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

Abstract ................................................................. ii
Acknowledgements ....................................................... iii

1. Introduction .......................................................... 1

2. Three-stage Risk-averse Stochastic Programming ............... 5
   2.1. Preliminaries ................................................... 5
   2.2. Three-Stage RASDP Formulations ........................... 7
   2.3. Generic Formulation ............................................ 11
       2.3.1. Problem Formulation with MSD Risk Measure .......... 11
       2.3.2. Problem Formulation with BRM Risk Measure .......... 14

3. Practical Examples .................................................. 18
   3.1. Supply Chain Production Planning ............................ 18
       3.1.1. Problem Formulation with BRM Risk Measure .......... 18
       3.1.2. Problem Formulation with MSD Risk Measure .......... 24
       3.1.3. Numerical Study ........................................... 26
   3.2. Portfolio Optimization Problem ............................... 29
       3.2.1. Problem Formulation with MSD Risk Measure .......... 29
       3.2.2. Problem Formulation with BRM Risk Measure .......... 34
       3.2.3. Numerical Study ........................................... 39
   3.3. Hydropower Energy Planning Problem ........................ 41
       3.3.1. Problem Formulation with MSD Risk Measure .......... 42
       3.3.2. Problem Formulation with BRM Risk Measure .......... 51
       3.3.3. Numerical Study ........................................... 56
4. Bundle Method for Bilevel Stochastic Programming ............... 59
   4.1. Bilevel Stochastic Programming Formulation .................... 59
   4.2. LPCC Formulation ............................................. 60
   4.3. Lower Bound for LPCC Formulation .............................. 61
   4.4. The Proximal Bundle Method ................................... 64
      4.4.1. Feasible Upper Bound .................................... 66
      4.4.2. Specialized Bundle Method Algorithm ...................... 68
   4.5. Bundle Method Computational Results ........................... 69

5. Using Bundle Method to Solve Supply Chain Problems with BRM
   Risk Measure ...................................................... 72
   5.1. Reformulation to the Standard Form ............................. 72
   5.2. Master Problem Adjustments ..................................... 73
   5.3. Dual of Master Problem ....................................... 75
   5.4. Numerical Study ............................................... 78

Bibliography ....................................................... 81
Chapter 1

Introduction

In this work, we study a broad class of risk-averse stochastic dynamic programs (RAS-DPs), 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 \(\sigma\)-algebra \(\mathcal{F}\) (collection of all events) on the sample space \(\Omega\) and probability measure \(P\) on \(\mathcal{F}\). Let \(\mathcal{L}\) be the space of all \(\mathcal{F}\)-measurable random variables and “\(\succeq\)” be a partial order on \(\mathcal{L}\) such that for \(Z, Z' \in \mathcal{L}\), \(Z \succeq Z'\) if and only if \(Z(\omega) \geq Z'(\omega)\) for all \(\omega \in \Omega\). A coherent risk measure on \(\mathcal{L}\) is defined as follows:

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

- **Monotonicity.** If \(Z \succeq W\), 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 $Z \in \mathcal{L}$ and $\gamma \in [0,1]$. The mean-upper semideviation risk is defined as:

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

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

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

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

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 $\alpha$.

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

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

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 $\beta$ corresponds to more risk aversion.

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

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

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 $Z, W \in \mathcal{L}$ 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\(_\gamma\), AVaR\(_\alpha\), and BRM\(_{\alpha\beta}\) 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 \(\Omega\) with \(S \in \mathcal{E}_2\). Also, let \(\mathcal{F}_2\) be the \(\sigma\)-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 \(X_3^\omega\) 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.](image-url)
In the first stage, the DM’s objective is to minimize a coherent risk measure \( \rho_1 \) of the total cost. We also consider linear constraints at each time stage. Let \( A_{11}, A^S_{2j}, \) and \( A^{\omega}_{3k}, \) 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 \( b_1, b^S_2, \) and \( b^\omega_3 \) be constant vectors of conforming size. The mathematical representation of the first-stage problem is as follows:

\[
\begin{align*}
\min \quad & \rho_1(Z_1 + Z_2 + Z_3) \\
\text{s.t.} \quad & A_{11}X_1 = b_1 \\
& A^S_{21}X_1 + A^S_{22}X^S_2 = b^S_2, \quad \forall S \in \mathcal{E}_2 \\
& A^\omega_{31}X_1 + A^\omega_{32}X^S_2 + A^\omega_{33}X^\omega_3 = b^\omega_3, \quad \forall S \in \mathcal{E}_2, \forall \omega \in S \\
& X_1, X^S_2, X^\omega_3 \geq 0, \quad \forall S \in \mathcal{E}_2, \forall \omega \in S.
\end{align*}
\] (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:

\[
\begin{align*}
\min \quad & \rho_2(Z_2 + Z_3|S) \\
\text{s.t.} \quad & A^S_{22}X^S_2 = b^S_2 - A^S_{21}X_1 \\
& A^\omega_{32}X^S_2 + A^\omega_{33}X^\omega_3 = b^\omega_3 - A^\omega_{31}X_1, \quad \forall \omega \in S \\
& X^S_2, X^\omega_3 \geq 0, \quad \forall \omega \in S.
\end{align*}
\] (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 \( X^S_2, \) with \( \omega \in S, \) is formulated as:

\[
\begin{align*}
\min \quad & Z_3(\omega) \\
\text{s.t.} \quad & A^\omega_{33}X^\omega_3 = b^\omega_3 - A^\omega_{31}X_1 - A^\omega_{32}X^S_2, \quad \forall \omega \in S \\
& X^\omega_3 \geq 0, \quad \forall \omega \in S.
\end{align*}
\] (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 (Ruszczynski 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 $X^{*}_{2,S}(X_1)$ 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 $((\bar{X}^{S_i}_2)^*, (\bar{X}^{S_i}_3)^*) \in X^{*}_{2,S_i}(X^{*}_1)$ such that $(X^{*}_1, \bar{X}^{S_i}_2, \bar{X}^{S_i}_3)$ remains optimal for problem (2.1), where $\bar{X}^{S_i}_j = ((X^{S_i}_1)^*, \ldots, (X^{S_i}_j)^*, (X^{S_i}_{j+1})^*, \ldots, (X^{S_i}_E)^*)$ for $j = 2, 3$ (i.e., in the first-stage optimal solution one replaces $(X^{S_i}_2)^*$ and $(X^{S_i}_3)^*$ by $(\bar{X}^{S_i}_2)^*$ and $(\bar{X}^{S_i}_3)^*$, 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:

\[
\begin{align*}
\min_{X_1, X_2, X_3} & \quad \rho_1(Z_1 + \rho_2(Z_2 + Z_3)) \\
\text{s.t.} & \quad A_{11}X_1 = b_1 \\
& \quad A_{21}^S X_1 + A_{22}^S X_2 = b_2^S, \quad \forall S \in \mathcal{E}_2 \\
& \quad A_{31}^S X_1 + A_{32}^S X_2 + A_{33}^S X_3 = b_3^S, \quad \forall S \in \mathcal{E}_2, \forall \omega \in S \\
& \quad X_1, X_2^S, X_3^\omega \geq 0, \quad \forall S \in \mathcal{E}_2, \forall \omega \in S.
\end{align*}
\]
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 \( X^S_2, X^S_3 \in X^*_S(X_1), \forall S \in \mathcal{E}_2 \) to problem (2.1) to formulate the problem as follows:

\[
\begin{align*}
\min_{X_1, X_2, X_3} & \quad \rho_1(Z_1 + Z_2 + Z_3) \\
\text{s.t.} & \quad A_{11}X_1 = b_1, \\
& \quad X_1 \geq 0, \\
& \quad X^S_2, X^S_3 \in X^*_S(X_1), \forall S \in \mathcal{E}_2.
\end{align*}
\]

(2.5)

where \( X^*_S(X_1) \) is the set of optimal solutions to the following recourse problem:

\[
\begin{align*}
\min_{X^S_2, X^S_3} & \quad \rho_2(Z_2 + Z_3|S) \\
\text{s.t.} & \quad A_{22}^S X^S_2 = b_2^S - A_{31}^S X_1, \\
& \quad A_{32}^S X^S_2 + A_{33}^S X^S_3 = b_3^S - A_{31}^S X_1, \quad \forall \omega \in S \\
& \quad X^S_2, X^S_3 \geq 0, \quad \forall \omega \in S.
\end{align*}
\]

(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 formu-
lation 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.

\[
Z_1 = c_1^T X_1,
\]

\[
Z_2^S = (c_2^S)^T X_2^S, \quad \forall S \in \mathcal{E}_2
\]

\[
Z_3^\omega = (c_3^\omega)^T X_3^\omega, \quad \forall S \in \mathcal{E}_2, \forall \omega \in S.
\]

(2.7)

Define \( \pi_i \) 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_{\gamma_1} and MSD_{\gamma_2} are the first-and second-stage risk measures, respec-
tively. 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 $i = 1, \ldots, N_1$
- **$\chi_i$**: Amount $W_i$ exceeds its overall mean in for OTC model second-stage scenario $i$, for $i = 1, \ldots, N_1$

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

$$Z_1 + \text{MSD}_{\gamma_1}(Z_2 + Z_3) = Z_1 + E(Z_2 + Z_3) + \gamma_1 E[Z_2 + Z_3 - E(Z_2 + Z_3)]_+$$

$$= Z_1 + \sum_{i=1}^{N_1} \pi_i Z_i^2 + \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} Z_{i,j}^3 + \gamma_1 \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} K_{i,j}$$

where

$$K_{i,j} \geq Z_{i}^2 + Z_{3}^{i,j} - \left( \sum_{i'=1}^{N_1} \pi_{i'} Z_{2'}^{i'} + \sum_{i'=1}^{N_1} \sum_{j'=1}^{N_2} \pi_{i'} \pi_{i',j'} Z_{3'}^{i',j'} \right), \forall i, \forall j$$

$$K_{i,j} \geq 0, \forall i, \forall j$$

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

$$Z_i^2 + \text{MSD}_{\gamma_2}(Z_3|i) = Z_i^2 + E(Z_3|i) + \gamma_2 E[Z_3 - E(Z_3)|i]$$

$$= Z_i^2 + \sum_{j=1}^{N_2} \pi_{i,j} Z_{3}^{i,j} + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}$$

where

$$Y_{i,j} \geq Z_{3}^{i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} Z_{3}^{i,j'}, \forall j$$

$$Y_{i,j} \geq 0, \forall j.$$  

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

\[
\min_{X_1, X_2, X_3} c_1^TX_1 + \sum_{i=1}^{N_1} \pi_i (e_2^i)^TX_2^i + \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} (e_3^{i,j})^TX_3^{i,j} + \gamma_1 \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} K_{i,j}
\]

s.t. 
\[
A_{11}X_1 = b_1
\]

\[
K_{i,j} \geq (e_2^i)^TX_2^i + (e_3^{i,j})^TX_3^{i,j} - \left( \sum_{i' = 1}^{N_1} \pi_{i'} (e_2^{i'})^TX_2^{i'} \right) + \sum_{i' = 1}^{N_1} \sum_{j' = 1}^{N_2} \pi_{i'} \pi_{i',j'} (e_3^{i',j'})^TX_3^{i',j'}, \quad \forall i, \forall j
\]

\[
X_1 \geq 0, \quad K_{i,j} \geq 0, \quad \forall i, \forall j
\]

\[
X_1, X_2, X_3 \in X^*_1(X_1), \quad \forall i.
\]

Second-stage problem, scenario \(i\) in the CTC model

\[
\min_{X_2^i, X_3^i} (e_2^i)^TX_2^i + \sum_{j=1}^{N_2} \pi_{i,j} (e_3^{i,j})^TX_3^{i,j} + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} \gamma_i
\]

s.t. 
\[
A_{22}^i X_2^i = b_2^i - A_{21}^i X_1^i,
\]

\[
A_{32}^i X_2^i + A_{33}^i X_3^i = b_3^i - A_{31}^i X_1^i, \quad \forall j
\]

\[
Y_{i,j} \geq (e_3^{i,j})^TX_3^{i,j} - \sum_{j' = 1}^{N_2} \pi_{i,j'} (e_3^{i',j'})^TX_3^{i',j'}, \quad \forall j
\]

\[
Y_{i,j} \geq 0, \quad X_2^i \geq 0, \quad X_3^{i,j} \geq 0, \quad \forall j.
\]

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

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

where

\[
\chi_i \geq W_i - \sum_{j' = 1}^{N_1} \pi_{i'} W_{i'} , \quad \forall i
\]

\[
W_i = Z_2^i + \sum_{j=1}^{N_2} \pi_{i,j} Z_3^{i,j} + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, \quad \forall i
\]

\[
Y_{i,j} \geq Z_3^{i,j} - \sum_{j' = 1}^{N_2} \pi_{i,j'} Z_3^{i',j'} , \quad \forall i, \forall j
\]

\[
\chi_i \geq 0, \quad Y_{i,j} \geq 0, \quad \forall i, \forall j.
\]
OTC model with nested MSD risk measure

\[
\min_{x_1, x_2, x_3} \quad c_1^T x_1 + \sum_{i=1}^{N_1} \pi_i w_i + \gamma_1 \sum_{i=1}^{N_1} \pi_i x_i
\]

s.t. \quad A_{11} x_1 = b_1

\[
A_{21}^i x_1 + A_{22}^i x_2 = b_2^i, \quad \forall i
\]

\[
A_{31}^i x_1 + A_{32}^i x_2 + A_{33} x_3^i = b_3^i, \quad \forall i, \forall j
\]

\[
\chi_i \geq w_i - \sum_{i'=1}^{N_1} \pi_{i'} w_{i'}, \quad \forall i
\]

\[
\chi_i \geq w_i - \sum_{i'=1}^{N_1} \pi_{i'} w_{i'}, \quad \forall i
\]

\[
W_i = (c_2^i)^T x_2^i + \sum_{j=1}^{N_2} \pi_{i,j} (c_3^i)^T x_3^i, \quad \forall i
\]

\[
Y_{i,j} \geq (c_3^i)^T x_3^i - \sum_{j'=1}^{N_2} \pi_{i,j'} (c_3^i)^T x_3^i, \quad \forall i, \forall j
\]

\[
X_1 \geq 0, \quad X_2 \geq 0, \quad X_3 \geq 0, \quad \chi_i \geq 0, \quad Y_{i,j} \geq 0, \quad \forall i, \forall j.
\]

### 2.3.2 Problem Formulation with BRM Risk Measure

Let \(\text{BRM}^{\alpha_1}_{\beta_1}\) and \(\text{BRM}^{\alpha_2}_{\beta_2}\) 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>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(g)</td>
<td>First-stage AVaR (\alpha_1) value for CTC model</td>
</tr>
<tr>
<td>(h_i)</td>
<td>AVaR (\alpha_2) value, in second-stage scenario (i = 1, \ldots, N_1)</td>
</tr>
<tr>
<td>(t)</td>
<td>Helper variable in AVaR (\alpha_1) definition in stage 1 for CTC model</td>
</tr>
<tr>
<td>(K_{i,j})</td>
<td>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)</td>
</tr>
<tr>
<td>(\tilde{t}_i)</td>
<td>Helper variable in AVaR (\alpha_2) definition in stage 2 scenario (i), for (i = 1, \ldots, N_1)</td>
</tr>
<tr>
<td>(Y_{i,j})</td>
<td>Positive part of difference between stage 3’s total cost in scenario ((i, j)) and (\tilde{t}_i), for (i = 1, \ldots, N_1) and (j = 1, \ldots, N_2)</td>
</tr>
<tr>
<td>(\omega_i)</td>
<td>Second stage cost plus (\text{BRM}^{\alpha_2}_{\beta_2}) value of third stage cost for OTC model, in second-stage scenario (i = 1, \ldots, N_1)</td>
</tr>
</tbody>
</table>
The first-stage objective function for the CTC model is:

\[ Z_1 + \text{BRM}_{\beta_1}^\alpha(Z_2 + Z_3) = Z_1 + (1 - \beta_1)\mathbb{E}(Z_2 + Z_3) + \beta_1 \text{AVaR}_{\alpha_1}(Z_2 + Z_3) \]

\[ = Z_1 + (1 - \beta_1)\left( \sum_{i=1}^{N_1} \pi_i Z^i_2 + \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} Z_{3}^{i,j} \right) + \beta_1 g, \]

where

\[ g \geq t + \frac{1}{\alpha_1} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} K_{i,j}, \]

\[ K_{i,j} \geq Z^i_2 + Z_{3}^{i,j} - t, \quad \forall i, \forall j \]

\[ K_{i,j} \geq 0, \quad \forall i, \forall j, \]

