Multivariate Risk Measures for Portfolio Risk Management

by

Huameng Jia

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Actuarial Science

Waterloo, Ontario, Canada, 2020

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In portfolio risk management, the main foci are to control the aggregate risk of the entire portfolio and to understand the contribution of each individual risk unit in the portfolio to the aggregate risk. When univariate risk measures are used to quantify the risks associated with a portfolio, there is usually a lack of consideration of correlations between individual risk units and the aggregate risk and of dependence among these risks. For this reason, multivariate risk measures defined by considering the joint distribution of risk units in the portfolio are more desirable. In this thesis, we define new multivariate risk measures by minimizing multivariate loss functions subject to various. constraints. With the proposed multivariate risk measures, we obtain risk measures for the entire portfolio and each individual risk unit in the portfolio at the same time.

In Chapter 2, we introduce a multivariate extension of Conditional Value-at-Risk (CVaR) based on a multivariate loss function associated with different risks related to portfolio risk management. We prove that the defined multivariate risk measure satisfies many desirable properties such as positive homogeneity, translation invariance and subadditivity. Then, we provide numerical illustrations with multivariate normal distribution to show the effects of the parameters in the model. After that, we also perform a comparison between our multivariate CVaR and other traditional univariate risk measures such as VaR and CVaR.

In Chapter 3, we define a multivariate risk measure for capital allocation purposes. Unlike most of the existing allocation principles that assume the total capital is exogenously given, we obtain the optimal total capital for the entire portfolio and the optimal capital allocation to all the individual risk units in the portfolio at the same time. In this chapter, we first discuss our model with a two-level organization/portfolio structure. Then, we move to a more complex three-level organization/portfolio structure. We find that many of the existing allocation principles can be seen as special or limiting cases of our model. In addition, our model can explain those allocation principles as solutions to optimization problems. Finally, we provide a numerical example for the two-level organization/portfolio structure model with two different error functions.

In Chapter 4, we introduce a multivariate shortfall risk measure induced by cumulative prospect theory (CPT) and give the corresponding risk allocations under the multivariate shortfall risk measure. To obtain this risk measure, we make an extension of previously studied univariate generalized shortfalls induced by CPT and incorporate the idea of systemic risk. In this study, we discuss the properties of the risk measure and conditions

for its existence and uniqueness. Also, we perform a simulation study and a comparison to a previously studied multivariate shortfall to show that our model can provide a more reasonable risk measure and allocation result.

Acknowledgements

I am most grateful to my supervisor Professor Jun Cai for his guidance throughout the years of my PhD study at the University of Waterloo. I started my PhD study without any research experience and Professor Cai provided me with great training on how to carryout research. Without his help, the completion of this dissertation is impossible.

My deep gratitude also goes to Professor Bin Li and Professor Chengguo Weng, both internal members of my PhD thesis examining committee, for their valuable advice on my dissertation. I am also very grateful to Professor Alan Huang (internal-external examiner from School of Accounting and Finance) and Professor Jiandong Ren (external examiner from Western University) for their insightful comments and suggestions.

In addition, I wish to acknowledge the help of Dr. Tom Moss Gamblin for proofreading my dissertation.

Finally, my special thanks go to my parents and my wife for their great love and support.

Dedication

This is dedicated to my parents. Because of their decision, I could immigrate to Canada and start my university education here. Thanks to their selfless support, I could concentrate on my PhD study in the past four years.

This is also dedicated to my wife. I was so lucky to meet her and marry her during my PhD study. Now, she is planning to start her PhD study and I hope that everything will go smoothly.

This is also dedicated to my grandparents who raised me. Although they only received limited education in their lives, they believed that education is the most important thing in life. They were very pleased when they knew I got admitted into the PhD program and really wanted to witness my graduation. Unfortunately, they passed away during my PhD study. They will be deeply missed.

Table of Contents

List of Tables

Chapter 1

Introduction

1.1 Background

In portfolio risk management, two of the most important questions are to quantify the risk for the entire portfolio and the risk units in the portfolio and to find the optimal allocation of the available capital to each risk unit in the portfolio. In this thesis, we use multivariate risk measure approaches to accomplish those two tasks.

1.1.1 Risk measures

To quantify risks, many risk measures have been developed, and their properties have been studied in detail. A risk measure ρ is a mapping from $\mathcal{X}$ to $\mathbb{R} = (-\infty, +\infty)$, where $\mathcal{X}$ is the set of loss random variables, namely, for any $X \in \mathcal{X}$, $\rho(X) \in (-\infty, +\infty)$. Some definitions and desirable properties of a risk measure are listed below:

A risk measure ρ defined on a convex cone $\mathcal{X}$ containing all the constants is *coherent* if the following four properties hold:

1. Monotonicity: $\rho(X) \leq \rho(Y)$ for all $X, Y \in \mathcal{X}$ with $X \leq Y$ almost surely (a.s.).

2. Translation-invariance: $\rho(X + c) = \rho(X) + c$ for all $X \in \mathcal{X}$ and all $c \in \mathbb{R}$.

3. Positive homogeneity: $\rho(\lambda X) = \lambda \rho(X)$ for all $\lambda \in \mathbb{R}_+ = [0, \infty)$ and all $X \in \mathcal{X}$.

4. Subadditivity: $\rho(X + Y) \leq \rho(X) + \rho(Y)$ for all $X, Y \in \mathcal{X}$.

A risk measure ρ defined on a convex set $\mathcal{X}$ that is closed under translation is *convex* if it satisfies monotonicity, translation-invariance and

5. Convexity: $\rho(aX + (1-a)Y) \leq a\rho(X) + (1-a)\rho(Y)$ for all $X, Y \in \mathcal{X}$ and $a \in (0, 1)$.

One of the popular risk measures used in risk management is Value at Risk (VaR). It has an intuitive interpretation and is easy to implement. However, it is also criticized by many researchers and regulators as not providing a proper measurement for the loss severity in a rare event, and as not satisfying the subadditivity property. Artzner et al. (1999) define the axiomatic properties of a desirable risk measure and introduce the concept of a coherent risk measure. These ideas play an essential role in modern quantitative risk management. Föllmer and Schied (2002) follow a similar idea by defining convex risk measure, and Rockafellar, Uryasev, et al. (2000) introduce Conditional Value at Risk (CVaR, also called expected shortfall). Following this, Bellini and Frittelli (2002) and Frittelli and Gianin (2002) incorporate expected utility theory in risk measures. Researchers also use distortion functions to model preference in risk management. For related literature, please refer to Wang (1995), Acerbi (2002), Belles-Sampera et al. (2014), Mao and Cai (2018) and so on. While univariate risk measures have many desirable properties, they may not be suitable for portfolio risk management. Univariate risk measures can define the risk of every single unit. However, as a univariate risk measure is determined based only on its single distribution, it is hard to incorporate the correlation between individual risk units within the same large portfolio under the same scenario. For example, a 99% VaR for unit 1 is very likely to have a different scenario than 99% VaR for unit 2 in the portfolio. It is also hard to determine the contribution of each risk unit to the aggregate risk of the portfolio under the same scenario.

In portfolio risk management, multivariate risk measures can help us to understand the risk structure of a portfolio and the correlation between the risk units in the portfolio better than univariate risk measures. Literature on multivariate risk measures can be found in Cai and Li (2005), Embrechts and Puccetti (2006), Lee and Prékopa (2013), Noyan and Rudolf (2013), Cousin and Di Bernardino (2014), Huerlimann (2014), Cossette et al. (2016), Landsman et al. (2016), Cai et al. (2017), Prékopa and Lee (2018), Herrmann et al. (2020), Shushi and Yao (2020), and references therein.

Most multivariate risk measures use a random vector, say $\mathbf{X} = (X_1, ..., X_n)$, to represent the portfolio, where each component of the random vector, X_i, for $i = 1, ..., n$, corresponds to each risk unit in the portfolio respectively. Then, this random vector is used to calculate the multivariate risk measure. During this process, the aggregate random variable, $S = \sum_{i=1}^{n} X_i$, is not directly involved, and the risk measure for the aggregate risk, S, is not obtained from these multivariate risk measures. In this thesis, we obtain the risk measures for the aggregate risk of the entire portfolio, S, and each risk unit in the portfolio at the same time by minimizing a multivariate objective function. In this way, we can confidently say the obtained multivariate risk measure, including the risk measure for the aggregate risk, is the optimal solution that satisfies the risk manager's objective or preference based on the given objective funtion. In this thesis, we provide two approaches. In the first approach, we derive the multivariate risk measure by including the aggregate risk random variable, S, in the multivariate objective function. The obtained multivariate risk measure has a dimension of $n + 1$ and is in the form of $(\rho_1(X_1), ..., \rho_n(X_n), \rho_S(S))$ where the first n components of the vector are the risk measures for the risk units in the portfolio and the last component of the vector is the risk measure for the aggregate risk of the entire portfolio. In the second approach, we combine the concept of systemic risk management with cumulative prospect theory (CPT). In systemic risk management, portfolio managers focus on analyzing the risk of failure of the entire system/portfolio caused by the failure of each risk unit in the system/portfolio. Systemic risk measures have been widely studied since the financial crisis of 2008. The literature on systemic risk measure studies includes Bartram et al. (2007), Chen et al. (2013), Kromer et al. (2016), Feinstein et al. (2017), Weber and Weske (2017), Acharya et al. (2017), Armenti et al. (2018), Biagini et al. (2019) and Brunnermeier and Cheridito (2019). In this approach, we use the acceptance set concept in systemic risk measures to define the acceptable monetary allocation, which is a vector of dimension n. When combining with CPT, we can represent the portfolio manager's risk appetite more accurately. Defining the risk measure for the entire portfolio to be the minimum of the sum of the n components of each acceptable monetary allocation, the acceptable monetary allocation that corresponds to this minimum is used to be the risk measure for each risk unit in the portfolio.

In comparison to the existing approaches for deriving multivariate risk measures, our approach can obtain risk measures for the entire portfolio and each risk unit in the portfolio at the same time as the optimal solution to a multivariate objective function. In this way, our approach can incorporate the correlations between all risk units and the aggregate risk. Our defined risk measure can help portfolio managers and shareholders to understand the relative importance of each individual risk in the system, information which can then be used to provide a guideline for more reasonable capital allocation method. Also, our approach can provide an intuitive meaning for the obtained risk measure, namely the optimal solution of the portfolio manager's objective based on the given objective function.

1.1.2 Capital allocation principles

Another important topic in portfolio risk management is finding an optimal capital allocation method. The portfolio manager/business owner needs to hold enough capital to make sure the portfolio/business can survive some extreme scenarios. At the same time, a portfolio/business may have many sub-portfolios/sub-business lines. In this situation, a proper allocation of the total capital to each sub-portfolio/sub-business line is important. This can be seen from different perspectives. For accounting purposes, capital allocation can help a portfolio manager to allocate costs and expenses. As pointed out by Dhaene, Tsanakas, et al. (2012), there are usually costs associated with holding capital, which may be either frictional costs or opportunity costs. After the capital allocation process, the portfolio manager can assign costs to each risk unit, and can also can allocate other direct and indirect expenses of portfolio management based on the allocated capital. For performance evaluation purposes, the portfolio manager can use the allocated capital to calculate the return on capital. For risk management, the portfolio manager can allocate regulatory or economic capital to each risk unit within the portfolio. The allocated capital can then be viewed as the risk exposure of each risk unit. In this case, the capital allocation process is also the risk allocation process. Furthermore, as pointed out in Hesselager and Andersson (2002), the risk allocation result can then be used to decide the premium for each risk unit in the insurance field. In all of the situations above, it is crucial to find the optimal allocation method.

In a capital allocation problem, we assume there are n individual risk units in a system with losses represented by X_1, X_2, ..., X_n, as in a multivariate risk measure problem. The aggregated loss S is given by $S = \sum_{i=1}^{n} X_i$. We have an additional variable K representing the total available capital, and variables K_i, for $i = 1, 2, ..., n$ representing the capital allocated to each individual risk unit. We can also extend this two-level structure portfolio to a three-level structure. In this case, each K_i will be considered as the allocated capital for each sub-portfolio X_i, and will then be distributed again to the next level. The goal is to find an optimal strategy for determining the allocation under some optimization criteria. In the past two decades, much research has been done on capital allocation strategies. The literature on capital allocations includes Overbeck (2000), Cummins (2000), Myers and Read Jr (2001), Denault (2001), Dhaene et al. (2003), Tsanakas (2004), Kalkbrener (2005), Sherris (2006), Tsanakas (2007), Dhaene et al. (2008), Dhaene, Tsanakas, et al.

(2012),Zaks and Tsanakas (2014), Cai and Wang (2020), and references therein.

Here, we give some of the commonly used capital allocation methods:

(a) Haircut: for given $\alpha \in (0,1)$ and $i = 1, 2, ..., n$,

$$K_i = \frac{F_{X_i}^{-1}(\alpha)}{\sum_{i=1}^{n} F_{X_i}^{-1}(\alpha)} K,$$

where $F_{X_i}^{-1}(\alpha)$ is the VaR at confidence level α for X_i.

(b) Covariance: for $i = 1, 2, ..., n$,

$$K_i = \frac{\text{Cov}(X_i, \sum_{i=1}^{n} X_i)}{\text{Var}(\sum_{i=1}^{n} X_i)} K.$$

(c) Conditional Tail Expectation (CTE): for given $\alpha \in (0,1)$ and $i = 1, 2, ..., n$,

$$K_i = \frac{\mathbb{E}[X_i | S > F_S^{-1}(\alpha)]}{\mathbb{E}[S | S > F_S^{-1}(\alpha)]} K,$$

where $S = \sum_{i=1}^{n} X_i$.

For all of the capital allocation methods above, the total capital K is calculated separately. Then, in the process of deriving the capital allocation method, we assume that the total capital K is given, obtaining the allocation method by minimizing objective functions that involve risk random variables X_i, for $i = 1, ..., n$, under some constraints. In this thesis, we provide a new approach to the capital allocation method. In our approach, we obtain the optimal total capital K and the allocation to each risk unit K_i simultaneously. To achieve this goal, we incorporate the aggregate loss random variable, S, in the multivariate objective function, and obtain the total capital, K, as part of our optimization process. In comparison to other approaches, our total capital K and allocated capital K_i are derived under the same framework, and they are consistent with the portfolio manager's objective. In additional to claiming the allocation method is optimal under the portfolio manager's objective, we can confidently claim that the total required capital is also optimal under the same objective. Moreover, our allocation method also generalizes many existing allocation methods and provides explanations for those allocation methods from the perspective of optimization problems.

1.2 Outline of the thesis

In Chapter 2, we extend the univariate CVaR to a multivariate CVaR. Our study is motivated by the multivariate geometric VaR and expectile risk measures studied in Chaudhuri (1996), Maume-Deschamps et al. (2017), and Herrmann et al. (2018). In this chapter, we propose a new multivariate CVaR (MCVaR) risk measure by defining a multivariate loss function from the perspective of systemic risk management. In contrast to Chaudhuri (1996), Maume-Deschamps et al. (2017), Herrmann et al. (2018), and most existing multivariate risk measures, our MCVaR will consider both individual and aggregate risks of portfolios, but will prioritize aggregate risks. In this chapter, we discuss the conditions for existence and uniqueness of a solution, following by desirable properties of the risk measure. We round out the chapter with a numerical illustration of our defined risk measure and a comparison with existing risk measures.

In Chapter 3, inspired by the idea of systemic risk measures, which consider the risk measure of the entire system and the contribution of the risk of each individual risk unit at the same time, we propose a new method of capital allocation which can simultaneously decide the optimal total capital and the capital allocation to each individual unit in the system. Our model extend the models proposed in Furman and Zitikis (2008) and Cai and Wang (2020), using additional components in the model to consider the risk of the entire system/portfolio. In this chapter, we discuss the optimal solution for a system with two or three layers of structures. We provide the conditions for existence and uniqueness of solutions. We also provide numerical illustrations of our proposed allocation and make a comparison with the results of existing allocation methods.

In Chapter 4, we extend the idea of generalized shortfalls induced by cumulative prospect theory (CPT) in Mao and Cai (2018) to multivariate risk measures with the concept of systemic risk, a common acceptance set technique used in systemic risk and the method used in Armenti et al. (2018). In systemic risk, we consider the risk of the entire system and how the failure of each individual unit in the system may lead to the failure of the entire system. In our model, we extend the univariate generalized shortfalls induced by CPT. This model can also be applied to the problem of capital allocation. In this chapter, we first review the concept of CPT and the model, univariate generalized shortfalls induced by CPT, followed by the technique used to extend the univariate model to our multivariate generalized shortfalls induced by CPT. Then, we discuss the existence and uniqueness of the solution. Also, we perform a simulation study based on Armenti et al. (2018)'s original study and make a comparison with Armenti et al. (2018)'s result to show why our model can provide a more reasonable result.

In Chapter 5, we provide a summarization for each chapter and for the whole thesis. In this part, we revisit our approach for defining new multivariate risk measures and the shared common approach for defining them.

Throughout this thesis, "increasing" means "non-decreasing" rather than "strictly increasing", and likewise "decreasing" means "non-increasing".

Chapter 2

A multivariate CVaR risk measure derived from minimizing a multivariate loss function

2.1 Introduction

In portfolio risk management, one important step is to quantify the risk of the entire portfolio and the risk of each individual component in the portfolio. To perform this task, many types of risk measures have been developed. Currently, the most commonly used risk measures are univariate risk measures due to their simplicity. Generally speaking, a univariate risk measure is a mapping from a set of random variables to real numbers. In insurance and finance risk management, a risk measure of a random variable X can provide not only an assessment of the severity of risk X, but also a guideline about how to prepare the required capital or to determine the insurance premium for risk X. When a decision maker faces a random vector or risk portfolio $\mathbf{X} = (X_1, ..., X_d)$, the decision maker wants to assess not only the severities of all individual risks, but also the severities of other risks in the portfolio, in particular, the severity of the aggregate risk $\sum_{i=1}^{d} X_i$ of the portfolio, which is the main concern in portfolio risk management as pointed out by Burgert and Rüschendorf (2006). A decision scheme of determining required capital or insurance premiums for risk portfolios often involves both individual and aggregate risks. To manage risk portfolios, many multivariate risk measures have been proposed from different perspectives. Many of these multivariate risk measures are extensions of the commonly used univariate VaR, CVaR, and expectile risk measures. To understand the ideas behind these extensions, we recall the definitions of the univariate VaR, CVaR, and expectile risk measures. More examples of univariate risk measures can be found in Mao and Cai (2018), Cai and Mao (2020), and the references therein. The value-at-risk (VaR) risk measure $\mathrm{VaR}_\alpha(X)$ of any random variable X with distribution F_X at confidence level $\alpha \in (0,1)$ is its quantile at level α and is defined by

$$\mathrm{VaR}_\alpha(X) = \inf\{x \in \mathbb{R} : F_X(x) \geq \alpha\}. \tag{2.1.1}$$

In addition, $\mathrm{VaR}_\alpha(X)$ is also the smallest minimizer to the optimization problem $\min\limits_{c\in\mathbb{R}} \mathbb{E}(f_\alpha(X-c))$, where the loss function $f_\alpha(t)$ is defined as

$$f_\alpha(t) = \alpha(t)_+ + (1-\alpha)(t)_- = \frac{1}{2}(|t| + (2\alpha-1)t), \tag{2.1.2}$$

where for any $x \in \mathbb{R}$, $x_+ = \max(x,0)$ and $x_- = \max(-x,0)$ satisfying $x = x_+ - x_-$ and $|x| = x_+ + x_-$. The expectile risk measure $e_\alpha(X)$ of a random variable X with $\mathbb{E}(X^2) < \infty$ at confidence level $\alpha \in (0,1)$ is the unique minimizer to the optimization problem $\min\limits_{c\in\mathbb{R}} \mathbb{E}(f_\alpha(X-c))$, where the loss function $f_\alpha(t)$ is defined as

$$f_\alpha(t) = \alpha((t)_+)^2 + (1-\alpha)((t)_-)^2 = \frac{1}{2}\,|t|\,(|t| + (2\alpha-1)t). \tag{2.1.3}$$

The conditional value-at-risk (CVaR) risk measure $\mathrm{CVaR}_\alpha(X)$ of a random variable X with $\mathbb{E}|X| < \infty$ at confidence level $\alpha \in (0,1)$ is the minimum value of an expected loss function $\mathbb{E}(f_{X,\alpha}(c))$, namely

$$\mathrm{CVaR}_\alpha(X) = \min_{c\in\mathbb{R}} \mathbb{E}(f_{X,\alpha}(c)) = \min_{c\in\mathbb{R}} \left\{c + \frac{1}{1-\alpha}\,\mathbb{E}(X-c)_+\right\}, \tag{2.1.4}$$

where the loss function $f_{x,\alpha}(t)$ is defined as

$$f_{x,\alpha}(t) = t + \frac{(x-t)_+}{1-\alpha} = t + \frac{|x-t| + (x-t)}{2(1-\alpha)}. \tag{2.1.5}$$

It is well known (see, for example, Rockafellar, Uryasev, et al. (2000)) that $\mathrm{VaR}_\alpha(X)$ is the smallest minimizer to optimization problem (2.1.4) and that

$$\mathrm{CVaR}_\alpha(X) = \mathrm{VaR}_\alpha(X) + \frac{1}{1-\alpha}\,\mathbb{E}(X - \mathrm{VaR}_\alpha(X))_+ \tag{2.1.6}$$

$$= \mathbb{E}(X|X > \mathrm{VaR}_\alpha(X)) \quad \text{(if } X \text{ has a continuous distribution).} \tag{2.1.7}$$

In fact, problem (2.1.4) can be viewed as the problem a regulator wishes to solve for minimizing a weighted sum of the total capital (premium) provided by shareholders (insureds) and the expected shortfall borne by the debtholder (insurer).

A multivariate VaR is defined in Cousin and Di Bernardino (2013) by extending the critical interval in (2.1.1) to a multidimensional critical set. Several multivariate conditional tail expectation (CTE) risk measures are introduced by extending the conditional expectation in (2.1.7) to multivariate forms conditioning on different extreme events associated with a random vector. See, for example, Cai and Li (2005), Embrechts and Puccetti (2006), Lee and Prékopa (2013), Noyan and Rudolf (2013), Cousin and Di Bernardino (2014), Huerlimann (2014), Cossette et al. (2016), Landsman et al. (2016), Cai et al. (2017), Prékopa and Lee (2018), Herrmann et al. (2020), Shushi and Yao (2020), and the references therein. For the purposes of portfolio risk management, another important way to define multivariate risk measures is to extend loss functions (2.1.2) and (2.1.3) to multivariate loss functions, and to define multivariate risk measures as optimal solutions that minimize the expectations of these multivariate loss functions. Indeed, Chaudhuri (1996) proposes a multivariate quantile/VaR by extending loss function (2.1.2) to a multivariate case. Maume-Deschamps et al. (2017) and Herrmann et al. (2018) discuss multivariate expectiles by extending loss function (2.1.3) to multivariate forms. Other multivariate risk measures proposed in this way can be found in Prékopa (2012), Torres et al. (2015), Meraklı and Küçükyavuz (2018), and the references therein.

For $\mathbf{x} = (x_1, ..., x_d)$, $\mathbf{y} = (y_1, ..., y_d) \in \mathbb{R}^d$, and $a \in \mathbb{R}$, $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^d x_i^2}$ is the Euclidean norm of $\mathbf{x}$, $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^d x_i y_i$ is the inner product of $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{x} + \mathbf{y} = (x_1 + y_1, ..., x_d + y_d)$, and $a\mathbf{x} = (ax_1, ..., ax_d)$. For a random vector $\mathbf{X} = (X_1, \ldots, X_d)$ and real vector $\mathbf{u} = (u_1, \ldots, u_d) \in B^d = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 < 1\}$, Chaudhuri (1996) defines a multivariate geometric quantile/VaR for $\mathbf{X}$ with preference vector $\mathbf{u}$, denoted by $\mathrm{MVaR}_{\mathbf{u}}(\mathbf{X}) = (\mathrm{MVaR}_1(\mathbf{X}), ..., \mathrm{MVaR}_d(\mathbf{X}))$, as the unique minimizer to the optimization problem $\min_{\mathbf{c} \in \mathbb{R}^d} \mathbb{E}(\Phi_{\mathbf{u}}(\mathbf{X} - \mathbf{c}))$ if such a unique minimizer exists, where the loss function $\Phi_{\mathbf{u}}(\mathbf{t})$ is defined as

$$\Phi_{\mathbf{u}}(\mathbf{t}) = \frac{1}{2}(\|\mathbf{t}\|_2 + \langle \mathbf{u}, \mathbf{t} \rangle), \qquad \mathbf{t} \in \mathbb{R}^d, \tag{2.1.8}$$

and $\mathrm{MVaR}_i(\mathbf{X})$ is the risk measure of X_i, $i = 1, ..., d$, and called the marginal VaR of X_i. Chaudhuri (1996) shows that if $\mathbf{X}$ has a joint density function, then there exists a unique $\mathbf{c}^* = (c_1^*, ..., c_d^*) = \mathrm{MVaR}_{\mathbf{u}}(\mathbf{X}) \in \mathbb{R}^d$ such that $\mathbb{E}(\Phi_{\mathbf{u}}(\mathbf{X} - \mathbf{c}^*)) = \min_{\mathbf{c} \in \mathbb{R}^d} \mathbb{E}(\Phi_{\mathbf{u}}(\mathbf{X} - \mathbf{c}))$, and $\mathbf{c}^*$ is the unique solution to the following system of equations:

$$\mathbb{E}\left(\frac{X_i - c_i}{\|\mathbf{X} - \mathbf{c}\|_2}\right) = -u_i, \quad i = 1, ..., d. \tag{2.1.9}$$

Following the idea of Chaudhuri (1996), Herrmann et al. (2018) define a multivariate geometric expectile $\rho_{\mathbf{u}}(\mathbf{X})$ for a random vector $\mathbf{X} = (X_1, ..., X_d)$ with preference vector $\mathbf{u} \in B^d$ as the unique minimizer to the optimization problem $\min_{\mathbf{c} \in \mathbb{R}^d} \mathbb{E}(\varphi_{\mathbf{u}}(\mathbf{X} - \mathbf{c}))$ if such a unique minimizer exists, where the loss function $\varphi_{\mathbf{u}}(\mathbf{t})$ is defined as

$$\varphi_{\mathbf{u}}(\mathbf{t}) = \frac{1}{2}\,\|\mathbf{t}\|_2\,(\|\mathbf{t}\|_2 + \langle \mathbf{u}, \mathbf{t}\rangle), \qquad \mathbf{t} \in \mathbb{R}^d.$$

Herrmann et al. (2018) show that if $\mathbb{E}(X_i)^2 < \infty$, $i = 1, .., d$, then there exists a unique $\mathbf{c}^* = (c_1^*, ..., c_d^*) = \rho_{\mathbf{u}}(\mathbf{X}) \in \mathbb{R}^d$ such that $\mathbb{E}(\varphi_{\mathbf{u}}(\mathbf{X} - \mathbf{c}^*)) = \min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(\varphi_{\mathbf{u}}(\mathbf{X} - \mathbf{c}))$.

Motivated by the multivariate geometric VaR and expectile risk measures studied in Chaudhuri (1996), Maume-Deschamps et al. (2017), and Herrmann et al. (2018), in this chapter we propose a new multivariate CVaR (MCVaR) risk measure by defining a multivariate loss function from the perspective of systemic risk management. In contrast to Chaudhuri (1996), Maume-Deschamps et al. (2017), Herrmann et al. (2018), and most existing multivariate risk measures, our MCVaR will consider both individual risks and aggregate risks of portfolios, but prioritize aggregate risks of portfolios.

In a standard approach to risk management of a set of risks, the risk measurement of the sum of risks is first determined by a univariate risk measure of the sum such as VaR or CVaR, which is then allocated among the set of risks by an allocation principle such as the haircut or CTE principles. Such a standard approach usually entails a top-down risk decomposition strategy. In a top-down approach, the risk measurement of the sum of risks depends only on the distribution of the sum or aggregate risk, and does not consider the relative importance of each individual risk to the aggregate risk and/or the dependence between each individual risk and the aggregate risk. In risk management of finance and insurance, the risk measurement of the sum of risks and of individual risks can be used to determine the required capital for a corporation and its subsidiaries, respectively, or to calculate the premiums for the combined risk and individual risks of an insurance portfolio, respectively. In a standard or top-down approach, the risk measurement of the sum of risks and the risk measurements of individual risks are determined in a two-step procedure. In the proposed multivariate CVaR approach, these risk measurements are determined simultaneously in a single-step procedure. Using the multivariate CVaR approach, a decision maker such as a regulator or an insurer can take into consideration the relative importance of each individual risk to the aggregate risk and/or the dependence between each individual risk and the aggregate risk when determining required capital for a corporation or premiums for an insurance portfolio. Generally speaking, such a one-step approach provides a way to determine risk measurements for a set of risks from a systemic point of view. In addition, the multivariate CVaR also provides a new way to determine the risk measurement of the sum of risks, which can be used in a standard or top-down approach to risk measurement of a set of risks.

In this chapter, we consider the risk measurement of a random vector or set of risks that represents the risk of a corporation or an insurance portfolio. We refer the systemic risk of the corporation or portfolio to the situation when the aggregate risk of the corporation or portfolio is greater than the required capital for the corporation or portfolio. Not only does the systemic risk of the corporation or portfolio depend on the aggregate risk of the corporation or portfolio, but it also links to the individual risks of the corporation or portfolio, the relative importance of each individual risk to the aggregate risk, the dependence between each individual risk and the aggregate risk, and the dependence among the individual risks. This motivates us to propose the multivariate CVaR risk measure from a systemic viewpoint or from the perspective of systemic risk management. A decision maker (a regulator or an insurer) may use this multivariate CVaR to determine the required capital for a corporation and its subsidiaries or to calculate premiums for the combined risk and individual risks of an insurance portfolio from a systemic point of view. In fact, determining the required capital for a corporation or premiums for an insurance portfolio is an important application of multivariate risk measures in risk management for finance and insurance.

The rest of this chapter is structured as follows. In Section 2.2, we give preliminary lemmas about convex optimization problems and consider a generalization of loss function (2.1.5) associated with the univariate CVaR. In Section 2.3, we extend this loss function (2.1.5) to a multivariate loss function considering the systemic risk of a portfolio and define a new multivariate CVaR (MCVaR) risk measure, which places emphasis on aggregate risks. In Section 2.4, we give sufficient conditions for the existence of the new MCVaR risk measure and obtain an expression for its solution. In Section 2.5, we discuss the properties of the new MCVaR including positive homogeneity, translation invariance, subadditivity, and monotonicity under some assumptions. In Section 2.6, we illustrate the MCVaR by numerical examples and explore the effect of dependence among individual risks on the MCVaR. Concluding remarks are given in Section 2.7.

2.2 Preliminary lemmas about convex optimization problems

In this section, we present preliminary lemmas about convex optimization problems and consider a generalization of loss function (2.1.5) associated with the univariate CVaR.

Definition 2.2.1. Let $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a function.

(i) f is said to be convex (strictly convex) on $\mathbb{R}^d$ if $f(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) \le (<) \lambda f(\mathbf{x})+(1-\lambda)f(\mathbf{y})$ for any $\mathbf{x},\mathbf{y}\in\mathbb{R}^d$ and any $\lambda\in[0,1]$.

(ii) f is said to be coercive on $\mathbb{R}^d$ if $\lim_{\|\mathbf{x}\|_2\to\infty} f(\mathbf{x})=\infty$.

In the following two lemmas, we recall some well-known results (see, for example, Niculescu and Persson (2006)) about convexity, coercivity, and solutions to convex optimization problems.

Lemma 2.2.1. *(i) If $f_i(x_i)$ is a convex (coercive) function of x_i on $\mathbb{R}$, $i=1,...,d$, then $\sum_{i=1}^d f_i(x_i)$ is a convex (coercive) function of $\mathbf{x}=(x_1,...,x_d)$ on $\mathbb{R}^d$.*

(ii) If $g(\mathbf{x})$ is a coercive function of $\mathbf{x}$ on $\mathbb{R}^d$ and $f(\mathbf{x})\ge g(\mathbf{x})$ for all $\mathbf{x}\in\mathbb{R}^d$, then $f(\mathbf{x})$ is also a coercive function of $\mathbf{x}$ on $\mathbb{R}^d$.

Lemma 2.2.2. *(i) If $f:\mathbb{R}^d\longrightarrow\mathbb{R}$ is a coercive and convex function on $\mathbb{R}^d$. Then there exists an element $\mathbf{x}^*=(x_1,...,x_d)\in\mathbb{R}^d$ such that $f(\mathbf{x}^*)=\inf_{\mathbf{x}\in\mathbb{R}^d} f(\mathbf{x})$, and $\mathbf{x}^*=(x_1,...,x_d)$ is a solution to the following system of inequalities:*

$$\frac{\partial^-}{\partial x_i} f(x_1,...,x_d)\le 0\le \frac{\partial^+}{\partial x_i} f(x_1,...,x_d),\quad i=1,...,d. \tag{2.2.1}$$

If, moreover, f is also differentiable, then $\mathbf{x}^*=(x_1,...,x_d)\in\mathbb{R}^d$ such that $f(\mathbf{x}^*)=\inf_{\mathbf{x}\in\mathbb{R}^d} f(\mathbf{x})$ is a solution to the following system of equations:

$$\frac{\partial}{\partial x_i} f(x_1,....,x_d)=0,\quad i=1,...,d. \tag{2.2.2}$$

(ii) If $f:\mathbb{R}^d\longrightarrow\mathbb{R}$ a coercive and strictly convex function on $\mathbb{R}^d$, then there exists a unique element $\mathbf{x}^*\in\mathbb{R}^d$ such that $f(\mathbf{x}^*)=\inf_{\mathbf{x}\in\mathbb{R}^d} f(\mathbf{x})$. In addition, if f also differentiable, then $\mathbf{x}^*=(x_1,...,x_d)\in\mathbb{R}^d$ such that $f(\mathbf{x}^*)=\inf_{\mathbf{x}\in\mathbb{R}^d} f(\mathbf{x})$ is the unique solution to the equations in (2.2.2).

Next, we discuss the properties of a generalization of loss function (2.1.5).

Lemma 2.2.3. *For random variable X with $\mathbb{E}|X|<\infty$ and real numbers a, b, v, define (expected) loss function $l_{X,b,v}(x)$ by*

$$\begin{aligned} l_{X,b,v}(x) &= ax+b\,\mathbb{E}|X-x|+v\,\mathbb{E}(X-x) && (2.2.3)\\ &= ax+\lambda\,\mathbb{E}(X-x)_+-\gamma\,\mathbb{E}(X-x)_-, && (2.2.4)\end{aligned}$$

where $\lambda=b+v$, $\gamma=v-b$, or equivalently, $b=\frac{\lambda-\gamma}{2}$, $v=\frac{\lambda+\gamma}{2}$. The following results hold.

(i) For any $b \geq 0$ and any $a, v \in \mathbb{R}$, $l_{X,b,v}(x)$ is a convex function of x on $\mathbb{R}$.

(ii) If $v - b < a < v + b$, or equivalently, if $\gamma < a < \lambda$, then $l_{X,b,v}(x)$ is coercive and convex on $\mathbb{R}$.

Proof. (i) Since $ax + v\,\mathbb{E}(X - x)$ is a linear function of x on $\mathbb{R}$ and $b\,\mathbb{E}|X - x|$ is a convex function of x on $\mathbb{R}$ for $b \geq 0$. Hence, $l_{X,b,v}(x)$ is a convex function of $x \in \mathbb{R}$ by (2.2.3).
(ii) Note that $|x| = 2(x)_+ - x$ and $(x)_- = (x)_+ - x$. Thus (2.2.3) and (2.2.4) have the equivalent expressions:

$$\begin{aligned} l_{X,b,v}(x) &= (a-\gamma)x + \gamma\mathbb{E}X + (\lambda-\gamma)\mathbb{E}(X-x)_+ \quad (2.2.5)\\ &= (a-\lambda)x + \lambda\,\mathbb{E}X - (\gamma-\lambda)\,\mathbb{E}(X-x)_-. \quad (2.2.6)\end{aligned}$$

By the monotone convergence theorem and $\mathbb{E}|X| < \infty$, we have $\lim_{x\to\infty}\mathbb{E}(X-x)_+ = 0$ and $\lim_{x\to-\infty}\mathbb{E}(X-x)_- = 0$. Thus, it follows from $\gamma < a < \lambda$ that $\lim_{x\to\infty} l_{X,b,v}(x) = \infty$ by (2.2.5) and that $\lim_{x\to-\infty} l_{X,b,v}(x) = \infty$ by (2.2.6). Hence, $l_{X,b,v}(x)$ is coercive. □

Lemma 2.2.4. *For $v - b < 1 < v + b$ or equivalently $\gamma < 1 < \lambda$, and any random variable X with $\mathbb{E}|X| < \infty$, denote mapping $\mathrm{GCVaR}_{b,v}(X)$ from X to $\mathbb{R}$ by*

$$\begin{aligned} \mathrm{GCVaR}_{b,v}(X) &= \min_{c\in\mathbb{R}}\{c + b\,\mathbb{E}|X-c| + v\,\mathbb{E}(X-c)\}\\ &= \min_{c\in\mathbb{R}}\{c + \lambda\,\mathbb{E}(X-c)_+ - \gamma\,\mathbb{E}(X-c)_-\}. \quad (2.2.7)\end{aligned}$$

Then,

$$\mathrm{GCVaR}_{b,v}(X) = \gamma\,\mathbb{E}(X) + (1-\gamma)\,\mathrm{CVaR}_\alpha(X), \quad (2.2.8)$$

where

$$\alpha = \frac{b+v-1}{2b} = \frac{\lambda-1}{\lambda-\gamma}. \quad (2.2.9)$$

Proof. Note that $(x)_- = (x)_+ - x$, we have

$$\begin{aligned} \mathrm{GCVaR}_{b,v}(X) &= \min_{c\in\mathbb{R}}\{(1-\gamma)c + (\lambda-\gamma)\,\mathbb{E}(X-c)_+ + \gamma\,\mathbb{E}X\}\\ &= \gamma\,\mathbb{E}X + (1-\gamma)\min_{c\in\mathbb{R}}\left\{c + \frac{\lambda-\gamma}{1-\gamma}\,\mathbb{E}(X-c)_+\right\}\\ &= \gamma\,\mathbb{E}X + (1-\gamma)\mathrm{CVaR}_\alpha(X),\end{aligned}$$

where the last equality follows from (2.1.4). We point out that (2.2.7) can be also obtained by using (2.19), (2.24), and Example 2.39 of Pflug and Romisch (2007). □

By setting $\gamma < 1 < \lambda$, it implies that we focus more on the shortfall risk. Furthermore, if $\gamma > 0$, the surplus will partially offset the shortfall risk. If $\gamma < 0$, we consider the surplus to be an additional risk.

2.3 A multivariate CVaR considering the systemic risk of a portfolio

Motivated by the multivariate loss functions used in Chaudhuri (1996) and Herrmann et al. (2018), we would extend loss function (2.1.5) for the univariate CVaR to a multivariate loss function considering the systemic risk of a portfolio/company. In doing so, for any $\mathbf{c} = (c_1, ..., c_d) \in \mathbb{R}^d$ and random vector $\mathbf{X} = (X_1, ..., X_d)$, define multivariate loss function $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ by

$$\begin{aligned} f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}) &= \sum_{i=1}^{d} c_i + \lambda\Big(\sum_{i=1}^{d} X_i - \sum_{i=1}^{d} c_i\Big)_+ - \gamma\Big(\sum_{i=1}^{d} X_i - \sum_{i=1}^{d} c_i\Big)_- \\ &\quad + \sum_{i=1}^{d}\Big(\lambda_i\Big(X_i - c_i\Big)_+ - \gamma_i\Big(X_i - c_i\Big)_-\Big) + \beta\|\mathbf{X} - \mathbf{c}\|_2, \qquad (2.3.1) \end{aligned}$$

where λ, γ, β, λ_i, γ_i, $i = 1, ..., d$, are real numbers, and $\mathbf{b} = (\lambda_1, ..., \lambda_d, \gamma_1, ..., \gamma_d)$. Practically, λ and λ_i for $i = 1, ..., d$ should be positive as shortfall risks should always be controlled. At the same time, γ and γ_i for $i = 1, ..., d$ can be both positive and negative as we can treat a surplus to be an additional risk or an offset to the shortfall risks depending on risk managers' preferences.

In this chapter, we denote $S(\mathbf{x})$ by the sum of the components of a vector $\mathbf{x}$, namely, for any $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{R}^n$, $S(\mathbf{x}) = \sum_{i=1}^{n} x_i$.

We first provide an interpretation of the loss function $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ in terms of capital requirements and then define a multivariate CVaR risk measure based on the loss function. In the loss function $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ with expression (2.3.1), c_i can be viewed a base capital for individual risk X_i; $\sum_{i=1}^{d} c_i$ is the total base capital for the aggregate risk $\sum_{i=1}^{d} X_i$ or for the portfolio/company; $\Big(\sum_{i=1}^{d} X_i - \sum_{i=1}^{d} c_i\Big)_+$ is the shortfall risk on the aggregate risk if the total base capital is not sufficient; $\Big(\sum_{i=1}^{d} X_i - \sum_{i=1}^{d} c_i\Big)_-$ is the surplus risk on the aggregate risk if the total base capital is over budgeted; $(X_i - c_i)_+$ is the shortfall risk on risk X_i if the base capital for it is not sufficient; $(X_i - c_i)_-$ is the surplus risk on risk X_i if the base capital for it is over budgeted; and $\|\mathbf{X} - \mathbf{c}\|_2$ is the Euclidean distance between the risk vector $\mathbf{X}$ and the base capital vector $\mathbf{c}$ and represents an overall deviation risk between the risk vector $\mathbf{X}$ and the base capital vector $\mathbf{c}$. In addition, the parameters λ, λ_i, γ, γ_i, and β represent the relative importances of the corresponding shortfall risks, surplus risks, and the overall deviation risk in the portfolio or reflects on decision maker's preferences on the corresponding shortfall risks, surplus risks, and the overall deviation risk in the portfolio.

