LPCC average time	36.90s	61.08s	101.56s	284.80s	264.38s	300s
Bundle average time	9.93s	11.11s	17.90s	33.46s	105.94s	180.67s
AD_1	-1.62%	-2.93%	-0.29%	-0.11%	0.06%	0.01%
AD_2	17.08%	28.15%	7.67%	13.21%	0.84%	2.62%
Fraction $AD_2 \leq 1\%$	30.95%	8.16%	18.75%	8.51%	77.08%	36.96%

Table 5.2: Computational results of bundle method.

In all problem categories, our algorithm converged consistently faster than the LPCC solution. Particularly, for larger problem categories IV, V, and VI, in a majority of cases, Gurobi could not provide a feasible solution and gap within the 300s time limit. However, our specialized bundle method converged in 33.46s, 105.94s, and 180.67s for Categories IV, V, and VI, respectively, and in all cases offered relatively a good feasible upper bound. The average relative difference between the upper and lower bounds were $AD_2 = 0.84\%$ for Category V and $AD_2 = 2.62\%$ for Category VI.

Notably, we observe that our proposed lower bound and the one found by Gurobi are not significantly different. In fact, on average, Gurobi offered slightly a better lower bound (see row AD_1) for relatively small problem instances (Categories I, II, II, and IV). This is because of our earlier observation that in supply chain planning problem, the relaxed LP solution provides a good lower bound, equipping Gurobi with a good starting point for its branch-and-bound method. However, note that for larger problems, Gurobi could not progress to find a feasible point, and we only know its lower bound is good from the upper bound found by our method. This observation indicates the utility of our "feasible" upper bound, which can also provide a reasonable course of action for the DM.

Bibliography

American Express Australia (2016), Technical report, Game Plan for Growth: 2016 American Express CFO Future-Proofing Survey. http://www.chieffutureofficer.com/.

Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1999), 'Coherent measures of risk', *Mathematical Finance* **9**(3), 203–228.

Belloni, A. & Sagastizábal, C. (2009), 'Dynamic bundle methods', *Mathematical Programming* **120**(2), 289–311.

Collado, R. A., Papp, D. & Ruszczyński, A. (2012), 'Scenario decomposition of risk-averse multistage stochastic programming problems', *Annals of Operations Research* **200**(1), 147–170.

de Oliveira, W. & Eckstein, J. (2015), A bundle method for exploiting additive structure in difficult optimization problems, Preprint 2015-05-4935, Optimization Online.

Dempe, S. (2000), 'A bundle algorithm applied to bilevel programming problems with non-unique lower level solutions', *Computational Optimization and Applications* **15**(2), 145–166.

Dempe, S. & Bard, J. F. (2001), 'Bundle trust-region algorithm for bilinear bilevel programming', *Journal of Optimization Theory and Applications* **110**(2), 265–288.

Eckstein, J., Eskandani, D. & Fan, J. (2016), 'Multilevel optimization modeling for risk-averse stochastic programming', *INFORMS Journal on Computing* **28**(1), 112–128.

Frangioni, A. (2002), 'Generalized bundle methods', *SIAM Journal on Optimization* **13**(1), 117–156.

Gülten, S. & Ruszczyński, A. (2015), 'Two-stage portfolio optimization with higher-order conditional measures of risk', *Annals of Operations Research* **229**(1), 409–427.

Kiwiel, K. C. (1991), 'Exact penalty functions in proximal bundle methods for constrained convex nondifferentiable minimization', *Mathematical Programming* **52**(1-3), 285–302.

Mäkelä, M. (2002), 'Survey of bundle methods for nonsmooth optimization', *Optimization methods and software* **17**(1), 1–29.

Osborne, M. J. et al. (2004), *An introduction to game theory*, Vol. 3, Oxford university press New York.

Pereira, M. V. & Pinto, L. M. (1991), 'Multi-stage stochastic optimization applied to energy planning', *Mathematical programming* **52**(1-3), 359–375.

Rockafellar, R. T., Uryasev, S. et al. (2000), 'Optimization of conditional value-at-risk', *Journal of Risk* **2**, 21–42.

Ruszczyński, A. (2006), *Nonlinear optimization*, Vol. 13, Princeton University Press.

Ruszczyński, A. (2010), 'Risk-averse dynamic programming for markov decision processes', *Mathematical programming* **125**(2), 235–261.

Schweitzer, M. E. & Cachon, G. P. (2000), 'Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence', *Management Science* **46**(3), 404–420.

Shapiro, A. (2012), 'Time consistency of dynamic risk measures', *Operations Research Letters* **40**(6), 436–439.

Shapiro, A., Dentcheva, D. & Ruszczyński, A. (2009), *Lectures on stochastic programming: modeling and theory*, SIAM.

Van Ackooij, W. & Frangioni, A. (2018), 'Incremental bundle methods using upper models', *SIAM Journal on Optimization* **28**(1), 379–410.