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

\[ Z^i_2 + \text{BRM}_{\beta_2}^\alpha(Z_3|i) = Z^i_2 + (1 - \beta_2)\mathbb{E}(Z_3|i) + \beta_2 \text{AVaR}_{\alpha_2}(Z_3|i) \]

\[ = Z^i_2 + (1 - \beta_2)\sum_{j=1}^{N_2} \pi_{i,j} Z_{3}^{i,j} + \beta_2 h_i, \]

where

\[ h_i \geq \bar{t}_i + \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, \]

\[ Y_{i,j} \geq Z_{3}^{i,j} - \bar{t}_i, \quad \forall j \]

\[ Y_{i,j} \geq 0, \quad \forall j. \]

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

**First-stage problem in CTC model**

\[ \min_{X_1, X_2, X_3} c^T X_1 + (1 - \beta_1)\left( \sum_{i=1}^{N_1} \pi_i (c^i_2)^T X^i_2 + \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} (c^i_3)^T X_{3}^{i,j} \right) + \beta_1 g \]

s.t.\[ A_{11} X_1 = b_1 \]

\[ g \geq t + \frac{1}{\alpha_1} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} K_{i,j}, \]

\[ K_{i,j} \geq (c^i_2)^T X^i_2 + (c^i_3)^T X_{3}^{i,j} - t, \quad \forall i, \forall j \]

\[ X_1 \geq 0, \quad K_{i,j} \geq 0, \quad \forall i, \forall j \]

\[ X^i_2, X_{3}^{i,j} \in \mathcal{X}_{2,i}^*(X_1), \quad \forall i. \]
Second-stage problem, scenario $i$ in CTC model

$$\min_{X^2, X^3} (c^i_2)^T X^2_i + (1 - \beta_2) \sum_{j=1}^{N_2} \pi_{i,j} (c^i_3)^T X^i_{3,j} + \beta_2 h_i$$

s.t.

$$A^i_{22} X^i_2 = b^i_2 - A^i_{21} X^i_1,$$

$$A^i_{32} X^i_2 + A^i_{33} X^i_{3,j} = b^i_3 - A^i_{31} X^i_1, \quad \forall j$$

(2.12)

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

$$Y_{i,j} \geq (c^i_3)^T X^i_{3,j} - \bar{t}_i, \quad \forall j$$

$$Y_{i,j} \geq 0, \quad X^i_2 \geq 0, \quad X^i_{3,j} \geq 0, \quad \forall j.$$  

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

$$Z_1 + \text{BRM}^1_{\beta_1} (Z_2 + \text{BRM}^2_{\beta_2} (Z_3)) = Z_1 + (1 - \beta_1) \sum_{i=1}^{N_1} \pi_i \omega_i + \beta_1 k,$$

where

$$k \geq \tau + \frac{1}{\alpha_1} \sum_{i=1}^{N_1} \pi_i \chi_i$$

$$\chi_i \geq \omega_i - \tau, \quad \forall i$$

$$\omega_i = Z^i_2 + (1 - \beta_2) \sum_{j=1}^{N_2} \pi_{i,j} Z^i_{3,j} + \beta_2 h_i, \quad \forall i$$

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

$$Y_{i,j} \geq Z^i_{3,j} - \bar{t}_i, \quad \forall j$$

$$\chi_i \geq 0, \quad Y_{i,j} \geq 0, \quad \forall i, \forall j.$$  

The “OTC” formulation is:

**OTC model with nested BRM Risk Measure**

$$\min_{X_1, X_2, X_3} c^T_1 X_1 + (1 - \beta_1) \sum_{i=1}^{N_1} \pi_i \omega_i + \beta_1 k$$

s.t.

$$A_{11} X_1 = b_1$$

$$A^i_{21} X_1 + A^i_{22} X^i_2 = b^i_2, \quad \forall i$$

$$A^i_{31} X_1 + A^i_{32} X^i_2 + A^i_{33} X^i_{3,j} = b^i_{3,j}, \quad \forall i, \forall j$$

(2.13)
\[ k \geq \tau + \frac{1}{\alpha_1} \sum_{i=1}^{N_1} \pi_i \chi_i \]

\[ \chi_i \geq \omega_i - \tau, \quad \forall i \]

\[ \omega_i = Z^i_2 + (1 - \beta_2) \sum_{j=1}^{N_2} \pi_{i,j} Z^{i,j}_3 + \beta_2 h_i, \quad \forall i \]

\[ h_i \geq \bar{t}_i + \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, \quad \forall i \]

\[ Y_{i,j} \geq Z^{i,j}_3 - \bar{t}_i, \quad \forall i, \forall j \]

\[ X_1 \geq 0, \quad X_2^i \geq 0, \quad X_3^{i,j} \geq 0, \quad \chi_i \geq 0, \quad Y_{i,j} \geq 0, \quad \forall i, \forall j. \]

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_\gamma and BRM_\alpha_\beta, 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 \( BRM_\beta^{\alpha} \) of the total cost function. The risk-neutral and AVaR_\alpha 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**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td>Number of Parts</td>
</tr>
<tr>
<td>$P_2$</td>
<td>Number of products</td>
</tr>
<tr>
<td>$c$</td>
<td>Vector of part prices, in $\mathbb{R}^{P_1}$</td>
</tr>
<tr>
<td>$z_{up}$</td>
<td>Vector of maximum quantities of parts that may be ordered, in $\mathbb{R}^{P_1}$</td>
</tr>
<tr>
<td>$\pi_i$</td>
<td>Probability of second-stage scenario $i = 1, \ldots, N_1$</td>
</tr>
<tr>
<td>$\pi_{i,j}$</td>
<td>Conditional probability of third-stage scenario $(i, j)$, given the occurrence of second-stage scenario $i$, for $i = 1, \ldots, N_1$ and $j = 1, \ldots, N_2$</td>
</tr>
<tr>
<td>$r$</td>
<td>Vector of revenues from products, in $\mathbb{R}^{P_2}$</td>
</tr>
<tr>
<td>$l$</td>
<td>Vector of unit shortfall costs for products, in $\mathbb{R}^{P_2}$</td>
</tr>
<tr>
<td>$H_{i,j}$</td>
<td>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 $j = 1, \ldots, N_2$</td>
</tr>
<tr>
<td>$M$</td>
<td>$P_1 \times P_2$ matrix indicating how many units of each part are required to build one unit of each product</td>
</tr>
<tr>
<td>$\alpha_l$</td>
<td>Risk aversion parameter for AVaR$_{\alpha_l}$ for stage $l = 1, 2$</td>
</tr>
<tr>
<td>$\beta_l$</td>
<td>Coefficient of AVaR$_{\alpha_l}$ in the risk measure for stage $l = 1, 2$</td>
</tr>
<tr>
<td>$1 - \beta_l$</td>
<td>Coefficient of expected value in the risk measure for stage $i = 1, l$</td>
</tr>
</tbody>
</table>

Table 3.2: **Fundamental model variables**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z$</td>
<td>Quantities of parts ordered, in $\mathbb{R}^{P_1}$</td>
</tr>
<tr>
<td>$g$</td>
<td>First-stage AVaR$_{\alpha}$ value for CTC model</td>
</tr>
<tr>
<td>$h_i$</td>
<td>AVaR$_{\alpha}$ value, in second-stage scenario $i = 1, \ldots, N_1$</td>
</tr>
<tr>
<td>$u_i$</td>
<td>Vector of amounts of each product made (in $\mathbb{R}^{P_2}$), in second-stage scenario $i = 1, \ldots, N_1$</td>
</tr>
</tbody>
</table>
Table 3.3: Auxiliary variables needed for AVaR

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_i^-$</td>
<td>Vector of product shortfalls below demand (in $\mathbb{R}^{P_2}$), in second-stage scenario $i = 1, \ldots, N_1$</td>
</tr>
<tr>
<td>$s_i^+$</td>
<td>Vector of product units made in excess of demand (in $\mathbb{R}^{P_2}$), in second-stage scenario $i = 1, \ldots, N_1$</td>
</tr>
<tr>
<td>$w_i$</td>
<td>Second stage cost plus $BRM^{\alpha_2}_{\beta_2}$ value for OTC model, in second-stage scenario $i = 1, \ldots, N_1$</td>
</tr>
<tr>
<td>$k$</td>
<td>First-stage AVaR$_\alpha$ value of $w$ for OTC model</td>
</tr>
</tbody>
</table>

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**

$$
\min \quad c^T z + (1 - \beta_1) \sum_{i=1}^{N_1} \pi_i \left( -r^T u_i + l^T s_i^- + \left( \sum_{j=1}^{N_2} \pi_{i,j} H_{i,j}^T s_i^+ \right) \right) + \beta_1 g \\
\text{s.t.} \quad z_{up} \geq z \geq 0 \\
g \geq t + \frac{1}{\alpha_1} \sum_{i,j} \pi_i \pi_{i,j} X_{i,j}, \\
X_{i,j} \geq -r^T u_i + l^T s_i^- + H_{i,j}^T s_i^+ - t, \quad X_{i,j} \geq 0, \quad \forall i, \forall j \\
(u_i, s_i^-, s_i^+) \in X_i^*(z), \quad \forall i.
$$
Second-stage problem, scenario $i$

$$\begin{align*}
\min & \quad -r^Tu_i + l^Ts_i^- + (1 - \beta_2)(\sum_j \pi_{i,j}H^T_{i,j})s_i^+ + \beta_2h_i \\
\text{s.t.} & \quad z \geq Mu_i, \\
& \quad s_i^+ \geq u_i - D_i, \quad s_i^+ \geq 0, \\
& \quad s_i^- \geq -u_i + D_i, \quad s_i^- \geq 0, \\
& \quad h_i \geq \bar{t}_i + \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j}Y_{i,j}, \\
& \quad Y_{i,j} \geq H^T_{i,j}s_i^+ - \bar{t}_i, \quad Y_{i,j} \geq 0, \quad \forall j.
\end{align*}$$  \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 $(u_i, s_i^-, s_i^+ \in X^*_i(z)$ 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

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_i$</td>
<td>Vector of Lagrange multipliers for constraints $Z \geq Mu_i$ below, for $i = 1,\ldots,N_1$</td>
</tr>
<tr>
<td>$\mu_i$</td>
<td>Vector of Lagrange multipliers for constraints $s_i^+ \geq u_i - D_i$ below, for $i = 1,\ldots,N_1$</td>
</tr>
<tr>
<td>$\epsilon_i$</td>
<td>Vector of Lagrange multipliers for constraints $s_i^- \geq -u_i + D_i$ below, for $i = 1,\ldots,N_1$</td>
</tr>
<tr>
<td>$\rho_i$</td>
<td>Scalar Lagrange multipliers for constraint $h_i \geq \bar{t}<em>i + \frac{1}{\alpha} \sum</em>{j=1}^{N_2} \pi_{i,j}Y_{i,j}$ below, for $i = 1,\ldots,N_1$</td>
</tr>
<tr>
<td>$\delta_{i,j}$</td>
<td>Scalar Lagrange multipliers for constraint $Y_{i,j} \geq H^T_{i,j}s_i^+ - \bar{t}_i$ below, for $i = 1,\ldots,N_1, j = 1,\ldots,N_2$</td>
</tr>
</tbody>
</table>
CTC formulation of supply chain problem with BRM risk measure

\[
\min \quad c^T z + (1 - \beta_1) \sum_{i=1}^{N_1} \pi_i \left( -r^T u_i + l^T s_i^+ + \left( \sum_{j=1}^{N_2} \pi_{i,j} H_{i,j}^T s_i^+ \right) \right) + \beta_1 g
\]

s.t. 
\[
\begin{align*}
z_{\text{up}} &\geq z \geq 0 \\
g &\geq t + \frac{1}{\alpha_1} \sum_{i,j} \pi_i \pi_{i,j} X_{i,j} \\
X_{i,j} &\geq -r^T u_i + l^T s_i^+ + H_{i,j}^T s_i^+ - t, \quad \forall i, \forall j \\
z &\geq Mu_i, \quad \forall i \\
s_i^+ &\geq u_i - D_i, \quad \forall i \\
s_i^- &\geq -u_i + D_i, \quad \forall i \\
h_i &\geq \bar{t}_i + \frac{1}{\alpha_2} \sum_j \pi_{i,j} Y_{i,j}, \quad \forall i, \forall j \\
Y_{i,j} &\geq H_{i,j}^T s_i^+ - \bar{t}_i, \quad \forall i, \forall j \\
- r^T + \lambda_i^T M + \mu_i^T - \epsilon_i^T &\geq 0, \quad \forall i \\
l^T - \epsilon_i^T &\geq 0, \quad \forall i \\
- \mu_i^T + \sum_{j=1}^{N_i} \left( 1 - \beta_2 \right) \pi_{i,j} + \delta_{i,j} \right) H_{i,j}^T &\geq 0, \quad \forall i, \forall j \\
- \delta_{i,j} + \frac{\rho_i}{\alpha_2} \pi_{i,j} &\geq 0, \quad \forall i, \forall j \\
\beta_2 - \rho_i &= 0, \quad \forall i \\
- \sum_{j=1}^{N_i} \delta_{i,j} + \rho_i &= 0, \quad \forall i \\
\lambda_i^T (z - Mu_i) &= 0, \quad \forall i \\
\mu_i^T (s_i^+ - u_i + D_i) &= 0, \quad \forall i \\
\epsilon_i^T (s_i^- + u_i - D_i) &= 0, \quad \forall i \\
\delta_{i,j} (Y_{i,j} - H_{i,j}^T s_i^+ + \bar{t}_i) &= 0, \quad \forall i, \forall j \\
\rho_i (h_i - \bar{t}_i - \frac{1}{\alpha_2} \sum_j \pi_{i,j} Y_{i,j}) &= 0, \quad \forall i, \forall j \\
\left( -r^T + \lambda_i^T M + \mu_i^T - \epsilon_i^T \right) u_i &= 0, \quad \forall i
\end{align*}
\]
\((l^T - \epsilon_i^T)s_i^- = 0, \forall i\)

\((-\mu_i^T + \sum_{j=1}^{N_i}((1 - \beta_2)\pi_{i,j} + \delta_{i,j})H_{i,j}^T)s_i^+ = 0, \forall i, \forall j\)

\((-\delta_{i,j} + \frac{\rho_i}{\alpha_2}\pi_{i,j})Y_{i,j} = 0, \forall i, \forall j\)

\(u_i, s_i^-, s_i^+, \chi_i, \mu_i, \epsilon_i, \rho_i \geq 0, \forall i\)

\(X_{i,j} \geq 0, Y_{i,j} \geq 0, \delta_{i,j} \geq 0, \forall i, \forall j\).