Hence, the loss function $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ with expression (2.3.1) can be understood in a similar way to the univariate CVaR and it can be viewed as the sum of the total base capital $S(\mathbf{c}) = \sum_{i=1}^{d} c_i$ for the aggregate risk plus the total backup/additional capital $\lambda\Big(S(\mathbf{X}) - S(\mathbf{c})\Big)_+ - \gamma\Big(S(\mathbf{X}) - S(\mathbf{c})\Big)_- + \sum_{i=1}^{d}\Big(\lambda_i\Big(X_i - c_i\Big)_+ - \gamma_i\Big(X_i - c_i\Big)_-\Big) + \beta\|\mathbf{X} - \mathbf{c}\|_2$. To buffer the impact of the shortfall risks, surplus risks, and overall deviation risk in the portfolio, it is reasonable to hold these backup/additional capital. However, holding too much capital reserves for a portfolio or a company would reduce the potential investment of the company. Hence, like the idea behind the univariate CVaR in (2.1.4), we would like to minimize the expected total required capital for the portfolio, namely to minimize $\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$, and then use $\min_{\mathbf{c}\in\mathbb{R}^d}\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ as the risk measure of the aggregate risk $S(\mathbf{X})$ or as the required capital for the aggregate risk $S(\mathbf{X})$. Thus, if there exists a unique $\mathbf{c}^* = (c_1^*, ..., c_d^*) = \arg\min_{\mathbf{c}\in\mathbb{R}^d}\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$, where c_i^* is the base capital for risk X_i or subunit i, then, the risk measure for the aggregate risk $S(\mathbf{X})$ of the portfolio, denoted by $\rho(S(\mathbf{X}))$, is

$$\begin{aligned}\rho(S(\mathbf{X})) &= \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}^*)) = \min_{\mathbf{c}\in\mathbb{R}^d}\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) \\ &= S(\mathbf{c}^*) + \lambda\,\mathbb{E}\Big(S(\mathbf{X}) - S(\mathbf{c}^*)\Big)_+ - \gamma\,\mathbb{E}\Big(S(\mathbf{X}) - S(\mathbf{c}^*)\Big)_- \\ &\quad + \sum_{i=1}^{d}\Big(\lambda_i\mathbb{E}(X_i - c_i^*)_+ - \gamma_i\mathbb{E}(X_i - c_i^*)_-\Big) + \beta\,\mathbb{E}\Big(\|\mathbf{X} - \mathbf{c}^*\|_2\Big). \qquad (2.3.2)\end{aligned}$$

Note that when $d = 1$, (2.3.2) is degenerated to

$$\begin{aligned}& c_1^* + \lambda\,\mathbb{E}\Big(X_1 - c_1^*\Big)_+ - \gamma\,\mathbb{E}\Big(X_1 - c_1^*\Big)_- + \lambda_1\mathbb{E}(X_1 - c_1^*)_+ - \gamma_1\mathbb{E}(X_1 - c_1^*)_- + \beta\,\mathbb{E}|X_1 - c_1^*| \\ = \;& c_1^* + (\lambda_1 + \lambda + \beta)\,\mathbb{E}\Big(X_1 - c_1^*\Big)_+ - (\gamma_1 + \gamma - \beta)\,\mathbb{E}\Big(X_1 - c_1^*\Big)_-,\end{aligned}$$

which means that the required capital for risk X_1 or subunit 1 is the base capital c_1^* plus the backup capital $(\lambda_1 + \lambda + \beta)\,\mathbb{E}(X_1 - c_1^*)_+ - (\gamma_1 + \gamma - \beta)\,\mathbb{E}(X_1 - c_1^*)_-$. Hence, for risk X_i, $i = 1, ..., d$, we can set that the required capital for risk X_i or subunit i is the base capital c_i^* plus the backup capital $(\lambda_i + \lambda + \beta)\,\mathbb{E}(X_i - c_i^*)_+ - (\gamma_i + \gamma - \beta)\,\mathbb{E}(X_i - c_i^*)_-$ under the preference parameter $\lambda_i + \lambda + \beta$ on the expected shortfall risk and the preference parameter $-(\gamma_i + \gamma - \beta)$ on the expected surplus risk. Thus, denote the risk measure of X_i by $\rho_i(\mathbf{X})$, then we can set that

$$\rho_i(\mathbf{X}) = c_i^* + (\lambda_i + \lambda + \beta)\,\mathbb{E}(X_i - c_i^*)_+ - (\gamma_i + \gamma - \beta)\,\mathbb{E}(X_i - c_i^*)_-, \quad i = 1, ..., d. \qquad (2.3.3)$$

Hence, for risk portfolio $\mathbf{X} = (X_1, ..., X_d)$, we could use (2.3.2) as the risk measure of the aggregate risk $S(\mathbf{X})$ and use (2.3.3) as the risk measure of X_i, and thus we obtain a multivariate risk measure $(\rho_1(\mathbf{X}), ..., \rho_d(\mathbf{X}), \rho(S(\mathbf{X})))$ for the risk portfolio $\mathbf{X} = (X_1, ..., X_d)$. In such a multivariate risk measure, the priority is to consider the aggregate risk of a portfolio, which is described in Burgert and Rüschendorf (2006) as a natural idea to take the aggregate risk of a portfolio as a main concern. Note that the real numbers $c_i^* = c_i^*(\mathbf{X})$ in (2.3.2) and (2.3.3), $i = 1, ..., d$, depend on $\mathbf{X}$ or its distribution.

Definition 2.3.1. For random vector $\mathbf{X} = (X_1, ..., X_d)$ and loss function $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ defined in (2.3.1), let $\mathbf{c}^* = (c_1^*, ..., c_d^*) \in \mathbb{R}^d$ be the unique minimizer to optimization problem $\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ if such a unique minimizer exists, namely

$$\mathbf{c}^* = (c_1^*, ..., c_d^*) = \arg\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})). \tag{2.3.4}$$

The multivariate CVaR (MCVaR) risk measure of $\mathbf{X}$, denoted by $\mathrm{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X})$, is a mapping from $\mathbf{X}$ to $\mathbb{R}^{d+1}$ and defined by

$$\mathrm{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X}) = \Big(\rho_1(\mathbf{X}), \ldots, \rho_d(\mathbf{X}), \rho(S(\mathbf{X})\Big), \tag{2.3.5}$$

where $\rho(S(\mathbf{X}))$ is the risk measure of the aggregate risk $S(\mathbf{X})$ defined by (2.3.2), and $\rho_i(\mathbf{X})$ is the risk measure of X_i defined by (2.3.3). □

In multivariate loss function (2.3.1), $\sum_{i=1}^d c_i$, $\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i\Big)_+$, $\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i\Big)_-$ can be viewed as the total capital (premium) provided by subsidiaries of a corporation (insureds of an insurer), the shortfall borne by the corporation (insurer), and the over-required capital to the corporation (the overcharged premium by the insurer), respectively, while $\Big(X_i - c_i\Big)_+$ and $\Big(X_i - c_i\Big)_-$ can be viewed as the shortfall borne by subsidiary (insured) i and the over-required capital (the overcharged premium) to subsidiary (insured) i, respectively. In addition, $\|\mathbf{X} - \mathbf{c}\|_2$ can be viewed as a penalty or cost faced by the corporation (insurer) due to the deviation between risk vector $\mathbf{X}$ and capital (premium) vector $\mathbf{c}$. Therefore, problem (2.3.4) can be seen as the problem a regulator (insurer) wishes to solve for minimizing a weighted sum of the total capital (premium) provided by subsidiaries (insureds), the expected shortfall borne by the corporation (insurer), the expected over-required capital to the corporation (the expected overcharged premium by the insurer), the expected shortfalls borne by subsidiaries, the expected overcharged premiums to insureds, and the expected penalty or cost of the corporation (insurer). This multivariate loss function (2.3.1) considers a set of risks from a systemic view.

Remark 2.3.1. If $\beta = 0$, $\lambda_i = 0$, $\gamma_i = 0$, $i = 1, ..., d$, the loss function $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ defined in (2.3.1) is reduced to

$$f_{\mathbf{X},\lambda,\gamma}(\mathbf{c}) \;=\; \sum_{i=1}^d c_i + \lambda\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i\Big)_+ - \gamma\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i\Big)_-. \tag{2.3.6}$$

Thus, if $\gamma < 1 < \lambda$, by Lemma 2.2.4, the risk measure $\rho(S(\mathbf{X}))$ of the aggregate risk $S(\mathbf{X})$ in the multivariate CVaR risk measure $\mathrm{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X})$ is reduced to

$$\begin{aligned}\rho(S(\mathbf{X})) &= \min_{\mathbf{c}\in\mathbb{R}^d}\Big\{\sum_{i=1}^d c_i + \lambda\mathbb{E}\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i\Big)_+ - \gamma\mathbb{E}\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i\Big)_-\Big\} \\ &= \min_{c\in\mathbb{R}}\Big\{c + \lambda\mathbb{E}\Big(S(\mathbf{X}) - c\Big)_+ - \gamma\mathbb{E}\Big(S(\mathbf{X}) - c\Big)_-\Big\} \\ &= \gamma\,\mathbb{E}(S(\mathbf{X})) + (1-\gamma)\,\mathrm{CVaR}_\alpha(S(\mathbf{X})),\end{aligned}$$

where $\alpha = \frac{\lambda-1}{\lambda-\gamma}$. This means that in this case, the risk measure $\rho(S(\mathbf{X}))$ of the aggregate risk $S(\mathbf{X})$ is a linear combination of the expected value of $S(\mathbf{X})$ and its CVaR at the confidence level $\frac{\lambda-1}{\lambda-\gamma}$. In addition, if $\gamma = 0$, the risk measure $\rho(S(\mathbf{X}))$ of the aggregate risk $S(\mathbf{X})$ is reduced to its CVaR at the confidence level $\frac{\lambda-1}{\lambda}$. By (2.3.6), we see that if the required capital for the aggregate risk of a portfolio or company is determined by CVaR, such a required capital only considers the shortfall risk on the aggregate risk. If $\mathrm{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X})$ is used to determine required capital or premiums of a portfolio, then the systemic risk of the portfolio, including the shortfall and surplus risks from individual risks and the aggregate risk and the overall deviation risk of the portfolio, is taken into consideration. □

2.4 Existence and solution expression of the MCVaR risk measure

In this section, we give sufficient conditions such that the MCVaR defined in Definition 2.3.1 is well defined. Then, we derive the solution expression of the MCVaR under certain conditions. To do so, we first study the properties of the objective function $\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ in (2.3.4).

Note that $x_+ = \frac{1}{2}(|x| + x)$ and $x_- = \frac{1}{2}(|x| - x)$. Thus, $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ defined in (2.3.1) has the following equivalent expression:

$$\begin{aligned}f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}) &= S(\mathbf{c}) + \beta_0\Big|S(\mathbf{X}) - S(\mathbf{c})\Big| + v_0\Big(S(\mathbf{X}) - S(\mathbf{c})\Big) \\ &\quad + \sum_{i=1}^d \beta_i\,|X_i - c_i| + \langle \mathbf{v}, \mathbf{X} - \mathbf{c}\rangle + \beta\|\mathbf{X} - \mathbf{c}\|_2, \end{aligned} \tag{2.4.1}$$

where

$$\beta_0 = \frac{\lambda-\gamma}{2}, \quad v_0 = \frac{\lambda+\gamma}{2}, \quad \beta_i = \frac{\lambda_i - \gamma_i}{2}, \quad v_i = \frac{\lambda_i+\gamma_i}{2}, \quad i = 1, ..., d, \tag{2.4.2}$$

and $\mathbf{v} = (v_1, ..., v_d)$, $\mathbf{u} = (\beta_1, ..., \beta_d, v_1, ..., v_d)$. The relationships in (2.4.2) are equivalent to

$$\lambda = v_0 + \beta_0, \quad \gamma = v_0 - \beta_0, \quad \lambda_i = v_i + \beta_i, \quad \gamma_i = v_i - \beta_i, \quad i = 1, ..., d. \tag{2.4.3}$$

Proposition 2.4.1. *Let $\mathbf{X} = (X_1, ..., X_d)$, $d \geq 2$, be a random vector with finite expectation, namely $\mathbb{E}|X_i| < \infty$, $i = 1, ..., d$, and $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ be the loss function defined in (2.3.1). Then, the following statements hold.*

(i) For any λ, γ, $\mathbf{b}$, and β in the loss function $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$, $\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ is a finite-valued function of $\mathbf{c} \in \mathbb{R}^d$, namely for any $\mathbf{c} \in \mathbb{R}^d$, $\Big|\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))\Big| < \infty$.

(ii) If $\beta \geq 0$, $\lambda \geq \gamma$, $\lambda_i \geq \gamma_i$, $i = 1, ..., d$, then $\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ is a convex function of $\mathbf{c} \in \mathbb{R}^d$.

(iii) If $\beta > 0$, $\lambda \geq \gamma$, $\lambda_i \geq \gamma_i$, $i = 1, ..., d$, and the distribution of $\mathbf{X}$ is not supported on a single straight line, then $\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ is a strictly convex function of $\mathbf{c} \in \mathbb{R}^d$.

(iv) If $\beta \geq 0$, $\lambda \geq \gamma$, $\lambda_i \geq \gamma_i$, $i = 1, ..., d$, and

$$\sum_{i=1}^{d} (\lambda_i + \gamma_i + \lambda + \gamma - 2)^2 < (2\beta + \beta^*)^2, \tag{2.4.4}$$

then $\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ is a coercive function of $\mathbf{c} \in \mathbb{R}^d$, where $\beta^* = \min\{\lambda_1 - \gamma_1, ..., \lambda_d - \gamma_d\}$.

(v) If $\beta \geq 0$, $\lambda \geq 0$, $\gamma = 0$, $\gamma_i < 1 < \lambda_i$, $i = 1, ..., d$, then $\mathbb{E}(f_{\mathbf{X},\lambda,0,\mathbf{b},\beta}(\mathbf{c}))$ is a coercive and convex function of $\mathbf{c} \in \mathbb{R}^d$.

Proof. For any $\mathbf{x} = (x_1, ..., x_d), \mathbf{y}, \mathbf{z} \in \mathbb{R}^d$ and $a, b \in \mathbb{R}$, the following inequalities hold:

$$|a + b| \leq |a| + |b|, \qquad \|\mathbf{x} + \mathbf{y}\|_2 \leq \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2, \qquad |\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|_2 \, \|\mathbf{y}\|_2, \tag{2.4.5}$$

and

$$|a - b| \geq |b| - |a|, \qquad \|\mathbf{x} - \mathbf{y}\|_2 \geq \|\mathbf{y}\|_2 - \|\mathbf{x}\|_2, \qquad \sum_{i=1}^{d} |x_i| \geq \|\mathbf{x}\|_2. \tag{2.4.6}$$

For any $\mathbf{c} = (c_1, ..., c_d)$, $\mathbf{v} = (v_1, ..., v_d) \in \mathbb{R}^d$, and any random vector $\mathbf{X} = (X_1, ..., X_d)$, by (2.4.1), we have

$$\begin{aligned} & f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}) \\ = \; & \sum_{i=1}^{d} (1 - v_i - v_0) c_i + \beta_0 \Big| S(\mathbf{X}) - S(\mathbf{c}) \Big| + \beta ||\mathbf{X} - \mathbf{c}||_2 + \sum_{i=1}^{d} (v_i + v_0) X_i + \sum_{i=1}^{d} \beta_i \, |X_i - c_i| \\ = \; & \langle \mathbf{v}^*, \, \mathbf{c} \rangle + \beta_0 \Big| S(\mathbf{X}) - S(\mathbf{c}) \Big| + \beta ||\mathbf{X} - \mathbf{c}||_2 + \langle \mathbf{w}^*, \, \mathbf{X} \rangle + \sum_{i=1}^{d} \beta_i \, |X_i - c_i|, \end{aligned} \tag{2.4.7}$$

where

$$\mathbf{v}^* = \mathbf{1} - \mathbf{v} - v_0\mathbf{1} = (1-v_0)\mathbf{1} - \mathbf{v}, \qquad \mathbf{w}^* = \mathbf{v} + v_0\mathbf{1}, \tag{2.4.8}$$

and $\mathbf{1} \in \mathbb{R}^d$ is a vector with all components of 1's.

(i) Applying the inequalities in (2.4.5) and (2.4.6) to (2.4.7), we obtain

$$\begin{aligned} |f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})| \;\; &\leq \;\; |\langle \mathbf{v}^*,\, \mathbf{c}\rangle| + |\beta_0| \sum_{i=1}^{n} |X_i| + |\beta_0|\, \Big|S(\mathbf{c})\Big| + |\beta| \sum_{i=1}^{d} |X_i| + |\beta|\, \|\mathbf{c}\|_2 \\ &\quad + \|\mathbf{w}^*\|_2 \sum_{i=1}^{d} |X_i| + \sum_{i=1}^{d} |\beta_i|\, |X_i| + \sum_{i=1}^{d} |\beta_i|\, |c_i|, \end{aligned}$$

which implies $\Big|\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))\Big| \leq \mathbb{E}\Big|f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})\Big| < \infty$ as $\mathbb{E}|X_i| < \infty$, $i = 1, ..., d$.

(ii) Note that $S(\mathbf{c})$, $\langle \mathbf{v}, \mathbf{X} - \mathbf{c}\rangle$, and $v_0\Big(S(\mathbf{X}) - S(\mathbf{c})\Big)$ are linear in $\mathbf{c}$, and thus are convex functions of $\mathbf{c}$ on $\mathbb{R}^d$. In addition, if $\beta \geq 0$, $\lambda \geq \gamma$, $\lambda_i \geq \gamma_i$, $i = 1, ..., d$, or equivalently if $\beta \geq 0$, $\beta_0 \geq 0$, $\beta_i \geq 0$, $i = 1, ..., d$, then $\beta_0\Big|S(\mathbf{X}) - S(\mathbf{c})\Big|$, $\beta||\mathbf{X} - \mathbf{c}||_2$, and $\sum_{i=1}^d \beta_i\, |X_i - c_i|$, are convex functions of $\mathbf{c}$ on $\mathbb{R}^d$ by the triangle inequalities. Hence, by the expression (2.4.1), $\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ is a convex function of $\mathbf{c} \in \mathbb{R}^d$.

(iii) Note that the assumptions that $\beta > 0$, $\lambda \geq \gamma$, $\lambda_i \geq \gamma_i$, $i = 1, ..., d$, are equivalent to $\beta > 0$, $\beta_0 \geq 0$, $\beta_i \geq 0$, $i = 1, ..., d$, By (2.4.1), if $\beta > 0$, then $\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ has the following equivalent expression:

$$\begin{aligned} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) \;\; &= \;\; S(\mathbf{c}) + \beta_0 \mathbb{E}\Big|S(\mathbf{X}) - S(\mathbf{c})\Big| + v_0 \mathbb{E}\Big(S(\mathbf{X}) - S(\mathbf{c})\Big) + \sum_{i=1}^{d} \beta_i\, \mathbb{E}|X_i - c_i| \\ &\quad + \mathbb{E}\Big(\langle \mathbf{v},\, \mathbf{X} - \mathbf{c}\rangle\Big) + \beta \mathbb{E}\|\mathbf{X} - \mathbf{c}\|_2. \end{aligned}$$

Note that functions $S(\mathbf{c})$, $v_0\mathbb{E}\Big(S(\mathbf{X}) - S(\mathbf{c})\Big)$, and $\mathbb{E}\Big(\langle \mathbf{v},\, \mathbf{X} - \mathbf{c}\rangle\Big)$ are linear in $\mathbf{c} \in \mathbb{R}^d$ and that $\beta_0\mathbb{E}\Big|S(\mathbf{X}) - S(\mathbf{c})\Big|$ and $\sum_{i=1}^d \beta_i\, \mathbb{E}|X_i - c_i|$ are convex in $\mathbf{c} \in \mathbb{R}^d$ if $\beta_0 \geq 0$ and $\beta_i \geq 0$, $i = 1, ..., d$. In addition, it follows from Theorem 2.17 of Kemperman (1987) that $\mathbb{E}\|\mathbf{X} - \mathbf{c}\|_2$ is a strictly convex function of $\mathbf{c} \in \mathbb{R}^d$ if $d \geq 2$ and the distribution of $\mathbf{X}$ is not supported on a single straight line. Thus, $\beta\, \mathbb{E}\|\mathbf{X} - \mathbf{c}\|_2$ is a strictly convex function of $\mathbf{c} \in \mathbb{R}^d$ if $\beta > 0$. Hence, $\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ is a strictly convex function of $\mathbf{c} \in \mathbb{R}^d$.

(iv) If $\beta \geq 0$, $\lambda \geq \gamma$, $\lambda_i \geq \gamma_i$, $i = 1, ..., d$, that is, $\beta \geq 0$, $\beta_0 \geq 0$, $\beta_i \geq 0$, $i = 1, ..., d$, applying the inequalities in (2.4.6) to (2.4.7) and noting that $|S(\mathbf{X}) - S(\mathbf{c})| \geq 0$ and

$\beta^* = \min\{\lambda_1 - \gamma_1, ..., \lambda_d - \gamma_d\} = 2\min\{\beta_1, ..., \beta_d\}$, we have

$$\begin{aligned} f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}) &\geq (\beta - \|\mathbf{v}^*\|_2)\,\|\mathbf{c}\|_2 - \beta\|\mathbf{X}\|_2 + \langle \mathbf{w}^*,\, \mathbf{X}\rangle + \frac{\beta^*}{2}\sum_{i=1}^d |c_i| - \sum_{i=1}^d \beta_i\,|X_i| \\ &\geq (\beta - \|\mathbf{v}^*\|_2)\,\|\mathbf{c}\|_2 - \beta\|\mathbf{X}\|_2 + \langle \mathbf{w}^*,\, \mathbf{X}\rangle + \frac{\beta^*}{2}\,\|\mathbf{c}\|_2 - \sum_{i=1}^d \beta_i\,|X_i| \\ &= \Big(\beta + \frac{\beta^*}{2} - \|\mathbf{v}^*\|_2\Big)\,\|\mathbf{c}\|_2 - \beta\|\mathbf{X}\|_2 + \langle \mathbf{w}^*,\, \mathbf{X}\rangle - \sum_{i=1}^d \beta_i\,|X_i|. \end{aligned}$$

Hence, we have

$$\begin{aligned} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) &\geq \Big(\beta + \frac{\beta^*}{2} - \|\mathbf{v}^*\|_2\Big)\,\|\mathbf{c}\|_2 - \beta\mathbb{E}\|\mathbf{X}\|_2 \\ &\quad + \mathbb{E}\Big(\langle \mathbf{w}^*,\, \mathbf{X}\rangle\Big) - \sum_{i=1}^d \beta_i\,\mathbb{E}|X_i|. \end{aligned} \tag{2.4.9}$$

Note that $\mathbb{E}\|\mathbf{X}\|_2$, $\mathbb{E}\Big(\langle \mathbf{w}^*,\, \mathbf{X}\rangle\Big)$, $\mathbb{E}|X_i|$, $i = 1, ..., d$, are all finite. Thus, if $\beta + \frac{\beta^*}{2} - \|\mathbf{v}^*\|_2 > 0$ or equivalently (2.4.4) holds, then (2.4.9) implies

$$\lim_{\|\mathbf{c}\|_2 \to \infty} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) = \infty,$$

which means that $\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ is coercive on $\mathbf{c} \in \mathbb{R}^d$.

(v) If $\beta \geq 0$, $\lambda \geq 0$, $\gamma = 0$, by (2.3.1), we have

$$\begin{aligned} \mathbb{E}(f_{\mathbf{X},\lambda,0,\mathbf{b},\beta}(\mathbf{c})) &= \sum_{i=1}^d \Big(c_i + \lambda_i \mathbb{E}(X_i - c_i)_+ - \gamma_i \mathbb{E}(X_i - c_i)_-\Big) \\ &\quad + \lambda\,\mathbb{E}\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i\Big)_+ + \beta\,\mathbb{E}\|\mathbf{X} - \mathbf{c}\|_2 \qquad (2.4.10) \\ &\geq \sum_{i=1}^d \Big(c_i + \lambda_i \mathbb{E}(X_i - c_i)_+ - \gamma_i \mathbb{E}(X_i - c_i)_-\Big). \qquad (2.4.11) \end{aligned}$$

Further, if $\gamma_i < 1 < \lambda_i$, $i = 1, ..., d$, by Lemma 2.2.3(ii), we have that $c_i + \lambda_i \mathbb{E}(X_i - c_i)_+ - \gamma_i \mathbb{E}(X_i - c_i)_-$ is a coercive and convex function of $c_i \in \mathbb{R}$, and thus, by Lemma 2.2.1(i), $\sum_{i=1}^d \Big(c_i + \lambda_i \mathbb{E}(X_i - c_i)_+ - \gamma_i \mathbb{E}(X_i - c_i)_-\Big)$ is a coercive and convex function of $\mathbf{c} \in \mathbb{R}^d$. Hence, by (2.4.11) and Lemma 2.2.1(ii), we have that $\mathbb{E}(f_{\mathbf{X},\lambda,0,\mathbf{b},\beta}(\mathbf{c}))$ is a coercive function of $\mathbf{c} \in \mathbb{R}^d$. In addition, by (2.4.10), $\mathbb{E}(f_{\mathbf{X},\lambda,0,\mathbf{b},\beta}(\mathbf{c}))$ is a convex function of $\mathbf{c} \in \mathbb{R}^d$ as both $\lambda\mathbb{E}\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i\Big)_+$ and $\beta\mathbb{E}\|\mathbf{X} - \mathbf{c}\|_2$ are convex functions of $\mathbf{c} \in \mathbb{R}^d$. □

Now, we make the following assumption, which gives sufficient conditions such that the new MCVaR risk measure in Definition 2.3.1 is well defined.

Assumption 2.4.1. For random vector $\mathbf{X}$ with finite expectation and loss function $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ defined in (2.3.1), assume that the following conditions hold for pair $(\mathbf{X},\ f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$:

(i) $\mathbf{X}$ has a continuous joint distribution function with a joint density function,

(ii) and that $\beta > 0$, $\lambda \geq \gamma$, $\lambda_i \geq \gamma_i$, $i = 1, ..., d$, and

$$\sum_{i=1}^{d}(\lambda_i + \gamma_i + \lambda + \gamma - 2)^2 < (2\beta + \beta^*)^2, \tag{2.4.12}$$

where $\beta^* = \min\{\lambda_1 - \gamma_1, ..., \lambda_d - \gamma_d\}$. □

Remark 2.4.2. Under Assumption 2.4.1, by Proposition 2.4.1(iii)-(iv), we know that $\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ is a coercive and strictly convex function of $\mathbf{c}$ on $\mathbb{R}^d$. Thus, by Lemma 2.2.2(ii), there exists a unique $\mathbf{c}^* = (c_1^*, ..., c_d^*) \in \mathbb{R}^d$ such that

$$\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) = \min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})). \tag{2.4.13}$$

Hence, Assumption 2.4.1 gives sufficient conditions such that the multivariate CVaR risk measure $\mathrm{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X})$ in Definition 2.3.1 is well defined. We also point out that the assumption that $\mathbf{X}$ has a joint density or has the absolute continuous distribution is also assumed in Chaudhuri (1996) to guarantee that the multivariate geometric quantile/VaR $\mathrm{MVaR}_{\mathbf{u}}(\mathbf{X})$ is well defined. □

Theorem 2.4.3. *Under Assumption 2.4.1, there exists a unique $\mathbf{c}^* = (c_1^*, ..., c_d^*) \in \mathbb{R}^d$ such that (2.4.13) holds, and moreover, $\mathbf{c}^* = (c_1^*, ..., c_d^*)$ is the unique solution to the following system of equations:*

$$\beta\, \mathbb{E}\left(\frac{X_i - c_i}{||\mathbf{X} - \mathbf{c}||_2}\right) = 1 - \lambda - \lambda_i + (\lambda_i - \gamma_i)F_{X_i}(c_i) + (\lambda - \gamma)F_{S(\mathbf{X})}\left(\sum_{i=1}^{d} c_i\right), \quad i = 1, ..., d. \tag{2.4.14}$$

Proof. Under Assumption 2.4.1, by Remark 2.4.2, there exists a unique $\mathbf{c}^* = (c_1^*, ..., c_d^*) \in \mathbb{R}^d$ such that (2.4.13) holds. By (2.3.1), we have

$$\begin{aligned}
&\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) \\
= \ &\sum_{i=1}^{d} c_i + \lambda\mathbb{E}\left(\sum_{i=1}^{d} X_i - \sum_{i=1}^{d} c_i\right)_+ - \gamma\mathbb{E}\left(\sum_{i=1}^{d} X_i - \sum_{i=1}^{d} c_i\right)_- + \ \beta\, \mathbb{E}\left(\sum_{i=1}^{d}(X_i - c_i)^2\right)^{\frac{1}{2}} \\
&+ \sum_{i=1}^{d} \lambda_i\mathbb{E}(X_i - c_i)_+ - \sum_{i=1}^{d} \gamma_i\mathbb{E}(X_i - c_i)_-.
\end{aligned} \tag{2.4.15}$$

Let $g_i(c_i) = \mathbb{E}(X_i - c_i)_+ = \int_{c_i}^{\infty} \bar{F}_X(x)dx$, $h_i(c_i) = \mathbb{E}(X_i - c_i)_- = \int_{-\infty}^{c_i} F_X(x)dx$, and

$$g(c_1, ..., c_d) = \mathbb{E}\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i\Big)_+ = \int_{\sum_{i=1}^d c_i}^{\infty} \bar{F}_{S(\mathbf{X})}(x)dx,$$
$$h(c_1, ..., c_d) = \mathbb{E}\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i\Big)_- = \int_{-\infty}^{\sum_{i=1}^d c_i} F_{S(\mathbf{X})}(x)dx.$$

Under Assumption 2.4.1(i), $g(c_1, ..., c_d)$, $h(c_1, ..., c_d)$, $g_i(c_i)$, $h_i(c_i)$ are differentiable with

$$\frac{\partial}{\partial c_i} g(c_1, ..., c_d) = -\bar{F}_{S(\mathbf{X})}\Big(\sum_{i=1}^d c_i\Big), \qquad \frac{\partial}{\partial c_i} h(c_1, ..., c_d) = F_{S(\mathbf{X})}\Big(\sum_{i=1}^d c_i\Big),$$
$$\frac{\partial}{\partial c_i} g_i(c_i) = -\bar{F}_{X_i}(c_i), \qquad \frac{\partial}{\partial c_i} h_i(c_i) = F_{X_i}(c_i).$$

Under Assumption 2.4.1(i), it is pointed out in Chaudhuri (1996) that $\mathbb{E}\Big(\sum_{i=1}^d (X_i - c_i)^2\Big)^{\frac{1}{2}}$ is differentiable with

$$\frac{\partial}{\partial c_i} \mathbb{E}\Big(\sum_{i=1}^d (X_i - c_i)^2\Big)^{\frac{1}{2}} = \frac{1}{2} \times \mathbb{E}\Big[\Big(\sum_{i=1}^d (X_i - c_i)^2\Big)^{-\frac{1}{2}} \times 2(X_i - c_i)(-1)\Big].$$

Thus, $\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ is differentiable with

$$\begin{aligned} & \frac{\partial}{\partial c_i} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) \\ = \; & 1 - \lambda_i \bar{F}_{X_i}(c_i) - \gamma_i F_{X_i}(c_i) - \lambda \bar{F}_{S(\mathbf{X})}\Big(\sum_{i=1}^d c_i\Big) - \gamma F_{S(\mathbf{X})}\Big(\sum_{i=1}^d c_i\Big) - \beta \mathbb{E}\Big(\frac{X_i - c_i}{||\mathbf{X} - \mathbf{c}||_2}\Big). \end{aligned}$$

By Lemma 2.2.2(ii), $\mathbf{c}^* = (c_1^*, ..., c_d^*)$ is the unique solution to equations $\frac{\partial}{\partial c_i}\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) = 0$, $i = 1, ..., d$, thus, we obtain

$$\begin{aligned} \beta\, \mathbb{E}\Big(\frac{X_i - c_i}{||\mathbf{X} - \mathbf{c}||_2}\Big) &= 1 - \lambda_i \bar{F}_{X_i}(c_i) - \gamma_i F_{X_i}(c_i) - \lambda \bar{F}_{S(\mathbf{X})}\Big(\sum_{i=1}^d c_i\Big) - \gamma F_{S(\mathbf{X})}\Big(\sum_{i=1}^d c_i\Big) \\ &= 1 - \lambda - \lambda_i + (\lambda_i - \gamma_i) F_{X_i}(c_i) + (\lambda - \gamma) F_{S(\mathbf{X})}\Big(\sum_{i=1}^d c_i\Big), \quad i = 1, ..., d, \end{aligned}$$

which yields (2.4.14). □

Remark 2.4.4. In Theorem 2.4.3, if $\lambda = \gamma$, $\lambda_i = \gamma_i$, $i = 1, ..., d$, the equations in (2.4.14) are reduced to

$$\mathbb{E}\Big(\frac{X_i - c_i}{||\mathbf{X} - \mathbf{c}||_2}\Big) = -\frac{\lambda + \lambda_i - 1}{\beta}, \quad i = 1, ..., d,$$

which are equivalent to those in (2.1.9) with $\mathbf{u} = (\frac{\lambda+\lambda_1-1}{\beta}, ..., \frac{\lambda+\lambda_d-1}{\beta})$. Thus, in this case, the minimizer $\mathbf{c}^*$ defined in (2.3.4) is reduced to the multivariate geometric quantile $\mathrm{MVaR}_{\mathbf{u}}(\mathbf{X})$, namely $c_i^* = \mathrm{MVaR}_i(\mathbf{X})$, $i = 1, ..., d$. The MCVaR risk measure of $\mathbf{X}$ is $(\rho_1(\mathbf{X}), ..., \rho_d(\mathbf{X}), \rho(S(\mathbf{X})))$ with the following expression:

$$\begin{aligned}
\rho_i(\mathbf{X}) &= (\lambda+\lambda_i)\,\mathbb{E}(X_i) + (1-\lambda-\lambda_i)\,\mathrm{MVaR}_i(\mathbf{X}) \\
&\quad + \beta\,\mathbb{E}\Big|X_i - \mathrm{MVaR}_i(\mathbf{X})\Big|, \qquad (2.4.16)\\
\rho(S(\mathbf{X})) &= \sum_{i=1}^{d}\Big((\lambda+\lambda_i)\,\mathbb{E}(X_i) + (1-\lambda-\lambda_i)\,\mathrm{MVaR}_i(\mathbf{X})\Big) \\
&\quad + \beta\,\mathbb{E}\Big(\sum_{i=1}^{d}\Big(X_i - \mathrm{MVaR}_i(\mathbf{X})\Big)^2\Big)^{\frac{1}{2}}. \qquad (2.4.17)
\end{aligned}$$

The expression (2.4.16) means that the required capital for risk X_i is a linear combination of the expected value of X_i and its marginal VaR in the portfolio plus a loading that is proportional to the expected absolute deviation between X_i and its marginal VaR in the portfolio. In addition, the expression (2.4.17) means that the required capital for the aggregate risk $S(\mathbf{X})$ is the total of all linear combinations of an individual risk and its marginal VaR in the portfolio plus a loading that is proportional to the expected Euclidean distance between the risk portfolio $\mathbf{X}$ and its MVaR. □

Next, we consider a special case of the loss function defined in (2.3.1) when $\lambda > 0$, $\gamma = 0$, $\beta = 0$, $\lambda_i > 1$, $\gamma_i = 0$, $i = 1, ..., d$. In this case, by (2.3.1), the loss function $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ is reduced to

$$f_{\mathbf{X},\lambda,\mathbf{\Lambda}}(\mathbf{c}) = \sum_{i=1}^{d} c_i + \lambda\Big(\sum_{i=1}^{d} X_i - \sum_{i=1}^{d} c_i\Big)_+ + \sum_{i=1}^{d}\lambda_i(X_i - c_i)_+, \qquad (2.4.18)$$

where $\mathbf{\Lambda} = (\lambda_1, ..., \lambda_d)$. The loss function $f_{\mathbf{X},\lambda,\mathbf{\Lambda}}(\mathbf{c})$ means that the total required capital for the aggregate risk is equal to the total base capital plus the additional capital on the shortfall risks of the aggregate risk and individual risks. This form (2.4.18) of multivariate loss function is similar to (2.1.5), in which the decision maker is concerned only about shortfall risks. By Proposition 2.4.1(v), $\mathbb{E}(f_{\mathbf{X},\lambda,\mathbf{\Lambda}}(\mathbf{c}))$ is a coercive and convex function of $\mathbf{c} \in \mathbb{R}^d$. Thus, by Lemma 2.2.2(i), there exists $\mathbf{c}^* \in \mathbb{R}^d$ such that

$$\mathbb{E}(f_{\mathbf{X},\lambda,\mathbf{\Lambda}}(\mathbf{c}^*)) = \min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\mathbf{\Lambda}}(\mathbf{c})). \qquad (2.4.19)$$

In the following theorem, we show under some conditions, there exists unique $\mathbf{c}^* \in \mathbb{R}^d$ such that (2.4.19) holds. Moreover, we give the explicit solution to the unique $\mathbf{c}^*$ satisfying (2.4.19).

Assumption 2.4.2. For random vector $\mathbf{X} = (X_1, ..., X_d)$ with finite expectation and loss function $f_{\mathbf{X},\lambda,\mathbf{\Lambda}}(\mathbf{c})$ defined in (2.4.18), assume that the following conditions hold for pair $(\mathbf{X},\ f_{\mathbf{X},\lambda,\mathbf{\Lambda}}(\mathbf{c}))$:

(i) Distribution functions $F_{S(\mathbf{X})}(x)$, $F_{X_i}(x)$, $i = 1, ..., d$, are continuous on $\mathbb{R}$, the left-continuous inverse functions $F_{X_i}^{-1}(q)$ of F_{X_i} are continuous on $(0, 1)$ with $\lim_{q\to 1} F_{X_i}^{-1}(q) = \infty$, $i = 1, ..., d$,

(ii) and that $\lambda > 1$, $\gamma = 0$, $\beta = 0$, $\lambda_i > 1$, $\gamma_i = 0$, $i = 1, ..., d$.

Theorem 2.4.5. *Under Assumption 2.4.2, there exists a unique $\mathbf{c}^* = (c_1^*, ..., c_d^*) \in \mathbb{R}^d$ such that (2.4.19) holds, and $\mathbf{c}^*$ has the following solution*

$$c_i^* = F_{X_i}^{-1}(H_{i,S(\mathbf{X})}(x_0)), \qquad i = 1, ..., d, \tag{2.4.20}$$

where $0 < H_{i,S(\mathbf{X})}(x_0) < 1$, $i = 1, ..., d$,

$$H_{i,S(\mathbf{X})}(x) = \frac{\lambda_i + \lambda - 1 - \lambda\, F_{S(\mathbf{X})}(x)}{\lambda_i} = \frac{\lambda_i - 1 + \lambda\, \bar{F}_{S(\mathbf{X})}(x)}{\lambda_i}, \tag{2.4.21}$$

and $x_0 = x_0(\mathbf{X}, \lambda, \mathbf{\Lambda})$ is the unique solution to equation

$$x - \sum_{i=1}^{d} F_{X_i}^{-1}\Big(H_{i,S(\mathbf{X})}(x)\Big) = 0, \qquad x > F_{S(\mathbf{X})}^{-1}\Big(\frac{\lambda - 1}{\lambda}\Big). \tag{2.4.22}$$

Proof. Under Assumption 2.4.2, by Proposition 2.4.1(v), we know that $h(\mathbf{c}) = \mathbb{E}(f_{\mathbf{X},\lambda,\mathbf{\Lambda}}(\mathbf{c}))$ is a coercive and convex function of $\mathbf{c} \in \mathbb{R}^d$. Thus, by Lemma 2.2.2(i), there exists $\mathbf{c}^* \in \mathbb{R}^d$ such that $h(\mathbf{c}^*) = \min_{\mathbf{c}\in\mathbb{R}^d} h(\mathbf{c})$. In this case, by (2.4.18),

$$h(\mathbf{c}) = \mathbb{E}(f_{\mathbf{X},\lambda,\mathbf{\Lambda}}(\mathbf{c})) = \sum_{i=1}^{d} c_i + \lambda\mathbb{E}\Big(\sum_{i=1}^{d} X_i - \sum_{i=1}^{d} c_i\Big)_+ + \sum_{i=1}^{d} \lambda_i\mathbb{E}\Big(X_i - c_i\Big)_+.$$

Under Assumption 2.4.2(i), $h(\mathbf{c})$ is differentiable. Thus, by Lemma 2.2.2(i), we know that $\mathbf{c}^* = (c_1^*, ..., c_d^*)$ is a solution to the system of equations

$$\frac{\partial}{\partial c_i} h(c_1, ..., c_d) = 0, \quad i = 1, ..., d. \tag{2.4.23}$$

By taking partial derivatives of $h(\mathbf{c})$ with respect to c_i, we easily see that the equations in (2.4.23) are reduced to

$$1 - \lambda - \lambda_i + \lambda_i F_{X_i}(c_i) + \lambda F_{S(\mathbf{X})}\Big(\sum_{i=1}^{d} c_i\Big) = 0, \qquad i = 1, ..., d,$$

which is equivalent to

$$
\begin{aligned}
F_{X_i}(c_i) &= \frac{\lambda_i + \lambda - 1 - \lambda\, F_{S(\mathbf{X})}\Big(\sum_{i=1}^d c_i\Big)}{\lambda_i} = \frac{\lambda_i - 1 + \lambda\, \bar{F}_{S(\mathbf{X})}\big(\sum_{i=1}^d c_i\big)}{\lambda_i} \\
&= H_{i,S(\mathbf{X})}\Big(\sum_{i=1}^d c_i\Big), \qquad i = 1, ..., d, \qquad (2.4.24)
\end{aligned}
$$

where function $H_{i,S(\mathbf{X})}(x)$ is defined in (2.4.21). Under the assumptions of $\lambda > 1$ and $\lambda_i > 1$, $i = 1, ..., d$, we have that $0 < \frac{\lambda-1}{\lambda} < 1$, $H_{i,S(\mathbf{X})}(x) > 0$, $i = 1, ..., d$, and that

$$
0 < H_{i,S(\mathbf{X})}(x) < 1,\; i = 1, ..., d \iff F_{S(\mathbf{X})}(x) > \frac{\lambda - 1}{\lambda} \iff x > F^{-1}_{S(\mathbf{X})}\Big(\frac{\lambda-1}{\lambda}\Big) \stackrel{\text{def}}{=} x^*,
$$

which means that if $\sum_{i=1}^d c_i > x^*$, then

$$
0 < H_{i,S(\mathbf{X})}\Big(\sum_{i=1}^d c_i\Big) < 1, \quad i = 1, ..., d. \qquad (2.4.25)
$$

Hence, if $\sum_{i=1}^d c_i > x^*$, by (2.4.24), we have

$$
c_i = F^{-1}_{X_i}\Big(H_{i,S(\mathbf{X})}\Big(\sum_{i=1}^d c_i\Big)\Big), \quad i = 1, ..., d, \qquad (2.4.26)
$$

which imply that

$$
\sum_{i=1}^d c_i = \sum_{i=1}^d F^{-1}_{X_i}\Big(H_{i,S(\mathbf{X})}\Big(\sum_{i=1}^d c_i\Big)\Big). \qquad (2.4.27)
$$

Let $g(x) = x - \sum_{i=1}^d F^{-1}_{X_i}\Big(H_{i,S(\mathbf{X})}(x)\Big)$ for $x \in (x^*, \infty)$. Note that $H_{i,S(\mathbf{X})}(x)$ is decreasing in x and $F^{-1}_{X_i}(x)$ is increasing for $i = 1, ..., d$, hence, $g(x)$ is strictly increasing as function x is strictly increasing. In addition, $g(x)$ is also continuous under Assumption 2.4.2(i). For any $i = 1, ..., d$,

$$
\lim_{x\to\infty} H_{i,S(\mathbf{X})}(x) = \frac{\lambda_i - 1}{\lambda_i}, \qquad \lim_{x\downarrow x^*} H_{i,S(\mathbf{X})}(x) = \frac{\lambda + \lambda_i - 1 - \lambda \times \frac{\lambda-1}{\lambda}}{\lambda_i} = 1.
$$

Hence, $\lim_{x\downarrow x^*} F^{-1}_{X_i}\Big(H_{i,S(\mathbf{X})}(x)\Big) = \infty$, $i = 1, ..., d$, which, together with $-\infty < x^* < \infty$, implies $\lim_{x\downarrow x^*} g(x) = -\infty$. Obviously, we have $\lim_{x\to\infty} g(x) = \infty$. Thus, there exists a unique $x_0 \in (x^*, \infty)$ such that $g(x_0) = 0$, which, together with (2.4.27), implies $\sum_{i=1}^d c_i = x_0$. Hence, by (2.4.26), we have (2.4.20). In addition, by (2.4.25), we have $0 < H_{i,S(\mathbf{X})}(x_0) < 1$, $i = 1, ..., d$. □

Remark 2.4.6. Under Assumption 2.4.2 and the notations in Theorem 2.4.5, denote $q_i = q_i(\mathbf{X}, \lambda, \mathbf{\Lambda}) = H_{i,S(\mathbf{X})}(x_0)$, $i = 1, ..., d$, and the multivariate CVaR risk measure $\text{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X})$ of $\mathbf{X}$ by $\text{MCVaR}_{\lambda,\mathbf{\Lambda}}(\mathbf{X}) = (\rho_1(\mathbf{X}), ..., \rho_d(\mathbf{X}), \rho(S(\mathbf{X})))$, then by (2.4.20), (2.3.3), (2.3.2), we have

$$\rho_i(\mathbf{X}) = F_{X_i}^{-1}(q_i) + (\lambda_i + \lambda)\,\mathbb{E}\Big(X_i - F_{X_i}^{-1}(q_i)\Big)_+, \qquad i = 1, ..., d, \tag{2.4.28}$$

$$\begin{aligned} \rho(S(\mathbf{X})) &= \sum_{i=1}^{d} F_{X_i}^{-1}(q_i) + \lambda\,\mathbb{E}\Big(\sum_{i=1}^{d} X_i - \sum_{i=1}^{d} F_{X_i}^{-1}(q_i)\Big)_+ \\ &\quad + \sum_{i=1}^{d} \lambda_i\,\mathbb{E}\Big(X_i - F_{X_i}^{-1}(q_i)\Big)_+. \end{aligned} \tag{2.4.29}$$

Expression (2.4.28) means that if the base capital for each individual risk is its VaR at confidence level q_i, then the required capital for the individual risk is the VaR plus an additional capital which is proportional to the expected shortfall risk at the preference parameter $\lambda_i + \lambda$. We point out that the confidence level q_i depends on the whole risk portfolio $\mathbf{X}$. In addition, expression (2.4.29) means the required capital for the aggregate risk is the total VaRs of individual risks at their own confidence levels plus the total additional capital for the shortfall risks from the aggregate risk and individual risks. □

2.5 Properties of the MCVaR risk measure

In this section, we discuss the properties of the new MCVaR. For risk vector $\mathbf{X} = (X_1, ..., X_d)$ with $\text{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X}) = (\rho_1(\mathbf{X}), ..., \rho_d(\mathbf{X}), \rho(S(\mathbf{X})))$, the risk measure $\rho(S(\mathbf{X}))$ of the aggregate risk $S(\mathbf{X})$, and the risk measure $\rho_i(\mathbf{X})$ of individual risk X_i, $i = 1, ..., d$, are all depending on the whole portfolio risk $\mathbf{X} = (X_1, ..., X_d)$. The priority in our MCVaR is the risk measure of the aggregate risk. In the following propositions, we show that under some conditions, $\text{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X})$ satisfies positive homogeneity, translation invariance, subadditivity, and monotonicity.