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**

\[
\begin{align*}
\min \quad & c^T z + (1 - \beta_1)\sum_{i=1}^N \pi_i w_i + \beta_1 k \\
\text{s.t.} \quad & z_{up} \geq z \geq 0 \\
\quad & z \geq Mu_i, \forall i \\
\quad & s_i^+ \geq u_i - D_i, \forall i \\
\quad & s_i^- \geq -u_i + D_i, \forall i \\
\quad & w_i = l^T s_i^- - r^T u_i + (1 - \beta_2)(\sum_{j=1}^{N_i} \pi_{i,j}H_{i,j}^T s_i^+) + \beta_2 h_i, \forall i \\
\quad & k \geq \tau + \frac{1}{\alpha_1} \sum_{i=1}^N \pi_i \chi_i, \forall i \\
\quad & \chi_i \geq w_i - t, \forall i \\
\quad & Y_{i,j} \geq H_{i,j}^T s_i^+ - \bar{t}_i, \forall i \\
\quad & h_i \geq \bar{t}_i + \frac{1}{\alpha_2} \sum_{j=1}^{N_i} \pi_{i,j} Y_{i,j}, \forall i, \forall j \\
\quad & \chi_i, Y_{i,j}, u_i, s_i^+, s_i^- \geq 0, \forall i, \forall j. 
\end{align*}
\]
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\(\gamma\). 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
\(\gamma_i\) Risk aversion parameter for MSD\(\gamma\) risk measure for stage 
l = 1, 2

Table 3.6: Fundamental model variables
\(V_i\) Second stage cost plus MSD\(\gamma_2\) value for OTC model, in second-stage scenario \(i = 1, \ldots, N\)

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 \(j = 1, \ldots, N_2\)
\(\chi_i\) Amount \(V_i\) exceeds its overall mean for OTC model in second stage scenario \(i\), for \(i = 1, \ldots, N\)

Table 3.8: Lagrange multiplier variables MSD
\(\lambda_i\) Vector of Lagrange multipliers for constraints \(z \geq Mu_i\) below, for \(i = 1, \ldots, N\)
\(\xi_i\) Vector of Lagrange multipliers for constraints \(s_i^- \geq -u_i + D_i\) below, for \(i = 1, \ldots, N\)
\(\mu_i\) Vector of Lagrange multipliers for constraints \(s_i^+ \geq u_i - D_i\) below, for \(i = 1, \ldots, N\)
\[ \delta_{i,j} \quad \text{Scalar Lagrange multiplier for each constraint} \quad Y_{i,j} \geq \left( H_{i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} H_{i,j'} \right) \quad \text{for} \quad i = 1, \ldots, N_1 \quad \text{and} \quad j = 1, \ldots, N_2 \]

**CTC formulation of supply chain problem with MSD risk measure**

\[
\begin{align*}
\min & \quad c^T z + \sum_{i=1}^{N_1} \pi_i \left( -r^T u_i + l^T s_i^- + \left( \sum_{j=1}^{N_2} \pi_{i,j} H_{i,j} \right) s_i^+ \right) + \gamma \sum_{i=1}^{N_1} \pi_i \left( \sum_{j=1}^{N_2} \pi_{i,j} X_{i,j} \right) \\
\text{s.t.} & \quad 0 \leq z \leq z_{up} \\
& \quad X_{i,j} \geq -r^T u_i + l^T s_i^- + H_{i,j}^T s_i^+ \\
& \quad + \sum_{i'=1}^{N_1} \pi_{i,i'} \left( -r^T u_{i'} + l^T s_{i'}^- + \left( \sum_{j'=1}^{N_2} \pi_{i',j'} H_{i',j'}^T \right) s_{i'}^+ \right), \quad \forall i, \forall j \\
& \quad z \geq Mu_i, \quad \forall i \\
& \quad s_i^+ \geq u_i - D_i, \quad \forall i \\
& \quad s_i^- \geq -u_i + D_i, \quad \forall i \\
& \quad Y_{i,j} \geq \left( H_{i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} H_{i,j'} \right) s_i^+, \quad \forall i, \forall j \\
& \quad M^T \lambda_i + \mu_i - \xi_i \geq r, \quad \forall i \\
& \quad \xi_i^T \leq l, \quad \forall i \\
& \quad -\mu_i + \sum_{j=1}^{N_2} \left( \pi_{i,j} \left( 1 - \sum_{j'=1}^{N_2} \delta_{i,j'} \right) + \delta_{i,j} \right) H_{i,j} \geq 0, \quad \forall i \\
& \quad \delta_{i,j} \leq \gamma 2 \pi_{i,j}, \quad \forall i, \forall j \\
& \quad \lambda_i^T (z - Mu_i) = 0, \quad \forall i \\
& \quad \mu_i^T (s_i^+ - u_i + D_i) = 0, \quad \forall i \\
& \quad \xi_i^T (s_i^- + u_i - D_i) = 0, \quad \forall i \\
& \quad \delta_{i,j} \left( Y_{i,j} - H_{i,j}^T s_i^+ + \sum_{j'=1}^{N_2} \pi_{i,j'} H_{i,j'}^T s_i^+ \right) = 0, \quad \forall i, \forall j \\
& \quad (-r + M^T \lambda_i + \mu_i - \xi_i)^T u_i = 0, \quad \forall i \\
\end{align*}
\]
(l - \xi_i)^T s_i^- = 0, \forall i

\left( -\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} \right)^T s_i^+ = 0, \forall i

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

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

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

OTC formulation of supply chain problem with nested MSD risk measure

\[
\min \ c^T z + \sum_{i=1}^{N_1} \pi_i V_i + \gamma_1 \left( \sum_{i=1}^{N_1} \pi_i \chi_i \right)
\]

s.t. 

\[
z_{up} \geq z \geq 0
\]

\[
z \geq Mu_i, \forall i
\]

\[
s_i^+ \geq u_i - D_i, \forall i
\]

\[
s_i^- \geq -u_i + D_i, \forall i
\]

\[
\chi_i \geq V_i - \sum_{i=1}^{N_1} \pi_i V_i, \forall i
\]

\[
V_i = l^T s_i^- - r^T u_i + \sum_{j=1}^{N_2} \pi_{i,j} H_{i,j}^T s_i^+ + \gamma_2 \left( \sum_{i=1}^{N_2} \pi_{i,j} Y_{i,j} \right), \forall i
\]

\[
Y_{i,j} \geq H_{i,j}^T s_i^+ - \sum_{j'=1}^{N_2} \pi_{i,j'} H_{i,j'}^T s_i^+, \forall i, \forall j
\]

\[
\chi_i, Y_{i,j}, u_i, s_i^+, s_i^- \geq 0, \forall i, \forall j.
\]

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 $\gamma = 0.3, 0.7, \text{and } 0.9$ for MSD$_\gamma$ and $(\alpha, \beta) = (0.01, 0.5), (0.05, 0.5), (0.25, 0.5), (0.05, 0.25), \text{and } (0.25, 0.75)$ for BRM$_\beta$. 
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 \(i \in \{\text{OTC, CTC}\}\), respectively. To measure the deviation of the CTC model from the OTC formulation, define the relative difference of the two models as

\[
\text{Diff} = \frac{\text{Obj}^{\text{CTC}}(\text{Sol}^{\text{OTC}}) - \text{Obj}^{\text{CTC}}(\text{Sol}^{\text{CTC}})}{\text{Obj}^{\text{CTC}}(\text{Sol}^{\text{OTC}})}.
\]

A larger difference translates to a worse solution suggested by the OTC approach and justifies the computational costs of the CTC model. Also, let \(\Delta_1 (\Delta_2)\) denote the maximum percentage difference between the first-stage (second-stage) optimal solutions of the OTC and CTC problems. Consequently, a higher value of \(\Delta_i\) corresponds to a higher deviation between the two models.

Tables 3.9 and 3.10 summarize the results for risk measures MSD\(_{\gamma}\) and BRM\(_{\alpha\beta}\), 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 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 \(\text{Obj}^{\text{CTC}}(\text{Sol}^{\text{OTC}}) = -0.0016\), and \(\text{Obj}^{\text{CTC}} = -17.4291\) 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 \(\gamma\), 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.
Table 3.9: Computational results of the supply chain problem with the MSD risk measure

<table>
<thead>
<tr>
<th></th>
<th>$\gamma = 0.3$</th>
<th>$\gamma = 0.7$</th>
<th>$\gamma = 0.9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diff</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$&lt; 0.01%$</td>
<td>54.42%</td>
<td>20.90%</td>
<td>13.56%</td>
</tr>
<tr>
<td>$0.01% \leq \text{Diff} &lt; 0.1%$</td>
<td>16.76%</td>
<td>12.28%</td>
<td>8.56%</td>
</tr>
<tr>
<td>$0.1% \leq \text{Diff} &lt; 1%$</td>
<td>22.76%</td>
<td>37.32%</td>
<td>35.40%</td>
</tr>
<tr>
<td>$1% \leq \text{Diff} &lt; 10%$</td>
<td>5.50%</td>
<td>25.18%</td>
<td>35.36%</td>
</tr>
<tr>
<td>$10% \leq \text{Diff}$</td>
<td>0.56%</td>
<td>4.32%</td>
<td>7.12%</td>
</tr>
<tr>
<td>Max Diff</td>
<td>3,333.66%</td>
<td>1,094,395%</td>
<td>2,596.47%</td>
</tr>
<tr>
<td>$\Delta_1$</td>
<td>8.84%</td>
<td>4.20%</td>
<td>6.33%</td>
</tr>
<tr>
<td>$\Delta_2$</td>
<td>19.18%</td>
<td>13.31%</td>
<td>13.92%</td>
</tr>
</tbody>
</table>

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 BRM$^{0.05}$ 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 Gultan & Ruszczynski (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$_\gamma$ 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>
<thead>
<tr>
<th>Table 3.11: Given Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
</tr>
<tr>
<td>Number of assets</td>
</tr>
</tbody>
</table>
\[ R_{i}^{1} \] 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 \( j = 1, \ldots, N_{2} \)

\( \pi_{i} \) 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 \( j = 1, \ldots, N_{2} \)

\( \epsilon \) Vector of relative transaction cost of securities in \( \mathbb{R}^{r} \)

\( \gamma_{i} \) Risk aversion parameter for MSD in 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 \( j = 1, \ldots, N_{2} \)
$\chi_i$ 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 as our risk measure, the first-stage objective function is:

$$\text{MSD}_{\gamma_1}(Z_3) = E(Z_3) + \gamma_1 E\left(\left( Z_3 - E(Z_3) \right)_+ \right) = -E(W) + \gamma_1 E\left(\left( E(W) - W \right)_+ \right)$$

where

$$X_{i,j} \geq \sum_{i' = 1}^{N_1} \sum_{j' = 1}^{N_2} \pi_{i,i'} \pi_{i',j'} W_{i',j'} - W_{i,j}, \quad i = 1, \ldots, N_1, \quad j = 1, \ldots, N_2$$

and the second-stage scenario $i$ objective function is:

$$\text{MSD}_{\gamma_2}(Z_3|i) = E(Z_3|i) + \gamma_2 E\left(\left( Z_3 - E(Z_3) \right)_+ | i \right) = -E(W|i) + \gamma_2 E\left(\left( E(W) - W \right)_+ | i \right)$$

where

$$Y_{i,j} \geq \sum_{j' = 1}^{N_2} \pi_{i,j'} W_{i,j'} - W_{i,j}, \quad j = 1, \ldots, N_2$$

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**

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

\text{(3.7)}
Second-stage problem, scenario $i$:

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

(3.8)

In the above problem, “$\circ$” is Hadamard product, element-wise product of two matrices with the same dimension. Also, $1$ 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**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_i$</td>
<td>Vector of Lagrange multipliers for constraints $y_i - (1 + R_{1,i}^1) \circ z - (1 - \epsilon) \circ b_i = 0$ for $i = 1, \ldots, N_1$</td>
</tr>
<tr>
<td>$\mu_i$</td>
<td>Scalar Lagrange multiplier for constraints $b_i^T 1 - s_i^T 1 = 0$, for $i = 1, \ldots, N_1$</td>
</tr>
<tr>
<td>$\rho_{i,j}$</td>
<td>Scalar Lagrange multiplier for constraints $W_{i,j} - (1 + R_{i,j}^2)^T y_i = 0$, for $i = 1, \ldots, N_1$, $j = 1, \ldots, N_2$</td>
</tr>
<tr>
<td>$\sigma_{i,j}$</td>
<td>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 $i = 1, \ldots, N_1$, $j = 1, \ldots, N_2$</td>
</tr>
</tbody>
</table>

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:

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

By replacing constraints $W_i \in X_i^*(z)$ 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$_\gamma$ risk measure

$$
\begin{align*}
\text{min} & \quad -\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} W_{i,j} + \gamma_1 \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j} \\
\text{s.t.} & \quad z^T 1 = 1, \quad z \geq 0, \quad \forall i \\
& \quad X_{i,j} \geq \sum_{j'=1}^{N_1} \sum_{j'=1}^{N_2} \pi_{i,j'} W_{i',j'} - W_{i,j}, \quad X_{i,j} \geq 0, \quad \forall i, \forall j \\
& \quad y_i = (1 + R_i^1) \circ z + (1 - \epsilon) \circ b_i - (1 + \epsilon) \circ s_i, \quad \forall i \\
& \quad b_i^T 1 - s_i^T 1 = 0, \quad \forall i \\
& \quad W_{i,j} = (1 + R_{i,j}^2)^T y_i, \quad \forall i, \forall j \\
& \quad Y_{i,j} \geq \sum_{j'=1}^{N_2} \pi_{i,j'} W_{i,j'} - W_{i,j}, \quad \forall i, \forall j \\
& \quad -\lambda_i + \sum_{j=1}^{N_2} \rho_{i,j} (1 + R_{i,j}^2) \geq 0, \quad \forall i \\
& \quad (1 - \epsilon) \circ \lambda_i - \mu_i 1 \geq 0, \quad \forall i \quad (3.9) \\
& \quad - (1 + \epsilon) \circ \lambda_i + \mu_i 1 \geq 0, \quad \forall i \\
& \quad -\pi_{i,j} - \rho_{i,j} - \sigma_{i,j} + \pi_{i,j} \sum_{j'=1}^{N_2} \sigma_{i,j'} = 0, \quad \forall i, \forall j \\
& \quad \gamma_2 \pi_{i,j} - \sigma_{i,j} \geq 0, \quad \forall i, \forall j \\
& \quad \sigma_{i,j} (Y_{i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} W_{i,j'} + W_{i,j}) = 0, \quad \forall i, \forall j \\
& \quad y_i^T (-\lambda_i + \sum_{j=1}^{N_2} \rho_{i,j} (1 + R_{i,j}^2)) = 0, \quad \forall i \\
& \quad b_i^T ((1 - \epsilon) \circ \lambda_i - \mu_i 1) = 0, \quad \forall i \\
& \quad s_i^T ((1 + \epsilon) \circ \lambda_i + \mu_i 1) = 0, \quad \forall i \\
& \quad Y_{i,j} (\gamma_2 \pi_{i,j} - \sigma_{i,j}) = 0, \quad \forall i, \forall j \\
& \quad y_i \geq 0, \ b_i \geq 0, \ s_i \geq 0, \ Y_{i,j} \geq 0, \ \sigma_{i,j} \geq 0, \quad \forall i, \forall j.
\end{align*}
$$
Above, “∀i” is a shorthand for “i = 1, . . . , N1”, and “∀j” is a shorthand for “j = 1, . . . , N2.” If the decision maker uses a nested risk measure, ρ1(ρ2(Z3)), for the first-stage objective function, the “OTC” formulation is:

**OTC formulation of portfolio problem with nested MSD risk measure**

\[
\begin{align*}
\text{min} & \quad \sum_{i=1}^{N_1} \pi_i V_i + \gamma_1 \left( \sum_{i=1}^{N_1} \pi_i \chi_i \right) \\
\text{s.t.} & \quad z^T \mathbf{1} = 1 \\
& \quad y_i = (1 + R^i_1) \circ z + (1 - \epsilon) \circ b_i - (1 + \epsilon) \circ s_i, \quad \forall i \\
& \quad b_i^T \mathbf{1} - s_i^T \mathbf{1} = 0, \quad \forall i \\
& \quad W_{i,j} = (1 + R^2_{i,j})^T y_i, \quad \forall i, \forall j \\
& \quad Y_{i,j} \geq \sum_{j'=1}^{N_2} \pi_{i,j'} W_{i,j'} - W_{i,j}, \quad \forall i, \forall j \\
& \quad V_i = -\left( \sum_{j=1}^{N_2} \pi_{i,j} W_{i,j} \right) + \gamma_2 \left( \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j} \right), \quad \forall i \\
& \quad \chi_i \geq V_i - \sum_{i=1}^{N_1} \pi_i V_i, \quad \forall i \\
& \quad z \geq \mathbf{0}, \quad b_i \geq \mathbf{0}, \quad d_i \geq \mathbf{0}, \quad \chi_i \geq \mathbf{0}, \quad Y_{i,j} \geq \mathbf{0}, \quad \forall i, \forall j.
\end{align*}
\]  

(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. 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**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>α_l</td>
<td>Risk aversion parameter for AVaR_{α_l} for stage l, 1, 2</td>
</tr>
<tr>
<td>β_l</td>
<td>Coefficient of AVaR_{α_l} in the risk measure for stage l, 1, 2</td>
</tr>
<tr>
<td>1 - β_l</td>
<td>Coefficient of expected value in the risk measure for stage l, 1, 2</td>
</tr>
</tbody>
</table>


Table 3.16: **Fundamental model variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g$</td>
<td>First-stage AVaR$_\alpha$ value for CTC model</td>
</tr>
<tr>
<td>$h_i$</td>
<td>AVaR$_\alpha$ value, in second-stage scenario $i = 1, \ldots, N$</td>
</tr>
<tr>
<td>$V_i$</td>
<td>Second stage cost plus BRM$_{\beta_2}^{\alpha}$ value for OTC model, in second-stage scenario $i = 1, \ldots, N$</td>
</tr>
<tr>
<td>$k$</td>
<td>First-stage AVaR$_{\alpha_1}$ value of $V_i$ for OTC model</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>Helper variable in AVaR definition in stage 1 for CTC model</td>
</tr>
<tr>
<td>$X_{i,j}$</td>
<td>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 $j = 1, \ldots, N_2$</td>
</tr>
<tr>
<td>$\bar{t}_i$</td>
<td>Helper variable in AVaR definition in stage 2 scenario $i$, for $i = 1, \ldots, N_1$</td>
</tr>
<tr>
<td>$K_{i,j}$</td>
<td>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 $j = 1, \ldots, N_2$</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Helper variable in AVaR$_{\alpha_1}$ definition in stage 1 for OTC model</td>
</tr>
<tr>
<td>$\chi_i$</td>
<td>Positive part of difference between $V_i$ and $\tau$, for $i = 1, \ldots, N_1$</td>
</tr>
</tbody>
</table>

The first stage objective function for CTC formulation is:

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

where

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

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

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

and the second stage objective function is:

$$\rho_2(Z_3|i) = (1 - \beta_2)E(Z_3|i) + \beta_2 \ast AVaR(Z_3|i) = (1 - \beta_2)\sum_{j=1}^{N_2} \pi_{i,j}(-W_{i,j}) + \beta_2 h_i$$
where,
\[ h_i \geq \bar{t}_i + \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j} K_{i,j}, \quad j = 1, \ldots, N_2, \]
\[ K_{i,j} \geq -W_{i,j} - \bar{t}_i, \quad j = 1, \ldots, N_2, \]
\[ K_{i,j} \geq 0, \quad j = 1, \ldots, N_2. \]

First-stage Problem

\[
\begin{align*}
\text{min} & \quad (1 - \beta_1) \left( \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} (-W_{i,j}) \right) + \beta_1 g \\
\text{s.t.} & \quad z^T 1 = 1 \\
& \quad g \geq t + \frac{1}{\alpha_1} \left( \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j} \right), \\
& \quad X_{i,j} \geq -W_{i,j} - t, \quad \forall i, \forall j \\
& \quad z \geq 0, \quad X_{i,j} \geq 0, \quad \forall i, \forall j \\
& \quad W_i \in X^*_{i}(z), \quad \forall i.
\end{align*}
\] (3.11)

Second-stage problem, scenario i:

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

In the above problem, “\( \circ \)” is Hadamard product, element-wise product of two matrices with the same dimension. Also, 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**

| \( \lambda_i \) | Vector of Lagrange multipliers for constraints \( y_i - (1 + R^1_i) \circ z \) for \( i = 1, \ldots, N_1 \) |
| \( \mu_i \) | 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 \( W_{i,j} - (1 + R^2_{i,j})^T y_i = 0 \), for \( i = 1, \ldots, N_1, j = 1, \ldots, N_2 \) |
| \( \theta_i \) | Scalar Lagrange multiplier for constraints \( h_i \geq \bar{t}_i + \frac{1}{\alpha_i} \sum_{j=1}^{N_2} \pi_{i,j} K_{i,j} \) 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:

\[
L_i = (1 - \beta_2) \sum_{j=1}^{N_2} \pi_{i,j}(-W_{i,j}) + \beta_2 h_i - \lambda_i(y_i - (1 + R^1_i) \circ z - (1 - \epsilon) \circ b_i) \\
+ (1 + \epsilon) \circ s_i - \mu_i(b_i^T \mathbf{1} - s_i^T \mathbf{1}) - \sum_{j=1}^{N_2} \rho_{i,j}(W_{i,j} - (1 + R^2_{i,j})^T y_i) \\
- \theta_i(h_i - \bar{t}_i - \frac{1}{\alpha_i} \sum_{j=1}^{N_2} \pi_{i,j} K_{i,j}) - \sum_{j=1}^{N_2} \sigma_{i,j}(K_{i,j} + W_{i,j} + \bar{t}_i) \\
= y_i^T(-\lambda_i + \sum_{j=1}^{N_2} \pi_{i,j}(1 + R^2_{i,j})) + b_i^T((1 - \epsilon) \circ \lambda_i - \mu_i \mathbf{1}) \\
+ s_i^T(-1 + \epsilon) \circ \lambda_i + \mu_1) + \sum_{j=1}^{N_2} W_{i,j}\left(-1 - (1 - \beta_2) \pi_{i,j} - \rho_{i,j} - \sigma_{i,j}\right) \\
+h_i(\beta_2 - \theta_i) + \bar{t}_i(\theta_i - \sum_{j=1}^{N_2} \sigma_{i,j}) + \sum_{j=1}^{N_2} K_{i,j}\left(\frac{1}{\alpha_i} \pi_{i,j} \theta_i - \sigma_{i,j}\right) \\
+ \lambda_i^T(1 + R^1_i) \circ z
\]

**CTC formulation of portfolio problem with BRM risk measure**

\[
\min \quad (1 - \beta_1) \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j}(-W_{i,j}) + \beta_1 g
\]

s.t. \( z^T \mathbf{1} = 1 \)

\[
g \geq t + \frac{1}{\alpha_1} \left(\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j}\right), \tag{3.13}
\]

\[
X_{i,j} \geq -W_{i,j} - t, \quad \forall i, \forall j
\]

\[
y_i = (1 + R^1_i) \circ z + (1 - \epsilon) \circ b_i - (1 + \epsilon) \circ s_i, \quad \forall i
\]
\begin{align*}
& b_i^T \mathbf{1} - s_i^T \mathbf{1} = 0, \\
& W_{i,j} = (1 + R_{i,j}^2) y_i, \\
& 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, \\
& -\lambda_i + \sum_{j=1}^{N_2} \rho_{i,j} (1 + R_{i,j}^2) \geq 0, \\
& (1 - \epsilon) \odot \lambda_i - \mu_i \mathbf{1} \geq 0, \\
& -(1 + \epsilon) \odot \lambda_i + \mu_i \mathbf{1} \geq 0, \\
& -(1 - \beta_2) \pi_{i,j} - \rho_{i,j} - \sigma_{i,j} \geq 0, \\
& \beta_2 - \theta_i = 0, \\
& \theta_i - \sum_{j=1}^{N_2} \sigma_{i,j} = 0, \\
& \frac{1}{\alpha_2} \pi_{i,j} \theta_i - \sigma_{i,j} \geq 0, \\
& \theta_i (h_i - \bar{t}_i - \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j} K_{i,j}) = 0, \\
& \sigma_{i,j} (K_{i,j} + W_{i,j} + \bar{t}_i) = 0, \\
& y_i^T (-\lambda_i + \sum_{j=1}^{N_2} \rho_{i,j} (1 + R_{i,j}^2)) = 0, \\
& b_i^T ((1 - \epsilon) \odot \lambda_i - \mu_i \mathbf{1}) = 0, \\
& s_i^T (-(1 + \epsilon) \odot \lambda_i + \mu_i \mathbf{1}) = 0, \\
& \sum_{j=1}^{N_2} K_{i,j} (\frac{1}{\alpha_2} \pi_{i,j} \theta_i - \sigma_{i,j}) = 0, \\
& z \geq 0, X_{i,j} \geq 0, \\
& y_i \geq 0, b_i, s_i \geq 0, K_{i,j} \geq 0, \\
& \theta_i, \sigma_{i,j} \geq 0,
\end{align*}

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

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

(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\(_{\gamma}\) and BRM\(_{\beta}\) with parameter values \(\gamma = 0.3, 0.7, \text{and} 0.9\) and \((\alpha, \beta) = (0.01, 0.5), (0.05, 0.5), (0.25, 0.5), (0.05, 0.25), \text{and} (0.25, 0.75)\). Tables 3.19 and 3.20 summarize the results for MSD\(_{\gamma}\) and BRM\(_{\beta}\) risk measures, respectively.

With the MSD risk measure, in approximately 81%, 58%, and 49% of the instances with \(\gamma\) 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
\[ \gamma = 0.3 \quad \gamma = 0.7 \quad \gamma = 0.9 \]

<table>
<thead>
<tr>
<th>Diff</th>
<th>( \gamma = 0.3 )</th>
<th>( \gamma = 0.7 )</th>
<th>( \gamma = 0.9 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.01%</td>
<td>80.98%</td>
<td>58.04%</td>
<td>48.76%</td>
</tr>
<tr>
<td>0.01% ≤ Diff &lt; 0.1%</td>
<td>6.04%</td>
<td>9.56%</td>
<td>9.02%</td>
</tr>
<tr>
<td>0.1% ≤ Diff &lt; 1%</td>
<td>11.62%</td>
<td>22.74%</td>
<td>27.3%</td>
</tr>
<tr>
<td>1% ≤ Diff &lt; 10%</td>
<td>1.36%</td>
<td>9.66%</td>
<td>14.92%</td>
</tr>
<tr>
<td>10% ≤ Diff</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Max Diff</td>
<td>2.14%</td>
<td>6.59%</td>
<td>9.03%</td>
</tr>
<tr>
<td>( \Delta_1 )</td>
<td>121.33%</td>
<td>99.67%</td>
<td>102.45%</td>
</tr>
<tr>
<td>( \Delta_2 )</td>
<td>40.70%</td>
<td>62.04%</td>
<td>58.20%</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Diff</th>
<th>( \beta = 0.5 )</th>
<th>( \beta = 0.5 )</th>
<th>( \beta = 0.75 )</th>
<th>( \beta = 0.5 )</th>
<th>( \beta = 0.75 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = 0.01 )</td>
<td>( \alpha = 0.05 )</td>
<td>( \alpha = 0.05 )</td>
<td>( \alpha = 0.25 )</td>
<td>( \alpha = 0.25 )</td>
<td></td>
</tr>
<tr>
<td>&lt; 0.01%</td>
<td>17.86%</td>
<td>19.88%</td>
<td>10.3%</td>
<td>30.1%</td>
<td>16.88%</td>
</tr>
<tr>
<td>0.01% ≤ Diff &lt; 0.1%</td>
<td>3.46%</td>
<td>4.5%</td>
<td>4.8%</td>
<td>6.9%</td>
<td>5.54%</td>
</tr>
<tr>
<td>0.1% ≤ Diff &lt; 1%</td>
<td>23.12%</td>
<td>26.14%</td>
<td>32.52%</td>
<td>28.48%</td>
<td>30.86%</td>
</tr>
<tr>
<td>1% ≤ Diff &lt; 10%</td>
<td>52.3%</td>
<td>47.78%</td>
<td>51.24%</td>
<td>34.24%</td>
<td>45.22%</td>
</tr>
<tr>
<td>10% ≤ Diff</td>
<td>3.26%</td>
<td>1.7%</td>
<td>1.14%</td>
<td>0.28%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Max Diff</td>
<td>34.18%</td>
<td>29.85%</td>
<td>37.05%</td>
<td>16.49%</td>
<td>43.50%</td>
</tr>
<tr>
<td>( \Delta_1 )</td>
<td>162.13%</td>
<td>102.31%</td>
<td>50.51%</td>
<td>112.47%</td>
<td>102.86%</td>
</tr>
<tr>
<td>( \Delta_2 )</td>
<td>84.066%</td>
<td>66.87%</td>
<td>56.53%</td>
<td>61.41%</td>
<td>105.52%</td>
</tr>
</tbody>
</table>

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 \( \gamma \)). 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$\gamma$ 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, 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

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>Set of thermal plants</td>
</tr>
<tr>
<td>$H$</td>
<td>Set of hydro plants</td>
</tr>
<tr>
<td>$c_1$</td>
<td>Vector of unit cost of generating electricity in thermal plants, in $\mathbb{R}^{</td>
</tr>
<tr>
<td>$c_{2,i}$</td>
<td>Vector of unit cost of generating electricity in thermal plants in second-stage scenario $i$, for $i = 1, \cdots, N_1$, in $\mathbb{R}^{</td>
</tr>
<tr>
<td>$c_{3,i,j}$</td>
<td>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 $\mathbb{R}^{</td>
</tr>
<tr>
<td>$c_l$</td>
<td>Unit cost of load-shedding ($$/GJ)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Production coefficient vector for hydro plants, in $\mathbb{R}^{</td>
</tr>
<tr>
<td>$\bar{g}$</td>
<td>Thermal generation capacity vector</td>
</tr>
<tr>
<td>$a_1$</td>
<td>Vector of hydrologic trend of plants in the first stage, in $\mathbb{R}^{</td>
</tr>
<tr>
<td>$a_{2,i}$</td>
<td>Vector of hydrologic trend of plants in second-stage scenario $i$, for $i = 1, \cdots, N_1$, in $\mathbb{R}^{</td>
</tr>
<tr>
<td>$a_{3,i,j}$</td>
<td>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}^{</td>
</tr>
</tbody>
</table>
\( v_1 \) Reservoir storage level of hydro plants in the beginning of first stage, in \( \mathbb{R}^{[H]} \)

\( \pi_i \) 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 \( i = 1, \ldots, N_1 \) and \( j = 1, \ldots, N_2 \)

\( M \) Incidence matrix of hydro plants

\( \bar{v} \) Vector of reservoir capacities, in \( \mathbb{R}^{[H]} \)

\( \bar{q} \) Limits on turbine outflow, in \( \mathbb{R}^{[H]} \)

\( q \) Lower bounds on total outflow, in \( \mathbb{R}^{[H]} \)

\( \gamma_l \) Risk aversion parameter for MSD\(_l\) risk measure for stage \( l = 1, 2 \)

Table 3.22: **Fundamental model variables**

\( G_1 \) Total generation of thermal plants in first stage, in \( \mathbb{R}^{[T]} \)

\( G_{2,i} \) 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 \( i = 1, \ldots, N_1 \)

\( L_{3,i,j} \) 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 \( W_i \)

\[ W_i \] Second stage cost plus MSD value of third stage cost for OTC model, in second stage scenario \( i \), for \( i = 1, \ldots, N_1 \) scenario \((i, j)\), for \( i = 1, \ldots, N_1 \) and \( j = 1, \ldots, N_2 \)

\[ \chi_i \] Amount \( W_i \) exceeds its overall mean in for OTC model second-stage scenario \( i \), for \( i = 1, \ldots, N_1 \)

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**

\[
\begin{align*}
\min & \quad c_i^T L_1 + \sum_{i=1}^{N_1} \pi_i c_i^T L_{2,i} + \sum_{j=1}^{N_2} \pi_{i,j} c_i^T L_{3,i,j} + c_i^T G_1 + \sum_{i=1}^{N_1} \pi_i c_{2,i}^T G_{2,i} + \sum_{j=1}^{N_2} \pi_{i,j} c_{3,i,j}^T G_{3,i,j} + \gamma \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_{i,j} X_{i,j} \\
\text{s.t.} & \quad X_{i,j} \geq c_i^T (L_{2,i} + L_{3,i,j}) + c_{2,i}^T G_{2,i} + c_{3,i,j}^T G_{3,i,j} - \sum_{i'=1}^{N_1} \pi_{i'} [c_{i'}^T L_{2,i'} + c_{2,i'}^T G_{2,i'} + \sum_{j'=1}^{N_2} \pi_{i',j'} (c_{3,i',j'} G_{3,i',j'} + c_{3,i',j'}^T G_{3,i',j'})]
\end{align*}
\]  

(3.15)
\[\begin{align*}
\rho^T Q_1 + 1^T G_1 + L_1 &= d_1 \\
V_2 &= v_1 + a_1 + M(Q_1 + S_1), \\
V_2 &\leq \bar{v}, \\
Q_1 &\leq \bar{q}, \\
Q_1 + S_1 &\geq q, \\
G_1 &\leq \bar{g}
\end{align*}\]