Proposition 2.5.1. *(Positive homogeneity) For random vector $\mathbf{X}$ and loss function $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ defined in (2.3.1), assume that minimization problem $\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ has a unique minimizer. Then for any $a > 0$,*

$$\text{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(a\mathbf{X}) = a\,\text{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X}), \tag{2.5.1}$$

which means that $\rho_i(a\mathbf{X}) = a\rho_i(\mathbf{X})$, $i = 1, ..., d$, and $\rho(S(a\mathbf{X})) = \rho(aS(\mathbf{X})) = a\rho(S(\mathbf{X}))$.

Proof. Denote by $\mathbf{c}^*(\mathbf{X}) = (c_1^*(\mathbf{X}), ..., c_d^*(\mathbf{X}))$ the unique minimizer of $\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$, that is,

$$\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}^*(\mathbf{X}))) = \min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})).$$

Note that $\mathbb{E}(f_{a\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(a\mathbf{c})) = a\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ for any $a > 0$ and $\mathbf{c} \in \mathbb{R}^d$ and that $\mathbf{c} \in \mathbb{R}^d \iff a\mathbf{c} \in \mathbb{R}^d$. Hence,

$$\begin{aligned} \min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{a\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) &= \min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{a\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(a\mathbf{c})) = a \min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) \\ &= a\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}^*(\mathbf{X}))) = \mathbb{E}(f_{a\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(a\mathbf{c}^*(\mathbf{X}))), \end{aligned}$$

which means that minimization problem $\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{a\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ has a unique minimizer $a\mathbf{c}^*(\mathbf{X})$. Thus, by (2.3.2) and (2.3.3), we can easily verify that for $i = 1, \ldots, d$

$$\begin{aligned} \rho_i(a\mathbf{X}) &= ac_i^*(\mathbf{X}) + (\lambda_i + \lambda + \beta)\, \mathbb{E}(aX_i - ac_i^*(\mathbf{X}))_+ - (\gamma_i + \gamma - \beta)\, \mathbb{E}(aX_i - ac_i^*(\mathbf{X}))_- \\ &= a\rho_i(\mathbf{X}) \end{aligned}$$

and

$$\begin{aligned} \rho(S(a\mathbf{X})) &= S(a\mathbf{c}^*(\mathbf{X})) + \lambda\, \mathbb{E}\Big(S(a\mathbf{X}) - S(a\mathbf{c}^*(\mathbf{X}))\Big)_+ - \gamma\, \mathbb{E}\Big(aS(\mathbf{X}) - aS(\mathbf{c}^*(\mathbf{X}))\Big)_- \\ &\quad + \sum_{i=1}^{d} \Big(\lambda_i \mathbb{E}(aX_i - ac_i^*(\mathbf{X}))_+ - \gamma_i \mathbb{E}(aX_i - ac_i^*(\mathbf{X}))_-\Big) + \beta\, \mathbb{E}\Big(\|a(\mathbf{X} - \mathbf{c}^*(\mathbf{X}))\|_2\Big) \\ &= a\rho(S(\mathbf{X})). \end{aligned}$$

□

Positive homogeneity of MCVaR provides a similar concept as positive homogeneity of a univariate risk measure, however, positive homogeneity of MCVaR is in terms of the whole business entity. For instance, if the risk size doubles for all risk units within the business entity, the risk measures for all units and the aggregate risk will double.

Proposition 2.5.2. *(Translation invariance) For random vector* $\mathbf{X}$ *and loss function* $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ *defined in* (2.3.1)*, assume that minimization problem* $\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ *has a unique minimizer. Then, for any real vector* $\mathbf{a} = (a_1, \ldots, a_d) \in \mathbb{R}^d$,

$$\text{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X} + \mathbf{a}) = \text{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X}) + (\mathbf{a}, S(\mathbf{a})), \tag{2.5.2}$$

which means that $\rho_i(\mathbf{X}+\mathbf{a}) = \rho_i(\mathbf{X}) + a_i$, $i = 1, ..., d$, *and* $\rho(S(\mathbf{X}+\mathbf{a})) = \rho(S(\mathbf{X}) + S(\mathbf{a})) = \rho(S(\mathbf{X})) + S(\mathbf{a})$.

Proof. Denote by $\mathbf{c}^*(\mathbf{X}) = (c_1^*(\mathbf{X}), ..., c_d^*(\mathbf{X}))$ the unique minimizer of $\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$, then

$$\mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}^*(\mathbf{X}))) = \min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})).$$

Note that $\mathbb{E}(f_{\mathbf{X}+\mathbf{a},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}+\mathbf{a})) = \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ for any $\mathbf{a}, \mathbf{c} \in \mathbb{R}^d$ and that $\mathbf{c} \in \mathbb{R}^d \Longleftrightarrow \mathbf{c} + \mathbf{a} \in \mathbb{R}^d$. Hence,

$$\begin{aligned}
\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X}+\mathbf{a},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) &= \min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X}+\mathbf{a},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}+\mathbf{a})) = \min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) \\
&= \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}^*(\mathbf{X}))) = \mathbb{E}(f_{\mathbf{X}+\mathbf{a},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}^*(\mathbf{X})+\mathbf{a})),
\end{aligned}$$

which means that minimization problem $\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X}+\mathbf{a},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ has a unique minimizer $c^*(\mathbf{X}) + \mathbf{a}$. Thus, by (2.3.2) and (2.3.3), for $i = 1, \ldots, d$

$$\begin{aligned}
\rho_i(\mathbf{X}+\mathbf{a}) &= c_i^*(\mathbf{X}) + a_i + (\lambda_i + \lambda + \beta)\,\mathbb{E}(X_i - c_i^*(\mathbf{X}))_+ - (\gamma_i + \gamma - \beta)\,\mathbb{E}(X_i - c_i^*(\mathbf{X}))_- \\
&= \rho_i(\mathbf{X}) + a_i
\end{aligned}$$

and

$$\begin{aligned}
\rho(S(\mathbf{X}+\mathbf{a})) &= S(\mathbf{c}^*(\mathbf{X}+\mathbf{a})) + \lambda\,\mathbb{E}\Big(S(\mathbf{X}) - S(\mathbf{c}^*(\mathbf{X}))\Big)_+ - \gamma\,\mathbb{E}\Big(S(\mathbf{X}) - S(\mathbf{c}^*(\mathbf{X}))\Big)_- \\
&\quad + \sum_{i=1}^{d} \Big(\lambda_i \mathbb{E}(X_i - c_i^*(\mathbf{X}))_+ - \gamma_i \mathbb{E}(X_i - c_i^*(\mathbf{X}))_-\Big) + \beta\,\mathbb{E}\Big(\|\mathbf{X} - \mathbf{c}^*\|_2\Big) \\
&= \rho(S(\mathbf{X})) + S(\mathbf{a}).
\end{aligned}$$

□

This translation invariance property of MCVaR means that if the risk of a subunit is increased by a constant amount, then the corresponding risk measure for each subunit will increase by the same amount. Furthermore, the risk measure of the aggregate risk will increase by the total of increased amounts.

In the following proposition, we discuss the subadditivity of the MCVaR. Unlike the subadditivity of a univariate risk measure, for a multivariate risk measure, the concept of subadditivity may not apply for any two random vectors since two random vectors may have different dimensions. From the purpose of portfolio risk management, we expect that the risk measure of the aggregate risk in an portfolio is not larger than the total of the risk measures of sub-portfolios. For instance, the premium of a combined policy is not larger than the total of premiums of individual policies or the required reserve on a merged company is not larger than the total of the required reserves on separated companies. In the following proposition, we show that MCVaR holds the subadditivity within a risk vector or a portfolio.

Proposition 2.5.3. *(Subadditivity within a risk vector) For random vector* $\mathbf{X} = (X_1, ..., X_d)$ *and loss function* $f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ *defined in* (2.3.1)*, assume that minimization problem*

$$\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) \tag{2.5.3}$$

has a unique minimizer. Then $MCVaR_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X}) = (\rho_1(\mathbf{X}), ..., \rho_d(\mathbf{X}), \rho(S(\mathbf{X})))$ *satisfies*

$$\rho(S(\mathbf{X})) \leq \sum_{i=1}^{d} \rho_i(\mathbf{X}). \tag{2.5.4}$$

Proof. Denote by $\mathbf{c}^*(\mathbf{X}) = (c_1^*(\mathbf{X}), ..., c_d^*(\mathbf{X}))$ the unique minimizer of $\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$. For the sake of simplicity, write $\mathbf{c}^*(\mathbf{X}) = (c_1^*(\mathbf{X}), ..., c_d^*(\mathbf{X}))$ as $\mathbf{c}^* = (c_1^*, ..., c_d^*)$. Note that $\|\mathbf{X} - \mathbf{c}^*\|_2 \leq \sum_{i=1}^d |X_i - c_i^*|$. Therefore, by (2.4.7), we have

$$\begin{aligned}
& \rho(S(\mathbf{X})) = \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) \\
= \;& \langle \mathbf{v}^*, \mathbf{c}^* \rangle + \beta_0 \mathbb{E}\Big|S(\mathbf{X}) - S(\mathbf{c}^*)\Big| + \beta \mathbb{E}\|\mathbf{X} - \mathbf{c}^*\|_2 + \mathbb{E}\Big(\langle \mathbf{w}^*, \mathbf{X} \rangle\Big) + \sum_{i=1}^{d} \beta_i \, \mathbb{E}|X_i - c_i^*| \\
\leq \;& \sum_{i=1}^{d} v_i^* c_i^* + \beta_0 \sum_{i=1}^{d} \mathbb{E}|X_i - c_i^*| + \beta \sum_{i=1}^{d} \mathbb{E}|X_i - c_i^*| + \sum_{i=1}^{d} w_i^* \, \mathbb{E}(X_i) + \sum_{i=1}^{d} \beta_i \, \mathbb{E}|X_i - c_i^*| \\
= \;& \sum_{i=1}^{d} v_i^* c_i^* + \sum_{i=1}^{d} w_i^* \, \mathbb{E}(X_i) + \sum_{i=1}^{d} (\beta + \beta_0 + \beta_i) \mathbb{E}|X_i - c_i^*| \\
= \;& \sum_{i=1}^{d} (1 - v_0 - v_i) c_i^* + \sum_{i=1}^{d} (v_0 + v_i) \, \mathbb{E}(X_i) + \sum_{i=1}^{d} (\beta + \beta_0 + \beta_i) \mathbb{E}|X_i - c_i^*| \\
= \;& \sum_{i=1}^{d} \Big(c_i^* + (\beta + \beta_0 + \beta_i) \mathbb{E}|X_i - c_i^*| + (v_0 + v_i) \, \mathbb{E}(X_i - c_i^*) \Big) \\
= \;& \sum_{i=1}^{d} \Big(c_i^* + (\lambda_i + \lambda + \beta) \, \mathbb{E}(X_i - c_i^*)_+ - (\gamma_i + \gamma - \beta) \, \mathbb{E}(X_i - c_i^*)_- \Big) \\
= \;& \sum_{i=1}^{d} \rho_i(\mathbf{X}),
\end{aligned}$$

where the last equality follows from (2.3.3). □

The subadditivity within a random vector means that if we manage the total risks of all subunits of a business entity at the enterprise level, the risk measure of the aggregate risk should be no larger than the total of the risk measures of the subunits. In the proposed MCVaR, we determine the risk measures of each subunit and the aggregate risk or the entire business entity simultaneously. However, the priority of our MCVaR is the risk measure of the aggregate risk or the entire business entity. In next proposition, we discuss the subadditivity of MCVaR for aggregate risks when risk vectors have the same dimension.

Proposition 2.5.4. *(Subadditivity among aggregate risks) For random vectors* $\mathbf{X}$*,* $\mathbf{Y}$*,* $\mathbf{X}+\mathbf{Y} \in \mathbb{R}^d$*, and loss function* $f_{\mathbf{Z},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})$ *defined in* (2.3.1)*, assume* $\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$*,* $\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{Y},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$*, and* $\min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X}+\mathbf{Y},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$ *have unique minimizers, respectively. Then, the risk measures* $\rho(S(\mathbf{X}))$*,* $\rho(S(\mathbf{Y}))$*, and* $\rho(S(\mathbf{X}+\mathbf{Y}))$ *in* $\mathrm{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X})$*,* $\mathrm{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{Y})$*, and* $\mathrm{MCVaR}_{\lambda,\gamma,\mathbf{b},\beta}(\mathbf{X}+\mathbf{Y})$*, respectively, satisfy*

$$\rho(S(\mathbf{X}+\mathbf{Y})) = \rho(S(\mathbf{X}) + S(\mathbf{Y})) \le \rho(S(\mathbf{X})) + \rho(S(\mathbf{Y})).$$

Proof. For any $\mathbf{c} = (c_1, ..., c_d)$, $\mathbf{c}' = (c'_1, ..., c'_d) \in \mathbb{R}^d$, by (2.4.7) and triangle inequalities, we have

$$\begin{aligned}
& f_{\mathbf{X}+\mathbf{Y},\beta_0,v_0,\mathbf{u},\beta}(\mathbf{c}+\mathbf{c}') \\
= \;& \langle \mathbf{v}^*, \mathbf{c}+\mathbf{c}'\rangle + \beta_0 \left| S(\mathbf{X}+\mathbf{Y}) - S(\mathbf{c}+\mathbf{c}')\right| + \beta \left\|\mathbf{X}+\mathbf{Y} - (\mathbf{c}+\mathbf{c}')\right\|_2 + \langle \mathbf{w}^*, \mathbf{X}+\mathbf{Y}\rangle \\
& + \sum_{i=1}^d \beta_i \left|X_i + Y_i - (c_i + c'_i)\right| \\
\le \;& \langle \mathbf{v}^*, \mathbf{c}\rangle + \langle \mathbf{v}^*, \mathbf{c}'\rangle + \beta_0 \left| S(\mathbf{X}) - S(\mathbf{c})\right| + \beta_0 \left| S(\mathbf{Y}) - S(\mathbf{c}')\right| \\
& + \beta \left\|\mathbf{X} - \mathbf{c}\right\|_2 + \beta \left\|\mathbf{Y} - \mathbf{c}'\right\|_2 + \langle \mathbf{w}^*, \mathbf{X}\rangle + \langle \mathbf{w}^*, \mathbf{Y}\rangle + \sum_{i=1}^d \beta_i \left|X_i - c_i\right| + \sum_{i=1}^d \beta_i \left|Y_i - c'_i\right| \\
= \;& f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}) + f_{\mathbf{Y},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}').
\end{aligned}$$

Hence, we have for any $\mathbf{c}, \mathbf{c}' \in \mathbb{R}^d$,

$$\mathbb{E}(f_{\mathbf{X}+\mathbf{Y},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}+\mathbf{c}')) \le \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c})) + \mathbb{E}(f_{\mathbf{Y},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}')). \tag{2.5.5}$$

Denote by $\mathbf{c}^*_{\mathbf{X}}$, $\mathbf{c}^*_{\mathbf{Y}}$, and $\mathbf{c}^*_{\mathbf{X}+\mathbf{Y}}$ the unique minimizers of minimization problems $\min\limits_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$, $\min\limits_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{Y},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$, and $\min\limits_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X}+\mathbf{Y},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$, respectively. Then, by the definition of MCVaR, we have $\rho(S(\mathbf{X}+\mathbf{Y})) = \mathbb{E}(f_{\mathbf{X}+\mathbf{Y},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}^*_{\mathbf{X}+\mathbf{Y}}))$. In addition, by the definition of $\mathbf{c}^*_{\mathbf{X}+\mathbf{Y}}$, we have for any $\mathbf{c} \in \mathbb{R}^d$, $\mathbb{E}(f_{\mathbf{X}+\mathbf{Y},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}^*_{\mathbf{X}+\mathbf{Y}})) \le \mathbb{E}(f_{\mathbf{X}+\mathbf{Y},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}))$. Hence,

$$\begin{aligned}
\mathbb{E}(f_{\mathbf{X}+\mathbf{Y},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}^*_{\mathbf{X}+\mathbf{Y}})) &\le \mathbb{E}(f_{\mathbf{X}+\mathbf{Y},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}^*_{\mathbf{X}} + \mathbf{c}^*_{\mathbf{Y}})) \\
&\le \mathbb{E}(f_{\mathbf{X},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}^*_{\mathbf{X}})) + \mathbb{E}(f_{\mathbf{Y},\lambda,\gamma,\mathbf{b},\beta}(\mathbf{c}^*_{\mathbf{Y}})) \\
&= \rho(S(\mathbf{X})) + \rho(S(\mathbf{Y})),
\end{aligned}$$

where the second inequality follows from (2.5.5). □

At the end of this section, we explore the monotonicity of the MCVaR. We notice that if random vectors $\mathbf{X}$ and $\mathbf{Y}$ have different dimensions, it does not make sense to compare the components of the two vectors. In addition, even if $\mathbf{X}$ and $\mathbf{Y}$ have the same dimension, the monotonicity of MCVaR may not hold for the components of two random vectors since the risk measure of one individual risk depends on the whole portfolio risk and its relative importance in the portfolio. However, in the next proposition, we will show that the monotonicity holds for the aggregate risks $S(\mathbf{X})$ and $S(\mathbf{Y})$ in the MCVaR under some special cases.

Proposition 2.5.5. *(Monotonicity on aggregate risks) For random vectors $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^d$, if Assumptions 2.4.2 holds for pairs $(\mathbf{X}, f_{\mathbf{X},\lambda,\mathbf{\Lambda}}(\mathbf{c}))$ and $(\mathbf{Y}, f_{\mathbf{Y},\lambda,\mathbf{\Lambda}}(\mathbf{c}))$, and $\mathbf{X} \le \mathbf{Y}$, then, the risk measures $\rho(S(\mathbf{X}))$ and $\rho(S(\mathbf{Y}))$ in the multivariate CVaR risk measures* $\mathrm{MCVaR}_{\lambda,\mathbf{\Lambda}}(\mathbf{X})$ *and* $\mathrm{MCVaR}_{\lambda,\mathbf{\Lambda}}(\mathbf{Y})$*, respectively, satisfy $\rho(S(\mathbf{X})) \le \rho(S(\mathbf{Y}))$.*

Proof. Under Assumptions 2.4.2, by Theorem 2.4.5, $\mathrm{MCVaR}_{\lambda,\mathbf{\Lambda}}(\mathbf{X})$ and $\mathrm{MCVaR}_{\lambda,\mathbf{\Lambda}}(\mathbf{Y})$ are well defined. Note that $\mathbf{X} = (X_1, ..., X_d) \le \mathbf{Y} = (Y_1, ..., Y_d)$ means that $X_i \le Y_i$, $i = 1, ..., d$. For any real number $c \in \mathbb{R}$, function $h(x) = (x - c)_+$ is increasing in $x \in \mathbb{R}$. In addition, for any $\mathbf{c} = (c_1, ..., c_d) \in \mathbb{R}^d$, by (2.4.18), we have

$$\begin{aligned}
\mathbb{E}(f_{\mathbf{X},\lambda,\mathbf{\Lambda}}(\mathbf{c})) &= \sum_{i=1}^{d} c_i + \lambda \mathbb{E}\Big(\sum_{i=1}^{d} X_i - \sum_{i=1}^{d} c_i\Big)_+ + \sum_{i=1}^{d} \lambda_i \mathbb{E}(X_i - c_i)_+ \\
&\le \sum_{i=1}^{d} c_i + \lambda \mathbb{E}\Big(\sum_{i=1}^{d} Y_i - \sum_{i=1}^{d} c_i\Big)_+ + \sum_{i=1}^{d} \lambda_i \mathbb{E}(Y_i - c_i)_+ \\
&= \mathbb{E}(f_{\mathbf{Y},\lambda,\mathbf{\Lambda}}(\mathbf{c})),
\end{aligned}$$

which implies $\rho(S(\mathbf{X})) = \min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{X},\lambda,\mathbf{\Lambda}}(\mathbf{c})) \le \min_{\mathbf{c}\in\mathbb{R}^d} \mathbb{E}(f_{\mathbf{Y},\lambda,\mathbf{\Lambda}}(\mathbf{c})) = \rho(S(\mathbf{Y}))$. □

2.6 Numerical illustrations of MCVaR

In this section, we use a multivariate risk portfolio to illustrate the proposed MCVaR and the effect of dependence among risks in a portfolio on MCVaR. We also compare MCVaR with VaR and CVaR if required capital is determined by these risk measures. In this section, risk portfolio $\mathbf{X} = (X_1, ..., X_d)$ is assumed to be a d-dimensional normal random vector with a d-dimensional normal distribution $\mathcal{N}_d(\boldsymbol{\mu}, \mathbf{\Sigma})$, where $\boldsymbol{\mu} = \mathbb{E}(\mathbf{X}) = (\mu_1, ..., \mu_d)$ is the mean vector, $\mu_i = \mathbb{E}(X_i)$, $i = 1, ..., d$, $\mathbf{\Sigma} = (Cov(X_i, X_j))_{i,j=1,...,d}$ is a positive-definite covariance matrix of $\mathbf{X}$, $Cov(X_i, X_j) = \rho_{ij}\, \sigma_i\, \sigma_j$, where $\sigma_i = \sqrt{Var(X_i)}$, $\sigma_j = \sqrt{Var(X_j)}$, and ρ_{ij} is the correlation coefficient between X_i and X_j satisfying $-1 < \rho_{ij} = \rho_{ji} < 1$ for $1 \le i < j \le d$. The joint density function of $(X_1, ..., X_d)$ is

$$f(\mathbf{x}) = f(x_1, ..., x_d) = \frac{1}{\sqrt{(2\pi)^d \det(\mathbf{\Sigma})}}\, e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})'}, \quad \mathbf{x} \in \mathbb{R}^d. \tag{2.6.1}$$

Thus, X_i and $S(\mathbf{X})$ have normal distributions $N(\mu_i, \sigma_i^2)$ and $N(\mu_S, \sigma_S^2)$, respectively, where $\mu_S = \sum_{i=1}^d \mu_i$, $\sigma_S = \sqrt{Var(S(\mathbf{X}))}$, and $Var(S(\mathbf{X})) = \sum_{i=1}^d \sigma_i^2 + 2\sum_{1\le i<j\le d} \rho_{ij}\,\sigma_i\,\sigma_j$. Let ϕ and Φ be the density and distribution functions of the standard normal distribution $N(0,1)$, respectively. If required capital for risks are determined by VaR at confidence level α, then

$$\begin{aligned} \mathrm{VaR}_\alpha(X_i) &= \mu_i + \sigma_i\,\Phi^{-1}(\alpha), \qquad i = 1, ..., d, && (2.6.2)\\ \mathrm{VaR}_\alpha(S(\mathbf{X})) &= \mu_S + \sigma_S\,\Phi^{-1}(\alpha). && (2.6.3) \end{aligned}$$

If required capital for risks are determined by CVaR at confidence level α, then

$$\mathrm{CVaR}_\alpha(X_i) = \mathbb{E}\left(X_i \,|\, X_i > \mathrm{VaR}_\alpha(X_i)\right) = \mu_i + \sigma_i \times \frac{\phi\left(\frac{\mathrm{VaR}_\alpha(X_i)-\mu_i}{\sigma_i}\right)}{1-\Phi\left(\frac{\mathrm{VaR}_\alpha(X_i)-\mu_i}{\sigma_i}\right)}, \quad i = 1, ..., d, \tag{2.6.4}$$

and

$$\mathrm{CVaR}_\alpha(S(\mathbf{X})) = \mathbb{E}\left(S(\mathbf{X}) \,|\, S(\mathbf{X}) > \mathrm{VaR}_\alpha(S(\mathbf{X}))\right) = \mu_S + \sigma_S \times \frac{\phi\left(\frac{\mathrm{VaR}_\alpha(S(\mathbf{X}))-\mu_S}{\sigma_S}\right)}{1-\Phi\left(\frac{\mathrm{VaR}_\alpha(S(\mathbf{X}))-\mu_S}{\sigma_S}\right)} \tag{2.6.5}$$

See Johnson et al. (1995) for these results about normal distributions. In the following two subsections, we use the above normal random vectors and Theorems 2.4.3 and 2.4.5 to illustrate MCVaR.

2.6.1 Applications of Theorem 2.4.3

In this subsection, we consider a special case of loss function (2.3.1) discussed in Theorem 2.4.3 when $\lambda > 0$, $\beta > 0$, $\gamma = 0$, $\lambda_i = 0$, $\gamma_i = 0$, $i = 1, ..., d$. In this case, loss function (2.3.1) is reduced to

$$f_{\mathbf{X},\lambda,\beta}(\mathbf{c}) = \sum_{i=1}^d c_i + \lambda\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i\Big)_+ + \beta\|\mathbf{X}-\mathbf{c}\|_2. \tag{2.6.6}$$

For loss function (2.6.6) and the normal random vector $\mathbf{X}$, Assumption 2.4.1 is reduced to the following condition:

$$\left|\frac{2-\lambda}{2}\right|\sqrt{d} < \beta. \tag{2.6.7}$$

By Theorem 2.4.3, the multivariate risk measure $\mathrm{MCVaR}_{\lambda,\beta}(\mathbf{X}) = (\rho_1(\mathbf{X}), ..., \rho_d(\mathbf{X}), \rho(S(\mathbf{X})))$ has the following expression:

$$\begin{aligned} \rho_i(\mathbf{X}) &= c_i^* + \lambda\,\mathbb{E}(X_i - c_i^*)_+ + \beta\,\mathbb{E}|X_i - c_i^*|, \quad i = 1, ..., d, && (2.6.8)\\ \rho(S(\mathbf{X})) &= \sum_{i=1}^d c_i^* + \lambda\,\mathbb{E}\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i^*\Big)_+ + \beta\,\mathbb{E}\Big(\sum_{i=1}^d (X_i - c_i^*)^2\Big)^{\frac{1}{2}}, && (2.6.9) \end{aligned}$$

where $(c_1^*, ..., c_d^*) = (c_1^*(\mathbf{X}, \lambda, \beta), ..., c_d^*(\mathbf{X}, \lambda, \beta))$ is the unique solution to the following system of equations:

$$\beta\,\mathbb{E}\Big(\frac{X_i - c_i}{\|\mathbf{X} - \mathbf{c}\|_2}\Big) = 1 - \lambda + \lambda\, F_{S(\mathbf{X})}\Big(\sum_{i=1}^{d} c_i\Big), \quad i = 1, ..., d. \tag{2.6.10}$$

In loss function (2.6.6), the additional capital is prepared for the shortfall risk $\Big(\sum_{i=1}^{d} X_i - \sum_{i=1}^{d} c_i\Big)_+$ of the aggregate risk and for the overall deviation $\|\mathbf{X} - \mathbf{c}\|_2$ of the portfolio.

Remark 2.6.1. We point out that in loss function (2.6.6), if $\lambda = 2$, then Assumption 2.4.1 or condition (2.6.7) is reduced to $\beta > 0$ and system (2.6.10) of equations has the unique solution $\mathbf{c} = \mathbb{E}(\mathbf{X})$ or $c_i = \mathbb{E}(X_i)$, $i = 1, ..., d$. To see that, for any normal random variable X, we have $F_X(\mathbb{E}(X)) = \mathbb{P}\{X \le \mathbb{E}(X)\} = \frac{1}{2}$. Thus, if $\lambda = 2$ and $c_i = \mathbb{E}(X_i)$, $i = 1, ..., d$, then the right-hand side of equation (2.6.10) is reduced to $1 - 2 + 2F_{S(\mathbf{X})}\Big(\mathbb{E}(S(\mathbf{X}))\Big) = 0$. Note that $\mathbf{X}$ has the joint normal density function (3.5.1), thus, $\mathbf{Y} = \mathbf{X} - \mathbb{E}(\mathbf{X})$ has the following joint normal density function:

$$g(\mathbf{y}) = g(y_1, ..., y_d) = \frac{1}{\sqrt{(2\pi)^d \det(\mathbf{\Sigma})}}\, e^{-\frac{1}{2}\mathbf{y}\,\mathbf{\Sigma}^{-1}\,\mathbf{y}'}, \quad \mathbf{y} \in \mathbb{R}^d, \tag{2.6.11}$$

which is an even symmetric function on $\mathbb{R}^d$ satisfying $g(\mathbf{y}) = g(y_1, ..., y_d) = g(-y_1, ..., -y_d) = g(-\mathbf{y})$ for any $\mathbf{y} = (y_1, ..., y_d) \in \mathbb{R}^d$. Hence, for any $i = 1, ..., d$,

$$\mathbb{E}\Big(\frac{X_i - \mathbb{E}(X_i)}{\|\mathbf{X} - \mathbb{E}(\mathbf{X})\|_2}\Big) = \mathbb{E}\Big(\frac{Y_i}{\|\mathbf{Y}\|_2}\Big) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{y_i}{\|\mathbf{y}\|_2} \times g(y_1, ..., y_d) dy_1 ... dy_d = 0,$$

which holds since $\frac{y_i}{\|\mathbf{y}\|_2}$ is odd symmetric on $\mathbb{R}^d$ satisfying $\frac{y_i}{\|\mathbf{y}\|_2} = -1 \times \frac{-y_i}{\|-\mathbf{y}\|_2}$ for any $\mathbf{y} = (y_1, ..., y_d) \in \mathbb{R}^d$. In fact, based on the above arguments, we can confirm that if $\lambda = 2$ and $\mathbf{Y} = \mathbf{X} - \mathbb{E}(\mathbf{X})$ has an even symmetric joint density function, then $\mathbf{c} = \mathbb{E}(\mathbf{X})$ is the unique solution to system (2.6.10) of equations. □

Next, we give numerical illustrations of $\mathrm{MCVaR}_{\lambda,\beta}(\mathbf{X})$. To do so, we set $d = 3$ or consider $\mathbf{X} = (X_1, X_2, X_3)$. Further, we assume that the marginal normal distributions of $\mathbf{X} = (X_1, X_2, X_3)$ have the following expectations and variances: $\mathbb{E}(X_1) = 130$, $Var(X_1) = 900$, $\mathbb{E}(X_2) = 150$, $Var(X_2) = 2500$, $\mathbb{E}(X_3) = 170$, and $Var(X_3) = 400$. In addition, to investigate effect of dependence among risks on MCVaR, we consider three cases of dependence: positive dependence, negative dependence, and mixed dependence. For each case of dependence, we calculate $\mathrm{MCVaR}_{\lambda,\beta}$ for different combinations of (λ, β) and compare these numerical results of $\mathrm{MCVaR}_{\lambda,\beta}$ with the corresponding results determined by $\mathrm{VaR}_{0.99}$ and $\mathrm{CVaR}_{0.99}$. In this section, numerical calculations are performed by using MATLAB.

(i) Positive dependence: Assume that risks in portfolio (X_1, X_2, X_3) are positively dependent and that the correlation coefficient of any two risks in the portfolio is positive with $\rho_{12} = 0.8$, $\rho_{13} = 0.2$, and $\rho_{23} = 0.3$. By solving system (2.6.10) of equations, we first obtain the numerical solutions (base capital) c_1^*, c_2^*, and c_3^*, and then calculate the expected shortfall $\mathbb{E}\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i^*\Big)_+$ of the aggregate risk and the expected overall deviation $\mathbb{E}\|\mathbf{X} - \mathbf{c}\|_2 = \mathbb{E}\Big(\sum_{i=1}^d (X_i - c_i^*)^2\Big)^{\frac{1}{2}}$ of the portfolio, for each of the different combinations of (λ, β) in Table 2.2. Next, by (3.5.2), (3.5.3), (3.5.4), (3.5.5), (2.6.8) and (2.6.9), we obtain the numerical results of $\mathrm{VaR}_{0.99}$, $\mathrm{CVaR}_{0.99}$, $\mathrm{MCVaR}_{\lambda,\beta}(\mathbf{X})$, which are presented in Table 2.1.

From Table 2.2, we observe that the expected shortfall of the aggregate risk is less than the expected overall deviation of the portfolio. For $\lambda = 2$, the base capital, the expected shortfall of the aggregate risk, and the expected overall deviation of the portfolio are independent of β, since in this case of $\lambda = 2$, by Remark 2.6.1, the base capital is equal to the mean vector of the portfolio, which is independent of β. For a fixed β (=2), the base capital and the expected overall deviation of the portfolio are increasing in λ while the expected shortfall of the aggregate risk is decreasing in λ, which means that there is a trade-off between the expected shortfall of the aggregate risk and the expected overall deviation of the portfolio.

From Table 2.1, we find that for a fixed λ (= 2), the required capital for the individual risks and the aggregate risk are increasing in β. For a fixed β (= 2), the required capital for the individual risks and the aggregate risk are increasing in λ. In addition, the required capital for the individual risks and the aggregate risk determined by $\mathrm{MCVaR}_{\lambda,\beta}(\mathbf{X})$ vary in the choices of (λ, β) and can produce less or larger required capital than those determined by VaR and CVaR.

(ii) Negative dependence: Assume that risks in portfolio (X_1, X_2, X_3) are negatively dependent and that the correlation coefficient of any two different risks in the portfolio is negative with $\rho_{12} = -0.4$, $\rho_{13} = -0.1$, and $\rho_{23} = -0.1$. Similarly to (i), we obtain the corresponding numerical results, which are reported in Tables 2.5 and 2.4.

From Tables 2.5 and 2.4, we find that the remarks on Tables 2.2 and 2.1 in case (i) still apply to Tables 2.5 and 2.4 in case (ii). However, by comparing Table 2.5 with Table 2.2, we notice that the expected shortfalls of the aggregate risk in case (ii) are less than those corresponding ones in case (i). Furthermore, by comparing $\mathrm{MCVaR}_{\lambda,\beta}(\mathbf{X})$ in Table 2.4 with those in Table 2.1, we observe that the required capital for the aggregate risk in case (ii) are less than those corresponding ones in case (i). These results are reasonable since in a portfolio with negatively dependent risks, the aggregate risk may be reduced due to the risk offsets from the negatively dependent risks. In general, the aggregate risk in a portfolio with negatively dependent risks is less risky than the aggregate risk in a portfolio with positively dependent risks. In addition, for $\lambda = 2$, the required capital for the individual risks in case (ii) are the same as those corresponding ones in case (i) since in this case of $\lambda = 2$, the base capital is equal to the mean vector of the portfolio and dependence has no impact on the required capital for the individual (marginal) risks.

(iii) Mixed dependence: Assume that risks in portfolio (X_1, X_2, X_3) are mixedly dependent and that some of the correlation coefficients of two risks in the portfolio are positive while some are negative with $\rho_{12} = 0.2$, $\rho_{13} = 0.8$, and $\rho_{23} = -0.3$. Similarly to (i), we obtain the corresponding numerical results, which are reported in Tables 2.8 and 2.7.

From Tables 2.8 and 2.7, we find that the remarks on Tables 2.2, 2.1, 2.5, and 2.4 in cases (i) and (ii) still apply to Tables 2.8 and 2.7 in case (ii). However, comparing Table 2.8 with Tables 2.2 and 2.5, we notice that the expected shortfalls of the aggregate risk in case (iii) are less than those corresponding ones in case (i) but larger than those corresponding ones in case (ii). Furthermore, by comparing $\text{MCVaR}_{\lambda,\beta}(\mathbf{X})$ in Table 2.7 with those in Tables 2.1 and 2.4, we observe that the required capital for the aggregate risk in case (iii) are less than those corresponding ones in case (i) but larger than those corresponding ones in case (ii). These results are also reasonable since portfolios with positively and negatively dependent risks are two extreme cases, the aggregate risk in the former portfolio is in the most risky situation, while the one in the latter portfolio is in the least risky situation.

2.6.2 Applications of Theorem 2.4.5

In this subsection, we consider a special case of loss function (2.4.18) discussed in Theorem 2.4.5 when $\lambda_i = \lambda_0$, $i = 1, ..., d$. In this case, loss function (2.4.18) is reduced to

$$f_{\mathbf{X},\lambda,\lambda_0}(\mathbf{c}) = \sum_{i=1}^{d} c_i + \lambda\Big(\sum_{i=1}^{d} X_i - \sum_{i=1}^{d} c_i\Big)_+ + \lambda_0 \sum_{i=1}^{d}(X_i - c_i)_+. \tag{2.6.12}$$

For loss function (2.6.12) and the normal random vector $\mathbf{X}$, Assumption 2.4.2 is reduced to

$$\lambda > 1, \qquad \lambda_0 > 1. \tag{2.6.13}$$

By Theorem 2.4.5, the multivariate risk measure $\text{MCVaR}_{\lambda,\lambda_0}(\mathbf{X}) = (\rho_1(\mathbf{X}), ..., \rho_d(\mathbf{X}), \rho(S(\mathbf{X})))$ has the following expression:

$$\rho_i(\mathbf{X}) = F_{X_i}^{-1}(q) + (\lambda + \lambda_0)\, \mathbb{E}\Big(X_i - F_{X_i}^{-1}(q)\Big)_+, \qquad i = 1, ..., d, \tag{2.6.14}$$

$$\rho(S(\mathbf{X})) = \sum_{i=1}^d F_{X_i}^{-1}(q) + \lambda\, \mathbb{E}\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d F_{X_i}^{-1}(q)\Big)_+ + \lambda_0 \sum_{i=1}^d \mathbb{E}\Big(X_i - F_{X_i}^{-1}(q)\Big)_+, \tag{2.6.15}$$

where $q = q(\mathbf{X}, \lambda, \lambda_0) = H_{\lambda,\lambda_0,S(\mathbf{X})}(x_0)$,

$$H_{\lambda,\lambda_0,S(\mathbf{X})}(x) = \frac{\lambda_0 + \lambda - 1 - \lambda\, F_{S(\mathbf{X})}(x)}{\lambda_i} = \frac{\lambda_0 - 1 + \lambda\, \bar{F}_{S(\mathbf{X})}(x)}{\lambda_0}, \tag{2.6.16}$$

and $x_0 = x_0(\mathbf{X}, \lambda, \lambda_0)$ is the unique solution to equation

$$x - \sum_{i=1}^d F_{X_i}^{-1}\Big(H_{\lambda,\lambda_0,S(\mathbf{X})}(x)\Big) = 0, \qquad x > F_{S(\mathbf{X})}^{-1}\Big(\frac{\lambda - 1}{\lambda}\Big). \tag{2.6.17}$$

In loss function (2.6.12), the additional capital is prepared for the shortfall risk $\Big(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i\Big)_+$ of the aggregate risk and for the total shortfall $\sum_{i=1}^d (X_i - c_i)_+$ of the individual risks. Note that by (2.6.17) and the definitions of x_0 and q, we have $x_0 = \sum_{i=1}^d F_{X_i}^{-1}(q)$, which means that $x_0 = x_0(\mathbf{X}, \lambda, \lambda_0)$ is the base capital for the aggregate risk.

Like Subsection 2.6.1, we set $d = 3$ or consider $\mathbf{X} = (X_1, X_2, X_3)$ and use the same distribution and dependence settings for (X_1, X_2, X_3) as those assumed in Subsection 2.6.1. For each of the three dependence cases, we calculate $\text{MCVaR}_{\lambda,\lambda_0}(\mathbf{X})$ for different combinations of (λ, λ_0) and compare these numerical results of $\text{MCVaR}_{\lambda,\lambda_0}(\mathbf{X})$ with the corresponding results determined by $\text{VaR}_{0.99}$ and $\text{CVaR}_{0.99}$.