\(L_1, G_1, X_{i,j}, Q_1, V_2, S_1 \geq 0, \quad \forall i, j\)

\((G_{2,i}, G_{3,i,j}, L_{2,i}, L_{3,i,j}) \in X_i^*(V_2), \quad \forall i\)

**Second-stage problem, scenario \(i\)**

\[
\begin{align*}
\min \quad & c_1^T L_{2,i} + c_{2,i}^T G_{2,i} + \sum_{j=1}^{N_2} \pi_{i,j} (c_1^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j}) + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j} \\
\text{s.t.} \quad & Y_{i,j} \geq c_1^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} (c_1^T L_{3,i,j'} + c_{3,i,j'}^T G_{3,i,j'}) \\
&\rho^T Q_{2,i} + 1^T G_{2,i} + L_{2,i} = d_{2,i} \\
&V_{3,i} = V_2 + a_{2,i} + M(Q_{2,i} + S_{2,i}), \quad V_{3,i} \leq \bar{v}, \\
&Q_{2,i} \leq \bar{q}, \\
&Q_{2,i} + S_{2,i} \geq q, \\
&G_{2,i} \leq \bar{g}
\end{align*}\]

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

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

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

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

\(Q_{3,i,j} + S_{3,i,j} \geq q, \quad \forall j\)

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

\(L_{2,i}, L_{3,i,j}, G_{2,i}, G_{3,i,j}, Y_{i,j} \geq 0, \quad \forall j\)

\(Q_{2,i}, S_{2,i}, Q_{3,i,j}, S_{3,i,j}, V_{3,i} \geq 0, \quad \forall j\)
Since the second-stage problems (3.16) are linear, they satisfy standard constraint qualification conditions and one can replace the set of constraints \((G_{2,i}, G_{3,i,j}, L_{2,i}, L_{3,i,j}) \in X_i^*(V_2)\) 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**

<table>
<thead>
<tr>
<th>(\lambda_{i,j})</th>
<th>Vector of Lagrange multipliers for constraints (Y_{i,j} \geq c_i^T L_{3,i,j} + c_{3,i,j} G_{3,i,j} - \sum_{j' = 1}^{N_2} \pi_{i,j'} (c_i^T L_{3,i,j'} + c_{3,i,j'} G_{3,i,j'})) below, for (i = 1, \ldots, N_1) and (j = 1, \ldots, N_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_{2,i})</td>
<td>Vector of Lagrange multipliers for constraints (\rho^T Q_{2,i} + 1^T G_{2,i} + L_{2,i} = d_{2,i}), for (i = 1, \ldots, N_1)</td>
</tr>
<tr>
<td>(\beta_{2,i})</td>
<td>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)</td>
</tr>
<tr>
<td>(\sigma_{2,i})</td>
<td>Vector of Lagrange multipliers for constraint (V_{3,i} \leq \tilde{v}) below, for (i = 1, \ldots, N_1)</td>
</tr>
<tr>
<td>(\theta_{2,i})</td>
<td>Vector of Lagrange multipliers for constraint (Q_{2,i} \leq \tilde{q}) below, for (i = 1, \ldots, N_1)</td>
</tr>
<tr>
<td>(\delta_{2,i})</td>
<td>Vector of Lagrange multipliers for constraint (Q_{2,i} + S_{2,i} \geq q) below, for (i = 1, \ldots, N_1)</td>
</tr>
<tr>
<td>(\mu_{2,i})</td>
<td>Vector of Lagrange multipliers for constraint (G_{2,i} \leq \bar{g}) below, for (i = 1, \ldots, N_1)</td>
</tr>
<tr>
<td>(\alpha_{3,i,j})</td>
<td>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)</td>
</tr>
<tr>
<td>(\beta_{3,i,j})</td>
<td>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)</td>
</tr>
<tr>
<td>(\sigma_{3,i,j})</td>
<td>Vector of Lagrange multipliers for constraint (V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \leq \tilde{v}) below, for (i = 1, \ldots, N_1) and (j = 1, \ldots, N_2)</td>
</tr>
</tbody>
</table>
θ_{3,i,j} \quad \text{Vector of Lagrange multipliers for constraint } Q_{3,i,j} \leq \bar{q} \text{ below, for } i = 1, \ldots, N_1 \text{ and } j = 1, \ldots, N_2

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

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

The second-stage Lagrangian function \( L_i(G_{2,i},G_{3,i,j},L_{2,i},L_{3,i,j},Q_{2,i},Q_{3,i,j},S_{3,i,j},\) \( \forall i) \) is:

\[
L_i = c_i^T L_{2,i} + c_i^T G_{2,i} + \sum_{j=1}^{N_2} \pi_{i,j} (c_i^T L_{3,i,j} + c_i^T G_{3,i,j}) + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}
- \sum_{j=1}^{N_2} \chi_{i,j}^T (Y_{i,j} - c_i^T L_{3,i,j} - c_i^T G_{3,i,j} + \sum_{j=1}^{N_2} \pi_{i,j'} (c_i^T L_{3,i,j'} + c_i^T G_{3,i,j'}))
- \alpha_{2,i}^T (\rho Q_{2,i} + 1 T G_{2,i} + L_{2,i} - d_{2,i}) - \beta_{2,i}^T (V_2 + a_{2,i} + M(Q_{2,i} + S_{2,i}) - V_3)
- \sigma_{2,i}^T (\bar{v} - V_3) - \theta_{2,i}^T (\bar{q} - Q_{3,i} + S_{3,i} - g) - \mu_{2,i}^T (\bar{g} - G_{2,i})
- \sum_{j=1}^{N_2} \alpha_{3,i,j}^T (\rho Q_{3,i,j} + G_{3,i,j} + L_{3,i,j} - d_{3,i,j})
- \sum_{j=1}^{N_2} \beta_{3,i,j}^T (V_3_s + a_{3,i,j} + M(Q_{3,i} + S_{3,i,j}))
- \sum_{j=1}^{N_2} \sigma_{3,i,j}^T (\bar{v} - V_3_s - a_{3,i,j} - M(Q_{3,i,j} + S_{3,i,j})) - \sum_{j=1}^{N_2} \theta_{3,i,j}^T (\bar{q} - Q_{3,i,j})
- \sum_{j=1}^{N_2} \mu_{3,i,j}^T (\bar{g} - G_{3,i,j})
\]

CTC formulation of hydropower energy planning problem with MSD

\[
\begin{align*}
\min & \quad c_i^T L_1 + c_i^T G_1 + \sum_{i=1}^{N_1} \pi_i (c_i^T L_{2,i} + c_i^T G_{2,i}) \\
& + \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} (c_i^T L_{3,i,j} + c_i^T G_{3,i,j}) + \gamma_1 \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j} \\
\text{s.t.} & \quad X_{i,j} \geq c_i^T L_{2,i} + c_i^T G_{2,i} + c_i^T G_{3,i,j} - \sum_{i'=1}^{N_1} \pi_{i'}^T \left(c_{i'}^T L_{2,i'} + c_{i'}^T G_{2,i'} + \sum_{j'=1}^{N_2} \pi_{i',j'} (c_{i'}^T L_{3,i',j'} + c_{i'}^T G_{3,i',j'}) \right), \quad \forall i, \forall j
\end{align*}
\]

\[
\begin{align*}
\rho^T Q_1 + 1 T G_1 + L_1 & = d_1, \\
V_2 & = v_1 + a_1 + M(Q_1 + S_1), \\
V_2 & \leq \bar{v}, \\
Q_1 & \leq \bar{q},
\end{align*}
\]
\[ Q_1 + S_1 \geq q, \]
\[ G_1 \leq \bar{g}, \]
\[ Y_{i,j} \geq c_i^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} (c_{i,j'}^T L_{3,i,j'} + c_{3,i,j'}^T G_{3,i,j'}), \quad \forall i, \forall j \]
\[ \rho^T Q_{2,i} + 1^T G_{2,i} + L_{2,i} = d_{2,i}, \quad \forall i \]
\[ V_{3,i} = V_2 + a_{2,i} + M(Q_{2,i} + S_{2,i}), \quad \forall i \]
\[ V_{3,i} \leq \bar{v}, \quad \forall i \]
\[ Q_{2,i} \leq \bar{q}, \quad \forall i \]
\[ Q_{2,i} + S_{2,i} \geq q, \quad \forall i \]
\[ G_{2,i} \leq \bar{g}, \quad \forall i \]
\[ \rho^T Q_{3,i,j} + G_{3,i,j} + L_{3,i,j} = d_{3,i,j}, \quad \forall i, \forall j \]
\[ V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \geq 0, \quad \forall i, \forall j \]
\[ V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \leq \bar{v}, \quad \forall i, j \]
\[ Q_{3,i,j} \leq \bar{q}, \quad \forall i, \forall j \]
\[ Q_{3,i,j} + S_{3,i,j} \geq q, \quad \forall i, \forall j \]
\[ G_{3,i,j} \leq \bar{g}, \quad \forall i, \forall j \]
\[ c_{2,i} - 1\alpha_{2,i} + \mu_{2,i} \geq 0, \quad \forall i \]
\[ \pi_{i,j} c_{3,i,j} + c_{3,i,j} \lambda_{i,j} - \pi_{i,j} c_3 \lambda_{i,j} \left( \sum_{j'=1}^{N_2} \lambda_{i,j'} \right) - \alpha_{3,i,j} + \mu_{3,i,j} \geq 0, \quad \forall i, \forall j \]
\[ c_i - \alpha_{2,i} \geq 0, \quad \forall i \]
\[ \pi_{i,j} c_i + c_i \lambda_{i,j} - \pi_{i,j} c_i \left( \sum_{j'=1}^{N_2} \lambda_{i,j'} \right) - \alpha_{3,i,j} \geq 0, \quad \forall i, j \]
\[ - \rho \alpha_{2,i} - M^T \beta_{2,i} + \theta_{2,i} - \delta_{2,i} \geq 0, \quad \forall i \]
\[ - M^T \beta_{2,i} - \delta_{2,i} \geq 0, \quad \forall i \]
\[ - \rho \alpha_{3,i,j} - M^T \beta_{3,i,j} + M^T \sigma_{3,i,j} + \theta_{3,i,j} - \delta_{3,i,j} \geq 0, \quad \forall i, \forall j \]
\[ - M^T \beta_{3,i,j} + M^T \sigma_{3,i,j} - \delta_{3,i,j} \geq 0, \quad \forall i, \forall j \]
\[ \beta_{2,i} + \sigma_{2,i} - \sum_{j=1}^{N_2} \beta_{3,i,j} + \sum_{j=1}^{N_2} \sigma_{3,i,j} \geq 0, \quad \forall i, \forall j \]
\begin{align*}
\gamma_2 \pi_{i,j} - \lambda & \geq 0, \quad \forall i, \forall j \\
\lambda_{i,j}^T (Y_{i,j} - c_{i,j}^T L_{3,i,j} - c_{3,i,j}^T G_{3,i,j} + \sum_{j'=1}^{N_2} \pi_{i,j'} (c_{i,j}^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, \quad \forall i, \forall j \\
\theta_{2,i}^T (\bar{q} - Q_{2,i}) & = 0, \quad \forall i \\
\delta_{2,i}^T (Q_{2,i} + S_{2,i} - q) & = 0, \quad \forall i \\
\mu_{2,i}^T (\bar{g} - G_{2,i}) & = 0, \quad \forall i \\
\beta_{3,i,j}^T (V_{3,i} + a_{3,i,j} + M (Q_{3,i,j} + S_{3,i,j})) & = 0, \quad \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, \quad \forall i, \forall j \\
\theta_{3,i,j}^T (\bar{q} - Q_{3,i,j}) & = 0, \quad \forall i, \forall j \\
\delta_{3,i,j}^T (Q_{3,i,j} + S_{3,i,j} - q) & = 0, \quad \forall i, \forall j \\
\mu_{3,i,j}^T (\bar{g} - G_{3,i,j}) & = 0, \quad \forall i \\
G_{2,i}^T (c_{2,i} - 1 \alpha_{2,i} + \mu_{2,i}) & = 0, \quad \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, \quad \forall i, \forall j \\
L_{2,i}^T (c_l - \alpha_{2,i}) & = 0, \quad \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, \quad \forall i, \forall j \\
Q_{2,i}^T (-\rho \alpha_{2,i} - M^T \beta_{2,i} + \theta_{2,i} - \delta_{2,i}) & = 0, \quad \forall i \\
S_{2,i}^T (-M^T \beta_{2,i} - \delta_{2,i}) & = 0, \quad \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, \quad \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, \quad \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, \quad \forall i, \forall j \\
Y_{i,j}^T (\gamma_2 \pi_{i,j} - \lambda_{i,j}) & = 0, \quad \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 \quad \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, \quad \forall i, \forall j
\end{align*}
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

\[
\begin{align*}
\min & \quad c_i^T L_1 + c_i^T G_1 + \sum_{i=1}^{N_1} \pi_i (W_i + \gamma_1 \chi_i) \\
\text{s.t.} & \quad \chi_i \geq W_i - \sum_{i'=1}^{N_1} \pi_{i'} W_{i'}, \quad \forall i \\
& \quad W_i = c_i^T L_{2,i} + c_{2,i}^T G_{2,i} + \sum_{j=1}^{N_2} \pi_{i,j} (c_i^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j}) + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, \quad \forall i \\
& \quad Y_{i,j} \geq c_i^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'} (c_i^T L_{3,i,j'} + c_{3,i,j'}^T G_{3,i,j'}), \quad \forall i, \forall j \\
& \quad \rho^T Q_1 + 1^T G_1 + L_1 = d_1, \\
& \quad V_2 = v_1 + a_1 + M(Q_1 + S_1), \\
& \quad V_2 \leq \bar{v}, \\
& \quad Q_1 \leq \bar{q}, \\
& \quad Q_1 + S_1 \geq q, \\
& \quad G_1 \leq \bar{g}, \\
& \quad \rho^T Q_{2,i} + 1^T G_{2,i} + L_{2,i} = d_{2,i}, \quad \forall i \\
& \quad V_{3,i} = V_2 + a_{2,i} + M(Q_{2,i} + S_{2,i}), \quad \forall i \\
& \quad V_{3,i} \leq \bar{v}, \quad \forall i \\
& \quad Q_{2,i} \leq \bar{q}, \quad \forall i \\
& \quad Q_{2,i} + S_{2,i} \geq q, \quad \forall i \\
& \quad G_{2,i} \leq \bar{g}, \quad \forall i
\end{align*}
\]
\[ \begin{align*}
\rho^T Q_{3,i,j} + G_{3,i,j} + L_{3,i,j} &= d_{3,i,j} \\
V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) &\geq 0, \\
V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) &\leq \bar{v}, \\
Q_{3,i,j} &\leq \bar{q}, \\
G_{3,i,j} &\leq \bar{g}, \\
L_{1,G_1,\chi_i,Q_1,V_2,S_1} &\geq 0, \\
L_{2,i,L_{3,i,j},G_{2,i},G_{3,i,j},Y_{i,j},Q_{2,i},S_{2,i},Q_{3,i,j},S_{3,i,j},V_{3,i}} &\geq 0,
\end{align*} \]

\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\(\alpha_{\beta}\). 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

| \(\alpha_l\) | Risk aversion parameter for AVaR\(\alpha_l\) for stage \(l = 1, 2\) |
| \(\beta_l\) | Coefficient of AVaR\(\alpha_l\) 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\(\alpha\) value for CTC model |
| \(h_{i}\) | AVaR\(\alpha\) value, in second-stage scenario \(i = 1, \ldots, N_1\) |
| \(\omega_{i}\) | Second stage cost plus BRM\(\alpha_{\beta_{\beta}}\) value for OTC model, in second-stage scenario \(i = 1, \ldots, N_1\) |
| \(k\) | First-stage AVaR\(\alpha_{\alpha_1}\) value of \(\omega_{i}\) 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 \(j = 1, \ldots, N_2\)

\( \tau \) Helper variable in AVaR\(^\alpha\) definition in stage 1 for OTC model

\( \chi_i \) Positive part of difference between \(\omega_i\) and \(\tau\), for \(i = 1, \ldots, N_1\)

Table 3.28: Lagrange multiplier variables for second-stage problems

\( \zeta_i \) 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\(^\beta\)

\[
\begin{align*}
\min & \quad c_1^T L_1 + c_1^T G_1 + (1 - \beta_1) \sum_{i=1}^{N_1} \left[ \pi_i (c_2^T L_{2,i} + c_2^T G_{2,i}) \right] \\
& \quad + \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} (c_3^T G_{3,i,j} + c_4^T L_{3,i,j}) \right] + \beta_1 g \\
\text{s.t.} & \quad g \geq t + \frac{1}{\alpha_1} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \pi_i \pi_{i,j} X_{i,j} \\
& \quad X_{i,j} \geq c_1^T (L_{2,i} + L_{3,i,j}) + c_2^T G_{2,i} + c_3^T G_{3,i,j} - t, \quad \forall i, \forall j \\
& \quad \rho^T Q_1 + \mathbf{1}^T G_1 + L_1 = d_1, \\
& \quad V_2 = v_1 + a_1 + M(Q_1 + S_1), \\
& \quad V_2 \leq \bar{v}, \\
& \quad Q_1 \leq \bar{q}, \\
& \quad Q_1 + S_1 \geq q, \\
& \quad G_1 \leq \bar{g}, \\
& \quad h_i \geq \bar{t}_i + \frac{1}{\alpha_2} \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, \quad \forall i 
\end{align*}
\] (3.19)
\[ \begin{align*}
Y_{i,j} & \geq c_i^T L_{3,i,j} + c_{3,i,j}^T G_{3,i,j} - \tilde{t}_i, & \forall i, \forall j \\
\rho^T Q_{2,i} + 1^T G_{2,i} + L_{2,i} & = d_{2,i}, & \forall i \\
V_{3,i} & = V_2 + a_{2,i} + M(Q_{2,i} + S_{2,i}), & \forall i \\
V_{3,i} & \leq \bar{v}, & \forall i \\
Q_{2,i} & \leq \bar{q}, & \forall i \\
Q_{2,i} + S_{2,i} & \geq \underline{q}, & \forall i \\
G_{2,i} & \leq \bar{g}, & \forall i \\
\rho^T Q_{3,i,j} + G_{3,i,j} + L_{3,i,j} & = d_{3,i,j}, & \forall i, \forall j \\
V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) & \geq 0, & \forall i, j \\
V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) & \leq \bar{v}, & \forall i, j \\
Q_{3,i,j} & \leq \bar{q}, & \forall i, \forall j \\
Q_{3,i,j} + S_{3,i,j} & \geq \underline{q}, & \forall i, \forall j \\
G_{3,i,j} & \leq \bar{g}, & \forall i, \forall j \\
c_{2,i} - \mathbf{1} \alpha_{2,i} + \mu_{2,i} & \geq 0, & \forall i \\
(1 - \beta_2) \pi_{i,j} c_{3,i,j} + c_{3,i,j} \lambda_{i,j} - \alpha_{3,i,j} + \mu_{3,i,j} & \geq 0, & \forall i, \forall j \\
c_{1} \alpha_{2,i} & \geq 0, & \forall i \\
(1 - \beta_2) \pi_{i,j} c_{1} + c_{1} \lambda_{i,j} - \alpha_{3,i,j} & \geq 0, & \forall i, j \\
- \rho \alpha_{2,i} - M^T \beta_{2,i} + \theta_{2,i} - \delta_{2,i} & \geq 0, & \forall i \\
- M^T \beta_{2,i} - \delta_{2,i} & \geq 0, & \forall i \\
- \rho \alpha_{3,i,j} - M^T \beta_{3,i,j} + M^T \sigma_{3,i,j} + \theta_{3,i,j} - \delta_{3,i,j} & \geq 0, & \forall i, \forall j \\
- M^T \beta_{3,i,j} + M^T \sigma_{3,i,j} - \delta_{3,i,j} & \geq 0, & \forall i, \forall j \\
\beta_{2,i} + \sigma_{2,i} - \sum_{j=1}^{N_2} \beta_{3,i,j} + \sum_{j=1}^{N_2} \sigma_{3,i,j} & \geq 0, & \forall i, \forall j \\
\frac{\zeta_i}{\alpha_2} \pi_{i,j} - \lambda & \geq 0, & \forall i, \forall j \\
\beta_2 - \zeta_i & = 0, & \forall i \\
\zeta_i - \sum_{j=1}^{N_2} \lambda_{i,j} & = 0, & \forall i
\end{align*} \]
\[
\begin{align*}
\zeta_i^T (h_i - \bar{t}_i - \frac{1}{\alpha} \sum_{j=1}^{N_z} \pi_{i,j} Y_{i,j}) &= 0, \\
\chi_{i,j}^T (Y_{i,j} - c_i^T L_{3,i,j} - c_{3,i,j}^T G_{3,i,j} + \bar{t}_i) &= 0, \\
\sigma_{2,i}^T (\bar{v} - V_{3,i}) &= 0, \\
\theta_{2,i}^T (\bar{q} - Q_{2,i}) &= 0, \\
\delta_{2,i}^T (Q_{2,i} + S_{2,i} - q) &= 0, \\
\mu_{2,i}^T (\bar{g} - G_{2,i}) &= 0, \\
\beta_{3,i,j}^T (V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j})) &= 0, \\
\sigma_{3,i,j}^T (\bar{v} - a_{3,i,j} - M(Q_{3,i,j} - S_{3,i,j})) &= 0, \\
\theta_{3,i,j}^T (\bar{q} - Q_{3,i,j}) &= 0, \\
\delta_{3,i,j}^T (Q_{3,i,j} + S_{3,i,j} - q) &= 0, \\
\mu_{3,i,j}^T (\bar{g} - G_{3,i,j}) &= 0, \\
G_{2,i}^T (c_{2,i} - 1\alpha_{2,i} + \mu_{2,i}) &= 0, \\
G_{3,i,j}^T ((1 - \beta_2)\pi_{i,j} c_{3,i,j} + c_{3,i,j} \lambda_{i,j} - \alpha_{3,i,j} + \mu_{3,i,j}) &= 0, \\
L_{2,i}^T (\alpha_1 - \alpha_{2,i}) &= 0, \\
L_{3,i,j}^T ((1 - \beta_2)\pi_{i,j} c_i + c_i \lambda_{i,j} - \alpha_{3,i,j}) &= 0, \\
Q_{2,i}^T (-\rho\alpha_{2,i} - M^T \beta_{2,i} + \theta_{2,i} - \delta_{2,i}) &= 0, \\
S_{2,i}^T (-M^T \beta_{2,i} - \delta_{2,i}) &= 0, \\
Q_{3,i,j}^T (-\rho\alpha_{3,i,j} - M^T \beta_{3,i,j} + M^T \sigma_{3,i,j} + \theta_{3,i,j} - \delta_{3,i,j}) &= 0, \\
S_{3,i,j}^T (-M^T \beta_{3,i,j} + M^T \sigma_{3,i,j} - \delta_{3,i,j}) &= 0, \\
V_{3,i}^T (\beta_{2,i} + \sigma_{2,i} - \sum_{j=1}^{N_z} \beta_{3,i,j} + \sum_{j=1}^{N_z} \sigma_{3,i,j}) &= 0, \\
Y_{i,j}^T \left( \frac{\zeta_i}{\alpha_2} \pi_{i,j} - \lambda \right) &= 0, \\
\lambda_{i,j}, \alpha_{2,i}, \beta_{2,i}, \sigma_{2,i}, \theta_{2,i}, \delta_{2,i}, \mu_{2,i} &\geq 0, \\
\alpha_{3,i,j}, \beta_{3,i,j}, \sigma_{3,i,j}, \theta_{3,i,j}, \delta_{3,i,j}, \mu_{3,i,j} &\geq 0, \\
L_1, G_1, X_{i,j}, Q_1, V_2, S_1 &\geq 0, \\
L_{2,i}, L_{3,i,j}, G_{2,i}, G_{3,i,j}, Y_{i,j}, Q_{2,i}, S_{2,i}, Q_{3,i,j}, S_{3,i,j}, V_3 &\geq 0.
\end{align*}
\]
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 BRM$^\alpha$**

$$\min \ c^T_l L_1 + c^T_1 G_1 + (1 - \beta_1) \sum_{i=1}^{N_1} \pi_i \omega_i + \beta_1 k$$

s.t.  
$$k \geq \tau + \frac{1}{\alpha} \sum_{i=1}^{N_1} \pi_i \chi_i$$

$$\chi_i \geq \omega_i - \tau, \quad \forall i$$

$$\omega_i = c^T_l L_{2,i} + c^T_{2,i} G_{2,i} + (1 - \beta_2)(\sum_{j=1}^{N_2} \pi_{i,j}(c^T_l L_{3,i,j} + c^T_{3,i,j} G_{3,i,j})) + \beta_2 h_i, \quad \forall i$$

$$h_i \geq \bar{t}_i + \frac{1}{\alpha} \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, \quad \forall i$$

$$Y_{i,j} \geq c^T_l L_{3,i,j} + c^T_{3,i,j} G_{3,i,j} - \bar{t}_i, \quad \forall i, \forall j$$

$$\rho^T Q_1 + 1^T G_1 + L_1 = d_1,$$

$$V_2 = v_1 + a_1 + M(Q_1 + S_1),$$

$$V_2 \leq \bar{v},$$

$$Q_1 \leq \bar{q},$$

$$Q_1 + S_1 \geq q,$$

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

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

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

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

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

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

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

$$\rho^T Q_{3,i,j} + G_{3,i,j} + L_{3,i,j} = d_{3,i,j} \quad \forall i, \forall j$$
\[ V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \geq 0, \quad \forall i, \forall j \]
\[ V_{3,i} + a_{3,i,j} + M(Q_{3,i,j} + S_{3,i,j}) \leq \bar{v}, \quad \forall i, \forall j \]
\[ Q_{3,i,j} \leq \bar{q}, \quad \forall i, \forall j \]
\[ Q_{3,i,j} + S_{3,i,j} \geq q, \quad \forall i, \forall j \]
\[ G_{3,i,j} \leq \bar{g}, \quad \forall i, \forall j \]
\[ Y_{i,j} \geq c_T^T L_{3,i,j} + c_T^T G_{3,i,j} - \sum_{j'=1}^{N_2} \pi_{i,j'}(c_T^T L_{3,i,j'} + c_T^T G_{3,i,j'}), \quad \forall i, \forall j \]
\[ W_i = c_T^T L_{2,i} + c_T^T G_{2,i} + \sum_{j=1}^{N_2} \pi_{i,j}(c_T^T L_{3,i,j} + c_T^T G_{3,i,j}) + \gamma_2 \sum_{j=1}^{N_2} \pi_{i,j} Y_{i,j}, \quad \forall i \]
\[ \chi_i \geq W_i - \sum_{i=1}^{N_1} \pi_i W_i, \quad \forall i \]
\[ L_1, G_1, \chi_i, Q_1, V_2, S_1 \geq 0, \quad \forall i \]
\[ L_2, i, L_{3,i,j}, G_{2,i}, L_{3,i,j}, Y_{i,j}, Q_2, i, S_2, i, Q_3, i, j, S_3, i, j, V_3, i \geq 0, \quad \forall i, \forall j. \]

### 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\( _\gamma \) 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\( ^\alpha_\beta \), giving a total of 8 risk measures.

Tables 3.29 and 3.30 summarize the results for instances with risk measure MSD\( _\gamma \) and BRM\( ^\alpha_\beta \), respectively. With MSD risk measure, the average solution time were 0.48s and 0.02s for CTC and OTC models, respectively. For 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 BRM\( ^{0.25}_{0.75} \). 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.

<table>
<thead>
<tr>
<th></th>
<th>$\gamma = 0.3$</th>
<th>$\gamma = 0.7$</th>
<th>$\gamma = 0.9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diff $&lt; 0.01%$</td>
<td>93.90%</td>
<td>90.70%</td>
<td>90.40%</td>
</tr>
<tr>
<td>$0.01% \leq \text{Diff} &lt; 0.1%$</td>
<td>1.10%</td>
<td>1.10%</td>
<td>1.00%</td>
</tr>
<tr>
<td>$0.1% \leq \text{Diff} &lt; 1%$</td>
<td>3.50%</td>
<td>3.20%</td>
<td>3.60%</td>
</tr>
<tr>
<td>$1% \leq \text{Diff} &lt; 10%$</td>
<td>1.50%</td>
<td>4.80%</td>
<td>4.40%</td>
</tr>
<tr>
<td>$10% \leq \text{Diff}$</td>
<td>0%</td>
<td>0.20%</td>
<td>0.60%</td>
</tr>
<tr>
<td>Max Diff</td>
<td>9.06%</td>
<td>14.19%</td>
<td>25.88%</td>
</tr>
</tbody>
</table>

Table 3.29: Computational results of the hydropower planning problem with MSD$_\gamma$.

<table>
<thead>
<tr>
<th></th>
<th>$\beta = 0.5$</th>
<th>$\beta = 0.5$</th>
<th>$\beta = 0.75$</th>
<th>$\beta = 0.5$</th>
<th>$\beta = 0.75$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha = 0.01$</td>
<td>$\alpha = 0.05$</td>
<td>$\alpha = 0.05$</td>
<td>$\alpha = 0.25$</td>
<td>$\alpha = 0.25$</td>
</tr>
<tr>
<td>Diff $&lt; 0.01%$</td>
<td>94.90%</td>
<td>94.90%</td>
<td>95.60%</td>
<td>91.40%</td>
<td>90.40%</td>
</tr>
<tr>
<td>$0.01% \leq \text{Diff} &lt; 0.1%$</td>
<td>0.10%</td>
<td>0.10%</td>
<td>0.20%</td>
<td>0.10%</td>
<td>0%</td>
</tr>
<tr>
<td>$0.1% \leq \text{Diff} &lt; 1%$</td>
<td>0.70%</td>
<td>1.00%</td>
<td>1.30%</td>
<td>1.40%</td>
<td>1.2%</td>
</tr>
<tr>
<td>$1% \leq \text{Diff} &lt; 10%$</td>
<td>3.00%</td>
<td>3.20%</td>
<td>1.40%</td>
<td>3.60%</td>
<td>4.00%</td>
</tr>
<tr>
<td>$10% \leq \text{Diff}$</td>
<td>1.30%</td>
<td>0.80%</td>
<td>1.5%</td>
<td>3.50%</td>
<td>4.4%</td>
</tr>
<tr>
<td>Max Diff</td>
<td>34.09%</td>
<td>26.34%</td>
<td>57.47%</td>
<td>46.89%</td>
<td>68.22%</td>
</tr>
</tbody>
</table>

Table 3.30: Computational results of the hydropower planning problem with BRM$_\alpha$.

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.

Table 3.31: The Diff value for 5 hydropower planning instances with MSD$\gamma$ and BRM$\beta$.

<table>
<thead>
<tr>
<th></th>
<th>$\gamma = 0.3$</th>
<th>$\gamma = 0.7$</th>
<th>$\gamma = 0.9$</th>
<th>$\alpha = 0.01$</th>
<th>$\alpha = 0.05$</th>
<th>$\alpha = 0.25$</th>
<th>$\alpha = 0.25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>9.06%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>B</td>
<td>0%</td>
<td>14.19%</td>
<td>8.43%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>C</td>
<td>0%</td>
<td>0%</td>
<td>25.88%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>4.45%</td>
</tr>
<tr>
<td>D</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>34.09%</td>
<td>0%</td>
<td>47.52%</td>
<td>0%</td>
</tr>
<tr>
<td>E</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>26.34%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>F</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>1.38%</td>
<td>0%</td>
<td>57.47%</td>
<td>0%</td>
</tr>
<tr>
<td>G</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>12.77%</td>
<td>0.41%</td>
<td>0%</td>
<td>46.89%</td>
</tr>
<tr>
<td>H</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>22.58%</td>
<td>0%</td>
<td>0%</td>
<td>68.22%</td>
</tr>
</tbody>
</table>

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:

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

(4.1)
in which $\mathcal{E}_2$ is the set of second-stage scenarios, $c_1 \in \mathbb{R}^{n_1}$, $c_2 = (c_2^{S_1}, \ldots, c_2^{S_{|\mathcal{E}_2|}}) \in \mathbb{R}^{|\mathcal{E}_2|n_2}$, and $b_1 \in \mathbb{R}^{m_1}$ are given, $X_1 \in \mathbb{R}^{n_1}$, and $X_2 = (X_2^{S_1}, \ldots, X_2^{S_{|\mathcal{E}_2|}}) \in \mathbb{R}^{|\mathcal{E}_2|n_2}$ are the decision variables, and given $X_1$ and $S \in \mathcal{E}_2$, $X_{2,S}^*(X_1)$ is the optimal solution set of the second-stage problem:

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

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 $X_{2,new}^S = \begin{bmatrix} X_2^S \\ X_3^S \end{bmatrix}$ and $A_{22,new}^S = \begin{bmatrix} A_{22}^S & 0 \\ A_{32}^S & A_{33}^S \end{bmatrix}$, $\forall S \in \mathcal{E}_2$.