(i) Positive dependence: $\rho_{12} = 0.8$, $\rho_{13} = 0.2$, and $\rho_{23} = 0.3$. By solving equation (2.6.17), we first obtain the numerical solution $x_0 = x_0(\mathbf{X}, \lambda, \lambda_0)$, and then get the base confidence level $q = q(\mathbf{X}, \lambda, \lambda_0)$, the base capital $x_0 = \sum_{i=1}^d F_{X_i}^{-1}(q)$ for the aggregate risk, the expected shortfall $\mathbb{E}\Big(\sum_{i=1}^d X_i - x_0\Big)_+$ of the aggregate risk, and the total expected shortfall $\sum_{i=1}^d \mathbb{E}(X_i - F_{X_i}^{-1}(q))_+$ of the individual risks, for each of the different combinations of (λ, λ_0) in Table 2.3. Next, by (2.6.14) and (2.6.15), we obtain the numerical results of $\text{MCVaR}_{\lambda,\lambda_0}(\mathbf{X})$, which are presented in Table 2.1.

From Table 2.3, we observe that the expected shortfall of the aggregate risk is less than the total expected shortfall of the individual risks. For a fixed λ_0 (=2), the base confidence level $q = q(\mathbf{X}, \lambda, \lambda_0)$ and the base capital for the aggregate risk are increasing in λ, hence, the expected shortfall of the aggregate risk and the total expected shortfall of the individual risks are decreasing in λ. For a fixed λ (=15), the base confidence level $q = q(\mathbf{X}, \lambda, \lambda_0)$ and the base capital for the aggregate risk are increasing in λ_0, hence, the expected shortfall of the aggregate risk and the total expected shortfall of the individual risks are decreasing in λ_0.

From Table 2.1, we find that for a fixed λ_0 $(= 2)$, the required capital for the individual risks and the aggregate risk are increasing in λ. For a fixed λ $(= 15)$, the required capital for the individual risks and the aggregate risk are increasing in λ_0. In addition, the required capital for the individual risks and the aggregate risk determined by $\text{MCVaR}_{\lambda,\lambda_0}(\mathbf{X})$ vary in the choices of (λ, λ_0) and can produce less or larger required capital than those determined by VaR and CVaR.

(ii) Negative dependence: $\rho_{12} = -0.4$, $\rho_{13} = -0.1$, and $\rho_{23} = -0.1$. Similarly to (i), we obtain the corresponding numerical results, which are reported in Tables 2.6 and 2.4.

From Table 2.6, we find that the remarks on Table 2.3 in case (i) still apply to Table 2.6 in case (ii). However, by comparing Table 2.6 with Table 2.3, we notice that the expected shortfalls of the aggregate risk in case (ii) are less than those corresponding ones in case (i). Furthermore, by comparing $\text{MCVaR}_{\lambda,\lambda_0}(\mathbf{X})$ in Table 2.4 with those in Table 2.1, we observe that the required capital for the aggregate risk in case (ii) are less than those corresponding ones in case (i). These results are consistent with those in cases (ii) of Subsection 2.6.1, and further indicate that the aggregate risk in a portfolio with negatively dependent risks is less risky than the aggregate risk in a portfolio with positively dependent risks.

From Table 2.4, we find that all the remarks on Table 2.1 in case (i) still apply to Table 2.4 in case (ii), expect that for a fixed λ $(= 15)$, the required capital for the individual risks are neither increasing nor decreasing in λ_0, which is different from case (i).

(iii) Mixed dependence: $\rho_{12} = 0.2$, $\rho_{13} = 0.8$, and $\rho_{23} = -0.3$. Similarly to (i), we obtain the corresponding numerical results, which are reported in Tables 2.9 and 2.7.

From Tables 2.9 and 2.7, we see that the remarks on Tables 2.3 and 2.1 in case (i) still apply to Tables 2.9 and 2.7 in case (iii). However, comparing Table 2.9 with Tables 2.3 and 2.6, we find that the base confidence levels $q = q(\mathbf{X}, \lambda, \lambda_0)$ in case (iii) are less than those corresponding ones in case (i) but larger than those corresponding ones in case (ii). Furthermore, by comparing Table 2.7 with Tables 2.1 and 2.4, we observe that the required capital for the aggregate risk in case (iii) are less than those corresponding ones in case (i) but larger than those corresponding ones in case (ii). These results are consistent with those in cases (iii) of Subsection 2.6.1.

2.7 Conclusions

The novel MCVaR risk measure proposed in this chapter balances the systemic risks resulting from individual risks and the aggregate risk of a portfolio, and gives first priority to the aggregate risk. The risk measures for individual risks and the aggregate risk in a portfolio resulting from this MCVaR depend not only on their own distributions, but also on correlations among the individual risks and the relative importance of the individual risks and the aggregate risk to a portfolio. This MCVaR is an objective-driven multivariate risk measure and it can minimize expected systemic risk in a risk portfolio.

	X_1	X_2	X_3	$S(\mathbf{X})$
$\mathrm{VaR}_{0.99}$	199.79	266.32	216.53	645.19
$\mathrm{CVaR}_{0.99}$	209.96	283.26	223.30	673.62
$\mathrm{MCVaR}_{\lambda=2,\,\beta=1}$	177.87	229.79	201.92	570.38
$\mathrm{MCVaR}_{\lambda=2,\,\beta=2}$	201.81	269.68	217.87	623.82
$\mathrm{MCVaR}_{\lambda=2,\,\beta=3}$	225.75	309.58	233.83	677.25
$\mathrm{MCVaR}_{\lambda=2,\,\beta=4}$	249.68	349.47	249.79	730.69
$\mathrm{MCVaR}_{\lambda=3,\,\beta=2}$	212.53	287.39	225.19	652.02
$\mathrm{MCVaR}_{\lambda=4,\,\beta=2}$	221.23	301.70	231.31	673.03
$\mathrm{MCVaR}_{\lambda=5,\,\beta=2}$	228.49	313.74	236.52	689.67
$\mathrm{MCVaR}_{\lambda=13,\,\lambda_0=2}$	189.60	249.33	209.73	617.63
$\mathrm{MCVaR}_{\lambda=14,\,\lambda_0=2}$	190.56	250.94	210.38	619.80
$\mathrm{MCVaR}_{\lambda=15,\,\lambda_0=2}$	191.46	252.44	210.98	621.81
$\mathrm{MCVaR}_{\lambda=16,\,\lambda_0=2}$	192.31	253.85	211.54	623.70
$\mathrm{MCVaR}_{\lambda=15,\,\lambda_0=3}$	192.08	253.47	211.39	625.80
$\mathrm{MCVaR}_{\lambda=15,\,\lambda_0=25}$	200.56	267.60	217.04	672.50
$\mathrm{MCVaR}_{\lambda=15,\,\lambda_0=50}$	205.64	276.11	220.43	694.84

Table 2.1: $\mathrm{VaR}_{0.99}$, $\mathrm{CVaR}_{0.99}$, $\mathrm{MCVaR}_{\lambda,\beta}$, and $\mathrm{MCVaR}_{\lambda,\lambda_0}$ with positively dependent risks

	c_1^*	c_2^*	c_3^*	$\mathbb{E}(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i^*)_+$	$\mathbb{E}\lVert\mathbf{X} - \mathbf{c}^*\rVert_2$
$(\lambda, \beta) = (2, 1)$	130	150	170	33.47	53.43
$(\lambda, \beta) = (2, 2)$	130	150	170	33.47	53.43
$(\lambda, \beta) = (2, 3)$	130	150	170	33.47	53.43
$(\lambda, \beta) = (2, 4)$	130	150	170	33.47	53.43
$(\lambda, \beta) = (3, 2)$	136.87	158.87	175.44	23.94	54.50
$(\lambda, \beta) = (4, 2)$	141.68	165.02	179.30	18.51	56.50
$(\lambda, \beta) = (5, 2)$	145.36	169.64	182.29	15.01	58.67

Table 2.2: Base capital, expected shortfalls of aggregate risks, and expected overall deviations with positively dependent risks based on model (2.6.6)

	q	$x_0 = \sum_{i=1}^d F_{X_i}^{-1}(q)$	$\mathbb{E}(\sum_{i=1}^d X_i - x_0)_+$	$\sum_{i=1}^d \mathbb{E}(X_i - F_{X_i}^{-1}(q))_+$
$(\lambda, \lambda_0) = (13, 2)$	0.9018	579.19	2.2448	4.6319
$(\lambda, \lambda_0) = (14, 2)$	0.9063	581.85	2.0856	4.3770
$(\lambda, \lambda_0) = (15, 2)$	0.9104	584.31	1.9468	4.1513
$(\lambda, \lambda_0) = (16, 2)$	0.9141	586.60	1.8247	3.9500
$(\lambda, \lambda_0) = (15, 3)$	0.9163	588.07	1.7502	3.8258
$(\lambda, \lambda_0) = (15, 25)$	0.9681	635.39	0.3990	1.2449
$(\lambda, \lambda_0) = (15, 50)$	0.9819	659.21	0.1716	0.6611

Table 2.3: Base confidence levels, expected shortfalls of aggregate risks, and total expected shortfalls of individual risks with positively dependent risks based on model (2.6.12)

	X_1	X_2	X_3	$S(\mathbf{X})$
$\mathrm{VaR}_{0.99}$	199.79	266.31	216.53	561.08
$\mathrm{CVaR}_{0.99}$	209.96	283.26	223.30	577.26
$\mathrm{MCVaR}_{\lambda=2,\,\beta=1}$	177.87	229.79	201.92	543.03
$\mathrm{MCVaR}_{\lambda=2,\,\beta=2}$	201.81	269.68	217.87	597.97
$\mathrm{MCVaR}_{\lambda=2,\,\beta=3}$	225.75	309.58	233.83	652.91
$\mathrm{MCVaR}_{\lambda=2,\,\beta=4}$	249.68	349.47	249.79	707.84
$\mathrm{MCVaR}_{\lambda=3,\,\beta=2}$	212.50	287.73	224.97	613.61
$\mathrm{MCVaR}_{\lambda=4,\,\beta=2}$	221.33	303.11	230.68	624.83
$\mathrm{MCVaR}_{\lambda=5,\,\beta=2}$	228.97	316.83	235.45	633.53
$\mathrm{MCVaR}_{\lambda=13,\,\lambda_0=2}$	207.41	279.02	221.61	566.07
$\mathrm{MCVaR}_{\lambda=14,\,\lambda_0=2}$	210.00	283.34	223.33	566.92
$\mathrm{MCVaR}_{\lambda=15,\,\lambda_0=2}$	212.52	287.54	225.02	567.71
$\mathrm{MCVaR}_{\lambda=16,\,\lambda_0=2}$	214.98	291.64	226.66	568.46
$\mathrm{MCVaR}_{\lambda=15,\,\lambda_0=3}$	211.28	285.46	224.19	578.36
$\mathrm{MCVaR}_{\lambda=15,\,\lambda_0=25}$	201.95	269.96	217.94	665.46
$\mathrm{MCVaR}_{\lambda=15,\,\lambda_0=50}$	205.93	276.58	220.62	692.09

Table 2.4: $\mathrm{VaR}_{0.99}$, $\mathrm{CVaR}_{0.99}$, $\mathrm{MCVaR}_{\lambda,\beta}$, and $\mathrm{MCVaR}_{\lambda,\lambda_0}$ with negatively dependent risks

	c_1^*	c_2^*	c_3^*	$\mathbb{E}(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i^*)_+$	$\mathbb{E}\|\mathbf{X} - \mathbf{c}^*\|_2$
$(\lambda, \beta) = (2, 1)$	130	150	170	19.05	54.94
$(\lambda, \beta) = (2, 2)$	130	150	170	19.05	54.94
$(\lambda, \beta) = (2, 3)$	130	150	170	19.05	54.94
$(\lambda, \beta) = (2, 4)$	130	150	170	19.05	54.94
$(\lambda, \beta) = (3, 2)$	134.21	155.48	174.04	12.97	55.49
$(\lambda, \beta) = (4, 2)$	137.03	159.10	176.75	9.76	56.46
$(\lambda, \beta) = (5, 2)$	139.12	161.75	178.77	7.78	57.48

Table 2.5: Base capital, expected shortfalls of aggregate risks, and expected overall deviations with negatively dependent risks based on model (2.6.6)

	q	$x_0 = \sum_{i=1}^d F_{X_i}^{-1}(q)$	$\mathbb{E}(\sum_{i=1}^d X_i - x_0)_+$	$\sum_{i=1}^d \mathbb{E}(X_i - F_{X_i}^{-1}(q))_+$
$(\lambda, \lambda_0) = (13, 2)$	0.7912	531.06	0.8785	11.7987
$(\lambda, \lambda_0) = (14, 2)$	0.7951	532.42	0.8190	11.5156
$(\lambda, \lambda_0) = (15, 2)$	0.7987	533.69	0.7671	11.2582
$(\lambda, \lambda_0) = (16, 2)$	0.8020	534.87	0.7213	11.0226
$(\lambda, \lambda_0) = (15, 3)$	0.8158	539.93	0.5510	10.0554
$(\lambda, \lambda_0) = (15, 25)$	0.9601	624.75	0.0015	1.6275
$(\lambda, \lambda_0) = (15, 50)$	0.9800	655.29	0.0001	0.7359

Table 2.6: Base confidence levels, expected shortfalls of aggregate risks, and total expected shortfalls of individual risks with negatively dependent risks based on model (2.6.12)

	X_1	X_2	X_3	$S(\mathbf{X})$
$\text{VaR}_{0.99}$	199.79	266.32	216.53	610.50
$\text{CVaR}_{0.99}$	209.96	283.26	223.30	633.88
$\text{MCVaR}_{\lambda=2,\,\beta=1}$	177.87	229.79	201.92	559.35
$\text{MCVaR}_{\lambda=2,\,\beta=2}$	201.73	269.58	217.81	613.67
$\text{MCVaR}_{\lambda=2,\,\beta=3}$	225.75	309.58	233.83	667.99
$\text{MCVaR}_{\lambda=2,\,\beta=4}$	249.53	349.28	249.69	722.32
$\text{MCVaR}_{\lambda=3,\,\beta=2}$	212.40	287.41	224.99	636.63
$\text{MCVaR}_{\lambda=4,\,\beta=2}$	221.02	302.13	230.86	653.50
$\text{MCVaR}_{\lambda=5,\,\beta=2}$	228.19	314.88	235.78	666.76
$\text{MCVaR}_{\lambda=13,\,\lambda_0=2}$	193.98	256.63	212.65	594.85
$\text{MCVaR}_{\lambda=14,\,\lambda_0=2}$	195.33	258.88	213.55	596.43
$\text{MCVaR}_{\lambda=15,\,\lambda_0=2}$	196.61	261.02	214.41	597.91
$\text{MCVaR}_{\lambda=16,\,\lambda_0=2}$	197.84	263.06	215.22	599.29
$\text{MCVaR}_{\lambda=15,\,\lambda_0=3}$	196.81	261.36	214.54	603.99
$\text{MCVaR}_{\lambda=15,\,\lambda_0=25}$	201.33	268.89	217.55	667.13
$\text{MCVaR}_{\lambda=15,\,\lambda_0=50}$	205.86	276.48	220.57	692.51

Table 2.7: $\text{VaR}_{0.99}$, $\text{CVaR}_{0.99}$, $\text{MCVaR}_{\lambda,\beta}$, and $\text{MCVaR}_{\lambda,\lambda_0}$ with mixedly dependent risks

	c_1^*	c_2^*	c_3^*	$\mathbb{E}(\sum_{i=1}^d X_i - \sum_{i=1}^d c_i^*)_+$	$\mathbb{E}\lVert\mathbf{X} - \mathbf{c}^*\rVert_2$
$(\lambda, \beta) = (2, 1)$	130	150	170	27.52	54.30
$(\lambda, \beta) = (2, 2)$	130	150	170	27.52	54.30
$(\lambda, \beta) = (2, 3)$	130	150	170	27.52	54.30
$(\lambda, \beta) = (2, 4)$	130	150	170	27.52	54.30
$(\lambda, \beta) = (3, 2)$	136.52	156.89	174.89	19.34	55.15
$(\lambda, \beta) = (4, 2)$	140.98	161.60	178.30	14.80	56.71
$(\lambda, \beta) = (5, 2)$	144.32	165.13	180.90	11.93	58.38

Table 2.8: Base capital, expected shortfalls of aggregate risks, and expected overall deviations with mixedly dependent risks based on model (2.6.6)

	q	$x_0 = \sum_{i=1}^d F_{X_i}^{-1}(q)$	$\mathbb{E}(\sum_{i=1}^d X_i - x_0)_+$	$\sum_{i=1}^d \mathbb{E}(X_i - F_{X_i}^{-1}(q))_+$
$(\lambda, \lambda_0) = (13, 2)$	0.8636	559.68	1.6427	6.9052
$(\lambda, \lambda_0) = (14, 2)$	0.8682	561.78	1.5288	6.6238
$(\lambda, \lambda_0) = (15, 2)$	0.8723	563.73	1.4294	6.3915
$(\lambda, \lambda_0) = (16, 2)$	0.8760	565.54	1.3419	6.1436
$(\lambda, \lambda_0) = (15, 3)$	0.8819	568.42	1.2119	5.7939
$(\lambda, \lambda_0) = (15, 25)$	0.9629	628.53	0.1050	1.4812
$(\lambda, \lambda_0) = (15, 50)$	0.9804	656.06	0.0276	0.7208

Table 2.9: Base confidence levels, expected shortfalls of aggregate risks, and total expected shortfalls of individual risks with mixedly dependent risks based on model (2.6.12)

Chapter 3

A new approach to determine the required capital for the aggregate risk and individual risks

3.1 Introduction

In portfolio risk management, after quantifying risk with risk measures, deciding how to allocate the available capital in the most efficient way is also an important topic. The rules for deciding capital allocation are usually guided by risk measures. In this chapter, we will develop a risk measure from the capital allocation perspective. Instead of discussing risk measure properties, we focus more on comparing currently existing allocation principles and with allocation principle based on our risk measure, and on how our model can be viewed as a generalization of many existing allocation principles. In this model, we use a similar idea to derive our multivariate risk measure as in the previous chapter: we simultaneously obtain the risk measure for the entire portfolio, which is used as the optimal total capital, and the risk measures for all the individual risk units in the portfolio which are used as the optimal allocation to the individual risk units.

Currently, in most optimal capital allocation studies, it is assumed that a given total capital K for a company will be allocated among its n (main) business lines with losses $X_1, ..., X_n$, respectively, and the predetermined total capital K is often determined by a decision maker based on the aggregate risk $S = \sum_{i=1}^{n} X_i$ of the company under a total capital criterion, such as $K = \mathrm{VaR}_q(S)$ (VaR criterion) or $K = \mathrm{CTE}_q(S) = \mathbb{E}(S|S > \mathrm{VaR}_q(S))$ (CTE criterion), where $0 < q < 1$. The VaR and CTE criteria are two important ways to determine the required total capital for insurance companies and financial institutions. In addition, under certain assumptions, $\mathrm{VaR}_q(S)$ and $\mathrm{CTE}_q(S)$ are the unique solutions to the optimization problem

$$\min_{x\in\mathbb{R}} \mathbb{E}\Big(\frac{\xi}{v}(S-x)^2\Big) \tag{3.1.1}$$

with respective choices of ξ and v, where ξ is a random variable and v is a real number such that $\frac{\xi}{v} \geq 0$ and $\frac{\mathbb{E}[\xi]}{v} > 0$. See, for instance, Furman and Zitikis (2008) and Cai and Wang (2020). In fact, it is easy to see that the unique minimizer of problem (3.1.1) is $\mathbb{E}(\frac{\xi}{v}S)$. If $\xi = \mathrm{I}_{\{S>\mathrm{VaR}_q(S)\}}$ and $v = \mathbb{P}\{S > \mathrm{VaR}_q(S)\}$, then $\mathrm{CTE}_q(S) = \mathbb{E}(\frac{\xi}{v}S)$. If $\xi = \mathrm{I}_{\{S\in A\}}$ and $v = \mathbb{P}\{S \in A\}$, where $A \subset \mathbb{R}$ is a set satisfying $\mathbb{E}(S|S \in A) = \mathrm{VaR}_q(S)$, then $\mathrm{VaR}_q(S) = \mathbb{E}(\frac{\xi}{v}S)$. The existence of such a set A is proved in Proposition 3.7 of Cai and Wang (2020) if S is a continuous random variable.

When a decision maker allocates the predetermined capital K among the n business lines, they use a criterion to allocate a capital K_i for business line i, $i = 1, ..., n$, so that $\sum_{i=1}^n K_i = K$. For example, the allocation $(K_1, ..., K_n)$ may be a solution that minimizes an expected loss function that concerns the decision maker. For instance, $(K_1, ..., K_n)$ could be a solution to the optimization problem

$$\begin{cases} \min_{(K_1,...,K_n)\in\mathbb{R}^n} \sum_{i=1}^n v_i\mathbb{E}\Big(\xi_i D\Big(\frac{X_i-K_i}{v_i}\Big)\Big) \\ \text{s.t.} \quad \sum_{i=1}^n K_i = K, \end{cases} \tag{3.1.2}$$

where ξ_i and v_i are non-negative random variables and positive real numbers, respectively, $i = 1, ..., n$, and D is a function. With suitable choices of ξ_i and v_i for, $i = 1, ..., n$ and of D, the optimal solution $(K_1, ..., K_n)$ to problem (3.1.2) can yield many interesting capital allocation principles including CTE, haircut, covariance and proportional allocation principles, and so on. See Dhaene, Tsanakas, et al. (2012) and Cai and Wang (2020) for details. In such a capital allocation scheme, the required total capital K and the corresponding allocation scheme are considered separately. If the required total capital K is determined by the CTE or VaR criterion, the deviation between the aggregate risk S and the required total capital K is the main concern for the decision maker. In this way, the required total capital K only depends on the distribution of the aggregate risk and ignores relationships among $X_1, ..., X_n, S$ such as dependences and covariances.

When determining the required total capital K for the aggregate risk S of a company and allocating the total capital K to business lines or individual risks of the company, the decision maker is concerned not only with the deviation between the total capital K and the aggregate risk S, but also with the total of the allocation deviations between the allocated capital K_i and the individual risk X_i, $i = 1, ..., n$. To balance the deviation between the total capital K and the aggregate risk S and the total of the allocation deviations between individual allocated capital K_i and risk X_i, $i = 1, ..., n$, in Section 3.2 we first consider how to determine the optimal required total capital K^* for a company and the corresponding optimal allocation scheme $(K_1^*, ..., K_n^*)$ among the n business lines of the company, which is formulated as the following optimization problem:

$$\begin{cases} \min_{(K,K_1,...,K_n)\in\mathbb{R}^{n+1}} \left\{(1-\alpha)v\,\mathbb{E}\Big(\xi D_1\Big(\frac{S-K}{v}\Big)\Big) + \alpha\sum_{i=1}^n v_i\,\mathbb{E}\Big(\xi_i\, D_{2i}\Big(\frac{X_i-K_i}{v_i}\Big)\Big)\right\} \\ \text{s.t.} \quad \sum_{i=1}^n K_i = K, \end{cases} \tag{3.1.3}$$

where $0 < \alpha < 1$, v and v_i, are positive real numbers, ξ and ξ_i, are non-negative random variables, $i = 1, ..., n$, and D_1 and D_{2i}, $i = 1, ..., n$ are real functions used to measure the deviation between the total capital K and the aggregate risk S and the allocation deviation between the allocated capital K_i and the individual risk X_i, respectively. We point out that in optimization problem (3.1.3), there are $n+1$ control variables $K, K_1, ..., K_n$ and a constraint $\sum_{i=1}^n K_i = K$.

As discussed in Zaks and Tsanakas (2014) and Cai and Wang (2020), in practice each (main) business line of a company may further have several sub-business lines, assuming that business line i has n_i sub-business lines with loss random vector $(X_{i1}, ..., X_{in_i})$, $i = 1, ..., n$, the capital K_i allocated to business line i needs to be further allocated among the n_i sub-business lines of business line i, say, a capital k_{ij} is allocated to X_{ij}, $j = 1, ..., n_i$, so that $\sum_{j=1}^{n_i} k_{ij} = K_i$, $i = 1, ..., n$. In this scenario, the decision maker is also concerned with the total of the allocation deviations between the capital k_{ij} allocated to sub-business line ij and the risk X_{ij}, $j = 1, ..., n_i$, $i = 1, ..., n$. To balance the deviations resulting from the three levels (company, main business lines, and sub-business lines), in Section 3.3 we discuss how to determine the optimal required total capital K^* for the company and the corresponding optimal allocated capital K_i^*, k_{ij}^*, $i = 1, ..., n$, $j = 1, ..., n_i$, among the two levels of main business lines and sub-business lines at the same time. Such an optimal allocation scheme can be formulated as follows: Let $\mathbf{K} = (K, K_1, ..., K_n)$, $\mathbf{k}_i = (k_{i1}, ..., k_{in_i})$, $i = 1, ..., n$, and consider the following optimization problem:

$$\begin{cases} \min_{(\mathbf{K},\mathbf{k_1},...,\mathbf{k_n})\in\mathbb{R}^d} & \Big\{\alpha_1 w\mathbb{E}\Big[\psi D_1\Big(\frac{S-K}{w}\Big)\Big] + \ \alpha_2\sum_{i=1}^n w_i\mathbb{E}\Big[\psi_i D_{2i}\Big(\frac{X_i-K_i}{w_i}\Big)\Big] \\ & \quad + \alpha_3\sum_{i=1}^n\sum_{j=1}^{n_i} w_{ij}\mathbb{E}\Big[\psi_{ij}D_{3ij}\Big(\frac{X_{ij}-k_{ij}}{w_{ij}}\Big)\Big]\Big\} \\ \text{s.t.} & \sum_{i=1}^n K_i = K, \quad \sum_{j=1}^{n_i} k_{ij} = K_i, \quad i = 1, ..., n, \end{cases} \tag{3.1.4}$$

where $d \triangleq 1 + n + \sum_{i=1}^n n_i$, the α_i and $0 < \alpha_i < 1$, $i = 1, 2, 3$, satisfy $\sum_{i=1}^3 \alpha_i = 1$, and in addition, w, w_i, and w_{ij} are positive real numbers and ψ, ψ_i, and ψ_{ij} are nonnegative random variables.

Note that in optimization problem (3.1.4), there are d control variables K, K_i, k_{ij}, $i = 1, ..., n$, $j = 1, ..., n_i$, and $1+n$ constraints $\sum_{i=1}^{n} K_i = K$ and $\sum_{j=1}^{n_i} k_{ij} = K_i$, $i = 1, ..., n$.

The rest of the chapter is organized as follows. In Section 3.2, we derive the unique solution $(K^*, K_1^*, ..., K_n^*)$ to problem (3.1.3) when $D_1(x) = x^2$ and $D_{2i}(x) = x^2$, $i = 1, ..., n$, namely the allocation deviations are measured by (weighted) squared errors. We show that many existing approaches for determining the required total capital, such as VaR and CTE criteria, and allocation schemes such as CTE, haircut, proportional and covariance allocation principles, and so on, are special or limiting cases of the unique solution $(K^*, K_1^*, ..., K_n^*)$ to optimization problem (3.1.3) when $D_1(x) = x^2$ and $D_{2i}(x) = x^2$, $i = 1, ..., n$. Moreover, these results clearly show the impact of a decision maker's attitude toward the allocation deviations at the three levels of company, main business lines, and sub-business lines on the optimal required total capital and the optimal allocation scheme. In Section 3.3, we extend the model in Section 3.2 to a three-level structure. We again use the squared error deviation function to measure the allocation deviation for all of the three levels and derive the unique solution for the three-level model. Furthermore, we show that under special conditions, this model can be reduced to the model in Zaks and Tsanakas (2014). In Section 3.4, we replace the squared error deviation function with an absolute error deviation function as the measurement of allocation deviation. Under this model, we discuss the conditions to guarantee the existence and uniqueness of the solution. In Section 3.5, we provide a numerical illustration of the allocation method with two-level model with both square error and absolute error deviation functions. We also compare our allocation method with current existing allocation methods such as the haircut and CTE principles. Concluding remarks are given in Section 3.6.

3.2 Optimal solutions based on weighted squared errors for a company with main business lines

In this section, we discuss optimal required total capital and optimal allocation scheme for a company with multiple business lines and consider problem (3.1.3) when $D_1(x) = x^2$ and $D_{2i} = x^2$, $i = 1, ..., n$, namely the allocation deviations are measured by weighted squared errors. In this case, problem (3.1.3) is reduced to the following problem:

$$\begin{cases} \min_{(K, K_1, ..., K_n) \in \mathbb{R}^{n+1}} \left\{ (1-\alpha)\mathbb{E}\left[\frac{\xi}{v}\left(S - K\right)^2\right] + \alpha \sum_{i=1}^{n} \mathbb{E}\left[\frac{\xi_i}{v_i}\left(X_i - K_i\right)^2\right]\right\} \\ \text{s.t.} \quad \sum_{i=1}^{n} K_i = K. \end{cases} \tag{3.2.1}$$

To guarantee that problem (3.2.1) has a unique solution $(K^*, K_1^*, ..., K_n^*)$, we assume that the following conditions hold.

Assumption 3.2.1. For problem (3.2.1), we assume that real numbers v and v_i, and random variables ξ and ξ_i, $i = 1, ..., n$, satisfy

$$\frac{\xi}{v} \geq 0, \quad \frac{\xi}{v_i} \geq 0, \quad \frac{\mathbb{E}[\xi]}{v} > 0, \quad \frac{\mathbb{E}[\xi_i]}{v_i} > 0, \quad i = 1, ..., n, \tag{3.2.2}$$

and that expectations $\mathbb{E}[\frac{\xi}{v} S^2]$, $\mathbb{E}[\frac{\xi}{v} S]$, $\mathbb{E}[\frac{\xi_i}{v_i} X_i^2]$, $\mathbb{E}[\frac{\xi_i}{v_i} X_i]$, $i = 1, ..., n$, exist. □

Denote the objective function in problem (3.2.1) by $J(K, K_1, ..., K_n)$, namely,

$$J(K, K_1, ..., K_n) = (1-\alpha)\mathbb{E}\Big[\frac{\xi}{v}\Big(S - K\Big)^2\Big] + \alpha \sum_{i=1}^{n} \mathbb{E}\Big[\frac{\xi_i}{v_i}\Big(X_i - K_i\Big)^2\Big]. \tag{3.2.3}$$

Note that by Assumption 3.2.1, we have

$$0 \leq \mathbb{E}\Big(\frac{\xi}{v}\Big(S - K\Big)^2\Big) = \mathbb{E}\Big(\frac{\xi}{v} S^2\Big) - 2K\, \mathbb{E}\Big(\frac{\xi}{v} S\Big) + K^2\, \mathbb{E}\Big(\frac{\xi}{v}\Big) < \infty$$

for any $K \in \mathbb{R}$. Similarly, by Assumption 3.2.1, we have $0 \leq \mathbb{E}\Big(\frac{\xi_i}{v_i}\Big(X_i - K_i\Big)^2\Big) < \infty$ for any $K_i \in \mathbb{R}$, $i = 1, ..., n$. Hence, the objective function $J(K, K_1, ..., K_n)$ for problem (3.2.1) is well defined.

Lemma 3.2.1. *Under Assumption 3.2.1, the objective function $J(K, K_1, ..., K_n)$ defined in (3.2.3) is a convex and coercive function of $(K, K_1, ..., K_n) \in \mathbb{R}^{n+1}$.*

Proof. Note that a quadratic function $ax^2 + bx + c$ is a convex and coercive function of $x \in \mathbb{R}$ when $a > 0$. Thus, under Assumption 3.2.1, $\mathbb{E}\Big[\frac{\xi}{v}\Big(S - K\Big)^2\Big]$ and $\mathbb{E}\Big[\frac{\xi_i}{v_i}\Big(X_i - K_i\Big)^2\Big]$ are convex and coercive functions of $K, K_i \in \mathbb{R}$, respectively, $i = 1, ..., n$.

It is well known that if $f_i(x_i)$ is a convex and coercive function of $x_i \in \mathbb{R}$, $i = 1, \ldots, m$, then $f(\mathbf{x}) = \sum_{i=1}^{m} f_i(x_i)$ is a convex and coercive function of $\mathbf{x} \in \mathbb{R}^m$, where $\mathbf{x} = (x_1, \ldots, x_m)$. Hence, $J(K, K_1, ..., K_n)$ defined in (3.2.3) is a convex and coercive function of $(K, K_1, ..., K_n) \in \mathbb{R}^{n+1}$. □

Theorem 3.2.2. *Under Assumption 3.2.1, problem (3.2.1) has the following unique solu-*

tion:

$$K^* = \tilde{\beta}(\alpha)\,\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} + \beta(\alpha)\sum_{i=1}^{n}\frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} \tag{3.2.4}$$

$$= \sum_{i=1}^{n}\frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} + \tilde{\beta}(\alpha)\Big(\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} - \sum_{i=1}^{n}\frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}\Big), \tag{3.2.5}$$

$$K_i^* = \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} + \beta_i(\alpha)\Big(\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} - \sum_{i=1}^{n}\frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}\Big) \tag{3.2.6}$$

$$= \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} + \frac{\frac{v_i}{\mathbb{E}[\xi_i]}}{\sum_{i=1}^{n}\frac{v_i}{\mathbb{E}[\xi_i]}}\Big(K^* - \sum_{i=1}^{n}\frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}\Big), \qquad i = 1, ..., n, \tag{3.2.7}$$

where

$$\beta(\alpha) = 1 - \tilde{\beta}(\alpha) = \frac{\alpha\frac{v}{\mathbb{E}[\xi]}}{\alpha\frac{v}{\mathbb{E}[\xi]} + (1-\alpha)\sum_{i=1}^{n}\frac{v_i}{\mathbb{E}[\xi_i]}} \tag{3.2.8}$$

and

$$\beta_i(\alpha) = \frac{(1-\alpha)\frac{v_i}{\mathbb{E}[\xi_i]}}{\alpha\frac{v}{\mathbb{E}[\xi]} + (1-\alpha)\sum_{i=1}^{n}\frac{v_i}{\mathbb{E}[\xi_i]}}. \tag{3.2.9}$$

Proof. By Lemma 3.2.1, we know that problem (3.2.1) is a constrained convex optimization problem and that the objective function $J(K, K_1, ..., K_n)$ of the problem is convex and coercive in $(K, K_1, ..., K_n) \in \mathbb{R}^{n+1}$. Hence, an optimal solution for problem (3.2.1) exists. Denote the Lagrangian of problem (3.2.1) by

$$L(K, K_1, ..., K_n, \lambda) = (1-\alpha)\mathbb{E}\Big(\frac{\xi}{v}\big(S-K\big)^2\Big) + \alpha\sum_{i=1}^{n}\mathbb{E}\Big(\frac{\xi_i}{v_i}\big(X_i-K_i\big)^2\Big) + \lambda\Big(\sum_{i=1}^{n}K_i - K\Big).$$

Then, under Assumption 3.2.1, $\mathbb{E}\Big(\frac{\xi}{v}\big(S-K\big)^2\Big) = \mathbb{E}\Big(\frac{\xi}{v}S^2\Big) - 2K\,\mathbb{E}\Big(\frac{\xi}{v}S\Big) + K^2\,\mathbb{E}\Big(\frac{\xi}{v}\Big)$ is differentiable with respect to K and

$$\frac{\partial\,\mathbb{E}\Big(\frac{\xi}{v}\big(S-K\big)^2\Big)}{\partial K} = -2\mathbb{E}\Big(\frac{\xi}{v}(S-K)\Big). \tag{3.2.10}$$

Similarly, we have that $\mathbb{E}\Big(\frac{\xi_i}{v_i}\big(X_i-K_i\big)^2\Big)$ is differentiable with respect to K_i and

$$\frac{\partial\,\mathbb{E}\Big(\frac{\xi_i}{v_i}\big(X_i-K_i\big)^2\Big)}{\partial K_i} = -2\mathbb{E}\Big(\frac{\xi_i}{v_i}(X_i-K_i)\Big). \tag{3.2.11}$$

Thus, the Lagrangian function $L(K, K_1, ..., K_n, \lambda)$ is differentiable with respect to each of $K, K_1, ..., K_n$. Hence, a solution to problem (3.2.1) is any solution to the following equations or any solution satisfying the following Karush-Kuhn-Tucker (KKT) conditions:

$$\left\{\begin{array}{l}\frac{\partial L}{\partial K}=0,\\ \frac{\partial L}{\partial K_i}=0, \qquad i=1,...,n,\\ \sum_{i=1}^n K_i=K.\end{array}\right. \tag{3.2.12}$$

By (3.2.10) and (3.2.11), we see that (3.2.12) is equivalent to

$$\left\{\begin{array}{l}-2(1-\alpha)\mathbb{E}\Big[\frac{\xi}{v}\Big(S-K\Big)\Big]-\lambda=0,\\ -2\alpha\mathbb{E}\Big[\frac{\xi_i}{v_i}\Big(X_i-K_i\Big)\Big]+\lambda=0, \qquad i=1,...,n,\\ \sum_{i=1}^n K_i=K.\end{array}\right. \tag{3.2.13}$$

The above equations are further equivalent to

$$\left\{\begin{array}{l}\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}-K=-\frac{\lambda v}{2(1-\alpha)\mathbb{E}[\xi]},\\ \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}-K_i=\frac{\lambda v_i}{2\alpha\mathbb{E}[\xi_i]}, \qquad i=1,...,n,\\ \sum_{i=1}^n K_i=K.\end{array}\right. \tag{3.2.14}$$

Summing two sides of the equations on the second line in (3.2.14) for $i=1,...,n$ and using the constraint on the third line in (3.2.14), we obtain

$$\sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}-K=\frac{\lambda}{2\alpha}\sum_{i=1}^n\frac{v_i}{\mathbb{E}[\xi_i]},$$

which, together with the equation on the first line in (3.2.14), yields

$$\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}-\sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}=-\frac{\lambda v}{2(1-\alpha)\mathbb{E}[\xi]}-\frac{\lambda}{2\alpha}\sum_{i=1}^n\frac{v_i}{\mathbb{E}[\xi_i]}.$$

From the above equation, we obtain

$$\begin{aligned}-\lambda &= \frac{2\alpha(1-\alpha)\mathbb{E}[\xi]}{\alpha v+(1-\alpha)\mathbb{E}[\xi]\sum_{i=1}^n\frac{v_i}{\mathbb{E}[\xi_i]}}\Big(\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}-\sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}\Big)\\ &= \frac{2(1-\alpha)\mathbb{E}[\xi]\beta(\alpha)}{v}\Big(\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}-\sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}\Big).\end{aligned} \tag{3.2.15}$$

Thus, using (3.2.15) and the equation on the first line in (3.2.14), we obtain

$$K=\tilde{\beta}(\alpha)\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}+\beta(\alpha)\sum_{i=1}^n\frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]},$$

which yields the first expression (3.2.4) for K^*. Then, the second expression (3.2.5) for K^* follows from (3.2.4) and $\beta(\alpha)=1-\tilde{\beta}(\alpha)$. Furthermore, using (3.2.15) and the equation on the second line in (3.2.14), we obtain

$$\begin{aligned} K_i &= \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} - \frac{\lambda v_i}{2\alpha \mathbb{E}[\xi_i]} \\ &= \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} + \frac{v_i}{2\alpha \mathbb{E}[\xi_i]} \times \frac{2(1-\alpha)\mathbb{E}[\xi]\beta(\alpha)}{v} \Big(\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} - \sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} \Big) \\ &= \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} + \beta_i(\alpha) \Big(\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} - \sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} \Big), \end{aligned}$$

which yields the first expression (3.2.6) for K_i^*. Furthermore, by (3.2.5), we have

$$\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} - \sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} = \frac{1}{\tilde{\beta}(\alpha)} \Big(K^* - \sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} \Big),$$

which, together with the first expression (3.2.6) for K_i^*, yields that the second expression (3.2.7) for K_i^*. It completes the proof of Theorem 3.2.2. □

Remark 3.2.3. By (3.2.8) and (3.2.9), we observe that $0 < \beta(\alpha)$, $\tilde{\beta}(\alpha)$, $\beta_i(\alpha) < 1$, $i = 1, ..., n$, and $\tilde{\beta}(\alpha) = \sum_{i=1}^{n} \beta_i(\alpha)$. From the first expression (3.2.4) for K^*, we see that the optimal required total capital K^* is between $\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}$ (the expected weighted aggregate risk) and $\sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}$ (the total of the expected weighted individual risks). From the first expression (3.2.6) for K_i^*, we see that the optimal capital K_i^* allocated to business lines i is equal to the base capital $\frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}$, which is the expected weighted risk for business line i, plus a loading or backup capital $\beta_i(\alpha)\Big(\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} - \sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}\Big)$, which is proportional to the difference between the expected weighted aggregate risk and the total of the expected weighted individual risks with the loading factor $\beta_i(\alpha)$. While from the second expression (3.2.5) for K^*, we see that the optimal required total capital K^* is equal to the total of the base capital for individual risks plus a loading or backup capital $\beta(\alpha)\Big(\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} - \sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}\Big)$, which is proportional to the difference between the expected weighted aggregate risk and the total of the expected weighted individual risks with the loading factor $\beta_i(\alpha) = \sum_{i=1}^{n} \beta_i(\alpha)$, which is the total of the loading factors for individual risks.

We also point out that the expected weighted aggregate risk $\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}$ is not necessarily equal to the total of the expected weighted individual risks $\sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}$. The optimal required total capital K^* for the company satisfies that if $\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} < \sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}$, then $\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} < K^* < \sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}$; if $\sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} < \frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}$, then $\sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} < K^* < \frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}$; and if $\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} = \sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}$, then $K^* = \frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}$. The relation between $\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}$ and $\sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}$ depends on the decision maker's attitudes toward the aggregate risks and individual risks, which are represented by the weighting factors ξ, ξ_i, $i = 1, ..., n$. □

3.2.1 Special or limiting cases of Theorem 3.2.2

In this subsection, we consider some special or limiting cases of Theorem 3.2.2 and show that Theorem 3.2.2 can yield many interesting results about how to determine the optimal required total capital of a company and how to allocate the optimal required total capital among its main business lines at the same time.

Proposition 3.2.4. *(i) In Theorem 3.2.2, if*

$$v_i = \mathbb{E}[\xi_i X_i], \quad i = 1, ..., n,$$

then the optimal allocation is reduced to the following proportional allocation principle

$$K_i^* = \frac{\frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}}{\sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}} \times K^*, \tag{3.2.16}$$

where the optimal required total capital K^* is given in Theorem 3.2.2.