### 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 $X_2^S \in X_{2,S}^*(X_1)$ 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:

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

So the KKT optimality conditions are:

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

Now, we obtain the following equivalent LPCC formulation:

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

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

**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 \(X^S\). For any \(Y_1, Y_2 \in \mathbb{R}^{n_1 \times |\mathcal{E}_2|}\), and \(\alpha \in [0, 1]\):

\[
\begin{align*}
    f^S(\alpha Y_1 + (1 - \alpha) Y_2) &= \min_{(X_1^S, X_2^S, \lambda^S) \in X^S} \frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, \alpha Y_1^S + (1 - \alpha) Y_2^S \rangle \\
    &= \min_{(X_1^S, X_2^S, \lambda^S) \in X^S} \left\{ \alpha \left( \frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y_1^S \rangle \right) \\
    &\quad + (1 - \alpha) \left( \frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y_2^S \rangle \right) \right\} \\
    \geq \alpha \min_{(X_1^S, X_2^S, \lambda^S) \in X^S} \left( \frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y_1^S \rangle \right) \\
    &\quad + (1 - \alpha) \left( \frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y_2^S \rangle \right) \\
    &= \alpha \min_{(X_1^S, X_2^S, \lambda^S) \in X^S} \left( \frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y_1^S \rangle \right) \\
    &\quad + (1 - \alpha) \left( \frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y_2^S \rangle \right) \\
    &= \alpha \min_{(X_1^S, X_2^S, \lambda^S) \in X^S} \left( \frac{1}{|\mathcal{E}_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y_1^S \rangle \right)
\end{align*}
\]
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):

$$\min_{X_1, X_2, \lambda} \sum_{S \in \mathcal{E}_2} \frac{c^T_S}{|\mathcal{E}_2|} X_1^S + \sum_{S \in \mathcal{E}_2} (c^S_2)^T X_2^S$$

s.t. \begin{align*}
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, & \forall S \in \mathcal{E}_2 \\
X_2^S &\geq 0, & \forall S \in \mathcal{E}_2.
\end{align*}

(4.5)

This problem is clearly equivalent to (4.3). Suppose $(X_1^*, X_2^*)$ 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

$$F = c^T_1 X_1^* + \sum_{S \in \mathcal{E}_2} (c^S_2)^T X_2^{S*}$$

$$= \sum_{S \in \mathcal{E}_2} \left( \frac{1}{|\mathcal{E}_2|} c^T_1 X_1^{S*} + (c^S_2)^T X_2^{S*} \right)$$

$$= \sum_{S \in \mathcal{E}_2} \left( \frac{1}{|\mathcal{E}_2|} c^T_1 X_1^{S*} + (c^S_2)^T X_2^{S*} \right) + \langle \mathcal{X}_1^*, Y \rangle$$

$$= \sum_{S \in \mathcal{E}_2} \left( \frac{1}{|\mathcal{E}_2|} c^T_1 X_1^{S*} + (c^S_2)^T X_2^{S*} + \langle X_1^{S*}, Y \rangle \right)$$

$$\geq \sum_{S \in \mathcal{E}_2} f^S(Y) = f(Y),$$

so $f^S$ is concave.
where the third equality follows since $K$ and $K^*$ are orthogonal complements, with $X_1^* = (X_1^{S_1^*}, \ldots, X_1^{S_{|E_2|}}) \in K$ 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., $Y = 0$). Then, solve subproblems (4.4) for all $S \in E_2$. 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 $X_1 \in K$. To see this, let $X^S$ be the feasible set of problem (4.4) and $X = \{(X_1, X_2, \lambda) | (X_1^S, X_2^S, \lambda^S) \in X^S, \forall S \in E_2\}$. Note that $X$ is the feasible set of all (4.5) constraints except for the nonanticipativity constraint. Then, Lagrangian relaxation of problem (4.5) obtained by dualizing $X_1 \in K$ is:

\[
\min_{(X_1, X_2, \lambda) \in X} \sum_{S \in E_2} \left( \frac{1}{|E_2|} c_1^T X_1^S + (c_2^S)^T X_2^S \right) + \langle X_1, Y \rangle \\
= \min_{(X_1, X_2, \lambda) \in X} \sum_{S \in E_2} \left( \frac{1}{|E_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y^S \rangle \right) \\
= \sum_{S \in E_2} \min_{(X_1^S, X_2^S, \lambda^S) \in X^S} \frac{1}{|E_2|} c_1^T X_1^S + (c_2^S)^T X_2^S + \langle X_1^S, Y^S \rangle \\
= \sum_{S \in E_2} f^S(Y) = f(Y).
\]

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 \( f^S(Y) \).

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

\[
\begin{align*}
  f^S(Y) & \leq \frac{c^T}{|E_2|} X_1^{S*} + (c_2^S)^T X_2^{S*} + \langle X_1^{S*}, Y^S \rangle \\
  &= \frac{c^T}{|E_2|} X_1^{S*} + (c_2^S)^T X_2^{S*} + \langle X_1^{S*}, Z^S \rangle + \langle X_1^{S*}, Y^S - Z^S \rangle \\
  &= f^S(Z) + \langle X_1^{S*}, Y^S - Z^S \rangle.
\end{align*}
\]

where the first inequality follows from the definition of \( f^S(Y) \) and the last equality follows from the lemma’s assumption. \( \blacksquare \)

For each \( S \in E_2 \) and iteration \( k \geq 1 \), define a set of indices \( B_k^S \subseteq \{1, \ldots, k\} \) and call it the *bundle*. From Lemma 7,

\[
\hat{f}_k^S(Y) \equiv \min_{j \in B_k^S} \{ f^S(Y_j) + \langle X_1^{S*}, Y^S - Y_j^S \rangle \}
\]

is a piecewise linear approximation of \( f^S(Y) \) with \((X_1^{S*}, X_2^{S*})\) 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.

\[
\begin{align*}
  \max_{r,Y} \quad & \sum_{S \in E_2} r^S - \frac{1}{2t_k} \| Y - \hat{Y}_k \|^2 \\
  \text{s.t.} \quad & r^S \leq f^S(Y_j) + \langle X_1^{S*}, Y^S - Y_j^S \rangle, \forall S \in E_2, \quad j \in B_k^S \\
  & Y = (Y^{S_1}, \ldots, Y^{S_{|E_2|}}) \in K^*.
\end{align*}
\]

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 \( S \in E_2 \), define:
\[ I^S_k = \{ j \mid r^{S^*} = f^S(Y_j) + \langle X^S_{ij}, Y^S_{k+1} - Y^S_j \rangle, \ j \in B^S_k \}, \]
\[ f^m_k(Y_{k+1}) = \sum_{S \in \mathcal{E}_2} r^{S^*}, \quad (4.8) \]
\[ \delta_k = f^m_k(Y_{k+1}) - f(\hat{Y}_k), \]
\[ \gamma_k = f(\bar{Y}_k) + \kappa \delta_k, \ \kappa \in (0, 1). \]

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

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 \( \psi^S_{k+1} \) equal to 1 when \( f^S(Y_{k+1}) \) is unbounded and equal to 0 otherwise. Consequently, \( \max_{S \in \mathcal{E}_2} \{ \psi^S_{k+1} \} = 1 \) 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 \( \max_{S \in \mathcal{E}_2} \{ \psi^S_{k+1} \} = 0 \), we have the value of \( f(Y_{k+1}) \) and we are able to compare \( f(Y_{k+1}) \) with \( \gamma_k \). 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:

\[ B^S_{k+1} = \begin{cases} 
I^S_k \cup \{ k + 1 \}, & \text{if } \psi^S_{k+1} = 0 \\
I^S_k, & \text{otherwise.} 
\end{cases} \quad (4.9) \]

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:

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

Next, we have:

\[
f^S(Y_j) + (X_{1j}^S, Y^S - Y_j^S) = f^S(Y_j) + (X_{1j}^S, Y^S - Y_j^S) + e_j^S - f^S(Y_j)
- (X_{1j}^S, \hat{Y}_k^S - Y_j^S) + f^S(\hat{Y}_k)
= f^S(\hat{Y}_k) + e_j^S + (X_{1j}^S, Y^S - Y_j^S - \hat{Y}_k^S + Y_j^S)
= f^S(\hat{Y}_k) + e_j^S + (X_{1j}^S, Y^S - \hat{Y}_k^S)
\]

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

\[
\max_{r,Y} \sum_{S \in \mathcal{E}_2} r^S - \frac{1}{2t_k} \|Y - \hat{Y}_k\|^2 \quad \text{s.t.} \quad r^S \leq f^S(\hat{Y}_k) + e_j^S + (X_{1j}^S, Y^S - \hat{Y}_k^S), \ \forall S \in \mathcal{E}_2, \ j \in \mathcal{B}_k^S \tag{4.12}
\]

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

\[
L(r, Y, \alpha, \beta) = \sum_{S \in \mathcal{E}_2} r^S - \frac{1}{2t_k} \|Y - \hat{Y}_k\|^2 - \sum_{S \in \mathcal{E}_2} \sum_{j \in \mathcal{B}_k^S} \alpha_j^S (r^S - f^S(\hat{Y}_k) - e_j^S - (X_{1j}^S, Y^S - \hat{Y}_k^S)) - \beta Y^S
\]

\[
= \sum_{S \in \mathcal{E}_2} r^S (1 - \sum_{j \in \mathcal{B}_k^S} \alpha_j^S) + \sum_{S \in \mathcal{E}_2} \left[ -\frac{1}{2t_k} \|Y^S - \hat{Y}_k^S\|^2 + \sum_{j \in \mathcal{B}_k^S} \alpha_j^S (X_{1j}^S, Y^S - \hat{Y}_k^S - \beta Y^S) \right] + \sum_{S \in \mathcal{E}_2} \sum_{j \in \mathcal{B}_k^S} \alpha_j^S (f^S(\hat{Y}_k) + e_j^S).
\]

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

\[
\nabla_{r^*} L = 1 - \sum_{j \in \mathcal{B}_k^S} \alpha_j^S = 0
\]

\[
\nabla_{Y^*} L = -\frac{1}{t_k} (Y^{S*} - \hat{Y}_k^{S*}) + \sum_{j \in \mathcal{B}_k^S} \alpha_j^S X_{1j}^{S*} - \beta = 0,
\]
for all $S \in \mathcal{E}_2$. From the last equation, we have:

$$Y^S = \hat{Y}^S_k + \sum_{j \in B_k} t_k \alpha_j^S X^*_j - t_k \beta.$$

Therefore,

$$\max_{r,Y} L(r,Y,\alpha,\beta) = \sum_{S \in \mathcal{E}_2} \left( \frac{1}{2t_k} \left\| \sum_{j \in B_k} t_k \alpha_j^S X^*_j - t_k \beta \right\|^2 \right)$$

$$+ \sum_{S \in \mathcal{E}_2} \sum_{j \in B_k} \alpha_j^S (\nu_{1j}^S + \sum_{S \in \mathcal{E}_2} \sum_{j \in B_k} t_k \alpha_j^S X^*_j - t_k \beta)$$

$$- \sum_{S \in \mathcal{E}_2} \beta (\hat{Y}^S_k + \sum_{j \in B_k} t_k \alpha_j^S X^*_j - t_k \beta) + \sum_{S \in \mathcal{E}_2} \sum_{j \in B_k} \sum_{j \in B_k} \alpha_j^S (f^S(\hat{Y}^S_k) + e_j^S)$$

$$= \sum_{S \in \mathcal{E}_2} \frac{t_k}{2} \left( \sum_{j \in B_k} \alpha_j^S \nu_{1j}^S + \beta, \sum_{j \in B_k} \alpha_j^S X^*_j - \beta \right)$$

$$- \sum_{S \in \mathcal{E}_2} \beta (\hat{Y}^S_k + \sum_{j \in B_k} t_k \alpha_j^S X^*_j - t_k \beta) + \sum_{S \in \mathcal{E}_2} \sum_{j \in B_k} \sum_{j \in B_k} \alpha_j^S (f^S(\hat{Y}^S_k) + e_j^S)$$

$$= \sum_{S \in \mathcal{E}_2} \frac{t_k}{2} \left( \sum_{j \in B_k} \alpha_j^S X^*_j - \beta \right)^2 - \sum_{S \in \mathcal{E}_2} \hat{Y}^S_k \beta$$

$$+ \sum_{S \in \mathcal{E}_2} \sum_{j \in B_k} \left( f^S(\hat{Y}^S_k) + e_j^S \right) \alpha_j^S.$$

The dual of master problem is:

$$\min_{\alpha,\beta} \sum_{S \in \mathcal{E}_2} \frac{t_k}{2} (\nu^S)^2 - \sum_{S \in \mathcal{E}_2} \hat{Y}^S_k \beta + \sum_{S \in \mathcal{E}_2} \sum_{j \in B_k} \sum_{j \in B_k} \alpha_j^S (f^S(\hat{Y}^S_k) + e_j^S) \alpha_j^S$$

s.t.

$$\sum_{j \in B_k} \alpha_j^S = 1, \forall S \in \mathcal{E}_2$$

$$\nu^S = \sum_{j \in B_k} X^*_j \alpha_j^S - \beta, \forall S \in \mathcal{E}_2$$

$$\alpha_j^S \geq 0, \forall S \in \mathcal{E}_2, j \in B_k.$$

In the master problem (4.12), constraint $Y \in K^*$ is the dual of constraint $X_1 \in K$ in problem (4.5). Also, note that in the master dual problem (4.13), $\beta$ is the Lagrange multiplier of constraint $Y \in K^*$. Accordingly, we can choose $\beta$ 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 $\beta$ satisfies these constraints. This procedure,
gives us the following optimization problem:

$$
\min_{\alpha, \beta, X_2} \sum_{s \in E_2} \frac{t_k}{2} (\nu^S)^2 - \sum_{s \in E_2} \hat{Y}_k^S \beta + \sum_{s \in E_2} \sum_{j \in \mathcal{B}_k^S} (f^S(\hat{Y}_k) + e_j^S) \alpha_j^S
$$

s.t.

$$\sum_{j \in \mathcal{B}_k^S} \alpha_j^S = 1, \forall s \in E_2$$

$$\nu^S = \sum_{j \in \mathcal{B}_k^S} X_{1j}^S \alpha_j^S - \beta, \forall s \in E_2$$

$$\alpha_j^S \geq 0, \forall s \in E_2, j \in \mathcal{B}_k^S$$

$$A_{11} \beta = b_1$$

$$A_{21}^S \beta + A_{22}^S X_2^S = b_2^S, \forall s \in E_2.$$  \hfill (4.14)

Let $\beta^*$ be the optimal value of $\beta$ in the above problem. Set $X_1 = \beta^*$ and solve all the second-stage problems (4.2) to find an optimal point $X_2^* \in \mathcal{X}_2^*(\beta^*)$. Note that $(\beta^*, X_2^*)$ 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) $(f^m_k(Y_{k+1}) - L)/L \leq \tau$

(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 \( B_0^S = \{0\}, \forall S \in E_2. \)

**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 \( f^m_k(Y_{k+1}), \delta_k, \) and \( \gamma_k \) based on (4.8).

**Step 4.** If STOP, Return \( L \) and \( U \).

**Step 5.** Calculate \( f(Y_{k+1}) \) and \( \psi^S_{k+1}, \forall S \in E_2. \)

**Step 6.** If \( \max_{S \in E_2} \psi^S_{k+1} = 0 \) 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 \( B_{k+1}^S, \forall S \in E_2 \) 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.
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 $\frac{L-L}{L}$. 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 \(|E_2|\), 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.

<table>
<thead>
<tr>
<th>Category</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n_1)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>(m_1)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>(</td>
<td>E_2</td>
<td>)</td>
<td>50</td>
<td>100</td>
<td>250</td>
<td>100</td>
</tr>
<tr>
<td>(n_2)</td>
<td>20</td>
<td>20</td>
<td>50</td>
<td>25</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>(m_2)</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Feasible instances</td>
<td>48</td>
<td>42</td>
<td>82</td>
<td>72</td>
<td>63</td>
<td>88</td>
</tr>
<tr>
<td>LPCC average time</td>
<td>300s</td>
<td>300s</td>
<td>300s</td>
<td>300s</td>
<td>300s</td>
<td>300s</td>
</tr>
<tr>
<td>Bundle average time</td>
<td>22s</td>
<td>45s</td>
<td>242s</td>
<td>52s</td>
<td>117s</td>
<td>298s</td>
</tr>
<tr>
<td>(AD_1)</td>
<td>7.75%</td>
<td>15.03%</td>
<td>81.39%</td>
<td>25.95%</td>
<td>24.25%</td>
<td>5.13%</td>
</tr>
<tr>
<td>(AD_2)</td>
<td>6.84%</td>
<td>6.43%</td>
<td>6.64%</td>
<td>55.81%</td>
<td>76.62%</td>
<td>111.10%</td>
</tr>
<tr>
<td>Fraction (AD_2 \leq 1%)</td>
<td>16.67%</td>
<td>9.52%</td>
<td>20.73%</td>
<td>8.33%</td>
<td>4.76%</td>
<td>2.27%</td>
</tr>
</tbody>
</table>

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

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

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

(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 $z_{up} \geq z$, so that $z + \hat{z} = z_{up}$. We replicate the first-stage decision variables $z, \hat{z}, t^+, \text{and} \ t^-$ for each second-stage scenario, denoting them by $z_i, \hat{z}_i, t_i^+, \text{and} \ t_i^-$, 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:

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

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 \( -\frac{\beta_i}{N_i} + Y_i^{t-} \), 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 \( (\frac{\beta_i}{N_i} + Y_i^{t+}) + (-\frac{\beta_i}{N_i} + Y_i^{t-}) = Y_i^{t+} + Y_i^{t-} < 0 \), 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 \( (\frac{\beta_i}{N_i} + Y_i^{t+}) < 0 \), then increasing \( t_i^+ \) can drive the subproblem value to \(-\infty\).

To address this case, we add constraints \( Y_i^{t+} \geq -\frac{\beta_i}{N_i} \), for all \( i \), to the master problem.

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

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

\[
\max_{r,Y} \sum_{i=1}^{N_1} r_i \left\| Y - \hat{Y}_k \right\|^2 \quad \text{s.t.} \quad \begin{align*}
    r^i &\leq f^i(\hat{Y}_k) + e^i_j + \langle X_{1,j}^*, Y^i - \hat{Y}_k, i \rangle, \quad \forall i \in \{1, \ldots, N_1\}, \ j \in \mathcal{B}_k^i \ \\
    Y &= (Y_1, \ldots, Y_{N_1}) \in K^* \\
    Y_i^{t+} + Y_i^{t-} &\geq 0, \quad \forall i \in \{1, \ldots, N_1\} \\
    Y_i^{t+} &\geq -\frac{\beta_i}{N_i}, \quad \forall i \in \{1, \ldots, N_1\} \\
    Y_i^{t-} &\geq \frac{\beta_i}{N_i} - \frac{\beta_i}{\alpha_i} \sum_j \pi_i \pi_{i,j}, \quad \forall i \in \{1, \ldots, N_1\}
\end{align*}
\] (5.3)

where \((X_{1,j}^*, X_{2,j}^*)\) is the optimal solution of \( f^i(Y_j) \) and \( e^i_j \) 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>
<thead>
<tr>
<th>Lagrange Multiplier</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_{j,i} )</td>
<td>Scalar Lagrange multipliers for constraints ( r^i \leq f^i(\hat{Y}<em>k) + e^i_j + \langle X^i</em>{j}, Y_i - \hat{Y}_{k,i} \rangle ) above, for ( i = 1, \ldots, N_1 ) and ( j \in B^i_k )</td>
</tr>
<tr>
<td>( \beta )</td>
<td>Vector of Lagrange multipliers for constraint ( Y_1 + \cdots + Y_{N_1} = 0 )</td>
</tr>
<tr>
<td>( \lambda_i )</td>
<td>Scalar Lagrange multipliers for constraints ( Y_i^{t^+} + Y_i^{t^-} \geq 0 ) above, for ( i = 1, \ldots, N_1 )</td>
</tr>
<tr>
<td>( \mu_i )</td>
<td>Scalar Lagrange multipliers for constraint ( Y_i^{t^+} \geq -\frac{\beta}{N_1} ) above, for ( i = 1, \ldots, N_1 )</td>
</tr>
<tr>
<td>( \delta_i )</td>
<td>Scalar Lagrange multipliers for constraint ( Y_i^{t^-} \geq \frac{\beta}{N_1} - \frac{\beta}{N_1} \sum_j \pi_i \pi_{i,j} ) above, for ( i = 1, \ldots, N_1 )</td>
</tr>
</tbody>
</table>

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

\[
\sum_{i=1}^{N_1} r^i - \frac{1}{2t_k} \|Y - \hat{Y}_k\|^2 - \sum_{i=1}^{N_1} \sum_{j \in B_k} \alpha_{j,i} (r^i - f^i(\hat{Y}_k) - e^i_j - \langle X_{ij}^*, Y_i - \hat{Y}_{k,i} \rangle)
\]

\[- \sum_{i=1}^{N_1} \beta Y_i - \sum_{i=1}^{N_1} \lambda_i (-Y_i^{t^+} - Y_i^{t^-}) - \sum_{i=1}^{N_1} \mu_i (-Y_i^{t^+} - \frac{\beta_1}{N_1})
\]

\[- \sum_{i=1}^{N_1} \delta_i (-Y_i^{t^-} + \frac{\beta_1}{N_1} - \frac{\beta_1}{\alpha_1} \sum_j \pi_i \pi_{i,j})
\]

\[= \sum_{i=1}^{N_1} r^i (1 - \sum_{j \in B_k} \alpha_{j,i}) + \sum_{i=1}^{N_1} \left[ - \frac{1}{2t_k} \|Y_i^{t^*} - \hat{Y}_{k,i}\|^2 + \sum_{j \in B_k} \alpha_{j,i} \langle z_{j,i}^*, Y_i^{t^*} - \hat{Y}_{k,i} \rangle - \hat{z}_{i}^* Y_i^{t^*} \right]
\]

\[+ \sum_{i=1}^{N_1} \left[ - \frac{1}{2t_k} (Y_i^{t^*} - \hat{Y}_{k,i})^2 + \sum_{j \in B_k} \alpha_{j,i} t_{j,i}^{t^*} (Y_i^{t^*} - \hat{Y}_{k,i}) - \beta^t Y_i^{t^*} + \lambda_i Y_i^{t^*} + \mu_i Y_i^{t^*} \right]
\]

\[+ \sum_{i=1}^{N_1} \left[ - \frac{1}{2t_k} (Y_i^{t^-} - \hat{Y}_{k,i})^2 + \sum_{j \in B_k} \alpha_{j,i} t_{j,i}^{t^-} (Y_i^{t^-} - \hat{Y}_{k,i}) - \beta^t Y_i^{t^-} + \lambda_i Y_i^{t^-} + \delta_i Y_i^{t^-} \right]
\]

\[+ \sum_{i=1}^{N_1} \sum_{j \in B_k} \alpha_{j,i} (f^i(\hat{Y}_k) + e^i_j) + \sum_{i=1}^{N_1} \mu_i \frac{\beta_1}{N_1} + \sum_{i=1}^{N_1} \delta_i (\frac{\beta_1}{N_1} + \frac{\beta_1}{\alpha_1} \sum_j \pi_i \pi_{i,j}).
\]

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

\[\nabla_{r^*} L = (1 - \sum_{j \in B_k^*} \alpha_{j,i}) = 0,
\]

\[\nabla_{Y_i^{t^*}} L = -\frac{1}{t_k} \langle Y_i^{t^*} - \hat{Y}_{k,i} \rangle + \sum_{j \in B_k^*} \alpha_{j,i} z_{j,i}^* - \hat{z}_{i}^* = 0,
\]

\[\nabla_{Y_i^{t^-}} L = -\frac{1}{t_k} \langle Y_i^{t^-} - \hat{Y}_{k,i} \rangle + \sum_{j \in B_k^*} \alpha_{j,i} z_{j,i}^* - \hat{z}_{i}^* = 0,
\]

\[\nabla_{(Y_i^{t^+})^*} L = -\frac{1}{t_k} \langle (Y_i^{t^+})^* - \hat{Y}_{k,i} \rangle + \sum_{j \in B_k^*} \alpha_{j,i} t_{j,i}^{t^*} - \beta^t + \lambda_i + \mu_i = 0,
\]

\[\nabla_{(Y_i^{t^-})^*} L = -\frac{1}{t_k} \langle (Y_i^{t^-})^* - \hat{Y}_{k,i} \rangle + \sum_{j \in B_k^*} \alpha_{j,i} t_{j,i}^{t^-} - \beta^t + \lambda_i + \delta_i = 0,
\]

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

\[Y_i^{t^*} = \hat{Y}_{k,i} + \sum_{j \in B_k^*} t_k \alpha_{j,i} z_{j,i}^* - t_k \hat{z}_{i}^*,
\]
By plugging these values in the Lagrangian function, \( \max_{r,Y} L(r, Y, \alpha, \beta, \lambda, \mu, \delta) \) is given by:

\[
\begin{align*}
\sum_{i=1}^{N_1} \left[ -\frac{1}{2t_k} \left( \sum_{j \in B_k^i} t_k \alpha_{j,i} z_{j,i}^t - t_k \beta z^t \right) \right]^2 + \sum_{j \in B_k^i} \alpha_{j,i} (z_{j,i}^t - \sum_{j \in B_k^i} t_k \alpha_{j,i} z_{j,i}^t - t_k \beta z^t) \\
- \beta^t (Y_{k,i} - \sum_{j \in B_k^i} t_k \alpha_{j,i} z_{j,i}^t - t_k \beta z^t) + \sum_{i=1}^{N_1} \left[ -\frac{1}{2t_k} \left( \sum_{j \in B_k^i} t_k \alpha_{j,i} t_{j,i}^{t^+} - t_k \beta t^+ + t_k \lambda_i + t_k \mu_i \right) \right]^2 \\
+ \sum_{j \in B_k^i} \alpha_{j,i} (t_{j,i}^{t^+} - \sum_{j \in B_k^i} t_k \alpha_{j,i} t_{j,i}^{t^+} - t_k \beta t^+ + t_k \lambda_i + t_k \mu_i) \\
+ (-\beta t^+ + \lambda_i + \mu_i) (Y_{k,i} - \sum_{j \in B_k^i} t_k \alpha_{j,i} t_{j,i}^{t^+} - t_k \beta t^+ + t_k \lambda_i + t_k \mu_i) \\
+ \sum_{i=1}^{N_1} \left[ -\frac{1}{2t_k} \left( \sum_{j \in B_k^i} t_k \alpha_{j,i} t_{j,i}^{t^-} - t_k \beta t^- + t_k \lambda_i + t_k \delta_i \right) \right]^2 \\
+ \sum_{j \in B_k^i} \alpha_{j,i} (t_{j,i}^{t^-} - \sum_{j \in B_k^i} t_k \alpha_{j,i} t_{j,i}^{t^-} - t_k \beta t^- + t_k \lambda_i + t_k \delta_i) \\
+ (-\beta t^- + \lambda_i + \delta_i) (Y_{k,i} - \sum_{j \in B_k^i} t_k \alpha_{j,i} t_{j,i}^{t^-} - t_k \beta t^- + t_k \lambda_i + t_k \delta_i) \\
+ \sum_{i=1}^{N_1} \sum_{j \in B_k^i} \alpha_{j,i} (f^j (Y_k) + e_j^i) + \sum_{i=1}^{N_1} \mu_i \frac{\beta_1}{N_1} + \sum_{i=1}^{N_1} \delta_i (\frac{\beta_1}{N_1} + \frac{\beta_1}{N_1} \sum_{j} \pi_i \pi_{i,j}) \\
- \sum_{i=1}^{N_1} \frac{t_k}{2} \left( \sum_{j \in B_k^i} z_{j,i}^s \alpha_{j,i} - \beta^2 \right)^2 - \sum_{i=1}^{N_1} \beta^2 Y_{k,i}^z + \sum_{i=1}^{N_1} \frac{t_k}{2} \left( \sum_{j \in B_k^i} z_{j,i}^s \alpha_{j,i} - \beta^2 \right)^2 - \sum_{i=1}^{N_1} \beta^2 Y_{k,i}^z \\
+ \sum_{i=1}^{N_1} \frac{t_k}{2} (\alpha_{j,i} t_{j,i}^{t^+} - \beta t^+ + \lambda_i + \mu_i)^2 + \sum_{i=1}^{N_1} (-\beta t^+ + \lambda_i + \mu_i) Y_{k,i}^{t^+}.
\end{align*}
\]
\[
\begin{align*}
&+ \sum_{i=1}^{N_1} \frac{t_k}{2}(a_{j,i} - t_{j,i} - \beta t + \lambda_i + \delta_i)^2 + \sum_{i=1}^{N_1} (-\beta t + \lambda_i + \delta_i)\hat{Y}_{k,i}^t \\
&+ \sum_{i=1}^{N_1} \sum_{j \in B_k^i} \alpha_{j,i}(f^i(\hat{Y}_k) + e_j^i) + \sum_{i=1}^{N_1} \mu_i + \sum_{i=1}^{N_1} \delta_i(-\beta + \lambda_i + \delta_i)\hat{Y}_{k,i}^t
\end{align*}
\]

Therefore, the dual of the master problem is:

\[
\begin{align*}
&\min_{\alpha, \beta, \lambda, \mu, \delta} \sum_{i=1}^{N_1} \left[ \nu_i z_i^2 + (\nu_i - \beta)z_i + \beta \hat{Y}_{k,i}^z + \beta \hat{Y}_{k,i}^t + (\nu_i - \beta)\hat{Y}_{k,i}^t \right] \\
&+ \sum_{i=1}^{N_1} \sum_{j \in B_k^i} \alpha_{j,i}(f^i(\hat{Y}_k) + e_j^i) + \sum_{i=1}^{N_1} \mu_i + \sum_{i=1}^{N_1} \delta_i(-\beta + \lambda_i + \delta_i)\hat{Y}_{k,i}^t
\end{align*}
\]

s.t.

\[
\begin{align*}
\sum_{j \in B_k^i} \alpha_{j,i} &= 1, \quad \forall i \\
\nu_i^z &= \sum_{j \in B_k^i} \hat{z}_{j,i} - \beta, \quad \forall i \\
\nu_i^z &= \sum_{j \in B_k^i} \hat{z}_{j,i} - \beta, \quad \forall i \\
\nu_i^{t+} &= \alpha_{j,i}t_{j,i}^{t+} - \beta^{t+} + \lambda_i + \mu_i, \quad \forall i \\
\nu_i^{t-} &= \alpha_{j,i}t_{j,i}^{t-} - \beta^{t-} + \lambda_i + \delta_i, \quad \forall i \\
\alpha_{i,j}, \beta, \lambda_i, \mu_i, \delta_i &\geq 0, \quad \forall i, j \in B_k^i
\end{align*}
\]

Let \( \beta^* \) be the optimal value of \( \beta \) in the above problem. Set \( X_1 = \beta^* \) and solve all the second-stage problems (5.2) to find an optimal point \( X_2^* \in X_2^*(\beta^*) \). By plugging \((\beta^*, X_2^*)\) 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 \( N_2 \) 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 \(BRM_{0.05}^{0.75}\) 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.

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

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


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