(ii) In Theorem 3.2.2, if

$$v = \mathbb{E}[\xi] \sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}, \qquad v_i = \mathbb{E}[\xi_i X_i], \qquad i = 1, ..., n, \tag{3.2.17}$$

then the optimal allocation is the proportional allocation principle (3.2.16) and the optimal required total capital K^* is given

$$K^* = (1-\alpha) \frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} + \alpha \sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}. \tag{3.2.18}$$

Proof. Under the condition that $v_i = \mathbb{E}[\xi_i X_i]$, $i = 1, ..., n$, by (3.2.7), we have

$$\begin{aligned} K_i^* &= \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} + \frac{\frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}}{\sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}} \left(K^* - \sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} \right) \\ &= \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} + \frac{\frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}}{\sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}} K^* - \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} \\ &= \frac{\frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}}{\sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}} K^*, \qquad i = 1, ..., n. \end{aligned}$$

Further, if the conditions in (3.2.17) holds, then $\sum_{i=1}^n \frac{v_i}{\mathbb{E}[\xi_i]} = \sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} = \frac{v}{\mathbb{E}[\xi]}$, and hence $\beta(\alpha)$ and $\beta_i(\alpha)$ defined in (3.2.8) and (3.2.9) are reduced to $\beta(\alpha) = 1 - \tilde{\beta}(\alpha) = \alpha$. Thus, (3.2.4) is reduced to (3.2.18). □

Remark 3.2.5. Proportional allocation principle (3.2.16) has the same proportional weights as the proportional allocation principle given in Proposition 3.11 of Cai and Wang (2020). However, in our proportional allocation principle (3.2.16), the optimal required total capital K^* is specified as well. We now discuss applications of Proposition 3.2.4 and give interesting samples of the optimal required total capital and the proportional allocation principle.

For any finite-valued mappings/risk measures $\rho(S)$, $\rho_i(X_i)$, $i = 1, ..., n$, if there exist sets A, $A_i \in \mathbb{R}$, $i = 1, ..., n$, satisfying

$$\mathbb{E}(S|S \in A) = \rho(S), \quad \mathbb{E}(X_i|X_i \in A_i) = \rho_i(X_i), \quad i = 1, ..., n,$$

then by setting

$$\xi = \frac{\mathrm{I}_{\{S \in A\}}}{\mathbb{P}\{S \in A\}}, \quad \xi_i = \frac{\mathrm{I}_{\{X_i \in A_i\}}}{\mathbb{P}\{X_i \in A_i\}}, \quad v_i = \mathbb{E}[\xi_i X_i] \quad i = 1, ..., n, \quad v = \sum_{i=1}^{n} v_i, \tag{3.2.19}$$

in Proposition 3.2.4, we see that $\mathbb{E}[\xi] = 1$, $\mathbb{E}[\xi_i] = 1$, $i = 1, ..., n$, the conditions in (3.2.17) hold, and that

$$\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} = \mathbb{E}[S|S \in A] = \rho(S), \quad \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} = \mathbb{E}[X_i|X_i \in A_i] = \rho_i(X_i), \quad i = 1, ..., n.$$

Thus, (3.2.16) is reduced to the following proportional principle in terms of risk measures:

$$K_i^* = \frac{\rho_i(X_i)}{\sum_{i=1}^{n} \rho_i(X_i)} \times K^*, \tag{3.2.20}$$

and the optimal required total capital K^* in 3.2.18 is reduced to

$$K^* = (1 - \alpha)\rho(S) + \alpha \sum_{i=1}^{n} \rho_i(X_i). \tag{3.2.21}$$

This proportional principle not only gives the optimal proportional weights, but also specifies the optimal required total capital K^*, which is equal to the weighted sum of the risk measure of the aggregate risk S and the total of the risk measures of individual risks. $\square$

Remark 3.2.6. Assume that $X_1, ..., X_n, S$ are continuous random variables.

(a) By Proposition 3.7 of Cai and Wang (2020), we know that there exist a set $A \subset \mathbb{R}$ satisfying $\mathbb{E}(S|S \in A) = \mathrm{VaR}_q(S)$, furthermore, we take $A_i = \{S > \mathrm{VaR}_q(S)\}$, $i = 1, ..., n$, thus, (3.2.20) and (3.2.21) are reduced to the following CTE allocation principle

$$K_i^* = \frac{\mathbb{E}(X_i|S > \mathrm{VaR}_q(S))}{\mathrm{CTE}_q(S)} \times K^*, \tag{3.2.22}$$

where

$$K^* = (1-\alpha)\mathrm{VaR}_q(S) + \alpha \mathrm{CTE}_q(S). \tag{3.2.23}$$

The optimal total capital criterion (3.2.23) and the corresponding CTE allocation principle (3.2.22) are very interesting. As pointed out in Belles-Sampera et al. (2014), if the required total capital for a company is determined by VaR, this capital may not be sufficient for buffering the impact of a potential right-tail risk. If the required total capital for a company is determined by CTE, this capital may be too conservative or too much. A reasonable required total capital for a company might be between VaR and CTE. Belles-Sampera et al. (2014) proposed a risk measure called GlueVaR, which is between VaR and CTE, for the purpose of determining the required capital. Total capital criterion (3.2.23) gives a simple way to determine the optimal required total capital. If the decision maker puts more weight (small α) on the aggregate risk, the required total capital will be close to VaR. If the decision maker puts more weight (large α) on individual risks, the required total capital will be close to CTE. Hence, using VaR or CTE as the reqired total capital for a company represents the decision maker's attitude towards the aggregate risk and individual risks. In general, our results (3.2.23) and (3.2.22) say the optimal scheme to allocate an optimal required total capital amounting between VaR and CTE, is the CTE allocation principle.

(b) In addition, by Proposition 3.7 of Cai and Wang (2020), we know that there exist sets A, $A_i \subset \mathbb{R}$, $i = 1, ..., n$, satisfying $\mathbb{E}(S|S \in A) = \mathrm{VaR}_q(S)$ and $\mathbb{E}(X_i|X_i \in A_i) = \mathrm{VaR}_q(X_i)$, $i = 1, ..., n$. Thus, (3.2.20) and (3.2.21) are reduced to the following haircut principle

$$K_i^* = \frac{\mathrm{VaR}_q(X_i)}{\sum_{i=1}^n \mathrm{VaR}_q(X_i)} \times K^*, \tag{3.2.24}$$

where

$$K^* = (1-\alpha)\mathrm{VaR}_q(S) + \alpha \sum_{i=1}^n \mathrm{VaR}_q(X_i). \tag{3.2.25}$$

(c) If we take sets $A = \{S > \mathrm{VaR}_q(S)\}$ and $A_i \subset \mathbb{R}$ so that $\mathbb{E}(S|S \in A) = \mathrm{CTE}_q(S)$ and $\mathbb{E}(X|X \in A_i) = \mathrm{VaR}_q(X_i)$, $i = 1, ..., n$, then, (3.2.20) and (3.2.21) are reduced to the following haircut principle

$$K_i^* = \frac{\mathrm{VaR}_q(X_i)}{\sum_{i=1}^n \mathrm{VaR}_q(X_i)} \times K^*, \tag{3.2.26}$$

where

$$K^* = (1-\alpha)\mathrm{CTE}_q(S) + \alpha \sum_{i=1}^n \mathrm{VaR}_q(X_i). \tag{3.2.27}$$

Note that the optimal required total capital K^* in (3.2.26) and (3.2.24) are different and they are determined by (3.2.27) and (3.2.25), respectively.

(d) If we take sets $A = \{S > \mathrm{VaR}_q(S)\}$ and $A_i = \{X_i > \mathrm{VaR}_q(X_i)\}$, $i = 1, ..., n$, then, (3.2.20) and (3.2.21) are reduced to the following CTE principle

$$K_i^* = \frac{\mathrm{CTE}_q(X_i)}{\sum_{i=1}^n \mathrm{CTE}_q(X_i)} \times K^*, \tag{3.2.28}$$

where

$$K^* = (1-\alpha)\mathrm{CTE}_q(S) + \alpha \sum_{i=1}^n \mathrm{CTE}_q(X_i). \tag{3.2.29}$$

□

Example 3.2.1. Similarly the assumptions used in Dhaene, Tsanakas, et al. (2012), in Theorem 3.2.2, if $\mathbb{E}[\xi] = 1$, $\mathbb{E}[\xi_i] = 1$, $\sum_{i=1}^n v_i = v$, $i = 1, ..., n$, then, by (3.2.8), we have $\beta(\alpha) = 1 - \tilde{\beta}(\alpha) = \alpha$, and the optimal required total capital and the optimal allocation are reduced to

$$\begin{aligned} K^* &= (1-\alpha)\,\mathbb{E}[\xi S] + \alpha \sum_{i=1}^n \mathbb{E}[\xi_i X_i] \\ K_i^* &= \mathbb{E}[\xi_i X_i] + \frac{v_i}{v}\left(K^* - \sum_{i=1}^n \mathbb{E}[\xi_i X_i]\right), \qquad i = 1, ..., n. \end{aligned}$$

We point out that the optimal allocation expressions for K_i^* are the main result of Theorem 1 in Zaks and Tsanakas (2014) when the required total capital is calculated as $(1-\alpha)\,\mathbb{E}[\xi S] + \alpha \sum_{i=1}^n \mathbb{E}[\xi_i X_i]$. □

Example 3.2.2. In Theorem 3.2.2, if $\xi_i = \xi$, $i = 1, ..., n$, then the optimal required total capital and the optimal allocation are reduced to

$$\begin{cases} K^* = \frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}, \\ K_i^* = \frac{\mathbb{E}[\xi X_i]}{\mathbb{E}[\xi]}, \quad i = 1, ..., n. \end{cases} \tag{3.2.30}$$

In this special case, we notice that the exposure parameters v and v_i and the preference parameter α do not affect the optimal solution. In fact, from the proof of Theorem 3.2.2, we find that if $\xi_i = \xi$, $i = 1, ..., n$, then, from (3.2.15), we have $\lambda = 0$, which, together with

(3.2.14), yields (3.2.30) and shows how the effects of the exposure parameters v and v_i and the preference parameter α disappear in the optimal solution. Intuitively, when $\xi_i = \xi$, $i = 1, ..., n$, the sum $\sum_{i=1}^n \xi_i X_i$ of the weighted losses from all the business lines is equal to the weighted loss ξS of the enterprise, and the decision maker's attitudes towards all the business lines and the enterprise are the same. Hence, the exposure parameters v and v_i and the preference parameter α do not affect the optimal solution.

In particular, if $\xi_i = \xi = \frac{\mathrm{I}_{\{S>\mathrm{VaR}_q(S)\}}}{\mathbb{P}\{S>\mathrm{VaR}_q(S)\}}$, $i = 1, ..., n$, then (3.2.30) is reduced to the well-known CTE principle:

$$\begin{cases} K^* &= \mathrm{CTE}_q(S), \\ K_i^* &= \mathbb{E}[X_i | S > \mathrm{VaR}_q(S)], \quad i = 1, ..., n. \end{cases} \tag{3.2.31}$$

In this principle, the required total capital is determined by CTE of the aggregate risk and the capital allocated to a business lines is the conditional expected risk of the business line conditioning on that the aggregate risk of a company exceeds its VaR. Hence, the CTE allocation principle is a special case of our model.

In addition, from (3.2.30), we find that many of the existing multivariate CTE risk measures can be viewed as the special cases of Theorem 3.2.2 when $\xi_i = \xi$, $i = 1, ..., n$. In fact, if $\xi = \xi_i = \mathrm{I}_A$, $i = 1, ..., n$, where A is an event associated with $X_1, ..., X_n, S$, then (3.2.30) is reduced to

$$\begin{cases} K^* &= \sum_{i=1}^n \mathbb{E}[X_i | A], \\ K_i^* &= \mathbb{E}[X_i | A], \quad i = 1, ..., n. \end{cases} \tag{3.2.32}$$

In the followings, we list several interesting examples of the event A.

(i) If

$$A = \{X_1 > \mathrm{VaR}_q(X_1), ..., X_n > \mathrm{VaR}_q(X_n)\}, \tag{3.2.33}$$

then (3.2.32) recovers the MCTE risk measure of Landsman et al. (2016). In this case, event A means that the losses of all the n business lines exceed the respective threshold levels that are determined by VaR_q.

(ii) If

$$A = \{X_1 > \mathrm{VaR}_q(X_1)\} \cup \cdots \cup \{X_n > \mathrm{VaR}_q(X_n)\}, \tag{3.2.34}$$

then (3.2.32) recovers the example of Cai et al. (2017). In this case, event A means that the loss of at least one of the n business lines exceeds the its threshold level that is determined by VaR_q.

(iii) Moreover, we can also consider the event A as the joint situation of the business lines and the company such as:

$$A = \{X_1 > \mathrm{VaR}_q(X_1), ..., X_n > \mathrm{VaR}_q(X_n), S > \mathrm{VaR}_q(S)\} \tag{3.2.35}$$

and

$$A = \{X_1 > \mathrm{VaR}_q(X_1)\} \cup \cdots \cup \{X_n > \mathrm{VaR}_q(X_n)\} \cup \{S > \mathrm{VaR}_q(S)\}. \tag{3.2.36}$$

All of the four extreme events of A involve all of the business lines or individual risks. In those scenarios, the optimal required total capital for the company is the conditional expectation of the aggregate risk of the company conditioning on occurrence of the extreme event A, and the optimal capital allocated to business line i or individual risk X_i is the conditional expected risk of this business line conditioning on occurrence of the same extreme event A. □

Example 3.2.3. (i) In Theorem 3.2.2, if

$$\xi = \mathrm{I}_{\{S>\mathrm{VaR}_q(S)\}}, \quad \xi_i = \mathrm{I}_{\{X_i>\mathrm{VaR}_{q_i}(X_i)\}}, \quad v_i = \mathbb{E}[\xi_i]\,\mathbb{E}[X_i] > 0, \quad v = \mathbb{E}[\xi]\,\mathbb{E}[S] > 0,$$

where $0 < q < 1$, $0 < q_i < 1$, $i = 1, ..., n$, then, the optimal required total capital and the optimal allocation are reduced to

$$\begin{cases} K^* &= (1-\alpha)\mathrm{CTE}_q(S) + \alpha \sum_{i=1}^n \mathrm{CTE}_{q_i}(X_i) \\ K_i^* &= \mathrm{CTE}_{q_i}(X_i) + (1-\alpha)\frac{\mathbb{E}[X_i]}{\mathbb{E}[S]} \times \Big(\mathrm{CTE}_q(S) - \sum_{i=1}^n \mathrm{CTE}_{q_i}(X_i)\Big), \quad i = 1, ..., n. \end{cases} \tag{3.2.37}$$

In this example, the optimal required total capital K^* is a weighted sum of the CTE of the aggregated risk at the required confidence level q for the company and the total of the CTEs of the individual risks at their own required confidence levels q_i, $i = 1, ..., n$. The optimal capital K_i^* allocated to business line i is equal to the base capital $\mathrm{CTE}_{q_i}(X_i)$, which is the CTE of the risk of business line i at the required confidence level q_i for business line i, plus a loading or backup capital $(1-\alpha)\frac{\mathbb{E}[X_i]}{\mathbb{E}[S]} \times \Big(\mathrm{CTE}_q(S) - \sum_{i=1}^n \mathrm{CTE}_{q_i}(X_i)\Big)$, which is proportional to the difference between the CTE of the aggregate risk and the total of the CTEs of individual risks with the loading factor $(1-\alpha)\frac{\mathbb{E}[X_i]}{\mathbb{E}[S]}$.

(ii) In Theorem 3.2.2, if

$$\xi = \mathrm{I}_{\{S>\mathrm{VaR}_q(S)\}}, \quad \xi_i = \mathrm{I}_{\{X_i>\mathrm{VaR}_{q_i}(X_i)\}}, \quad v_i = \mathbb{E}[\xi_i]\,Cov(X_i,\, S) > 0, \quad v = \mathbb{E}[\xi]\,Var(S) > 0,$$

where $0 < q < 1$, $0 < q_i < 1$, $i = 1, ..., n$, then, the optimal required total capital and the optimal allocation are reduced to

$$\begin{cases} K^* &= (1-\alpha)\mathrm{CTE}_q(S) + \alpha \sum_{i=1}^n \mathrm{CTE}_{q_i}(X_i), \\ K_i^* &= \mathrm{CTE}_{q_i}(X_i) + (1-\alpha)\frac{Cov(X_i,S)}{Var(S)} \times \Big(\mathrm{CTE}_q(S) - \sum_{i=1}^n \mathrm{CTE}_{q_i}(X_i)\Big), \quad i = 1, ..., n. \end{cases} \tag{3.2.38}$$

(iii) In general, let $\rho(\mathbf{X}, S)$, $h(\mathbf{X}, S)$, $\rho_i(\mathbf{X}, S)$ $h_i(\mathbf{X}, S)$, $i = 1, ..., n$, be mappings/risk measures from $(\mathbf{X}, S)$ to $\mathbb{R}$, where $\mathbf{X} = (X_1, ..., X_n)$ and $\sum_{i=1}^n h_i(\mathbf{X}, S) = h(\mathbf{X}, S)$. Thus, in Theorem 3.2.2, if

$$\xi = \mathrm{I}_{\{S \in A\}}, \quad v = \mathbb{E}[\xi]\, h(\mathbf{X}, S) > 0, \quad \xi_i = \mathrm{I}_{\{X_i \in A_i\}}, \quad v_i = \mathbb{E}[\xi_i]\, h_i(\mathbf{X}, S) > 0, \quad i = 1, ..., n,$$

where the sets $A, A_i \in \mathbb{R}$, $i = 1, ..., n$, satisfy

$$\mathbb{E}(S|S \in A) = \rho(\mathbf{X}, S), \quad \mathbb{E}(X_i|X_i \in A_i) = \rho_i(\mathbf{X}, S), \quad i = 1, ..., n,$$

then, the optimal required total capital and the optimal allocation are reduced to

$$\begin{cases} K^* &= (1-\alpha)\rho(\mathbf{X}, S) + \alpha \sum_{i=1}^n \rho_i(\mathbf{X}, S), \\ K_i^* &= \rho_i(\mathbf{X}, S) + (1-\alpha)\frac{h_i(\mathbf{X}, S)}{h(\mathbf{X}, S)} \times \left(\rho(\mathbf{X}, S) - \sum_{i=1}^n \rho_i(\mathbf{X}, S)\right), \quad i = 1, ..., n. \end{cases} \tag{3.2.39}$$

□

Example 3.2.4. (i) In Theorem 3.2.2, if $\alpha \to 0$, then $\beta(\alpha) \to 0$, $\tilde{\beta}(\alpha) \to 1$, $\beta_i(\alpha) \to \frac{\frac{v_i}{\mathbb{E}[\xi_i]}}{\sum_{i=1}^n \frac{v_i}{\mathbb{E}[\xi_i]}} \triangleq \beta_i(0)$, and

$$\begin{cases} K^* &\to \frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}, \\ K_i^* &\to \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]} + \frac{\frac{v_i}{\mathbb{E}[\xi_i]}}{\sum_{i=1}^n \frac{v_i}{\mathbb{E}[\xi_i]}}\left(\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} - \sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}\right), \quad i = 1, ..., n. \end{cases} \tag{3.2.40}$$

In this limiting case of $\alpha \to 0$, from (3.2.1), we see that the main concern of the decision maker is the deviation between the aggregate risk and the required total capital. The optimal required total capital K^* is approximating to the expected weighted aggregated risk and the optimal capital K_i^* allocated to main business line i is approximating to the base capital $\frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}$, which is the expected weighted risk for business line i, plus a loading $\beta_i(0)\left(\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]} - \sum_{i=1}^n \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}\right)$, which is proportional to the difference between the expected weighted aggregate risk and the total of the expected weighted individual risks with the loading factor $\beta_i(0)$. By (3.2.40), we see Theorem 1 of Dhaene, Tsanakas, et al. (2012) is the limiting case of Theorem 3.2.2 when $\alpha \to 0$ and the given total capital K in Theorem 1 of Dhaene, Tsanakas, et al. (2012) is determined by the expected weighted aggregate risk $\frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}$. In addition, we notice that the exposure parameter v does not affect the optimal solution in this limiting case. In fact, from Theorem 3.2.2, we find that the exposure parameter v affect the optimal solution only through the weights $\beta(\alpha)$, $\tilde{\beta}(\alpha)$, $\tilde{\beta}_i(\alpha)$, $i = 1, ..., n$. Obviously, when $\alpha \to 0$, the exposure parameter v disappears in all the weights $\beta(\alpha)$, $\tilde{\beta}(\alpha)$, $\tilde{\beta}_i(\alpha)$, $i = 1, ..., n$, in the limiting case.

(ii) In Theorem 3.2.2, if $\alpha \to 1$, then $\beta(\alpha) \to 1$, $\tilde{\beta}(\alpha) \to 0$, $\beta_i(\alpha) \to 0$, and

$$\begin{cases} K^* & \to \sum_{i=1}^{n} \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}, \\ K_i^* & \to \frac{\mathbb{E}[\xi_i X_i]}{\mathbb{E}[\xi_i]}, \quad i = 1, ..., n. \end{cases} \tag{3.2.41}$$

In this limiting case of $\alpha \to 1$, from (3.2.1), we see that the main concern of the decision maker is the total of the allocation deviations among individual risks and the corresponding allocated capital. The optimal required total capital K^* is approximating to the total of the expected weighted individual risks and the optimal capital K_i^* allocated to main business line i is approximating to the expected weighted risk of the business line. In this limiting case, we notice that all the exposure parameters v, v_i, $i = 1, ..., n$, have no effect in the optimal solution in the limiting case. In fact, from Theorem 3.2.2, we find that the exposure parameters v, v_i, $i = 1, ..., n$, affect the optimal solution only through the weights $\beta(\alpha)$, $\tilde{\beta}(\alpha)$, $\tilde{\beta}_i(\alpha)$, $i = 1, ..., n$. Obviously, when $\alpha \to 1$, the exposure parameters v, v_i, $i = 1, ..., n$, disappear in all the weights $\beta(\alpha)$, $\tilde{\beta}(\alpha)$, $\tilde{\beta}_i(\alpha)$, $i = 1, ..., n$, in the limiting case. □

3.3 Optimal solutions based on weighted squared errors for a company with main business lines and sub-business lines

In this section, we discuss optimal required total capital and optimal allocation scheme for a company with main business lines and sub-business lines and consider problem (3.1.4) when $D_1(x) = x^2$ and $D_{2i} = x^2$, $i = 1, ..., n$, namely the allocation deviations are measured by (weighted) squared errors. In this case, problem (3.1.4) is reduced to the following problem:

$$\begin{cases} \min_{(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n}) \in \mathbb{R}^d} & \Big\{ \alpha_1 \mathbb{E}\Big[\frac{\psi}{w}\Big(S - K\Big)^2\Big] + \alpha_2 \sum_{i=1}^{n} \mathbb{E}\Big[\frac{\psi_i}{w_i}(X_i - K_i)^2\Big] \\ & \quad + \alpha_3 \sum_{i=1}^{n} \sum_{j=1}^{n_i} \mathbb{E}\Big[\frac{\psi_{ij}}{w_{ij}}(X_{ij} - k_{ij})^2\Big]\Big\} \\ \text{s.t.} & \sum_{i=1}^{n} K_i = K, \quad \sum_{j=1}^{n_i} k_{ij} = K_i, \quad i = 1, ..., n. \end{cases} \tag{3.3.1}$$

To guarantee that problem (3.3.1) has a unique solution $(\mathbf{K}^*, \mathbf{k}_1^*, ..., \mathbf{k}_n^*)$, we assume that the following conditions hold.

Assumption 3.3.1. For problem (3.3.1), we assume that real numbers w, w_i, w_{ij}, and random variables ψ, ψ_i, ψ_{ij}, $i = 1, ..., n$, $j = 1, ..., n_i$, satisfy

$$\frac{\psi}{w} \geq 0, \quad \frac{\psi_i}{w_i} \geq 0, \quad \frac{\psi_{ij}}{w_{ij}} \geq 0, \quad \frac{\mathbb{E}[\psi]}{w} > 0, \quad \frac{\mathbb{E}[\psi_i]}{w_i} > 0, \quad \frac{\mathbb{E}[\psi_{ij}]}{w_{ij}} > 0, \quad i = 1, ..., n, \ j = 1, ..., n_i, \tag{3.3.2}$$

and expectations $\mathbb{E}[\frac{\psi}{w}S^2]$, $\mathbb{E}[\frac{\psi}{w}S]$, $\mathbb{E}[\frac{\psi_i}{w_i}X_i^2]$, $\mathbb{E}[\frac{\psi_i}{w_i}X_i]$, $\mathbb{E}[\frac{\psi_{ij}}{w_{ij}}X_{ij}^2]$, $\mathbb{E}[\frac{\xi_{ij}}{w_{ij}}X_{ij}]$, $i = 1, ..., n$, $j = 1, ..., n_i$, exist. □

Denote the objective function in problem (3.3.1) by $J(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n})$, namely,

$$J(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n}) = \alpha_1 \mathbb{E}\left[\frac{\psi}{w}\Big(S-K\Big)^2\right] + \alpha_2 \sum_{i=1}^{n} \mathbb{E}\left[\frac{\psi_i}{w_i}(X_i - K_i)^2\right] + \alpha_3 \sum_{i=1}^{n}\sum_{j=1}^{n_i} \mathbb{E}\left[\frac{\psi_{ij}}{w_{ij}}(X_{ij} - k_{ij})^2\right]. \tag{3.3.3}$$

Lemma 3.3.1. *Under Assumption 3.3.1, the objective function $J(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n})$ defined in (3.3.3) is a convex and coercive function of $(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n})$ in $\mathbb{R}^d$, where $d = 1 + n + \sum_{i=1}^n n_i$.*

Proof. Clearly, under Assumption 3.3.1, $\mathbb{E}\left[\frac{\psi}{w}\Big(S-K\Big)^2\right]$, $\mathbb{E}\left[\frac{\psi_i}{w_i}\Big(X_i-K_i\Big)^2\right]$, and $\mathbb{E}\left[\frac{\psi_{ij}}{w_{ij}}\Big(X_{ij}-k_{ij}\Big)^2\right]$ are convex and coercive functions of K, K_i, k_{ij}, respectively. Hence, $J(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n})$ defined in (3.3.3) is a convex and coercive function of $(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n})$ on $\mathbb{R}^d$. □

Theorem 3.3.2. *Under Assumption 3.3.1, problem (3.3.1) has the following unique solution:*

$$K^* = \delta(\alpha_1,\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} + \tilde{\delta}(\alpha_1,\alpha_2,\alpha_3)\sum_{i=1}^{n}\Big(\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\Big),$$

$$\begin{aligned} K_i^* &= \beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} \\ &\quad + \theta_i(\alpha_1,\alpha_2,\alpha_3)\Big(\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} - \sum_{i=1}^{n}\Big(\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\Big)\Big), \\ &= \beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} \\ &\quad + \frac{\theta_i(\alpha_1,\alpha_2,\alpha_3)}{\delta(\alpha_1,\alpha_2,\alpha_3)}\Big(K^* - \sum_{i=1}^{n}\Big(\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\Big)\Big), \quad i = 1, ..., n, \end{aligned}$$

$$k_{ij}^* = \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} + \frac{\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}{\sum_{j=1}^{n_i}\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}\Big(K_i^* - \sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\Big), \quad i = 1, ..., n, \quad j = 1, ..., n_i,$$

where

$$\tilde{\delta}(\alpha_1,\alpha_2,\alpha_3) = 1 - \delta(\alpha_1,\alpha_2,\alpha_3) = \frac{\alpha_2\frac{w}{\mathbb{E}(\psi)}}{\alpha_2\frac{w}{\mathbb{E}(\psi)} + \alpha_1\sum_{i=1}^{n}\beta_i(\alpha_2,\alpha_3)\frac{w_i}{\mathbb{E}(\psi_i)}}, \tag{3.3.4}$$

$$\tilde{\beta}_i(\alpha_2,\alpha_3) = 1 - \beta_i(\alpha_2,\alpha_3) = \frac{\alpha_3\frac{w_i}{\mathbb{E}(\psi_i)}}{\alpha_3\frac{w_i}{\mathbb{E}(\psi_i)} + \alpha_2\sum_{j=1}^{n_i}\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}, \tag{3.3.5}$$

$$\theta_i(\alpha_1,\alpha_2,\alpha_3) = \frac{\beta_i(\alpha_2,\alpha_3)}{\alpha_2\frac{w}{\mathbb{E}(\psi)} + \alpha_1\sum_{i=1}^{n}\beta_i(\alpha_2,\alpha_3)\frac{w_i}{\mathbb{E}(\psi_i)}} \times \frac{\alpha_1 w_i}{\mathbb{E}[\psi_i]}. \tag{3.3.6}$$

Proof. By Lemma 3.3.1, we know that problem (3.3.1) is a constrained convex optimization problem and that the objective function of the problem is convex and coercive. Hence, optimal solutions for problem (3.3.1) do exist. Denote the Lagrangian of problem (3.3.1) by

$$L(\mathbf{K},\mathbf{k_1},...,\mathbf{k_n},\lambda,\lambda_1,...,\lambda_n) = \alpha_1\mathbb{E}\Big[\frac{\psi}{w}\Big(S-K\Big)^2\Big] + \alpha_2\sum_{i=1}^{n}\mathbb{E}\Big[\frac{\psi_i}{w_i}\,(X_i-K_i)^2\Big]$$
$$+\,\alpha_3\sum_{i=1}^{n}\sum_{j=1}^{n_i}\mathbb{E}\Big[\frac{\psi_{ij}}{w_{ij}}\,(X_{ij}-k_{ij})^2\Big] + \lambda\Big(\sum_{i=1}^{n}K_i-K\Big) + \sum_{i=1}^{n}\lambda_i\Big(K_i-\sum_{j=1}^{n_i}k_{ij}\Big).$$

Under Assumption 3.3.1, it is easy to see that the Lagrangian function $L(\mathbf{K},\mathbf{k_1},...,\mathbf{k_n},\lambda,\lambda_1,...,\lambda_n)$ is differentiable. Hence, a solution to problem (3.3.1) is any solution to the following equations or any solution satisfying the following Karush-Kuhn-Tucker (KKT) conditions:

$$\begin{cases} \frac{\partial L}{\partial K}=0, \\ \frac{\partial L}{\partial K_i}=0, \quad i=1,...,n, \\ \frac{\partial L}{\partial k_{ij}}=0, \quad i=1,...,n,\ j=1,...,n_i, \\ \sum_{i=1}^{n}K_i=K, \quad \sum_{j=1}^{n_i}k_{ij}=K_i, \quad i=1,...,n. \end{cases} \tag{3.3.7}$$

It is easy to see that (3.3.7) is equivalent to

$$\begin{cases} -2\alpha_1\mathbb{E}\Big[\frac{\psi}{w}\Big(S-K\Big)\Big]-\lambda=0, \\ -2\alpha_2\mathbb{E}\Big[\frac{\psi_i}{w_i}\Big(X_i-K_i\Big)\Big]+\lambda+\lambda_i=0, \quad i=1,...,n, \\ -2\alpha_3\mathbb{E}\Big[\frac{\psi_{ij}}{w_{ij}}\Big(X_i-k_{ij}\Big)\Big]-\lambda_i=0, \quad i=1,...,n, \quad j=1,...,n_i \\ \sum_{i=1}^{n}K_i=K, \quad \sum_{j=1}^{n_i}k_{ij}=K_i, \quad i=1,...,n. \end{cases} \tag{3.3.8}$$

The above equations are further equivalent to

$$\begin{cases} \frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]}-K=-\frac{\lambda w}{2\alpha_1\mathbb{E}[\psi]}, \\ \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]}-K_i=\frac{(\lambda+\lambda_i)w_i}{2\alpha_2\mathbb{E}[\psi_i]}, \quad i=1,...,n, \\ \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}-k_{ij}=-\frac{\lambda_i w_{ij}}{2\alpha_3\mathbb{E}[\psi_{ij}]}, \quad i=1,...,n, \quad j=1,...,n_i, \\ \sum_{i=1}^{n}K_i=K, \quad \sum_{j=1}^{n_i}k_{ij}=K_i, \quad i=1,...,n. \end{cases} \tag{3.3.9}$$

Summing two sides of the equations on the third line in (3.3.9) for $j=1,...,n_i$ and using the constraint on the fourth line in (3.3.9), we obtain

$$\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}-K_i=-\frac{\lambda_i}{2\alpha_3}\sum_{j=1}^{n_i}\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}, \tag{3.3.10}$$

which yields

$$\lambda_i=-\frac{2\alpha_3\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}}{\sum_{j=1}^{n_i}\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}+\frac{2\alpha_3K_i}{\sum_{j=1}^{n_i}\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}=-\frac{2\alpha_3h_i}{g_i}+\frac{2\alpha_3K_i}{g_i}, \tag{3.3.11}$$

where

$$h_i = \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]}, \qquad g_i = \sum_{j=1}^{n_i} \frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}.$$

Substituting (3.3.11) into the second line of (3.3.9), we get

$$\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} - K_i = \frac{\lambda w_i}{2\alpha_2 \mathbb{E}[\psi_i]} - \frac{h_i \alpha_3 w_i}{g_i \alpha_2 \mathbb{E}[\psi_i]} + \frac{\alpha_3 w_i}{g_i \alpha_2 \mathbb{E}(\psi_i)} K_i.$$

Rearranging the above equation, we get,

$$\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} - \left(1 + \frac{\alpha_3 w_i}{g_i \alpha_2 \mathbb{E}(\psi_i)}\right) K_i = \frac{\lambda w_i}{2\alpha_2 \mathbb{E}[\psi_i]} - \frac{h_i \alpha_3 w_i}{g_i \alpha_2 \mathbb{E}[\psi_i]}. \tag{3.3.12}$$

Note that

$$\frac{1}{1 + \frac{\alpha_3 w_i}{g_i \alpha_2 \mathbb{E}(\psi_i)}} = \frac{\alpha_2 \mathbb{E}(\psi_i) \sum_{j=1}^{n_i} \frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}{\alpha_3 w_i + \alpha_2 \mathbb{E}(\psi_i) \sum_{j=1}^{n_i} \frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}} = \beta_i(\alpha_2, \alpha_3) = 1 - \tilde{\beta}_i(\alpha_2, \alpha_3),$$

where $\beta_i(\alpha_2, \alpha_3)$ and $\tilde{\beta}_i(\alpha_2, \alpha_3)$ are defined in (3.3.5). Thus, equation (3.3.12) is reduced to

$$\beta_i(\alpha_2, \alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} - K_i = \beta_i(\alpha_2, \alpha_3) \left(\frac{\lambda w_i}{2\alpha_2 \mathbb{E}[\psi_i]} - \frac{h_i \alpha_3 w_i}{g_i \alpha_2 \mathbb{E}[\psi_i]} \right), \quad i = 1, ..., n.$$

Summing two sides of the above equations for $i = 1, ..., n$ and using the constraint on the fourth line in (3.3.9), we obtain

$$\sum_{i=1}^{n} \beta_i(\alpha_2, \alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} - K = \sum_{i=1}^{n} \beta_i(\alpha_2, \alpha_3) \left(\frac{\lambda w_i}{2\alpha_2 \mathbb{E}[\psi_i]} - \frac{h_i \alpha_3 w_i}{g_i \alpha_2 \mathbb{E}[\psi_i]} \right),$$

which, together with the equation on the first line in (3.3.9), yields

$$\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} - \sum_{i=1}^{n} \beta_i(\alpha_2, \alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} = -\frac{\lambda w}{2\alpha_1 \mathbb{E}[\psi]} - \sum_{i=1}^{n} \beta_i(\alpha_2, \alpha_3) \left(\frac{\lambda w_i}{2\alpha_2 \mathbb{E}[\psi_i]} - \frac{h_i \alpha_3 w_i}{g_i \alpha_2 \mathbb{E}[\psi_i]} \right) \tag{3.3.13}$$

Thus, solving (3.3.13) for λ, we get

$$\begin{aligned}
\lambda &= -\left(\frac{w}{2\alpha_1\mathbb{E}(\psi)} + \sum_{i=1}^{n}\beta_i(\alpha_2,\alpha_3)\frac{w_i}{2\alpha_2\mathbb{E}(\psi_i)}\right)^{-1} \times \\
&\quad \left(\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} - \sum_{i=1}^{n}\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} - \sum_{i=1}^{n}\beta_i(\alpha_2,\alpha_3)\frac{h_i\alpha_3 w_i}{g_i\alpha_2\mathbb{E}[\psi_i]}\right) \\
&= -\left(\frac{2\alpha_1\mathbb{E}(\psi)}{w + 2\alpha_1\mathbb{E}(\psi)\sum_{i=1}^{n}\beta_i(\alpha_2,\alpha_3)\frac{w_i}{2\alpha_2\mathbb{E}(\psi_i)}}\right) \times \\
&\quad \left(\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} - \sum_{i=1}^{n}\left(\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \beta_i(\alpha_2,\alpha_3)\frac{\alpha_3 w_i\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}}{\alpha_2\mathbb{E}[\psi_i]\sum_{j=1}^{n_i}\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}\right)\right) \\
&= -\left(\frac{2\alpha_1\mathbb{E}(\psi)}{w + 2\alpha_1\mathbb{E}(\psi)\sum_{i=1}^{n}\beta_i(\alpha_2,\alpha_3)\frac{w_i}{2\alpha_2\mathbb{E}(\psi_i)}}\right) \times \\
&\quad \left(\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} - \sum_{i=1}^{n}\left(\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\right)\right), \qquad (3.3.14)
\end{aligned}$$

where, the last equality follows from

$$\frac{\tilde{\beta}_i(\alpha_2,\alpha_3)}{\beta_i(\alpha_2,\alpha_3)} = \frac{\alpha_3 w_i}{\alpha_2\mathbb{E}[\psi_i]\sum_{j=1}^{n_i}\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}.$$

Next, from (3.3.10) and the second line of (3.3.9), we get

$$\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} - \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} = \frac{(\lambda+\lambda_i)w_i}{2\alpha_2\mathbb{E}[\psi_i]} + \lambda_i\sum_{j=1}^{n_i}\frac{w_{ij}}{2\alpha_3\mathbb{E}[\psi_{ij}]},$$

which implies

$$\begin{aligned}
\lambda_i &= \left(\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} - \sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} - \frac{\lambda w_i}{2\alpha_2\mathbb{E}[\psi_i]}\right) \times \left(\frac{w_i}{2\alpha_2\mathbb{E}[\psi_i]} + \sum_{j=1}^{n_i}\frac{w_{ij}}{2\alpha_3\mathbb{E}[\psi_{ij}]}\right)^{-1} \\
&= \left(\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} - \sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} - \frac{\lambda w_i}{2\alpha_2\mathbb{E}[\psi_i]}\right) \times \frac{2\alpha_2\mathbb{E}[\psi_i]}{w_i + 2\alpha_2\mathbb{E}[\psi_i]\sum_{j=1}^{n_i}\frac{w_{ij}}{2\alpha_3\mathbb{E}[\psi_{ij}]}}. \qquad (3.3.15)
\end{aligned}$$

Substituting equation (3.3.14) into the first line of equation (3.3.9), we get

$$\begin{aligned}
K &= \frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} - \left(\frac{w}{w + 2\alpha_1\mathbb{E}(\psi)\sum_{i=1}^{n}\beta_i(\alpha_2,\alpha_3)\frac{w_i}{2\alpha_2\mathbb{E}(\psi_i)}}\right) \\
&\quad \times \left(\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} - \sum_{i=1}^{n}\left(\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\right)\right) \\
&= \delta(\alpha_1,\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} + \tilde{\delta}(\alpha_1,\alpha_2,\alpha_3)\sum_{i=1}^{n}\left(\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\right),
\end{aligned}$$

which yields the expression for K^* in Theorem 3.3.2, where $\delta(\alpha_1, \alpha_2, \alpha_3)$ and $\tilde{\delta}(\alpha_1, \alpha_2, \alpha_3)$ are defined in (3.3.4).

Next, we derive the expressions for K_i^* in Theorem 3.3.2. From the second line of equation (3.3.9), we get

$$K_i = \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} - \frac{(\lambda + \lambda_i) w_i}{2\alpha_2 \mathbb{E}[\psi_i]}. \tag{3.3.16}$$

By (3.3.15), we have

$$\begin{aligned}
\lambda + \lambda_i &= \lambda + \Big(\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} - \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]} - \frac{\lambda w_i}{2\alpha_2 \mathbb{E}[\psi_i]}\Big) \times \Big(\frac{2\alpha_2 \mathbb{E}[\psi_i]}{w_i + \alpha_2 \mathbb{E}[\psi_i] \sum_{j=1}^{n_i} \frac{w_{ij}}{\alpha_3 \mathbb{E}[\psi_{ij}]}}\Big) \\
&= \lambda\Big(1 - \frac{w_i}{w_i + \alpha_2 \mathbb{E}[\psi_i] \sum_{j=1}^{n_i} \frac{w_{ij}}{\alpha_3 \mathbb{E}[\psi_{ij}]}}\Big) \\
&\quad + \Big(\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} - \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]}\Big) \times \Big(\frac{2\alpha_2 \mathbb{E}[\psi_i]}{w_i + \alpha_2 \mathbb{E}[\psi_i] \sum_{j=1}^{n_i} \frac{w_{ij}}{\alpha_3 \mathbb{E}[\psi_{ij}]}}\Big),
\end{aligned}$$

which, together with (3.3.16), implies

$$\begin{aligned}
K_i &= \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} - \frac{w_i}{2\alpha_2 \mathbb{E}[\psi_i]} \lambda \left(1 - \frac{w_i}{w_i + \alpha_2 \mathbb{E}[\psi_i] \sum_{j=1}^{n_i} \frac{w_{ij}}{\alpha_3 \mathbb{E}[\psi_{ij}]}}\right) \\
&\quad - \frac{w_i}{2\alpha_2 \mathbb{E}[\psi_i]} \left(\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} - \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]}\right) \times \left(\frac{2\alpha_2 \mathbb{E}[\psi_i]}{w_i + \alpha_2 \mathbb{E}[\psi_i] \sum_{j=1}^{n_i} \frac{w_{ij}}{\alpha_3 \mathbb{E}[\psi_{ij}]}}\right) \\
&= \left(1 - \frac{w_i}{w_i + \alpha_2 \mathbb{E}[\psi_i] \sum_{j=1}^{n_i} \frac{w_{ij}}{\alpha_3 \mathbb{E}[\psi_{ij}]}}\right) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \frac{w_i}{w_i + \alpha_2 \mathbb{E}[\psi_i] \sum_{j=1}^{n_i} \frac{w_{ij}}{\alpha_3 \mathbb{E}[\psi_{ij}]}} \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]} \\
&\quad - \frac{w_i}{2\alpha_2 \mathbb{E}[\psi_i]} \lambda \left(1 - \frac{w_i}{w_i + \alpha_2 \mathbb{E}[\psi_i] \sum_{j=1}^{n_i} \frac{w_{ij}}{\alpha_3 \mathbb{E}[\psi_{ij}]}}\right) \\
&= \beta_i(\alpha_2, \alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2, \alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]} - \frac{w_i \lambda \beta_i(\alpha_2, \alpha_3)}{2\alpha_2 \mathbb{E}[\psi_i]}.
\end{aligned}$$

By (3.3.14), we have

$$
\begin{aligned}
-\frac{w_i\lambda\beta_i(\alpha_2,\alpha_3)}{2\alpha_2\mathbb{E}[\psi_i]} &= \left(\frac{\alpha_1\mathbb{E}(\psi)}{w+\alpha_1\mathbb{E}(\psi)\sum_{i=1}^n\beta_i(\alpha_2,\alpha_3)\frac{w_i}{\alpha_2\mathbb{E}(\psi_i)}}\right)\beta_i(\alpha_2,\alpha_3)\frac{w_i}{\alpha_2\mathbb{E}[\psi_i]} \\
&\quad\times\left(\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]}-\sum_{i=1}^n\left(\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_iX_i]}{\mathbb{E}[\psi_i]}+\tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\right)\right) \\
&= \tilde{\delta}(\alpha_1,\alpha_2,\alpha_3)\beta_i(\alpha_2,\alpha_3)\frac{\alpha_1\mathbb{E}(\psi)w_i}{\alpha_2\mathbb{E}[\psi_i]w} \\
&\quad\times\left(\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]}-\sum_{i=1}^n\left(\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_iX_i]}{\mathbb{E}[\psi_i]}+\tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\right)\right) \\
&= \theta_i(\alpha_1,\alpha_2,\alpha_3)\times\left(\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]}-\sum_{i=1}^n\left(\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_iX_i]}{\mathbb{E}[\psi_i]}+\tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\right)\right),
\end{aligned}
$$

where

$$
\theta_i(\alpha_1,\alpha_2,\alpha_3)=\tilde{\delta}(\alpha_1,\alpha_2,\alpha_3)\,\beta_i(\alpha_2,\alpha_3)\,\frac{\alpha_1\mathbb{E}(\psi)w_i}{\alpha_2\mathbb{E}[\psi_i]w}=\frac{\beta_i(\alpha_2,\alpha_3)}{\alpha_2\frac{w}{\mathbb{E}(\psi)}+\alpha_1\sum_{i=1}^n\beta_i(\alpha_2,\alpha_3)\frac{w_i}{\mathbb{E}(\psi_i)}}\times\frac{\alpha_1w_i}{\mathbb{E}[\psi_i]},
$$

which is the definition $\theta_i(\alpha_1,\alpha_2,\alpha_3)$ given in (3.3.6). Thus, we get

$$
\begin{aligned}
K_i &= \beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_iX_i]}{\mathbb{E}[\psi_i]}+\tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} \\
&\quad+\theta_i(\alpha_1,\alpha_2,\alpha_3)\Big(\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]}-\sum_{i=1}^n\Big(\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_iX_i]}{\mathbb{E}[\psi_i]}+\tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\Big)\Big),
\end{aligned}
$$

which yields the first expression for K_i^* in Theorem 3.3.2.

By replacing $\tilde{\delta}(\alpha_1,\alpha_2,\alpha_3)$ by $1-\delta(\alpha_1,\alpha_2,\alpha_3)$ in the expression of K^* in Theorem 3.3.2, we obtain

$$
\begin{aligned}
&\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]}-\sum_{i=1}^n\Big(\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_iX_i]}{\mathbb{E}[\psi_i]}+\tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\Big) \\
&= \frac{1}{\delta(\alpha_1,\alpha_2,\alpha_3)}\Big(K^*-\sum_{i=1}^n\Big(\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_iX_i]}{\mathbb{E}[\psi_i]}+\tilde{\beta}_i(\alpha_2,\alpha_3)\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\Big)\Big),
\end{aligned}
$$

which, together with the first expression of K_i^* in Theorem 3.3.2, yields the second expression of K_i^* in Theorem 3.3.2.

Next, by the third line of equation (3.3.9) and equation (3.3.11), we get

$$\begin{aligned} k_{ij} &= \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} - \frac{\lambda_i w_{ij}}{2\alpha_3 \mathbb{E}[\psi_{ij}]} \\ &= \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} - \Big(\frac{\sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}}{\sum_{j=1}^{n_i} \frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}} - \frac{K_i}{\sum_{j=1}^{n_i} \frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}} \Big) \frac{w_{ij}}{\mathbb{E}[\psi_{ij}]} \\ &= \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} + \frac{\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}{\sum_{j=1}^{n_i} \frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}} \Big(K_i - \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} \Big), \end{aligned}$$

which yields the expression for k_{ij}^* in Theorem 3.3.2. □

Remark 3.3.3. From Theorem 3.3.2, we observe that $0 < \delta(\alpha_1, \alpha_2, \alpha_3)$, $\tilde{\delta}(\alpha_1, \alpha 2, \alpha_3) < 1$ and that $0 < \beta_i(\alpha_2, \alpha_3)$, $\tilde{\beta}_i(\alpha_2, \alpha_3)$, $\theta_i(\alpha_1, \alpha_2, \alpha_3) < 1$, $i = 1, ..., n$. From the expressions of K^*, K_i^*, and k_{ij}^* given in Theorem 3.3.2, we see that the optimal required total capital K^* for the company is between the expected weighted aggregate risk of the company and the weighted sum of the expected weighted losses of main business and sub-business lines. We point out that the expected weighted aggregate risk of a company is not necessarily equal to the total of the expected weighted risks of the main business and sub-business lines of the company. If fact, the optimal required total capital K^* for the company satisfies that if $\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} < \sum_{i=1}^n \Big(\beta_i(\alpha_2, \alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2, \alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} \Big)$, then

$$\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} < K^* < \sum_{i=1}^n \Big(\beta_i(\alpha_2, \alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2, \alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} \Big);$$

if $\sum_{i=1}^n \Big(\beta_i(\alpha_2, \alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2, \alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} \Big) < \frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]}$, then

$$\sum_{i=1}^n \Big(\beta_i(\alpha_2, \alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2, \alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} \Big) < K^* < \frac{\mathbb{E}[\xi S]}{\mathbb{E}[\xi]};$$

and if $\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} = \sum_{i=1}^n \Big(\beta_i(\alpha_2, \alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2, \alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} \Big)$, then

$$K^* = \frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]}.$$

The relation between $\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]}$ and $\sum_{i=1}^n \Big(\beta_i(\alpha_2, \alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2, \alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} \Big)$ depends on the decision maker's attitudes toward the aggregate risks of the company and individual risks of the main business lines and sub-business lines, which are represented by the weighting factors ψ, ψ_i, ψ_{ij}, $i = 1, ..., n$, $j = 1, ..., n_i$.

Moreover, from the expression for K^* in Theorem 3.3.2, we observe that at the top level, the optimal required capital K^* for the enterprise is the weighed sum of the adjusted expected losses from all the three levels: the enterprise, the main business lines, and the sub-business lines. At the second level, from the second expression for K_i^* in Theorem 3.3.2, we find that the optimal capital K_i^* allocated to main business line i is equal to a base capital, which is a weighed sum of the adjusted expected losses of all the main business lines and the sub-business lines in the enterprise, plus a loading or a back up capital, which is proportional to the difference between the required total capital K^* for the enterprise and the weighed sum of the adjusted expected losses from all the main business lines and the sub-business lines in the enterprise, with a loading factor $\frac{\theta_i(\alpha_1,\alpha_2,\alpha_3)}{\delta(\alpha_1,\alpha_2,\alpha_3)}$. At the bottom level or the level of the sub-business lines, the optimal capital k_{ij}^* allocated to sub-business lines ij is equal to a base capital, which is the adjusted expected loss of sub-business line ij, plus a loading or a back up capital, which is proportional to the difference between the capital allocated to main business line i and the total of the adjusted expected losses from all the sub-business lines in main business line i, with a loading factor $\frac{\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}{\sum_{j=1}^{n_i}\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}$. □

3.3.1 Special or limiting cases of Theorem 3.3.2

In this subsection, we consider some special or limiting cases of Theorem 3.3.2 and show Theorem 3.3.2 can yield many interesting results about how to determine the optimal required total capital of a company and how to allocate the optimal required total capital among the main business lines and their sub-business lines at the same time.

Proposition 3.3.4. *(i) In Theorem 3.3.2, if*

$$w_i = \mathbb{E}[\psi_i X_i], \qquad w_{ij} = \mathbb{E}[\psi_{ij} X_{ij}], \qquad i = 1, ..., n, \quad j = 1, ..., n_i, \tag{3.3.17}$$

then, the optimal allocations among the main business lines and the sub-business lines are respectively reduced to the following proportional allocation principles

$$K_i^* = \frac{\beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]}}{\sum_{i=1}^n \beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]}} \times K^*, \quad i = 1, ..., n, \tag{3.3.18}$$

$$k_{ij}^* = \frac{\frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]}}{\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]}} \times K_i^*, \quad i = 1, ..., n, \quad j = 1, ..., n_i, \tag{3.3.19}$$

where the optimal required total capital K^* and the weights $\beta_i(\alpha_2,\alpha_3)$, $i = 1, ..., n$, are given in Theorem 3.3.2.

(ii) In Theorem 3.3.2, if

$$\begin{cases} w_i = \mathbb{E}[\psi_i X_i], \quad w_{ij} = \mathbb{E}[\psi_{ij} X_{ij}], \quad i = 1, ..., n, \quad j = 1, ..., n_i, \\ w = \mathbb{E}[\psi] \sum_{i=1}^n \beta_i(\alpha_2,\alpha_3)\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]}, \end{cases} \tag{3.3.20}$$

then, the optimal allocations among the main business lines and the sub-business lines are reduced to the proportional allocation principles (3.3.18) *and* (3.3.19)*, respectively, and the optimal required total capital* K^* *is given by*

$$K^* = \frac{\alpha_1}{\alpha_1+\alpha_2} \times \frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} + \frac{\alpha_2}{\alpha_1+\alpha_2} \sum_{i=1}^{n} \left(\beta_i(\alpha_2,\alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2,\alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]} \right) \tag{3.3.21}$$

where $\beta_i(\alpha_2,\alpha_3)$*,* $\tilde{\beta}_i(\alpha_2,\alpha_3)$*,* $i = 1, ..., n$*, are given in Theorem 3.3.2.*

Proof. By (3.3.4), (3.3.5), and (3.3.6), we have

$$\beta_i(\alpha_2,\alpha_3) = \tilde{\beta}_i(\alpha_2,\alpha_3) \times \frac{\alpha_2 \sum_{j=1}^{n_i} \frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}{\alpha_3 \frac{w_i}{\mathbb{E}(\psi_i)}}$$

and

$$\frac{\theta_i(\alpha_1,\alpha_2,\alpha_3)}{\delta(\alpha_1,\alpha_2,\alpha_3)} = \frac{\beta_i(\alpha_2,\alpha_3) \times \frac{w_i}{\mathbb{E}[\psi_i]}}{\sum_{i=1}^{n} \beta_i(\alpha_2,\alpha_3) \frac{w_i}{\mathbb{E}(\psi_i)}} = \frac{\tilde{\beta}_i(\alpha_2,\alpha_3) \times \sum_{j=1}^{n_i} \frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}{\sum_{i=1}^{n} \tilde{\beta}_i(\alpha_2,\alpha_3) \times \sum_{j=1}^{n_i} \frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}.$$

(i) Under the conditions in (3.3.17), by the second expression for K_i^* in Theorem 3.3.2, we have

$$\begin{aligned}
K_i^* &= \beta_i(\alpha_2,\alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2,\alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]} + \frac{\theta_i(\alpha_1,\alpha_2,\alpha_3)}{\delta(\alpha_1,\alpha_2,\alpha_3)} \times K^* \\
&\quad - \frac{\theta_i(\alpha_1,\alpha_2,\alpha_3)}{\delta(\alpha_1,\alpha_2,\alpha_3)} \sum_{i=1}^{n} \left(\beta_i(\alpha_2,\alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} \right) - \frac{\theta_i(\alpha_1,\alpha_2,\alpha_3)}{\delta(\alpha_1,\alpha_2,\alpha_3)} \sum_{i=1}^{n} \left(\tilde{\beta}_i(\alpha_2,\alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]} \right), \\
&= \frac{\theta_i(\alpha_1,\alpha_2,\alpha_3)}{\delta(\alpha_1,\alpha_2,\alpha_3)} \times K^* + \beta_i(\alpha_2,\alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2,\alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]} \\
&\quad - \frac{\beta_i(\alpha_2,\alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}(\psi_i)}}{\sum_{i=1}^{n} \beta_i(\alpha_2,\alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}(\psi_i)}} \sum_{i=1}^{n} \beta_i(\alpha_2,\alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} \\
&\quad - \frac{\tilde{\beta}_i(\alpha_2,\alpha_3) \times \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]}}{\sum_{i=1}^{n} \tilde{\beta}_i(\alpha_2,\alpha_3) \times \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]}} \sum_{i=1}^{n} \left(\tilde{\beta}_i(\alpha_2,\alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]} \right), \\
&= \frac{\theta_i(\alpha_1,\alpha_2,\alpha_3)}{\delta(\alpha_1,\alpha_2,\alpha_3)} \times K^* \\
&= \frac{\beta_i(\alpha_2,\alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]}}{\sum_{i=1}^{n} \beta_i(\alpha_2,\alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]}}, \qquad i = 1, ..., n.
\end{aligned}$$

Moreover, under the conditions in (3.3.17), by the expression for k_{ij}^* in Theorem 3.3.2, we have

$$
\begin{aligned}
k_{ij}^* &= \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} + \frac{\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}}{\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}}\left(K_i^* - \sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\right) \\
&= \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} + \frac{\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}}{\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}} K_i^* - \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]} \\
&= \frac{\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}}{\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}} K_i^*, \quad i = 1, ..., n, \quad j = 1, ..., n_i.
\end{aligned}
$$

(ii) If the conditions in (3.3.20) hold, then by (3.3.4), we see that $\tilde{\delta}(\alpha_1, \alpha_2, \alpha_3)$ is reduced to

$$
\tilde{\delta}(\alpha_1, \alpha_2, \alpha_3) = 1 - \delta(\alpha_1, \alpha_2, \alpha_3) = \frac{\alpha_2}{\alpha_2 + \alpha_1},
$$

which, together with the expression for K^* in Theorem 3.3.2, yields (3.3.21). □

Remark 3.3.5. The above proposition is an extension of the proportional allocation principle for a two-level company in Proposition 3.2.4 to a three-level organization structure. We point out that besides the conditions in (3.3.20), if

$$
\mathbb{E}[\psi_i X_i] = \mathbb{E}[\psi_i] \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}, \quad i = 1, ..., n, \tag{3.3.22}
$$

then the proportional allocation principles (3.3.18) and (3.3.19) and the optimal required total capital (3.3.21) are reduced to

$$
\begin{aligned}
K_i^* &= \frac{\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]}}{\sum_{i=1}^{n}\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]}} \times K^*, \quad i = 1, ..., n, \\
k_{ij}^* &= \frac{\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}}{\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}} \times K_i^*, \quad i = 1, ..., n, \quad j = 1, ..., n_i, \\
K^* &= \frac{\alpha_1}{\alpha_1 + \alpha_2} \times \frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} + \frac{\alpha_2}{\alpha_1 + \alpha_2} \sum_{i=1}^{n}\left(\frac{\alpha_2}{\alpha_2 + \alpha_3} \times \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \frac{\alpha_3}{\alpha_2 + \alpha_3} \times \sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\right) \\
&= \frac{\alpha_1}{\alpha_1 + \alpha_2} \times \frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]} + \frac{\alpha_2}{\alpha_1 + \alpha_2} \sum_{i=1}^{n}\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]}
\end{aligned}
$$

respectively. If fact, the conditions in (3.3.22) imply

$$
\tilde{\beta}_i(\alpha_2, \alpha_3) = 1 - \beta_i(\alpha_2, \alpha_3) = \frac{\alpha_3}{\alpha_3 + \alpha_2}, \quad i = 1, ..., n.
$$

For instance, in Proposition 3.3.4, let sets A, A_i, A_{ij} satisfy

$$\mathbb{E}(S|S\in A)=\mathrm{VaR}_q(S),\ A_i=\{X_i>\mathrm{VaR}_q(X_i)\},\ A_{ij}=\{X_i>\mathrm{VaR}_q(X_i)\},\ i=1,...,n,\ j=1,...,n_i,$$

and let

$$\psi=\frac{\mathrm{I}_{\{S\in A\}}}{\mathbb{P}\{S\in A\}},\quad \psi_i=\frac{\mathrm{I}_{\{X_i\in A_i\}}}{\mathbb{P}\{X_i\in A_i\}},\quad \psi_{ij}=\frac{\mathrm{I}_{\{X_{ij}\in A_{ij}\}}}{\mathbb{P}\{X_{ij}\in A_{ij}\}},\quad i=1,...,n,\quad j=1,...,n_i,$$

and

$$w=\frac{\alpha_2}{\alpha_2+\alpha_3}\sum_{i=1}^{n}\mathbb{E}[\psi_i X_i],\quad w_i=\mathbb{E}[\psi_i X_i],\quad w_{ij}=\mathbb{E}[\psi_{ij}X_{ij}],\quad i=1,...,n,\quad j=1,...,n_i,$$

thus $\mathbb{E}[\psi]=1$, $\mathbb{E}[\psi_i]=1$, $\mathbb{E}[\psi_{ij}]=1$, $\mathbb{E}[\psi S]=\mathrm{VaR}_q(S)$, $\mathbb{E}[\psi_i X_i]=\mathbb{E}[X_i|X_i>\mathrm{VaR}_q(X_i)]=\mathrm{CTE}_q(X_i)$, $\mathbb{E}[\psi_{ij}X_{ij}]=\mathbb{E}[X_{ij}|X_i>\mathrm{VaR}_q(X_i)]$, $i=1,...,n$, $j=1,...,n_i$, and

$$\mathbb{E}[\psi_i X_i]=\mathbb{E}[X_i|X_i>\mathrm{VaR}_q(X_i)]=\sum_{j=1}^{n_i}\mathbb{E}[X_{ij}|X_i>\mathrm{VaR}_q(X_i)]=\mathbb{E}[\psi_i]\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]},\qquad i=1,...,n,$$

which mean that the conditions in (3.3.22) hold and $\beta_i(\alpha_2,\alpha_3)=\frac{\alpha_2}{\alpha_2+\alpha_3}$. Hence, all the conditions in (3.3.20) and (3.3.22) hold. Thus, the proportional allocation principles (3.3.18) and (3.3.19) and the optimal required total capital (3.3.21) are reduced to

$$\begin{aligned}
K_i^* &= \frac{\mathrm{CTE}_q(X_i)}{\sum_{i=1}^n \mathrm{CTE}_q(X_i)}\times K^*,\quad i=1,...,n,\\
k_{ij}^* &= \frac{\mathbb{E}[X_{ij}|X_i>\mathrm{VaR}_q(X_i)]}{\mathrm{CTE}_q(X_i)}\times K_i^*,\quad i=1,...,n,\quad j=1,...,n_i,\\
K^* &= \frac{\alpha_1}{\alpha_1+\alpha_2}\mathrm{VaR}_q(S)+\frac{\alpha_2}{\alpha_1+\alpha_2}\sum_{i=1}^{n}\mathbb{E}[X_i|X_i>\mathrm{VaR}_q(X_i)]\\
&= \frac{\alpha_1}{\alpha_1+\alpha_2}\mathrm{VaR}_q(S)+\frac{\alpha_2}{\alpha_1+\alpha_2}\sum_{i=1}^{n}\mathrm{CTE}_q(X_i).
\end{aligned}$$

respectively. Note that the above example, we assume $X_i=\sum_{j=1}^{n_i}X_{ij}$, $i=1,...,n$, $j=1,...,n_i$. □

Example 3.3.1. Similarly to the assumptions used in Zaks and Tsanakas (2014), in Theorem 3.3.2, if $\mathbb{E}[\psi]=1$, $\mathbb{E}[\psi_i]=1$, $\mathbb{E}[\psi_{ij}]=1$, $\sum_{i=1}^n w_i=w$, $\sum_{i=1}^n w_{ij}=w_i$, $i=1,...,n$, $j=1,...,n_i$, then, by (3.3.5), (3.3.4), and (3.3.6), we have

$$\begin{aligned}
\tilde{\beta}_i(\alpha_2,\alpha_3) &= 1-\beta_i(\alpha_2,\alpha_3) = \frac{\alpha_3}{\alpha_3+\alpha_2},\\
\tilde{\delta}(\alpha_1,\alpha_2,\alpha_3) &= 1-\delta(\alpha_1,\alpha_2,\alpha_3) = \alpha_2+\alpha_3,\\
\theta_i(\alpha_1,\alpha_2,\alpha_3) &= \frac{\alpha_1 w_i}{w}.
\end{aligned}$$

Thus, the optimal solution in Theorem 3.3.2 are reduced to

$$\begin{aligned}
K^* &= \alpha_1\mathbb{E}[\psi S] + \alpha_2\sum_{i=1}^{n}\mathbb{E}[\psi_i X_i] + \alpha_3\sum_{i=1}^{n}\sum_{j=1}^{n_i}\mathbb{E}[\psi_{ij}X_{ij}],\\
K_i^* &= \frac{\alpha_2}{\alpha_2+\alpha_3}\mathbb{E}[\psi_i X_i] + \frac{\alpha_3}{\alpha_3+\alpha_2}\sum_{j=1}^{n_i}\mathbb{E}[\psi_{ij}X_{ij}]\\
&\quad + \frac{\alpha_1 w_i}{w}\left(K^* - \frac{\alpha_2}{\alpha_2+\alpha_3}\sum_{i=1}^{n}\mathbb{E}[\psi_i X_i] - \frac{\alpha_3}{\alpha_2+\alpha_3}\sum_{i=1}^{n}\sum_{j=1}^{n_i}\mathbb{E}[\psi_{ij}X_{ij}]\right), \quad i = 1, ..., n,\\
k_{ij}^* &= \mathbb{E}[\psi_{ij}X_{ij}] + \frac{w_{ij}}{w_i}\left(K_i^* - \sum_{j=1}^{n_i}\mathbb{E}[\psi_{ij}X_{ij}]\right), \quad i = 1, ..., n, \quad j = 1, ..., n_i.
\end{aligned}$$

We point out that the optimal allocation expressions for K_i^* and k_{ij} are the main results of section 3 in Zaks and Tsanakas (2014) when the required total capital is calculated as $K^* = \alpha_1\mathbb{E}[\psi S] + \alpha_2\sum_{i=1}^{n}\mathbb{E}[\psi_i X_i] + \alpha_3\sum_{i=1}^{n}\sum_{j=1}^{n_i}\mathbb{E}[\psi_{ij}X_{ij}]$. □

Note that $\alpha_1, \alpha_2, \alpha_3$ represent the decision maker's attitudes towards the allocation deviations of at the levels of the company, the main business lines, and the sub-business lines, respectively. In the rest of this section, we consider the following three limiting cases of Theorem 3.3.2.

Limiting case 1: $\alpha_3 \to 0$, $\frac{\alpha_3}{\alpha_2} \to 0$. In this case, we have $\alpha_1 + \alpha_2 \to 1$ and $\alpha_2 > \alpha_3$, which mean that the main concern of the decision maker is the allocation deviations among the main business lines and the derivation between the aggregate risk and the required total capital. The allocation deviations among the sub-business lines are less important.

Limiting case 2: $\alpha_1 \to 1$ and $\frac{\alpha_3}{\alpha_2} \to 0$. In this case, we have $\alpha_2+\alpha_3 \to 0$ and $\alpha_2 > \alpha_3$, which mean that the main concern of the decision maker is the derivation between the aggregate risk and the required total capital. The allocation deviations among the main business lines the sub-business lines are less important.

Limiting case 3: $\alpha_1 \to 0$ and $\frac{\alpha_1}{\alpha_2} \to 0$. In this case, we have $\alpha_2+\alpha_3 \to 1$ and $\alpha_2 > \alpha_1$, which mean that the main concern of the decision maker is the allocation deviations among the main business lines the sub-business lines. The derivation between the aggregate risk and the required total capital are less important.

Proposition 3.3.6. *In Theorem 3.3.2, if* $\alpha_3 \to 0$, $\frac{\alpha_3}{\alpha_2} \to 0$, *then* $\tilde{\beta}_i(\alpha_2, \alpha_3) \to 0$, $\beta_i(\alpha_2, \alpha_3) \to$

1,

$$\begin{aligned}
\tilde{\delta}(\alpha_1,\alpha_2,\alpha_3) &\to \frac{\alpha_2\frac{w}{\mathbb{E}(\psi)}}{\alpha_2\frac{w}{\mathbb{E}(\psi)}+\alpha_1\sum_{i=1}^n\frac{w_i}{\mathbb{E}(\psi_i)}},\\
\delta(\alpha_1,\alpha_2,\alpha_3) &\to \frac{\alpha_1\sum_{i=1}^n\frac{w_i}{\mathbb{E}(\psi_i)}}{\alpha_2\frac{w}{\mathbb{E}(\psi)}+\alpha_1\sum_{i=1}^n\frac{w_i}{\mathbb{E}(\psi_i)}},\\
\theta_i(\alpha_1,\alpha_2,\alpha_3) &\to \frac{\frac{\alpha_1 w_i}{\mathbb{E}[\psi_i]}}{\alpha_2\frac{w}{\mathbb{E}(\psi)}+\alpha_1\sum_{i=1}^n\frac{w_i}{\mathbb{E}(\psi_i)}},
\end{aligned}$$

and the optimal required total capital and the optimal allocations have the following limits:

$$\begin{aligned}
K^* &\to \frac{\alpha_1\sum_{i=1}^n\frac{w_i}{\mathbb{E}(\psi_i)}}{\alpha_2\frac{w}{\mathbb{E}(\psi)}+\alpha_1\sum_{i=1}^n\frac{w_i}{\mathbb{E}(\psi_i)}}\times\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]}+\frac{\alpha_2\frac{w}{\mathbb{E}(\psi)}}{\alpha_2\frac{w}{\mathbb{E}(\psi)}+\alpha_1\sum_{i=1}^n\frac{w_i}{\mathbb{E}(\psi_i)}}\times\sum_{i=1}^n\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]},\\
K_i^* &\to K_i^*(\alpha_1,\alpha_2,0)\triangleq\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]}+\frac{\frac{\alpha_1 w_i}{\mathbb{E}[\psi_i]}}{\alpha_2\frac{w}{\mathbb{E}(\psi)}+\alpha_1\sum_{i=1}^n\frac{w_i}{\mathbb{E}(\psi_i)}}\times\left(\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]}-\sum_{i=1}^n\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]}\right),\quad i=1,...,n,\\
k_{ij}^* &\to \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}+\frac{\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}{\sum_{j=1}^{n_i}\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}\left(K_i^*(\alpha_1,\alpha_2,0)-\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\right),\quad i=1,...,n,\quad j=1,...,n_i.
\end{aligned}$$

Note that in the proof of Theorem 3.3.2 and its proof, it is assumed that $0<\alpha_i<1$, $i=1,2,3$. If fact, the condition of $\alpha_i>0$, $i=1,2,3$ is used in the proof of Theorem 3.3.2. Hence, we could not obtain Theorem 3.2.2 by simply setting $\alpha_3=0$ in Theorem 3.3.2 and its proof. Although the proof of Theorem 3.2.2 is similar to that for Theorem 3.3.2 and the optimal required total capital K^* and the optimal allocation $(K_1^*,...,K_n^*)$ are equal to the limiting forms of K^* and K_i^* in Theorem 3.3.2, we keep the proof of Theorem 3.2.2 since the proof of Theorem 3.2.2 is helpful for one to understand the proof of Theorem 3.3.2.

Proposition 3.3.7. *In Theorem 3.3.2, if* $\alpha_1\to 1$ *and* $\frac{\alpha_3}{\alpha_2}\to 0$*, then*

$$\beta_i(\alpha_2,\alpha_3)\to 1,\ \tilde{\beta}_i(\alpha_2,\alpha_3)\to 0,\ \tilde{\delta}(\alpha_1,\alpha_2,\alpha_3)\to 0,\ \delta(\alpha_1,\alpha_2,\alpha_3)\to 1,\ \theta_i(\alpha_1,\alpha_2,\alpha_3)\to\frac{\frac{w_i}{\mathbb{E}(\psi_i)}}{\sum_{i=1}^n\frac{w_i}{\mathbb{E}(\psi_i)}},$$

and the optimal required total capital and the optimal allocations have the following limits:

$$\begin{cases}
K^* \to \frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]},\\
K_i^* \to K_i^*(1,0,0)\triangleq\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]}+\frac{\frac{w_i}{\mathbb{E}(\psi_i)}}{\sum_{i=1}^n\frac{w_i}{\mathbb{E}(\psi_i)}}\left(\frac{\mathbb{E}[\psi S]}{\mathbb{E}[\psi]}-\sum_{i=1}^n\frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]}\right)\\
k_{ij}^* \to \frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}+\frac{\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}{\sum_{j=1}^{n_i}\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}\left(K_i^*(1,0,0)-\sum_{j=1}^{n_i}\frac{\mathbb{E}[\psi_{ij}X_{ij}]}{\mathbb{E}[\psi_{ij}]}\right),\quad i=1,...,n,\quad j=1,...,n_i,
\end{cases}$$

In this proposition, as $\alpha_1\to 1$, the optimal required total capital at the enterprise level is completely decided by the weighted aggregate loss of the enterprise.

Proposition 3.3.8. *In Theorem 3.3.2, if* $\alpha_1 \to 0$ *and* $\frac{\alpha_1}{\alpha_2} \to 0$*, then* $\tilde{\delta}(\alpha_1, \alpha_2, \alpha_3) \to 1$*,* $\delta(\alpha_1, \alpha_2, \alpha_3) \to 0$*,* $\theta_i(\alpha_1, \alpha_2, \alpha_3) \to 0$*, and the optimal required total capital and the optimal allocation are reduced to*

$$\begin{aligned}
K^* &\to \sum_{i=1}^{n} \Big(\beta_i(\alpha_2, \alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2, \alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]} \Big), \\
K_i^* &\to \beta_i(\alpha_2, \alpha_3) \frac{\mathbb{E}[\psi_i X_i]}{\mathbb{E}[\psi_i]} + \tilde{\beta}_i(\alpha_2, \alpha_3) \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]} \triangleq K_i^*(0, \alpha_2, \alpha_3), \quad i = 1, ..., n, \\
k_{ij}^* &\to \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]} + \frac{\frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}}{\sum_{j=1}^{n_i} \frac{w_{ij}}{\mathbb{E}[\psi_{ij}]}} \Big(K_i^*(0, \alpha_2, \alpha_3) - \sum_{j=1}^{n_i} \frac{\mathbb{E}[\psi_{ij} X_{ij}]}{\mathbb{E}[\psi_{ij}]} \Big), \quad i = 1, ..., n, \quad j = 1, ..., n_i.
\end{aligned}$$

3.4 Optimal solutions based on weighted absolute errors for a company with main business lines and sub-business lines

In this section, we discuss optimal required total capital and optimal allocation scheme for a company with main business lines and sub-business lines and consider problem (3.1.4) when $D_1(x) = |x|$ and $D_{2i} = |x|$, $i = 1, ..., n$, namely the allocation deviations are measured by (weighted) absolute errors. In this case, problem (3.1.4) is reduced to the following problem:

$$\begin{cases}
\min_{(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n}) \in \mathbb{R}^d} & \Big\{ \alpha_1 \mathbb{E}\Big[\psi \Big| S - K \Big|\Big] + \alpha_2 \sum_{i=1}^{n} \mathbb{E}\Big[\psi_i \Big| X_i - K_i \Big|\Big] \\
& \quad + \alpha_3 \sum_{i=1}^{n} \sum_{j=1}^{n_i} \mathbb{E}\Big[\psi_{ij} \Big| X_{ij} - k_{ij} \Big|\Big] \Big\} \\
\text{s.t.} & \sum_{i=1}^{n} K_i = K, \quad \sum_{j=1}^{n_i} k_{ij} = K_i, \quad i = 1, ..., n.
\end{cases} \tag{3.4.1}$$

To guarantee that problem (3.4.1) has an optimal solution $(\mathbf{K}^*, \mathbf{k}_1^*, ..., \mathbf{k}_n^*)$, we assume that the following conditions hold.

Assumption 3.4.1. For problem (3.4.1), we assume that random variables ψ, ψ_i, ψ_{ij}, $i = 1, ..., n$, $j = 1, ..., n_i$, satisfy

$$\psi \geq 0, \quad \psi_i \geq 0, \quad \psi_{ij} \geq 0, \quad \mathbb{E}[\psi] > 0, \quad \mathbb{E}[\psi_i] > 0, \quad \mathbb{E}[\psi_{ij}] > 0, \quad i = 1, ..., n, \quad j = 1, ..., n_i,$$

and that expectations $\mathbb{E}(\psi|S|)$, $\mathbb{E}(\psi_i|X_i|)$, $\mathbb{E}(\psi_{ij}|X_{ij}|)$, $i = 1, ..., n$, $j = 1, ..., n_i$, exist. $\square$

Denote the objective function in problem (3.4.1) by $J(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n})$, namely,

$$J(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n}) = \alpha_1 \mathbb{E}\Big[\psi \Big| S - K \Big|\Big] + \alpha_2 \sum_{i=1}^{n} \mathbb{E}\Big[\psi_i \Big| X_i - K_i \Big|\Big] + \alpha_3 \sum_{i=1}^{n} \sum_{j=1}^{n_i} \mathbb{E}\Big[\psi_{ij} \Big| X_{ij} - k_{ij} \Big|\Big] \tag{3.4.2}$$

To solve problem (3.4.1), we follow the similar approach used in the proof of Theorem 4.2 of Cai and Wang (2020). We first recall the inverses of distribution functions. For the distribution function $F(x) = \Pr\{X \leq x\}$ of a random variable X, the left-continuous and right-continuous inverse functions of F are respectively defined as $F^{-1}(q) = \inf\{x \in \mathbb{R} : F(x) \geq q\}$, $F^{-1+}(q) = \sup\{x \in \mathbb{R} : F(x) \leq q\}$, and the p-mixed inverse function $F^{-1(p)}$ of the distribution function F is defined as

$$F^{-1(p)}(q) = pF^{-1}(q) + (1-p)F^{-1+}(q),$$

where $0 \leq p \leq 1$ and $0 < q < 1$. By the definition of $F^{-1(p)}$, we see that if $0 < q < 1$, then

$$[F^{-1}(q),\ F^{-1+}(q)] = \{F^{-1(p)}(q),\ p \in [0,1]\}. \tag{3.4.3}$$

For problem (3.4.1), we define functions $F, F_i, F_{ij}, G, G_i, G_{ij}$ on $\mathbb{R}$ as

$$\begin{cases} F(x) &= \frac{\mathbb{E}[\psi \mathrm{I}_{\{S \leq x\}}]}{\mathbb{E}[\psi]}, \qquad F_i(x) = \frac{\mathbb{E}[\psi_i \mathrm{I}_{\{X_i \leq x\}}]}{\mathbb{E}[\psi_i]}, \qquad F_{ij}(x) = \frac{\mathbb{E}[\psi_{ij} \mathrm{I}_{\{X_{ij} \leq x\}}]}{\mathbb{E}[\psi_{ij}]}, \\ G(x) &= \frac{\mathbb{E}[\psi \mathrm{I}_{\{S < x\}}]}{\mathbb{E}[\psi]}, \qquad G_i(x) = \frac{\mathbb{E}[\psi_i \mathrm{I}_{\{X_i < x\}}]}{\mathbb{E}[\psi_i]}, \qquad G_{ij}(x) = \frac{\mathbb{E}[\psi_{ij} \mathrm{I}_{\{X_{ij} < x\}}]}{\mathbb{E}[\psi_{ij}]}. \end{cases} \tag{3.4.4}$$

It is easy to see that under Assumption 3.4.1, F, F_i and F_{ij} are distribution functions and that $G(x) = \lim_{y \to x^-} F(y) = F(x-)$, $G_i(x) = \lim_{y \to x^-} F_i(y) = F_i(x-)$, $G_{ij}(x) = \lim_{y \to x^-} F_{ij}(y) = F_{ij}(x-)$, which mean that G, G_i, and G_{ij} are left-continuous distribution functions. In addition, we have that $0 \leq G(x) \leq F(x) \leq 1$, $0 \leq G_i(x) \leq F_i(x) \leq 1$, and $0 \leq G_{ij}(x) \leq F_{ij}(x) \leq 1$. Furthermore, for distributions F, F_i, F_{ij}, $i = 1, ..., n$, $j = 1, ..., n_i$, denote their p-mixed inverses by

$$F^{-1(p)}(q) = pF^{-1}(q) + (1-p)F^{-1+}(q), \tag{3.4.5}$$

$$F_i^{-1(p)}(q) = pF_i^{-1}(q) + (1-p)F_i^{-1+}(q), \tag{3.4.6}$$

$$F_{ij}^{-1(p)}(q) = pF_{ij}^{-1}(q) + (1-p)F_{ij}^{-1+}(q). \tag{3.4.7}$$

Moreover, for left-continuous distribution functions G, G_i and G_{ij}, denote their right-continuous inverses by $G^{-1+}(q) = \sup\{x \in \mathbb{R} : G(x) \leq q\}$, $G_i^{-1+}(q) = \sup\{x \in \mathbb{R} : G_i(x) \leq q\}$, and $G_{ij}^{-1+}(q) = \sup\{x \in \mathbb{R} : G_{ij}(x) \leq q\}$, $0 < q < 1$, $i = 1, ..., n$, $j = 1, ..., n_i$. Thus, by Lemma 4.1 of Cai and Wang (2020), we have that for any $0 < q < 1$,

$$G^{-1+}(q) = F^{-1+}(q), \qquad G_i^{-1+}(q) = F_i^{-1+}(q), \qquad G_{ij}^{-1+}(q) = F_{ij}^{-1+}(q), \tag{3.4.8}$$

and that for any $x \in \mathbb{R}$ and any $0 < q < 1$,

$$G(x) \leq q \leq F(x) \iff F^{-1}(q) \leq x \leq F^{-1+}(q), \tag{3.4.9}$$

$$G_i(x) \leq q \leq F_i(x) \iff F_i^{-1}(q) \leq x \leq F_i^{-1+}(q), \tag{3.4.10}$$

$$G_{ij}(x) \leq q \leq F_{ij}(q) \iff F_{ij}^{-1}(q) \leq x \leq F_{ij}^{-1+}(q). \tag{3.4.11}$$

Lemma 3.4.1. *Let $g(c) = \mathbb{E}\Big(\xi|X-c|\Big)$, where $\xi \geq 0$ and X are random variables. Assume $g(c) < +\infty$ for any $c \in \mathbb{R}$. Then,*

$$g'_{-}(c) = -\mathbb{E}\Big[\xi \mathrm{I}_{\{X \geq c\}}\Big] + \mathbb{E}\Big[\xi \mathrm{I}_{\{X < c\}}\Big], \tag{3.4.12}$$

and

$$g'_{+}(c) = -\mathbb{E}\Big[\xi \mathrm{I}_{\{X > c\}}\Big] + \mathbb{E}\Big[\xi \mathrm{I}_{\{X \leq c\}}\Big]. \tag{3.4.13}$$

Proof. The result follows from Lemma A.2 of Mao and Cai (2018) by the relation $|x| = (x)_+ + (x)_-$.

Theorem 3.4.2. *Under Assumption 3.4.1, solutions to problem (3.4.1) do exist and a solution $(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n})$ to problem (3.4.1) has the following expression:*

$$\begin{cases} K = F^{-1(p)}\Big(\dfrac{\mathbb{E}[\psi] + \frac{\lambda}{\alpha_1}}{2\mathbb{E}[\psi]}\Big), \\ K_i = F_i^{-1(p_i)}\Big(\dfrac{\mathbb{E}[\psi_i] - \frac{\lambda+\lambda_i}{\alpha_2}}{2\mathbb{E}[\psi_i]}\Big), \quad i = 1, \ldots, n, \\ k_{ij} = F_{ij}^{-1(p_{ij})}\Big(\dfrac{\mathbb{E}[\psi_{ij}] + \frac{\lambda_i}{\alpha_3}}{2\mathbb{E}[\psi_{ij}]}\Big), \quad i = 1, ..., n, \quad j = 1, \ldots, n_i, \end{cases} \tag{3.4.14}$$

for any p, p_i, $p_{ij} \in [0, 1]$, and any λ, $\lambda_i \in \mathbb{R}$, $i = 1, ..., n$, $j = 1, ..., n_i$, providing that the parameters p, λ, p_i, λ_i, p_{ij}, $i = 1, ..., n$, $j = 1, ..., n_i$, satisfy the following equations and inequalities:

$$\begin{cases} \displaystyle\sum_{i=1}^{n} F_i^{-1(p_i)}\Big(\dfrac{\mathbb{E}[\psi_i] - \frac{\lambda+\lambda_i}{\alpha_2}}{2\mathbb{E}[\psi_i]}\Big) = F^{-1(p)}\Big(\dfrac{\mathbb{E}[\psi] + \frac{\lambda}{\alpha_1}}{2\mathbb{E}[\psi]}\Big), \\ \displaystyle\sum_{j=1}^{n_i} F_{ij}^{-1(p_{ij})}\Big(\dfrac{\mathbb{E}[\psi_{ij}] + \frac{\lambda_i}{\alpha_3}}{2\mathbb{E}[\psi_{ij}]}\Big) = F_i^{-1(p_i)}\Big(\dfrac{\mathbb{E}[\psi_i] - \frac{\lambda+\lambda_i}{\alpha_2}}{2\mathbb{E}[\psi_i]}\Big), \quad i = 1, \ldots, n, \end{cases} \tag{3.4.15}$$

and

$$\begin{cases} -\mathbb{E}[\psi] < \dfrac{\lambda}{\alpha_1} < \mathbb{E}[\psi], \\ \underline{M} < -\dfrac{\lambda + \lambda_i}{\alpha_2} < \overline{M}, \quad i = 1, \ldots, n, \\ \underline{m} < \dfrac{\lambda_i}{\alpha_3} < \overline{m}, \quad i = 1, ..., n, \end{cases} \tag{3.4.16}$$

where

$$\begin{cases} \overline{M} = \min\{\mathbb{E}[\psi_i],\ i = 1, ..., n\}, \\ \underline{M} = \max\{-\mathbb{E}[\psi_i],\ i = 1, ..., n\}, \\ \overline{m} = \min\{\mathbb{E}[\psi_{ij}],\ i = 1, ..., n,\ j = 1, ..., n_i\}, \\ \underline{m} = \max\{-\mathbb{E}[\psi_{ij}],\ i = 1, ..., n,\ j = 1, ..., n_i\}. \end{cases} \tag{3.4.17}$$

Proof. Denote the Lagrangian of problem (3.4.1) by

$$L(\mathbf{K}, \mathbf{k}_1, ..., \mathbf{k}_n, \lambda, \lambda_1, ..., \lambda_n) = J(\mathbf{K}, \mathbf{k}_1, ..., \mathbf{k}_n) + \lambda\Big(\sum_{i=1}^{n} K_i - K\Big) + \sum_{i=1}^{n} \lambda_i\Big(K_i - \sum_{j=1}^{n_i} k_{ij}\Big).$$

Under Assumption 3.4.1, it ie easy to see that $\mathbb{E}\Big[\psi\Big|S-K\Big|\Big]$, $\mathbb{E}\Big[\psi_i\Big|X_i-K_i\Big|\Big]$, and $\mathbb{E}\Big[\psi_{ij}\Big|X_{ij}-k_{ij}\Big|\Big]$ are convex and coercive functions of K, K_i, k_{ij}, respectively. Hence, $J(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n})$ defined in (3.4.2) is a convex and coercive function of $(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n})$ on $\mathbb{R}^d$. Thus, $L(\mathbf{K}, \mathbf{k}_1, ..., \mathbf{k}_n, \lambda, \lambda_1, ..., \lambda_n)$ is a convex and coercive function of $(\mathbf{K}, \mathbf{k_1}, ..., \mathbf{k_n})$ on $\mathbb{R}^d$. Hence, optimal solutions to the constrained convex optimization problem (3.4.1) do exist and $(\mathbf{K}, \mathbf{k}_1, ..., \mathbf{k}_n)$ is an optimal solution to problem (3.4.1) if it satisfies the following Karush-Kuhn-Tucker (KKT) conditions:

$$\begin{cases} 0 \in \Big[\dfrac{\partial^- L}{\partial K}, \ \ \dfrac{\partial^+ L}{\partial K}\Big], \\ 0 \in \Big[\dfrac{\partial^- L}{\partial K_i}, \ \ \dfrac{\partial^+ L}{\partial K_i}\Big], \ i = 1, \dots, n, \\ 0 \in \Big[\dfrac{\partial^- L}{\partial k_{ij}}, \ \ \dfrac{\partial^+ L}{\partial k_{ij}}\Big], \ i = 1, ..., n, \ j = 1, ..., n_i, \\ \displaystyle\sum_{i=1}^{n} K_i = K, \ \ \sum_{j=1}^{n_i} k_{ij} = K_i, \ i = 1, ..., n. \end{cases} \Longleftrightarrow \begin{cases} \dfrac{\partial^- L}{\partial K} \leq 0, \ \ \dfrac{\partial^+ L}{\partial K} \geq 0, \\ \dfrac{\partial^- L}{\partial K_i} \leq 0, \ \ \dfrac{\partial^+ L}{\partial K_i} \geq 0, \ i = 1, \dots, n, \\ \dfrac{\partial^- L}{\partial k_{ij}} \leq 0, \ \ \dfrac{\partial^+ L}{\partial k_{ij}} \geq 0, \ i = 1, ..., n, \ j = 1, ..., n_i, \\ \displaystyle\sum_{i=1}^{n} K_i = K, \ \ \sum_{j=1}^{n_i} k_{ij} = K_i, \ i = 1, ..., n. \end{cases}$$

By applying Lemma 3.4.1, we see that the above KKT conditions are reduced to the following system of inequalities and equations:

$$\begin{cases} \alpha_1 \, \mathbb{E}\Big[\psi \, \mathrm{I}_{\{S \leq K\}}\Big] \geq \alpha_1 \, \mathbb{E}\Big[\psi \, \mathrm{I}_{\{S > K\}}\Big] + \lambda, \\ \alpha_2 \, \mathbb{E}\Big[\psi_i \, \mathrm{I}_{\{X_i \leq K_i\}}\Big] \geq \alpha_2 \, \mathbb{E}\Big[\psi_i \, \mathrm{I}_{\{X_i > K_i\}}\Big] - \lambda - \lambda_i, \quad i = 1, ..., n, \\ \alpha_3 \, \mathbb{E}\Big[\psi_{ij} \, \mathrm{I}_{\{X_{ij} \leq k_{ij}\}}\Big] \geq \alpha_3 \, \mathbb{E}\Big[\psi_{ij} \, \mathrm{I}_{\{X_{ij} > k_{ij}\}}\Big] + \lambda_i, \quad i = 1, ..., n, \quad j = 1, ..., n_i, \\ \alpha_1 \, \mathbb{E}\Big[\psi \, \mathrm{I}_{\{S \geq K\}}\Big] \geq \alpha_1 \, \mathbb{E}\Big[\psi \, \mathrm{I}_{\{S < K\}}\Big] - \lambda, \\ \alpha_2 \, \mathbb{E}\Big[\psi_i \, \mathrm{I}_{\{X_i \geq K_i\}}\Big] \geq \alpha_2 \, \mathbb{E}\Big[\psi_i \, \mathrm{I}_{\{X_i < K_i\}}\Big] + \lambda + \lambda_i, \quad i = 1, \dots, n, \\ \alpha_3 \, \mathbb{E}\Big[\psi_{ij} \, \mathrm{I}_{\{X_{ij} \geq k_{ij}\}}\Big] \geq \alpha_3 \, \mathbb{E}\Big[\psi_{ij} \, \mathrm{I}_{\{X_{ij} < k_{ij}\}}\Big] - \lambda_i, \quad i = 1, \dots, n, \quad j = 1, ..., n_i, \\ \displaystyle\sum_{i=1}^{n} K_i = K, \quad \sum_{j=1}^{n_i} k_{ij} = K_i, \quad i = 1, ..., n. \end{cases} \tag{3.4.18}$$

Note that $\mathrm{I}_{\{x>a\}} = 1 - \mathrm{I}_{\{x \leq a\}}$ and $\mathrm{I}_{\{x \geq a\}} = 1 - \mathrm{I}_{\{x < a\}}$ for any $x, a \in \mathbb{R}$. The above system

(3.4.18) is equivalent to

$$\begin{cases} 2\alpha_1\,\mathbb{E}\Big[\psi\,\mathrm{I}_{\{S\le K\}}\Big] \ge \alpha_1\,\mathbb{E}\Big[\psi\Big] + \lambda, \\ 2\alpha_2\,\mathbb{E}\Big[\psi_i\,\mathrm{I}_{\{X_i\le K_i\}}\Big] \ge \alpha_2\mathbb{E}[\psi_i] - \lambda - \lambda_i, \quad i=1,...,n, \\ 2\alpha_3\,\mathbb{E}\Big[\psi_{ij}\,\mathrm{I}_{\{X_{ij}\le k_{ij}\}}\Big] \ge \alpha_3\,\mathbb{E}\Big[\psi_{ij}\Big] + \lambda_i, \quad i=1,...,n, \quad j=1,...,n_i, \\ \alpha_1\,\mathbb{E}\Big[\psi\Big] \ge 2\alpha_1\,\mathbb{E}\Big[\psi\,\mathrm{I}_{\{S<K\}}\Big] - \lambda, \\ \alpha_2\,\mathbb{E}\Big[\psi_i\Big] \ge 2\alpha_2\,\mathbb{E}\Big[\psi_i\,\mathrm{I}_{\{X_i<K_i\}}\Big] + \lambda + \lambda_i, \ i=1,\dots,n, \\ \alpha_3\,\mathbb{E}\Big[\psi_{ij}\Big] \ge 2\alpha_3\,\mathbb{E}\Big[\psi_{ij}\,\mathrm{I}_{\{X_{ij}<k_{ij}\}}\Big] - \lambda_i, \quad i=1,\dots,n, \quad j=1,...,n_i, \\ \sum_{i=1}^{n} K_i = K, \quad \sum_{j=1}^{n_i} k_{ij} = K_i, \quad i=1,...,n. \end{cases} \tag{3.4.19}$$

By the definitions of functions F, G, F_i, G_i, F_{ij}, and G_{ij}, we see that system (3.4.19) is equivalent to

$$\begin{cases} G(K) \le \dfrac{\mathbb{E}[\psi] + \frac{\lambda}{\alpha_1}}{2\mathbb{E}[\psi]} \le F(K), \\ G_i(K_i) \le \dfrac{\mathbb{E}[\psi_i] - \frac{\lambda+\lambda_i}{\alpha_2}}{2\mathbb{E}[\psi_i]} \le F_i(K_i), \quad i=1,\dots,n, \\ G_{ij}(k_{ij}) \le \dfrac{\mathbb{E}[\psi_{ij}] + \frac{\lambda_i}{\alpha_3}}{2\mathbb{E}[\psi_{ij}]} \le F_{ij}(k_{ij}), \quad i=1,...,n, \quad j=1,\dots,n_i, \\ \sum_{i=1}^{n} K_i = K, \quad \sum_{j=1}^{n_i} k_{ij} = K_i, \quad i=1,...,n. \end{cases} \tag{3.4.20}$$

Hence, if

$$\begin{cases} 0 < \dfrac{\mathbb{E}[\psi] + \frac{\lambda}{\alpha_1}}{2\mathbb{E}[\psi]} < 1, \\ 0 < \dfrac{\mathbb{E}[\psi_i] - \frac{\lambda+\lambda_i}{\alpha_2}}{2\mathbb{E}[\psi_i]} < 1, \quad i=1,\dots,n, \\ 0 < \dfrac{\mathbb{E}[\psi_{ij}] + \frac{\lambda_i}{\alpha_3}}{2\mathbb{E}[\psi_{ij}]} < 1, \quad i=1,...,n, \quad j=1,...,n_i, \end{cases} \tag{3.4.21}$$

by (3.4.9), (3.4.10) and (3.4.11), we know that system (3.4.20) is equivalent to

$$\begin{cases} F^{-1}\Big(\dfrac{\mathbb{E}[\psi] + \frac{\lambda}{\alpha_1}}{2\mathbb{E}[\psi]}\Big) \le K \le F^{-1+}\Big(\dfrac{\mathbb{E}[\psi] + \frac{\lambda}{\alpha_1}}{2\mathbb{E}[\psi]}\Big), \\ F_i^{-1}\Big(\dfrac{\mathbb{E}[\psi_i] - \frac{\lambda+\lambda_i}{\alpha_2}}{2\mathbb{E}[\psi_i]}\Big) \le K_i \le F_i^{-1+}\Big(\dfrac{\mathbb{E}[\psi_i] - \frac{\lambda+\lambda_i}{\alpha_2}}{2\mathbb{E}[\psi_i]}\Big), \quad i=1,\dots,n, \\ F_{ij}^{-1}\Big(\dfrac{\mathbb{E}[\psi_{ij}] + \frac{\lambda_i}{\alpha_3}}{2\mathbb{E}[\psi_{ij}]}\Big) \le k_{ij} \le F_{ij}^{-1+}\Big(\dfrac{\mathbb{E}[\psi_i] + \frac{\lambda_i}{\alpha_3}}{2\mathbb{E}[\psi_{ij}]}\Big), \quad i=1,\dots,n, \quad j=1,...,n_i, \\ \sum_{i=1}^{n} K_i = K, \quad \sum_{j=1}^{n_i} k_{ij} = K_i, \quad i=1,...,n. \end{cases} \tag{3.4.22}$$

Note that conditions in (3.4.21) are equivalent to those in (3.4.16). Hence, by (3.4.22) and (3.4.3), we see that any $(\mathbf{K}, \mathbf{k}_1, ..., \mathbf{k}_n)$ with expression (3.4.14) satisfies the KKT conditions in (3.4.18), providing that the parameters p, λ, p_i, λ_i, p_{ij}, $i = 1, ..., n$, $j = 1, ..., n_i$, satisfy the equations in (3.4.15) and the inequalities in (3.4.16). It completes the proof of Theorem 3.4.2. □

From Theorem 3.4.2, we know that optimal solutions to problems (3.4.1) are not unique. However, under some additional conditions, problems (3.4.1) has the following unique solution:

Theorem 3.4.3. *Under the assumptions and notations of Theorem 3.4.2, if $F^{-1}(q)$, $F_i^{-1}(q)$ and $F_{ij}^{-1}(q)$ are continuous and strictly increasing in $q \in (0, 1)$, $i = 1, ..., n$, $j = 1, ..., n_i$, $G_1(\underline{b}+) > G_2(\underline{b}+)$ and $G_1(\overline{b}-) < G_2(\overline{b}-)$, where*

$$\begin{cases} G_1(x) = \displaystyle\sum_{i=1}^{n} F_i^{-1}\Big(\frac{\mathbb{E}[\psi_i] - \frac{q_i(x)}{\alpha_2}}{2\mathbb{E}[\psi_i]}\Big), \\ G_2(x) = F^{-1}\Big(\dfrac{\mathbb{E}[\psi] + \frac{x}{\alpha_1}}{2\mathbb{E}[\psi]}\Big), \\ \ q_i(x) = x + f_i^{-1}(-x), \\ \ f_i(x) = x + \alpha_2\Big(2\mathbb{E}[\psi_i]\, F_i\Big(g_i\Big(\dfrac{x}{\alpha_3}\Big)\Big) - \mathbb{E}[\psi_i]\Big), \\ \ g_i(x) = \displaystyle\sum_{j=1}^{n_i} F_{ij}^{-1}\Big(\frac{\mathbb{E}[\psi_{ij}] + x}{2\mathbb{E}[\psi_{ij}]}\Big), \end{cases} \tag{3.4.23}$$

and

$$\begin{cases} \underline{b} = \max\{-\alpha_1\mathbb{E}[\psi],\ -f_i(\alpha_3\,\overline{m}),\ q_i^{-1}(-\alpha_2\,\overline{M}),\quad i = 1, ..., n\}, \\ \overline{b} = \min\{\alpha_1\mathbb{E}[\psi],\ -f_i(\alpha_3\,\underline{m}),\ q_i^{-1}(-\alpha_2\,\underline{M}),\quad i = 1, ..., n\}, \end{cases} \tag{3.4.24}$$

then, problem (3.4.1) has the following unique solution:

$$\begin{cases} K^* = F^{-1}\Big(\dfrac{\mathbb{E}[\psi] + \frac{\lambda^*}{\alpha_1}}{2\mathbb{E}[\psi]}\Big), \\ K_i^* = F_i^{-1}\Big(\dfrac{\mathbb{E}[\xi_i] - \frac{\lambda^* + f_i^{-1}(-\lambda^*)}{\alpha_2}}{2\mathbb{E}[\psi_i]}\Big), \quad i = 1, \ldots, n, \\ k_{ij}^* = F_{ij}^{-1}\Big(\dfrac{\mathbb{E}[\psi_{ij}] + \frac{f_i^{-1}(-\lambda^*)}{\alpha_3}}{2\mathbb{E}[\psi_{ij}]}\Big), \quad i = 1, ..., n,\ j = 1, \ldots, n_i, \end{cases} \tag{3.4.25}$$

where $\lambda^* \in (\underline{b},\ \overline{b})$ is the unique solution to equation

$$\sum_{i=1}^{n} F_i^{-1}\Big(\frac{\mathbb{E}[\psi_i] - \frac{\lambda^* + f_i^{-1}(-\lambda^*)}{\alpha_2}}{2\mathbb{E}[\psi_i]}\Big) = F^{-1}\Big(\frac{\mathbb{E}[\psi] + \frac{\lambda^*}{\alpha_1}}{2\mathbb{E}[\psi]}\Big). \tag{3.4.26}$$

Proof. If $F^{-1}(q)$, $F_i^{-1}(q)$ and $F_{ij}^{-1}(q)$ are continuous and strictly increasing on $(0,1)$, then we know that $F^{-1(p)}(q) = F^{-1}(q)$, $F_i^{-1(p)}(q) = F_i^{-1}(q)$ and $F_{ij}^{-1(p)}(q) = F_{ij}^{-1}(q)$ for any $p \in [0,1]$ and $q \in (0,1)$, and that $F(x)$, $F_i(x)$ and $F_{ij}(x)$ are continuous on $\mathbb{R}$. Moreover, $g_i(x)$ defined in (3.4.23) and $F_i^{-1}\left(\frac{\mathbb{E}[\psi_i]+x}{2\mathbb{E}[\psi_i]}\right)$ are continuous and strictly increasing respectively on $x \in (\underline{m}, \overline{m})$ and $x \in (\underline{M}, \overline{M})$. Thus, by (3.4.14), (3.4.15), and (3.4.16), we see that an optimal solution $(\mathbf{K}, \mathbf{k}_1, ..., \mathbf{k}_n)$ for problem (3.4.1) has the following expression:

$$\begin{cases} K = F^{-1}\Big(\dfrac{\mathbb{E}[\psi] + \frac{\lambda}{\alpha_1}}{2\mathbb{E}[\psi]}\Big), \\ K_i = F_i^{-1}\Big(\dfrac{\mathbb{E}[\psi_i] - \frac{\lambda+\lambda_i}{\alpha_2}}{2\mathbb{E}[\psi_i]}\Big), \quad i = 1, \ldots, n, \\ k_{ij} = F_{ij}^{-1}\Big(\dfrac{\mathbb{E}[\psi_{ij}] + \frac{\lambda_i}{\alpha_3}}{2\mathbb{E}[\psi_{ij}]}\Big), \quad i = 1, ..., n, \quad j = 1, \ldots, n_i, \end{cases} \tag{3.4.27}$$

where the parameters λ, λ_i, $i = 1, ..., n$, are solutions to the following system of equations and inequalities:

$$\begin{cases} \displaystyle\sum_{i=1}^{n} F_i^{-1}\Big(\dfrac{\mathbb{E}[\psi_i] - \frac{\lambda+\lambda_i}{\alpha_2}}{2\mathbb{E}[\psi_i]}\Big) = F^{-1}\Big(\dfrac{\mathbb{E}[\psi] + \frac{\lambda}{\alpha_1}}{2\mathbb{E}[\psi]}\Big), \\ \displaystyle\sum_{j=1}^{n_i} F_{ij}^{-1}\Big(\dfrac{\mathbb{E}[\psi_{ij}] + \frac{\lambda_i}{\alpha_3}}{2\mathbb{E}[\psi_{ij}]}\Big) = F_i^{-1}\Big(\dfrac{\mathbb{E}[\psi_i] - \frac{\lambda+\lambda_i}{\alpha_2}}{2\mathbb{E}[\psi_i]}\Big), \quad i = 1, \ldots, n. \\ -\mathbb{E}[\psi] < \dfrac{\lambda}{\alpha_1} < \mathbb{E}[\psi], \\ \underline{M} < -\dfrac{\lambda + \lambda_i}{\alpha_2} < \overline{M}, \quad i = 1, \ldots, n, \\ \underline{m} < \dfrac{\lambda_i}{\alpha_3} < \overline{m}, \quad i = 1, ..., n. \end{cases} \tag{3.4.28}$$

By the definition of g_i in (3.4.23), the second line of (3.4.28) is rewritten as

$$g_i\Big(\frac{\lambda_i}{\alpha_3}\Big) = F_i^{-1}\Big(\frac{\mathbb{E}[\psi_i] - \frac{\lambda+\lambda_i}{\alpha_2}}{2\mathbb{E}[\psi_i]}\Big), \quad i = 1, ..., n. \tag{3.4.29}$$

Note that distribution function $F_i(x)$ is increasing and continuous in $x \in \mathbb{R}$, hence $F_i(F_i^{-1}(y)) = y$ for any $0 < y < 1$. Thus, it follows from (3.4.29) that

$$F_i\Big(g_i\Big(\frac{\lambda_i}{\alpha_3}\Big)\Big) = \frac{\mathbb{E}[\psi_i] - \frac{\lambda+\lambda_i}{\alpha_2}}{2\mathbb{E}[\psi_i]}, \quad i = 1, ..., n. \tag{3.4.30}$$

By (3.4.30) and the definition of f_i in (3.4.23), we obtain

$$-\lambda = \lambda_i + \alpha_2\Big(2\mathbb{E}[\psi_i]\, F_i\Big(g_i\Big(\frac{\lambda_i}{\alpha_3}\Big)\Big) - \mathbb{E}[\psi_i]\Big) = f_i(\lambda_i), \quad i = 1, ..., n. \tag{3.4.31}$$

Note that $F_i\Big(g_i\Big(\frac{x}{\alpha_3}\Big)\Big)$ is continuous and increasing on $x \in (\alpha_3\,\underline{m},\ \alpha_3\,\overline{m})$. Hence, $f_i(x)$ is continuous and strictly increasing on $(\alpha_3\,\underline{m},\ \alpha_3\,\overline{m})$. Thus, $f_i^{-1}(x)$ is continuous and strictly increasing on $(f_i(\alpha_3\,\underline{m}),\ f_i(\alpha_3\,\overline{m}))$. Therefore, (3.4.31) implies

$$\lambda_i = f_i^{-1}(-\lambda), \quad i = 1, ..., n. \tag{3.4.32}$$

Hence, the condition on the fifth line of (3.4.28) is equivalent to

$$\underline{m} < \frac{f_i^{-1}(-\lambda)}{\alpha_3} < \overline{m}, \quad i = 1, ..., n,$$

which is further equivalent to $-f_i(\alpha_3\,\overline{m}) < \lambda < -f_i(\alpha_3\,\underline{m})$, $i = 1, ..., n$. Hence, λ satisfies

$$\underline{c} < \lambda < \bar{c}, \tag{3.4.33}$$

where $\underline{c} = \max\{-f_i(\alpha_3\,\overline{m}),\ i = 1, ..., n\}$ and $\bar{c} = \min\{-f_i(\alpha_3\,\underline{m}),\ i = 1, ..., n\}$.

It is well-known that any increasing function has non-negative derivatives almost everywhere on its domain. Hence, by the definition of f_i in (3.4.23) and noticing that F_i and g_i are increasing, we have

$$f_i'(x) = 1 + 2\alpha_2 \mathbb{E}[\psi_i]\, F_i'\Big(g_i\Big(\frac{x}{\alpha_3}\Big)\Big)\, g_i'\Big(\frac{x}{\alpha_3}\Big)\,\frac{1}{\alpha_3} > 1.$$

Thus, by the definition of q_i in (3.4.23) and the derivative rule for an inverse function, we have

$$q_i'(x) = 1 + \frac{1}{f_i'(f_i^{-1}(-x)) \times (-1)} > 0.$$

Therefore, $q_i(x) = x + f_i^{-1}(-x)$ is continuous and strictly increasing on $(-f_i(\alpha_3\,\overline{m}),\ -f_i(\alpha_3\,\underline{m}))$. Now, the condition on the fourth line of (3.4.28) is reduced to

$$\underline{M} < -\frac{q_i(\lambda)}{\alpha_2} < \overline{M}, \quad i = 1, ..., n, \tag{3.4.34}$$

which is equivalent to $-\alpha_2\,\overline{M} < q_i(\lambda) < -\alpha_2\,\underline{M},\ \ i = 1, ..., n$. Thus, it follows from (3.4.34) that

$$q_i^{-1}(-\alpha_2\,\overline{M}) < \lambda < q_i^{-1}(-\alpha_2\,\underline{M}), \quad i = 1, ..., n,$$

which imply

$$\underline{d} < \lambda < \bar{d}, \tag{3.4.35}$$

where $\underline{d} = \max\{q_i^{-1}(-\alpha_2\,\overline{M}),\ i = 1, ..., n\}$ and $\bar{d} = \min\{q_i^{-1}(-\alpha_2\,\underline{M}),\ i = 1, ..., n\}$.

In addition, the condition on the first line of (3.4.28) is reduced to

$$G_1(\lambda) = \sum_{i=1}^{n} F_i^{-1}\Big(\frac{\mathbb{E}[\psi_i] - \frac{q_i(\lambda)}{\alpha_2}}{2\mathbb{E}[\psi_i]}\Big) = F^{-1}\Big(\frac{\mathbb{E}[\psi] + \frac{\lambda}{\alpha_1}}{2\mathbb{E}[\psi]}\Big) = G_2(\lambda). \tag{3.4.36}$$

Note that the intersection of the domains of $q_i(\lambda)$, $i = 1, ..., n$, is $\bigcap_{i=1}^{n}(-f_i(\alpha\overline{m}),\ -f_i(\alpha\underline{m})) = (\underline{c},\ \bar{c})$. In addition, the intersection of the domains of $F_i^{-1}\Big(\frac{\mathbb{E}[\psi_i]+x}{2\mathbb{E}[\psi_i]}\Big)$, $i = 1, ..., n$, is $(\underline{M},\ \overline{M})$. Hence, $-\frac{q_i(\lambda)}{\alpha_2}$, $i = 1, ..., n$, in function $G_1(\lambda)$ should satisfy

$$\underline{M} < -\frac{q_i(\lambda)}{\alpha_2} < \overline{M}, \quad i = 1, ..., n,$$

which are equivalent to (3.4.34) or (3.4.35). Furthermore, $\frac{\lambda}{\alpha_1}$ in $G_2(\lambda)$ should satisfy $-\mathbb{E}[\psi] < \frac{\lambda}{\alpha_1} < \mathbb{E}[\psi]$, which is equivalent to

$$-\alpha_1\mathbb{E}[\psi] < \lambda < \alpha_1\mathbb{E}[\psi]. \tag{3.4.37}$$

Let $\underline{b} = \max\{-\alpha_1\mathbb{E}[\psi],\ \underline{c},\ \underline{d}\}$ and $\bar{b} = \min\{\alpha_1\mathbb{E}[\psi],\ \bar{c},\ \bar{d}\}$, which are respectively equivalent to the definitions of $\underline{b}$ and $\bar{b}$ in (3.4.24). Note that $G_1(\lambda)$ is continuous and strictly decreasing while $G_2(\lambda)$ is continuous and strictly increasing in $\lambda \in (\underline{b},\ \bar{b})$. Thus, if $G_1(\underline{b}+) > G_2(\underline{b}+)$ and $G_1(\bar{b}-) < G_2(\bar{b}-)$, then there exists a unique $\lambda^* \in (\underline{b},\ \bar{b})$ satisfying $G_1(\lambda^*) = G_2(\lambda^*)$, which means that equation (3.4.26) has the unique solution λ^*. Thus, by (3.4.32), we have $\lambda_i^* = f_i^{-1}(-\lambda^*)$, hence problem (3.4.1) has the unique solution (3.4.25). □

If $\alpha_3 = 0$, problem (3.1.4) is reduced to the following problem:

$$\begin{cases} \min_{(K,K_1,...,K_n)\in\mathbb{R}^{n+1}} & \Big\{(1-\alpha)\mathbb{E}\Big[\psi\Big|S-K\Big|\Big] + \ \alpha\sum_{i=1}^{n}\mathbb{E}\Big[\psi_i\Big|X_i-K_i\Big|\Big]\Big\} \\ \text{s.t.} & \sum_{i=1}^{n} K_i = K, \end{cases} \tag{3.4.38}$$

where $0 < \alpha < 1$. Solutions to problem (3.4.38) will give the optimal required total capital and optimal allocation scheme for a company with main business lines.

Assumption 3.4.2. For problem (3.4.38), we assume that random variables ψ, ψ_i, $i = 1, ..., n$, satisfy

$$\psi \geq 0, \quad \psi_i \geq 0, \quad \mathbb{E}[\psi] > 0, \quad \mathbb{E}[\psi_i] > 0, \quad i = 1, ..., n,$$

and that expectations $\mathbb{E}(\psi|S|)$, $\mathbb{E}(\psi_i|X_i|)$, $i = 1, ..., n$, exist. □

Following the same arguments used in the proofs for Theorem 3.4.2 and Theorem 3.4.3, we easily obtain the general solution and the unique solution to problem (3.4.38) in the following two theorems.

Theorem 3.4.4. *Under Assumption 3.4.2 and the notations of Theorem 3.4.2, solutions to problem (3.4.38) do exist and a solution $(K, K_1, ..., K_n)$ to problem (3.4.38) has the following expression:*

$$\begin{cases} K = F^{-1(p)}\Big(\dfrac{\mathbb{E}[\psi] + \frac{\lambda}{1-\alpha}}{2\mathbb{E}[\psi]}\Big), \\ K_i = F_i^{-1(p_i)}\Big(\dfrac{\mathbb{E}[\psi_i] - \frac{\lambda}{\alpha}}{2\mathbb{E}[\psi_i]}\Big), \quad i = 1, \dots, n, \end{cases} \tag{3.4.39}$$

for any $p, p_i \in [0,1]$, $i = 1, ..., n$, and any $\lambda \in \mathbb{R}$, providing that the parameters p, λ, p_i, $i = 1, ..., n$, satisfy the following system of equations and inequalities:

$$\begin{cases} \displaystyle\sum_{i=1}^{n} F_i^{-1(p_i)}\Big(\dfrac{\mathbb{E}[\psi_i] - \frac{\lambda}{\alpha}}{2\mathbb{E}[\psi_i]}\Big) = F^{-1(p)}\Big(\dfrac{\mathbb{E}[\psi] + \frac{\lambda}{1-\alpha}}{2\mathbb{E}[\psi]}\Big), \\ -\mathbb{E}[\psi] < \dfrac{\lambda}{1-\alpha} < \mathbb{E}[\psi], \\ \underline{M} < -\dfrac{\lambda}{\alpha} < \overline{M}, \quad i = 1, \dots, n. \end{cases} \tag{3.4.40}$$

Theorem 3.4.5. *Under the assumptions and notations of Theorem 3.4.4, if $F^{-1}(q)$ and $F_i^{-1}(q)$ are continuous and strictly increasing in $q \in (0,1)$, $i = 1, ..., n$, $H_1(\underline{\lambda}+) > H_2(\underline{\lambda}+)$ and $H_1(\overline{\lambda}-) < H_2(\overline{\lambda}-)$, where*

$$\begin{cases} H_1(x) = \displaystyle\sum_{i=1}^{n} F_i^{-1}\Big(\dfrac{\mathbb{E}[\psi_i] - \frac{x}{\alpha}}{2\mathbb{E}[\psi_i]}\Big), \\ H_2(x) = F^{-1}\Big(\dfrac{\mathbb{E}[\psi] + \frac{x}{1-\alpha}}{2\mathbb{E}[\psi]}\Big), \end{cases} \tag{3.4.41}$$

and

$$\begin{cases} \underline{\lambda} = \max\{-(1-\alpha)\mathbb{E}[\psi],\ -\alpha\,\overline{M}\}, \\ \overline{\lambda} = \min\{(1-\alpha)\mathbb{E}[\psi],\ \ -\alpha\,\underline{M}\}, \end{cases} \tag{3.4.42}$$

then, problem (3.4.38) has the following unique solution:

$$\begin{cases} K^* = F^{-1}\Big(\dfrac{\mathbb{E}[\psi] + \frac{\lambda^*}{1-\alpha}}{2\mathbb{E}[\psi]}\Big), \\ K_i^* = F_i^{-1}\Big(\dfrac{\mathbb{E}[\psi_i] - \frac{\lambda^*}{\alpha}}{2\mathbb{E}[\psi_i]}\Big), \quad i = 1, \dots, n, \end{cases} \tag{3.4.43}$$

where $\lambda^*(= \lambda^*(\alpha)) \in (\underline{\lambda},\ \overline{\lambda})$ is the unique solution to equation

$$\sum_{i=1}^{n} F_i^{-1}\Big(\dfrac{\mathbb{E}[\psi_i] - \frac{\lambda^*}{\alpha}}{2\mathbb{E}[\psi_i]}\Big) = F^{-1}\Big(\dfrac{\mathbb{E}[\psi] + \frac{\lambda^*}{1-\alpha}}{2\mathbb{E}[\psi]}\Big). \tag{3.4.44}$$

3.5 Numerical illustrations

Now, we would like to provide some numerical examples based on model (3.2.1) or Theorem 3.2.2 and model (3.4.38) or Theorem 3.4.5. We will illustrate the result with multivariate normal distribution.

Let $\mathbf{X} = (X_1, ..., X_d)$ be a d-dimensional normal random vector with a d-dimensional normal distribution $\mathcal{N}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = \mathbb{E}(\mathbf{X}) = (\mu_1, ..., \mu_d)$ is the mean vector, $\mu_i = \mathbb{E}(X_i)$, $i = 1, ..., d$, $\boldsymbol{\Sigma} = (Cov(X_i, X_j))_{i,j=1,...,d}$ is a positive-definite covariance matrix of $\mathbf{X}$, $Cov(X_i, X_j) = \rho_{ij}\,\sigma_i\,\sigma_j$, where $\sigma_i = \sqrt{Var(X_i)}$, $\sigma_j = \sqrt{Var(X_j)}$, and ρ_{ij} is the correlation coefficient between X_i and X_j satisfying $-1 < \rho_{ij} = \rho_{ji} < 1$ for $1 \le i < j \le d$. The joint density function of $(X_1, ..., X_d)$ is

$$f(\mathbf{x}) = f(x_1, ..., x_d) = \frac{1}{\sqrt{(2\pi)^d \det(\boldsymbol{\Sigma})}}\, e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})'}, \quad \mathbf{x} \in \mathbb{R}^d. \tag{3.5.1}$$

Thus, X_i and $S(\mathbf{X})$ have normal distributions $N(\mu_i, \sigma_i^2)$ and $N(\mu_S, \sigma_S^2)$, respectively, where $\mu_S = \sum_{i=1}^d \mu_i$, $\sigma_S = \sqrt{Var(S(\mathbf{X}))}$, and $Var(S(\mathbf{X})) = \sum_{i=1}^d \sigma_i^2 + 2\sum_{1\le i<j\le d} \rho_{ij}\,\sigma_i\,\sigma_j$. Let ϕ and Φ be the density and distribution functions of the standard normal distribution $N(0, 1)$, respectively. If required capital for risks is determined by VaR at confidence level α, then

$$\begin{aligned} \mathrm{VaR}_\alpha(X_i) &= \mu_i + \sigma_i\,\Phi^{-1}(\alpha), \quad i = 1, ..., d, & (3.5.2)\\ \mathrm{VaR}_\alpha(S(\mathbf{X})) &= \mu_S + \sigma_S\,\Phi^{-1}(\alpha). & (3.5.3) \end{aligned}$$

If required capital for risks is determined by CVaR at confidence level α, then

$$\mathrm{CTE}_\alpha(X_i) = \mathbb{E}\left(X_i \,|\, X_i > \mathrm{VaR}_\alpha(X_i)\right) = \mu_i + \sigma_i \times \frac{\phi\left(\frac{\mathrm{VaR}_\alpha(X_i)-\mu_i}{\sigma_i}\right)}{1-\Phi\left(\frac{\mathrm{VaR}_\alpha(X_i)-\mu_i}{\sigma_i}\right)}, \quad i = 1, ..., d, \tag{3.5.4}$$

and

$$\mathrm{CTE}_\alpha(S(\mathbf{X})) = \mathbb{E}\left(S(\mathbf{X}) \,|\, S(\mathbf{X}) > \mathrm{VaR}_\alpha(S(\mathbf{X}))\right) = \mu_S + \sigma_S \times \frac{\phi\left(\frac{\mathrm{VaR}_\alpha(S(\mathbf{X}))-\mu_S}{\sigma_S}\right)}{1-\Phi\left(\frac{\mathrm{VaR}_\alpha(S(\mathbf{X}))-\mu_S}{\sigma_S}\right)} \tag{3.5.5}$$

See Johnson et al. (1995) for these results about normal distributions.

Here we consider $d = 3$. We assume that the marginal normal distributions of $\mathbf{X} = (X_1, X_2, X_3)$ have the following expectations and variances: $\mathbb{E}(X_1) = 130$, $Var(X_1) = 900$, $\mathbb{E}(X_2) = 150$, $Var(X_2) = 2500$, $\mathbb{E}(X_3) = 170$, and $Var(X_3) = 400$. Further more, by affine transformation of multivariate normal distribution, the joint distribution of (X_i, S) where

$i = 1, 2, 3$ is $\mathcal{N}_2(\mathbf{A}_i\boldsymbol{\mu}, \mathbf{A}_i\mathbf{\Sigma}\mathbf{A}_i^T)$ where $\mathbf{A}_i$ is a constant $2\times d$ matrix such that $(X_i, S) = \mathbf{A}_i\mathbf{X}$. In our case, we set $d = 1$ and we can derivate that

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad \mathbf{A}_3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}. \tag{3.5.6}$$

We use the above normal risk vector (X_1, X_2, X_3) and the allocation principles of (3.2.22), (3.2.24), and (3.2.26) to illustrate the applications of model (3.2.1) and Theorem (3.2.2) and show the influences of dependence on the optimal allocations. With the definition of ξ, ξ_i, v and v_i in (3.2.19), the model in (3.2.1) is reduced to

$$\begin{cases} \min_{(K,K_1,\dots,K_n)\in\mathbb{R}^{n+1}} \left\{(1-\alpha)\frac{\mathbb{E}\left[\left(S-K\right)^2|A\right]}{\sum_{i=1}^n \mathbb{E}(X_i|A_i)} + \alpha\sum_{i=1}^n \frac{\mathbb{E}\left[\left(X_i-K_i\right)^2|A_i\right]}{\mathbb{E}(X_i|A_i)}\right\} \\ \text{s.t.} \quad \sum_{i=1}^n K_i = K. \end{cases} \tag{3.5.7}$$

with different definition of event A and A_i, $i = 1, 2, 3$.

Then, we set $q = 0.99$ in (3.2.22), (3.2.24), (3.2.26), and (3.2.28), thus, the CTE principle (3.2.22) is reduced to

$$K_i^* = \frac{\mathbb{E}(X_i \mid S > \mathrm{VaR}_{0.99}(S))}{\mathrm{CTE}_{0.99}(S)} \times K^* \tag{3.5.8}$$

with $K^* = (1-\alpha)\mathrm{VaR}_{0.99}(S) + \alpha\mathrm{CTE}_{0.99}(S)$, and by Proposition 3.7 of Cai and Wang (2020), we know that there exist sets $A \subset \mathbb{R}$, satisfying $\mathbb{E}(S|S \in A) = \mathrm{VaR}_{0.99}(S)$ and $A_i = \{S > \mathrm{VaR}_{0.99}(S)\}$, $i = 1, ..., n$; the haircut principle (3.2.24) is reduced to

$$K_i^* = \frac{\mathrm{VaR}_{0.99}(X_i)}{\sum_{i=1}^n \mathrm{VaR}_{0.99}(X_i)} \times K^* \tag{3.5.9}$$

with $K^* = (1-\alpha)\mathrm{VaR}_{0.99}(S) + \alpha\sum_{i=1}^n \mathrm{VaR}_{0.99}(X_i)$, and by Proposition 3.7 of Cai and Wang (2020), we know that there exist sets A, $A_i \subset \mathbb{R}$, $i = 1, ..., n$, satisfying $\mathbb{E}(S|S \in A) = \mathrm{VaR}_{0.99}(S)$ and $\mathbb{E}(X_i|X_i \in A_i) = \mathrm{VaR}_{0.99}(X_i)$, $i = 1, ..., n$; the haircut principle (3.2.26) is reduced to

$$K_i^* = \frac{\mathrm{VaR}_{0.99}(X_i)}{\sum_{i=1}^n \mathrm{VaR}_{0.99}(X_i)} \times K^* \tag{3.5.10}$$

with $K^* = (1-\alpha)\mathrm{CTE}_{0.99}(S) + \alpha\sum_{i=1}^n \mathrm{VaR}_{0.99}(X_i)$, and $A = \{S > \mathrm{VaR}_{0.99}(S)\}$, and by Proposition 3.7 of Cai and Wang (2020), we know that there exist sets $A_i \subset \mathbb{R}$, $i = 1, ..., n$, satisfying $\mathbb{E}(X_i|X_i \in A_i) = \mathrm{VaR}_{0.99}(X_i)$, $i = 1, ..., n$; and, the CTE principle (3.2.28) is reduced to

$$K_i^* = \frac{\mathrm{CTE}_{0.99}(X_i)}{\sum_{i=1}^n \mathrm{CTE}_{0.99}(X_i)} \times K^* \tag{3.5.11}$$

with $K^* = (1-\alpha)\text{CTE}_{0.99}(S) + \alpha\sum_{i=1}^{n}\text{CTE}_{0.99}(X_i)$, and $A = \{S > \text{VaR}_{0.99}(S)\}$ and $A_i = \{S > \text{VaR}_{0.99}(S)\}$, $i = 1, ..., n$.

With the joint density function of (X_i, S), $\mathbb{E}(X_i|S > \text{VaR}_{0.99}(S))$ in (3.5.8) is calculated by the following formula

$$\begin{aligned}\mathbb{E}(X_i|S > \text{VaR}_{0.99}(S)) &= \frac{1}{\mathbb{P}\{S > \text{VaR}_{0.99}(S)\}}\int_{-\infty}^{\infty}\int_{\text{VaR}_{0.99}(S)}^{\infty} x f_{X_i,S}(x,s)dsdx \\ &= \frac{1}{0.01}\int_{-\infty}^{\infty}\int_{\text{VaR}_{0.99}(S)}^{\infty} x f_{X_i,S},(x,s)dsdx\end{aligned}$$

where $f_{X_i,S}$, $i = 1, 2, 3$ is the joint density function of X_i and S.

In the following numerical results, we will illustrate the allocation result for positive dependence, negative dependence and mixed dependence. The numerical results are calculated with Matlab and shown in the tables below.

Note that, for all case (i) (ii) and (iii), we have $\text{VaR}_{0.99}(X_1) = 199.79$, $\text{VaR}_{0.99}(X_2) = 266.32$, $\text{VaR}_{0.99}(X_3) = 216.53$.

(i) Positive dependence: Assume that risks in portfolio (X_1, X_2, X_3) are positively dependent and that the correlation coefficient of any two risks in the portfolio is positive with $\rho_{12} = 0.8$, $\rho_{13} = 0.2$, and $\rho_{23} = 0.3$.

In this case, we have $\sigma_S^2 = 7040$, $\rho_{X_1,S} = 0.8820$, $\rho_{X_2,S} = 0.9535$, $\rho_{X_3,S} = 0.4886$. Also, we have $\text{VaR}_{0.99}(S) = 645.19$, $\text{CTE}_{0.99}(S) = 673.62$, $\mathbb{E}(X_1|S > \text{VaR}_{0.99}(S)) = 200.52$, $\mathbb{E}(X_2|S > \text{VaR}_{0.99}(S)) = 277.06$, $\mathbb{E}(X_3|S > \text{VaR}_{0.99}(S)) = 196.05$. The optimal allocation results are presented in the following table.

		K^*	K_1^*	K_2^*	K_3^*
CTE principle (3.5.8)	$\alpha = 0.05$	646.61	192.48	265.95	188.19
	$\alpha = 0.10$	648.03	192.90	266.53	188.60
	$\alpha = 0.50$	659.41	196.29	271.21	191.91
	$\alpha = 0.90$	670.78	199.67	275.89	195.22
	$\alpha = 0.95$	672.20	200.09	276.47	195.63
Haircut principle (3.5.9)	$\alpha = 0.05$	647.06	189.38	252.44	205.24
	$\alpha = 0.10$	648.94	189.93	253.17	205.84
	$\alpha = 0.50$	663.91	194.31	259.01	210.59
	$\alpha = 0.90$	678.89	198.69	264.86	215.34
	$\alpha = 0.95$	680.76	199.24	265.59	215.93
Haircut principle (3.5.10)	$\alpha = 0.05$	674.07	197.29	262.98	213.81
	$\alpha = 0.10$	674.53	197.42	263.15	213.95
	$\alpha = 0.50$	678.13	198.47	264.56	215.10
	$\alpha = 0.90$	681.73	199.53	265.97	216.24
	$\alpha = 0.95$	682.18	199.66	266.14	216.38
CTE principle (3.5.11)	$\alpha = 0.05$	675.77	198.02	267.15	210.60
	$\alpha = 0.10$	677.91	198.64	268.00	211.27
	$\alpha = 0.50$	695.07	203.67	274.78	216.62
	$\alpha = 0.90$	712.23	208.70	281.56	221.97
	$\alpha = 0.95$	714.38	209.33	282.41	222.64

Table 3.1: Capital allocation result of model (3.2.1) for multivariate normal distribution with correlation $\rho_{12} = 0.8$, $\rho_{13} = 0.2$, and $\rho_{23} = 0.3$

(ii) Negative dependence: Assume that risks in portfolio (X_1, X_2, X_3) are negatively dependent and that the correlation coefficient of any two different risks in the portfolio is negative with $\rho_{12} = -0.4$, $\rho_{13} = -0.1$, and $\rho_{23} = -0.1$.

In this case, we have $\sigma_S^2 = 2280$, $\rho_{X_1,S} = 0.1675$, $\rho_{X_2,S} = 0.7539$, $\rho_{X_3,S} = 0.2513$. Also, we have $\mathrm{VaR}_{0.99}(S) = 561.08$, $\mathrm{CTE}_{0.99}(S) = 577.26$, $\mathbb{E}(X_1|S > \mathrm{VaR}_{0.99}(S)) = 143.40$, $\mathbb{E}(X_2|S > \mathrm{VaR}_{0.99}(S)) = 250.47$, $\mathbb{E}(X_3|S > \mathrm{VaR}_{0.99}(S)) = 183.40$. The optimal allocation results are presented in the following table.

		K^*	K_1^*	K_2^*	K_3^*
CTE principle (3.5.8)	$\alpha = 0.05$	561.89	139.58	243.80	178.51
	$\alpha = 0.10$	562.70	139.78	244.15	178.77
	$\alpha = 0.50$	569.18	141.39	246.96	180.83
	$\alpha = 0.90$	575.64	142.99	249.77	182.88
	$\alpha = 0.95$	576.45	143.20	250.12	183.14
Haircut principle (3.5.9)	$\alpha = 0.05$	567.16	165.99	221.27	179.90
	$\alpha = 0.10$	573.24	167.77	223.64	181.83
	$\alpha = 0.50$	621.86	182.00	242.61	197.25
	$\alpha = 0.90$	670.48	196.23	261.58	212.67
	$\alpha = 0.95$	676.56	198.01	263.95	214.60
Haircut principle (3.5.10)	$\alpha = 0.05$	582.53	170.49	227.26	184.77
	$\alpha = 0.10$	587.80	172.03	229.32	186.45
	$\alpha = 0.50$	629.95	184.37	245.76	199.82
	$\alpha = 0.90$	672.10	196.71	262.21	213.18
	$\alpha = 0.95$	677.37	198.25	264.26	214.86
	$\alpha = 0.90$	681.73	199.53	265.97	216.24
	$\alpha = 0.95$	682.18	199.66	266.14	216.38
CTE principle (3.5.11)	$\alpha = 0.05$	584.23	171.19	230.96	182.07
	$\alpha = 0.10$	591.19	173.23	233.71	184.24
	$\alpha = 0.50$	646.89	189.55	255.73	201.60
	$\alpha = 0.90$	702.60	205.88	277.76	218.96
	$\alpha = 0.95$	709.56	207.92	280.51	221.13

Table 3.2: Capital allocation result of model (3.2.1) for multivariate normal distribution with correlation $\rho_{12} = -0.4$, $\rho_{13} = -0.1$, and $\rho_{23} = -0.1$

(iii) Mixed dependence: Assume that risks in portfolio (X_1, X_2, X_3) are mixedly dependent and that some of the correlation coefficients of two risks in the portfolio are positive while some are negative with $\rho_{12} = 0.2$, $\rho_{13} = 0.8$, and $\rho_{23} = -0.3$.

In this case, we have $\sigma_S^2 = 4760$, $\rho_{X_1,S} = 0.8117$, $\rho_{X_2,S} = 0.7247$, $\rho_{X_3,S} = 0.4203$. Also, we have $\mathrm{VaR}_{0.99}(S) = 610.50$, $\mathrm{CTE}_{0.99}(S) = 633.88$, $\mathbb{E}(X_1|S > \mathrm{VaR}_{0.99}(S)) = 194.90$, $\mathbb{E}(X_2|S > \mathrm{VaR}_{0.99}(S)) = 246.58$, $\mathbb{E}(X_3|S > \mathrm{VaR}_{0.99}(S)) = 192.41$. The optimal allocation results are presented in the following table.

		K^*	K_1^*	K_2^*	K_3^*
CTE principle (3.5.8)	$\alpha = 0.05$	611.67	188.07	237.94	185.66
	$\alpha = 0.10$	612.84	188.43	238.39	186.02
	$\alpha = 0.50$	622.19	191.30	242.03	188.86
	$\alpha = 0.90$	631.54	194.18	245.67	191.70
	$\alpha = 0.95$	632.71	194.54	246.12	192.05
Haircut principle (3.5.9)	$\alpha = 0.05$	614.11	179.73	239.58	194.79
	$\alpha = 0.10$	617.71	180.79	240.99	195.93
	$\alpha = 0.50$	646.57	189.23	252.25	205.09
	$\alpha = 0.90$	675.42	197.68	263.50	214.24
	$\alpha = 0.95$	679.03	198.73	264.91	215.38
Haircut principle (3.5.10)	$\alpha = 0.05$	636.32	186.23	248.25	201.84
	$\alpha = 0.10$	638.76	186.95	249.20	202.61
	$\alpha = 0.50$	658.26	192.66	256.81	208.79
	$\alpha = 0.90$	677.76	198.36	264.42	214.98
	$\alpha = 0.95$	680.20	199.08	265.37	215.75
	$\alpha = 0.90$	681.73	199.53	265.97	216.24
	$\alpha = 0.95$	682.18	199.66	266.14	216.38
CTE principle (3.5.11)	$\alpha = 0.05$	638.01	186.95	252.22	198.84
	$\alpha = 0.10$	642.14	188.16	253.86	200.12
	$\alpha = 0.50$	675.20	197.85	266.93	210.43
	$\alpha = 0.90$	708.26	207.53	279.99	220.73
	$\alpha = 0.95$	712.39	208.75	281.63	222.02

Table 3.3: Capital allocation result of model (3.2.1) for multivariate normal distribution with correlation $\rho_{12} = 0.2$, $\rho_{13} = 0.8$, and $\rho_{23} = -0.3$

Next, we illustrate the applications of model (3.4.38) and Theorem 3.4.5 by setting $\psi = \frac{\mathrm{I}_{\{S \in A\}}}{\mathbb{P}\{S \in A\}}$, $\psi_i = \frac{\mathrm{I}_{\{X_i \in A_i\}}}{\mathbb{P}\{X_i \in A_i\}}$, $i = 1, ..., n$, in model (3.4.38). Then, the model is reduced to

$$\begin{cases} \min_{(K,K_1,...,K_n)\in\mathbb{R}^{n+1}} & \Big\{(1-\alpha)\mathbb{E}\Big[\Big|S-K\Big|\,|A\Big] + \ \alpha \sum_{i=1}^n \mathbb{E}\Big[\Big|X_i - K_i\Big|\,|A_i\Big]\Big\} \\ \text{s.t.} & \sum_{i=1}^n K_i = K, \end{cases} \tag{3.5.12}$$

Thus, by (3.4.43), the optimal allocation $(K^*, K_1^*, ..., K_n^*)$ for model (3.4.38) is reduced to

$$\begin{cases} K^* = F^{-1}\Big(\dfrac{1+\frac{\lambda^*}{1-\alpha}}{2}\Big) = F^{-1}(q^*), \\ K_i^* = F_i^{-1}\Big(\dfrac{1-\frac{\lambda^*}{\alpha}}{2}\Big) = F_i^{-1}(q_0^*), \quad i = 1,\dots,n, \end{cases} \tag{3.5.13}$$

where $F(x) = \mathbb{P}\{S \le x \,|\, A\}$, $F_i(x) = \mathbb{P}\{X_i \le x \,|\, A_i\}$, $S = X_1 + X_2 + X_3$,

$$q^* = \frac{1+\frac{\lambda^*}{1-\alpha}}{2}$$

and

$$q_0^* = \frac{1-\frac{\lambda^*}{\alpha}}{2},$$

and $\lambda^*(=\lambda^*(\alpha)) \in (\underline{\lambda},\ \overline{\lambda})$ is the unique solution to equation

$$\sum_{i=1}^{n} F_i^{-1}\Big(\frac{1-\frac{\lambda^*}{\alpha}}{2}\Big) = F^{-1}\Big(\frac{1+\frac{\lambda^*}{1-\alpha}}{2}\Big).$$

We consider the following four choices of the sets A and A_i.

(a) Take $A \subset \mathbb{R}$ so that $\mathbb{E}(S|S \in A) = \mathrm{VaR}_{0.99}(S)$ and let $A_i = \{S > \mathrm{VaR}_{0.99}(S)\}$, $i = 1,...,n$. By Proposition 3.7 of Cai and Wang (2020), if $\mathrm{VaR}_{0.99}(S) > \mathbb{E}(S)$, then there exists a unique $x^* \in \mathbb{R}$ so that $A = (x^*, \infty)$ and $\mathbb{E}[S|S \in A] = \mathbb{E}[S|S > x^*] = \mathrm{VaR}_{0.99}(S)$. In this case, the distribution functions $F(x)$ and $F_i(x)$ are reduced to

$$F(x) = \mathbb{P}\{S \le x \,|\, A\} = \mathbb{P}\{S \le x \,|\, S > x^*\} = \begin{cases} 0 & x \le x^*, \\ \frac{F_S(x)-F_S(x^*)}{1-F_S(x^*)}, & x > x^*, \end{cases} \tag{3.5.14}$$

and

$$\begin{aligned} F_i(x) &= \mathbb{P}\{X_i \le x \,|\, S > \mathrm{VaR}_{0.99}(S)\} = \frac{\mathbb{P}\{X_i \le x\,,\, S > \mathrm{VaR}_{0.99}(S)\}}{\mathbb{P}\{S > \mathrm{VaR}_{0.99}(S)\}} \\ &= \frac{1}{0.01}\int_{-\infty}^{x}\int_{\mathrm{VaR}_{0.99}(S)}^{\infty} f_{X_i,S}(t,s)\,ds\,dt, \qquad -\infty < x < \infty, \end{aligned} \tag{3.5.15}$$

where $f_{X_i,S}$, $i = 1, 2, 3$ is the joint density function of X_i and S.

(b) Let $A = \{S > \mathrm{VaR}_{0.99}(S)\}$ and take $A_i \subset \mathbb{R}$ so that $\mathbb{E}(X|X \in A_i) = \mathrm{VaR}_{0.99}(X_i)$, $i = 1,...,n$. By Proposition 3.7 of Cai and Wang (2020), if $\mathrm{VaR}_{0.99}(X_i) > \mathbb{E}(X_i)$, then there exists a unique $x_i^* \in \mathbb{R}$ so that $A_i = (x_i^*, \infty)$ and $\mathbb{E}[X_i|X_i \in A_i] = \mathbb{E}[X_i|X_i >$

$x_i^*] = \mathrm{VaR}_{0.99}(X_i)$. In this case, the distribution functions $F(x)$ and $F_i(x)$ are reduced to

$$\begin{aligned} F(x) &= \mathbb{P}\{S \le x \,|\, S \in A\} = \mathbb{P}\{S \le x \,|\, S > \mathrm{VaR}_{0.99}(S)\} \\ &= \begin{cases} 0 & x \le \mathrm{VaR}_{0.99}(S), \\ \frac{F_S(x) - F_S(\mathrm{VaR}_{0.99}(S))}{1 - F_S(\mathrm{VaR}_{0.99}(S))}, & x > \mathrm{VaR}_{0.99}(S), \end{cases} \\ &= \begin{cases} 0 & x \le \mathrm{VaR}_{0.99}(S), \\ \frac{F_S(x) - 0.99}{0.01}, & x > \mathrm{VaR}_{0.99}(S), \end{cases} \end{aligned} \tag{3.5.16}$$

and

$$F_i(x) = \mathbb{P}\{X_i \le x \,|\, A_i\} = \mathbb{P}\{X_i \le x \,|\, X_i > x_i^*\} = \begin{cases} 0 & x \le x_i^*, \\ \frac{F_{X_i}(x) - F_{X_i}(x_i^*)}{1 - F_{X_i}(x_i^*)}, & x > x_i^*. \end{cases} \tag{3.5.17}$$

(c) Take sets $A, A_i \subset \mathbb{R}$ so that $\mathbb{E}(S|S \in A) = \mathrm{VaR}_{0.99}(S)$, $\mathbb{E}(X|X \in A_i) = \mathrm{VaR}_{0.99}(X_i)$, $i = 1, ..., n$. In this case, the distribution functions $F(x) = \mathbb{P}\{S \le x \,|\, A\} = \mathbb{P}\{S \le x \,|\, S > x^*\}$ and $F_i(x) = \mathbb{P}\{X_i \le x \,|\, A_i\} = \mathbb{P}\{X_i \le x \,|\, X_i > x_i^*\}$ are given in (3.5.14) and (3.5.17), respectively.

(d) Tabe $A = \{S > \mathrm{VaR}_{0.99}(S)\}$, $A_i = \{X_i > \mathrm{VaR}_{0.99}(X_i)\}$, $i = 1, ..., n$. In this case, the distribution function $F(x) = \mathbb{P}\{S \le x \,|\, A\} = \mathbb{P}\{S \le x \,|\, S > \mathrm{VaR}_{0.99}(S)\}$ is given in (3.5.16) and the distribution function $F_i(x) = \mathbb{P}\{X_i \le x \,|\, A_i\} = \mathbb{P}\{X_i \le x \,|\, X_i > \mathrm{VaR}_{0.99}(X_i)\}$ is given by

$$\begin{aligned} F_i(x) &= \mathbb{P}\{X_i \le x \,|\, X_i \in A_i\} = \mathbb{P}\{X_i \le x \,|\, X_i > \mathrm{VaR}_{0.99}(X_i)\} \\ &= \begin{cases} 0 & x \le \mathrm{VaR}_{0.99}(X_i), \\ \frac{F_{X_i}(x) - F_{X_i}(\mathrm{VaR}_{0.99}(X_i))}{1 - F_{X_i}(\mathrm{VaR}_{0.99}(X_i))}, & x > \mathrm{VaR}_{0.99}(X_i), \end{cases} \\ &= \begin{cases} 0 & x \le \mathrm{VaR}_{0.99}(X_i), \\ \frac{F_{X_i}(x) - 0.99}{0.01}, & x > \mathrm{VaR}_{0.99}(X_i), \end{cases} \end{aligned} \tag{3.5.18}$$

Then, we use the same normal risk vector (X_1, X_2, X_3) as that used in the applications of model (3.2.1) and Theorem 3.2.2. We obtain the optimal allocation $(K^*, K_1^*, K_2^*, K_3^*)$ for each of the four choices of sets A, A_1, A_2, A_3 and for each of the three dependence settings by Matlab. The numerical results for the three choices of sets A, A_1, A_2, A_3 and one dependence setting are presented in one table. In particular, in the following tables, the first five-row numerical results are for the choice (a) of sets A, A_1, A_2, A_3, the second five-row numerical results are for the choice (b) of sets A, A_1, A_2, A_3, the third five-row numerical results are for the choice (c) of sets A, A_1, A_2, A_3, and the fourth five-row numerical results are for the choice (d) of sets A, A_1, A_2, A_3,

Note that $\mathbb{E}(X_1) = 130$, $\mathbb{E}(X_2) = 150$, $\mathbb{E}(X_3) = 170$, $\mathbb{E}[S] = 450$, $\mathrm{VaR}_{0.99}(X_1) = 199.79$, $\mathrm{VaR}_{0.99}(X_2) = 266.32$, $\mathrm{VaR}_{0.99}(X_3) = 216.53$.. For both choices of (b) and (c) for sets A, A_1, A_2, A_3, we find that $x_1^* = 188.41$, $x_2^* = 247.35$, and $x_3^* = 208.94$. In fact, the vale of x_i^* only depends on the distribution of X_i, $i = 1, 2, 3$.

(i) Positive dependence: Assume that risks in portfolio (X_1, X_2, X_3) are positively dependent and that the correlation coefficient of any two risks in the portfolio is positive with $\rho_{12} = 0.8$, $\rho_{13} = 0.2$, and $\rho_{23} = 0.3$. In this case, we have $\mathrm{VaR}_{0.99}(S)) = 645.19$, $\mathrm{CTE}_{0.99}(S) = 673.62$, and $x^* = 613.36$.

From Theorem 3.4.5 , we know that one of the condition to guarantee the uniqueness of the solution is that $H_1(\underline{\lambda}+) > H_2(\underline{\lambda}+)$ and $H_1(\overline{\lambda}-) < H_2(\overline{\lambda}-)$. Therefore, we have derived equation for H_1 and H_2 here. In the following tables, we do have the cases that such condition is not satisfied. We have identified the cases that the condition is not satisfied. In those situations, we list one of the possible solution in the tables.

As we don't have $q_0^* = 0$ for case (a), we only derived the H_1 and H_2 function for case (b) and (c).

For case (b), we have

$$F(x) = \frac{F_S(x) - 0.99}{0.01} = \frac{1 + \frac{y}{1-\alpha}}{2}, \quad x > \mathrm{VaR}_{0.99}(S). \tag{3.5.19}$$

Therefore,

$$F_S(x) = 0.01\frac{1 + \frac{y}{1-\alpha}}{2} + 0.99, \tag{3.5.20}$$

and

$$x = H_2(y) = F_S^{-1}\left(0.01\frac{1 + \frac{y}{1-\alpha}}{2} + 0.99\right), \tag{3.5.21}$$

where F_S^{-1} is the inverse function of cdf function of S. Similarly, we have

$$x = H_1(y) \quad = \quad F_{X_1}^{-1}\left((1 - F_{X_1}(x_1^*))\frac{1 - \frac{y}{\alpha}}{2} + F_{X_1}(x_1^*)\right) \tag{3.5.22}$$

$$+ \quad F_{X_2}^{-1}\left((1 - F_{X_2}(x_2^*))\frac{1 - \frac{y}{\alpha}}{2} + F_{X_2}(x_2^*)\right) \tag{3.5.23}$$

$$+ \quad F_{X_3}^{-1}\left((1 - F_{X_3}(x_3^*))\frac{1 - \frac{y}{\alpha}}{2} + F_{X_3}(x_3^*)\right), \tag{3.5.24}$$

where $F_{X_i}^{-1}$ is the inverse function of cdf function of X_i for $i = 1, 2, 3$.

For case (c), we have same H_1 as case (b), and for H_2, we have

$$x = H_2(y) = F_S^{-1}\left((1 - F_S(x^*))\frac{1 + \frac{y}{1-\alpha}}{2} + F_S(x^*)\right), \tag{3.5.25}$$

	$\underline{\lambda}$	$\overline{\lambda}$	λ^*	q^*	q_0^*	K^*	K_1^*	K_2^*	K_3^*
$\alpha = 0.05$	-0.05	0.05	0.0238	0.5125	0.2617	637.91	189.94	263.23	184.73
$\alpha = 0.10$	-0.10	0.10	0.0465	0.5258	0.2674	638.80	190.21	263.55	185.04
$\alpha = 0.50$	-0.50	0.50	0.1612	0.6612	0.3388	649.43	193.42	267.35	188.66
$\alpha = 0.90$	-0.10	0.10	0.0621	0.8103	0.4655	666.78	198.64	273.65	194.49
$\alpha = 0.95$	-0.05	0.05	0.0325	0.8250	0.4829	669.11	199.34	274.51	195.26
$\alpha = 0.05$	-0.05	0.05	0.0091	0.5048	0.4089	666.40	194.92	258.20	213.28
$\alpha = 0.10$	-0.10	0.10	0.0174	0.5096	0.4132	666.69	195.01	258.34	213.34
$\alpha = 0.50$	-0.50	0.50	0.0506	0.5506	0.4494	669.20	195.76	259.60	213.84
$\alpha = 0.90$	-0.10	0.10	0.0190	0.5949	0.4895	672.16	196.65	261.08	214.43
$\alpha = 0.95$	-0.05	0.05	0.0101	0.6006	0.4947	672.56	196.77	261.28	214.51
$\alpha = 0.05$	-0.05	0.05	0.05	0.5263	0	638.83	186.45	245.21	207.17
$\alpha = 0.10$	-0.10	0.10	0.1	0.5555	0	640.88	187.17	245.97	207.74
$\alpha = 0.50$	-0.50	0.50	0.2391	0.7391	0.2609	657.39	192.22	253.69	211.48
$\alpha = 0.90$	-0.10	0.10	0.0663	0.8316	0.4632	670.20	196.06	260.10	214.04
$\alpha = 0.95$	-0.05	0.05	0.0340	0.8397	0.4821	671.60	196.48	260.80	214.32
$\alpha = 0.05$	-0.05	0.05	0.05	0.5263	0	667.69	194.90	261.07	211.76
$\alpha = 0.10$	-0.10	0.10	0.10	0.5555	0	669.52	195.47	261.68	212.37
$\alpha = 0.50$	-0.50	0.50	0.2966	0.7966	0.2034	691.05	202.31	270.52	218.21
$\alpha = 0.90$	-0.10	0.10	0.0761	0.8803	0.4577	704.77	206.43	277.38	220.95
$\alpha = 0.95$	-0.05	0.05	0.0387	0.8869	0.4796	706.20	206.86	278.10	221.24

Table 3.4: Capital allocation result of model (3.4.38) for multivariate normal distribution with correlation $\rho_{12} = 0.8$, $\rho_{13} = 0.2$, and $\rho_{23} = 0.3$ for case a, b, and c (from top to bottom separated by horizontal line)

(ii) Negative dependence: Assume that risks in portfolio (X_1, X_2, X_3) are negatively dependent and that the correlation coefficient of any two different risks in the portfolio is negative with $\rho_{12} = -0.4$, $\rho_{13} = -0.1$, and $\rho_{23} = -0.1$. In this case, we have $\mathrm{VaR}_{0.99}(S)) = 561.08$, $\mathrm{CTE}_{0.99}(S) = 577.26$, and $x^* = 542.97$.

	$\underline{\lambda}$	$\overline{\lambda}$	λ^*	q^*	q_0^*	K^*	K_1^*	K_2^*	K_3^*
$\alpha = 0.05$	-0.05	0.05	0.0096	0.5050	0.4043	556.65	136.22	241.75	178.69
$\alpha = 0.10$	-0.10	0.10	0.0190	0.5105	0.4052	556.86	136.29	241.83	178.74
$\alpha = 0.50$	-0.50	0.50	0.0814	0.5814	0.4186	559.72	137.31	243.01	179.40
$\alpha = 0.90$	-0.10	0.10	0.0550	0.7748	0.4695	570.51	141.13	247.48	181.91
$\alpha = 0.95$	-0.05	0.05	0.0312	0.8116	0.4836	573.48	142.18	248.71	182.59
$\alpha = 0.05$	-0.05	0.05	0.05	0.5263	0	573.88	164.98	222.89	186.02
$\alpha = 0.10$	-0.10	0.10	0.1	0.5555	0	574.93	165.16	224.07	185.70
$\alpha = 0.50$	-0.50	0.50	0.4975	0.9977	0.0025	644.81	188.44	247.40	208.96
$\alpha = 0.90$	-0.10	0.10	0.0999	0.9998	0.4445	668.85	195.66	259.43	213.77
$\alpha = 0.95$	-0.05	0.05	0.0499	0.9998	0.4737	670.97	196.29	260.49	214.19
$\alpha = 0.05$	-0.05	0.05	0.05	0.5263	0	557.46	149.15	229.15	179.15
$\alpha = 0.10$	-0.10	0.10	0.01	0.5555	0	558.63	149.54	229.54	179.54
$\alpha = 0.50$	-0.50	0.50	0.4991	0.9991	0.00088	612.16	188.41	247.35	208.94
$\alpha = 0.90$	-0.10	0.10	0.09999771	0.9999886	0.4445	668.85	195.66	259.43	213.77
$\alpha = 0.95$	-0.05	0.05	0.049947	0.99999111	0.4737	670.97	196.29	260.49	214.19
$\alpha = 0.05$	-0.05	0.05	0.05	0.5263	0	574.12	163.55	229.80	180.30
$\alpha = 0.10$	-0.10	0.10	0.10	0.5555	0	574.93	163.99	229.92	181.02
$\alpha = 0.50$	-0.50	0.50	0.499945	0.999945	0.000055224	682.64	199.79	266.32	216.53
$\alpha = 0.90$	-0.10	0.10	0.0999989	0.9999947	0.4444	703.92	206.18	276.96	220.78
$\alpha = 0.95$	-0.05	0.05	0.499996	0.999996	0.4737	705.80	206.74	277.90	221.16

Table 3.5: Capital allocation result of model (3.4.38) for multivariate normal distribution with correlation $\rho_{12} = -0.4$, $\rho_{13} = -0.1$, and $\rho_{23} = -0.1$ for case a, b, and c (from top to bottom separated by horizontal line)

(iii) Mixed dependence: Assume that risks in portfolio (X_1, X_2, X_3) are mixedly dependent and that some of the correlation coefficients of two risks in the portfolio are positive while some are negative with $\rho_{12} = 0.2$, $\rho_{13} = 0.8$, and $\rho_{23} = -0.3$. In this case, we have $\text{VaR}_{0.99}(S)) = 610.50$, $\text{CTE}_{0.99}(S) = 633.88$, and $x^* = 584.3282$.

	$\underline{\lambda}$	$\overline{\lambda}$	λ^*	q^*	q_0^*	K^*	K_1^*	K_2^*	K_3^*
$\alpha = 0.05$	-0.05	0.05	0.0154	0.5081	0.3458	604.27	187.14	232.01	185.12
$\alpha = 0.10$	-0.10	0.10	0.0304	0.5169	0.3482	604.75	187.26	232.25	185.24
$\alpha = 0.50$	-0.50	0.50	0.1196	0.6196	0.3804	611.01	188.87	235.33	186.81
$\alpha = 0.90$	-0.10	0.10	0.0603	0.8013	0.4665	627.14	193.02	243.27	190.85
$\alpha = 0.95$	-0.05	0.05	0.0325	0.8249	0.4829	630.16	193.80	244.75	191.61
$\alpha = 0.05$	-0.05	0.05	0.05	0.5263	0	629.00	183.31	241.69	204.00
$\alpha = 0.10$	-0.10	0.10	0.1	0.5555	0	630.50	183.79	242.26	204.45
$\alpha = 0.50$	-0.50	0.50	0.3329	0.8329	0.1671	652.45	190.73	251.22	210.49
$\alpha = 0.90$	-0.10	0.10	0.0853	0.9265	0.4526	669.44	195.83	259.72	213.89
$\alpha = 0.95$	-0.05	0.05	0.0433	0.9329	0.4772	671.24	196.37	260.62	214.25
$\alpha = 0.05$	-0.05	0.05	0.05	0.5263	0	605.27	165.09	245.09	195.09
$\alpha = 0.10$	-0.10	0.10	0.10	0.5555	0	606.96	165.65	245.65	195.65
$\alpha = 0.50$	-0.50	0.50	0.4211	0.9211	0.0789	648.21	189.46	249.10	209.64
$\alpha = 0.90$	-0.10	0.10	0.0942	0.9710	0.4477	669.08	195.72	259.54	213.82
$\alpha = 0.95$	-0.05	0.05	0.0474	0.9737	0.4751	671.08	196.32	260.54	214.22
$\alpha = 0.05$	-0.05	0.05	0.05	0.5263	0	629.00	182.02	247.92	199.06
$\alpha = 0.10$	-0.10	0.10	0.10	0.5555	0	630.50	182.49	248.60	199.41
$\alpha = 0.50$	-0.50	0.50	0.4652	0.9652	0.0348	683.96	200.19	266.98	216.79
$\alpha = 0.90$	-0.10	0.10	0.0977	0.9884	0.4457	704.00	206.20	277.00	220.80
$\alpha = 0.95$	-0.05	0.05	0.049	0.9896	0.4742	705.84	206.75	277.92	221.17

Table 3.6: Capital allocation result of model (3.4.38) for multivariate normal distribution with correlation $\rho_{12} = 0.2$, $\rho_{13} = 0.8$, and $\rho_{23} = -0.3$ for case a, b, and c (from top to bottom separated by horizontal line)

Note that, for the cases where $q_0^* = 0$, we have checked that the condition of $H_1(\overline{\lambda}-) < H_2(\overline{\lambda}-)$ in Theorem 3.4.5 is not satisfied. In those cases, the solutions are not unique and we present one of them in the table.

Remark 3.5.1. *From both models in* (3.5.7) *and* (3.5.12) *and all of the allocation principles, we found that our total capital K increases as we put more weight on individual unit level risk. The intuition behind this is that if we focus on the aggregated risk on enterprise level, the risks that are from different individual risk units may offset each other which lead to the reduction of total capital. If we focus more on the individual risk unit level, then we ignore the offset effect which can lead to more total required capital.*

Example 3.5.1. When $\psi = \gamma > 0$, $\psi_i = \gamma_i > 0$, $\psi_{ij} = \gamma_{ij} > 0$ $i = 1, ..., n$, $j = 1, ..., n_i$, are positive real numbers, problem (3.4.1) is reduced to the following optimal allocation problem:

$$\begin{cases} \min_{\mathbf{K},\mathbf{k_1},...,\mathbf{k_n}} & \Big\{\alpha_1\gamma\mathbb{E}\Big[\Big|S-K\Big|\Big] + \alpha_2\,\gamma_i\sum_{i=1}^{n}\mathbb{E}\Big[\Big|X_i-K_i\Big|\Big] \\ & \quad + \alpha_3\sum_{i=1}^{n}\sum_{j=1}^{n_i}\gamma_{ij}\mathbb{E}\Big[\Big|X_{ij}-k_{ij}\Big|\Big]\Big\} \\ \text{s.t.} & \sum_{i=1}^{n} K_i = K, \quad \sum_{j=1}^{n_i} k_{ij} = K_i, \quad i = 1, ..., n, \end{cases} \tag{3.5.26}$$

In this case, functions defined in (3.4.4) are reduced to $F(x) = F_S(x)$, $F_i(x) = F_{X_i}(x)$, $F_{ij}(x) = F_{X_{ij}}(x)$, $G(x) = F_S(x-)$, $G_i(x) = F_{X_i}(x-)$, $G_{ij}(x) = F_{X_{ij}}(x-)$. The parameters γ, γ_i, and γ_{ij} in this case because penalty factors at enterprise, subsidiary and business line level respectively.

Example 3.5.2. When $\psi = \frac{\gamma\,\mathrm{I}_A}{\Pr\{A\}}$, $\psi_i = \frac{\gamma_i\,\mathrm{I}_{A_i}}{\Pr\{A_i\}}$, and $\psi_{ij} = \frac{\gamma_{ij}\,\mathrm{I}_{A_{ij}}}{\Pr\{A_{ij}\}}$, where A, A_i and A_{ij} are extreme tail events associated with X, X_i and X_{ij}, respectively, $\gamma > 0$, $\gamma_i > 0$, $\gamma_{ij} > 0$, are positive real numbers, problem (3.4.1) is reduced to

$$\begin{cases} \min_{\mathbf{K},\mathbf{k_1},...,\mathbf{k_n}} & \Big\{\alpha_1\gamma\mathbb{E}\Big[\Big|S-K\Big|\Big|A\Big] + \alpha_2\,\gamma_i\sum_{i=1}^{n}\mathbb{E}\Big[\Big|X_i-K_i\Big|\Big|A_i\Big] \\ & \quad + \alpha_3\sum_{i=1}^{n}\sum_{j=1}^{n_i}\gamma_{ij}\mathbb{E}\Big[\Big|X_{ij}-k_{ij}\Big|\Big|A_{ij}\Big]\Big\} \\ \text{s.t.} & \sum_{i=1}^{n} K_i = K \\ & \sum_{j=1}^{n_i} k_{ij} = K_i, \quad i = 1, ..., n, \end{cases} \tag{3.5.27}$$

In this case, functions defined in (3.4.4) are reduced to $F(x) = \mathbb{P}\{S \le x \,|\, A\}$, $F_i(x) = \mathbb{P}\{X_i \le x \,|\, A_i\}$, $F_{ij}(x) = \mathbb{P}\{X_{ij} \le x \,|\, A_{ij}\}$, $G(x) = \mathbb{P}\{S < x \,|\, A\}$, $G_i(x) = \mathbb{P}\{X_i < x \,|\, A_i\}$, $G_{ij}(x) = \mathbb{P}\{X_{ij} < x \,|\, A_{ij}\}$. Similarly to Example 3.5.1, the parameters γ, γ_i, and γ_{ij} in this case because penalty factors at enterprise, subsidiary and business line level respectively. The examples of the sets A, A_i and A_{ij} can be those used in 3.2.33, 3.2.34, 3.2.35, 3.2.36.

3.6 Conclusions

In the traditional capital allocation problem, we usually have a given total capital and try to find an optimal way to allocate the given capital to all individual risk units. In this chapter, we proposed a new method of capital allocation based on a similar idea as in the previous chapter. We are trying to decide the optimal total capital and the optimal allocated capital to each individual risk unit at the same time. As in the previous chapter, this allocation method depends not only on the distributions of each individual risks, but also on correlations among the individual risks and the relative importance of the individual risks and the aggregate risk to a portfolio. This study also provides an explanation for previously studied allocation methods as the result of optimization problems, providing an intuitive explanation for why those allocation methods are reasonable.

Chapter 4

A multivariate shortfall risk measure based on cumulative prospect theory

4.1 Introduction

In the previous two chapters, we have developed two multivariate risk measures with applications in capital allocations that are motivated by portfolio risk management purposes. In this chapter, we will present another multivariate risk measure that can be very useful for portfolio risk management from another point of view. In this model, we combine the acceptance set concept, distortion risk measures, and cumulative prospect theory (CPT). Before we move to the details of our model, we would like to provide some background on the involvement of risk measures and how those concepts that we mentioned above are developed and used in risk management.

Risk measures provide numerical references on risks' severity and frequency and guide risk managers in preparing appropriate capital based on regulatory requirements and their own risk appetites. Therefore, obtaining an adequate risk measure that can represent a risk position accurately is a crucial question in quantitative risk management. Risk measure develops from a starting point of univariate risk measures.

For a long time, the financial industry has used VaR as a risk measure due to its simplicity. However, VaR has some serious weaknesses such as its tendency to underestimate the severity of a rare event and the fact that it does not satisfy subadditivity. In Artzner et al. (1999), the well-known concept of coherent risk measure is brought up. In this paper, the desirable properties of a risk measure are defined and since then they have played an essential role in modern quantitative risk management. In Föllmer and Schied

(2002), the condition of positive homogeneity is relaxed and the concept of convex measure is defined. This concept is also widely used in this research area. As time has passed, univariate risk measures have become more and more sophisticated. For example, there are many simple risk measures like the most commonly used risk measure, VaR. There are also more complicated univariate risk measures such as expectile and Conditional Value at Risk (CVaR/expected shortfall) by Rockafellar, Uryasev, et al. (2000). For more complicated models, such as those in Bellini and Frittelli (2002) and Frittelli and Gianin (2002), expected utility theory is also involved. Furthermore, per th Ellsberg paradox (see Bellini and Frittelli (2002)), we know that expected utility theory may not always create appropriate models to represent risk managers' preferences, and that models that involve a distortion function can appear. For these topics, please refer to the distortion risk measure as generalized by Wang (1995), spectral measures of risk in Acerbi (2002), the GlueVaR distortion risk measure in Belles-Sampera et al. (2014), generalized quantiles based on RDEU theory in Mao and Cai (2018), and so on.

Although many sophisticated univariate risk measures have been developed, one of the shared weaknesses of univariate risk measures is that they define the risk of every single unit one at a time, making it hard to incorporate the correlation between individual risk units within the same large portfolio and their contribution to the aggregated risk of the portfolio under the same scenario. However, the correlation structure is especially important for portfolio risk management, as it can help us to understand the contribution to the aggregated risk from each individual risk unit. Therefore, multivariate risk measures have been developed. One of the most common ways to define a multivariate risk measure is to extend a univariate risk measure. Many difficulties can arise with this procedure. For example, some univariate risk measures use a coefficient, α, to represent the risk level. When the univariate risk measure is extended to a multivariate risk measure, how we define the risk level can present a challenge. From existing contexts, there are a few ways to define the multivariate risk level. Cousin and Di Bernardino (2013) extend CTE to multivariate CTE by defining a critical layer and a level set. Cai and Li (2005), Cai et al. (2017) and Landsman et al. (2016) extend univariate risk measures to multivariate ones by defining a level for each component of the random vector and applying a univariate risk measure at the corresponding level to each component. Prékopa (2012) uses a concept called "p-level efficient point" to represent the level. Chaudhuri (1996) and Herrmann et al. (2018) use the norm of a directional vector to represent the level.

Besides defining the level explicitly, there is also a way to define the level implicitly. In this approach, we need to employ the concept of an acceptance set, first used in Artzner et al. (1999), who use the concept of an acceptance set and its properties to define coherent risk measure. This concept gives the risk measure an intuitive and practical explanation and a solid mathematical definition at the same time. Usually, regulatory bodies or corporate management teams have a risk appetite that can be represented by an acceptance set $\mathcal{A}$. We say a risk random variable X defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is acceptable if $X \in \mathcal{A}$. Otherwise, a capital c needs to be prepared for the extreme situation so that the remaining risk, $X - c$, is acceptable, i.e., $X - c \in \mathcal{A}$. Since that paper, the acceptance set idea has been widely used in defining risk measures including the convex risk measure in Föllmer and Schied (2002), systemic risk measure in Biagini et al. (2019), and Feinstein et al. (2017). In Frittelli and Gianin (2002), although an acceptance set is not explicitly used, the idea of an acceptable position is mentioned. It has already been shown in Mao and Cai (2018) that in the univariate case, with this method, the risk level can be implicitly defined via the choice of preference functions and distortion functions. In the multivariate case, Armenti et al. (2018) also use this method.

In this chapter, we extend the idea of generalized shortfall induced by cumulative prospect theory (CPT) in Mao and Cai (2018) to a multivariate risk measure with the acceptance set concept used in systemic risk measurement and the method employed by Armenti et al. (2018). The concept of systemic risk involves considering the risk of the entire system and how the failure of each individual unit in the system may lead to systemic failure. In our model, we extend the univariate generalized shortfall induced by CPT based on this concept. Although we use an approach similar to CPT, we provide a different explanation for the utility functions used in our model. We would call this function a preference function since instead of using the utility function to describe people's behaviour, we use it to represent stakeholders' risk preferences. For example, a shareholder who worries more about the default probability of the company may choose a preference function that puts more weight on undercapitalization scenarios than the overcapitalization ones. This would provide a more reasonable meaning for a risk measure as normally used to represent people's risk appetites and preferences, instead describing people's behaviours from a observer's perspective. With this model, we can obtain a risk measure for the whole risk system and risk allocation to the risk units in the system at the same time. This model can be applied to the problem of capital allocation.

The rest of this chapter is structured as follows. In Section 4.2, we start with preliminary theory that includes the notations that we are going to use later, the definition of a coherent risk measure, a distortion risk measure, and univariate generalized shortfall induced by the CPT model. In Section 4.3, we discuss the properties of the risk set, the set of acceptable monetary allocation, and multivariate generalized shortfalls induced by CPT. We also discuss the existence and uniqueness of multivariate generalized shortfalls induced by CPT with the properties of objective functions. In Section 4.4 we provide some numerical illustrations with possible selections of distortion functions and preference functions. Finally, in Section 4.5, we provide a conclusion for this chapter.

4.2 Preliminaries

4.2.1 Basic notation

First, we introduce some notations that we are going to use in this chapter. Let $(\Omega, \mathcal{F}, P)$ be the probability space and let $\mathcal{L}^0(\mathbb{R}^d) := \mathcal{L}^0(\Omega, \mathcal{F}; \mathbb{R}^d)$ be the space of random vectors with dimension d. For $\mathbf{X}$ and $\mathbf{Y} \in \mathbb{R}^d$, we say $\mathbf{X} \geq \mathbf{Y}$ or $\mathbf{X} > \mathbf{Y}$ if $P(\mathbf{X} \geq \mathbf{Y}) = 1$ or $P(\mathbf{X} > \mathbf{Y}) = 1$. Let $\mathbf{X} = (X_1, \ldots, X_d) \in \mathcal{L}^0(\mathbb{R}^d)$ be a random vector of financial losses, which means that the negative values of X_k are used to represent the profits. Let $\mathcal{U}$ denote the set of all continuous increasing functions on $\mathbb{R}$. Let $\mathcal{C}$ denote the set of all continuous functions on $\mathbb{R}$. Let $\mathcal{H}$ be the set of all distortion functions $h : [0,1] \to [0,1]$ which are increasing with $h(0) = 0$, $h(1) = 1$ and have no jumps at 0 and 1. Let h^* be the dual distortion function of h defined by $h^*(x) = 1 - h(1-x)$. Let $X^+ = \max(X, 0)$ and $X^- = -\min(X, 0)$. For a function v defined on $\mathbb{R}$, $v_+(x) = v(x)$ on $\mathbb{R}_+$ and $v_-(x) = v(-x)$ on $\mathbb{R}_-$. Let $\frac{\partial^+}{\partial x} f(x) = f'_+(x)$ and $\frac{\partial^-}{\partial x} f(x) = f'_-(x)$.

Definition 4.2.1. A risk measure ρ defined on a convex cone $\mathcal{X}$ containing all the constants is coherent if the following four properties hold:
(P1) Monotonicity: $\rho(X) \leq \rho(Y)$ for all $X, Y \in \mathcal{X}$ with $X \leq Y$ almost surely (a.s.).
(P2) Translation-invariance: $\rho(X + c) = \rho(X) + c$ for all $X \in \mathcal{X}$ and all $c \in \mathbb{R}$.
(P3) Positive homogeneity: $\rho(\lambda X) = \lambda \rho(X)$ for all $\lambda \in \mathbb{R}_+ = [0, \infty)$ and all $X \in \mathcal{X}$.
(P4) Subadditivity: $\rho(X + Y) \leq \rho(X) + \rho(Y)$ for all $X, Y \in \mathcal{X}$.
A risk measure ρ defined on a convex set $\mathcal{X}$ that is closed under translation is convex if it satisfies (P1) monotonicity, (P2) translation-invariance and
(P5) Convexity: $\rho(aX + (1-a)Y) \leq a\rho(X) + (1-a)\rho(Y)$ for all $X, Y \in \mathcal{X}$ and $a \in (0, 1)$.

Similar to univariate risk measures, we discuss these properties for our multivariate risk measures. However, in multivariate risk measures, some property may be not properly defined such as subadditivity. For example, we have two companies A and B with n and m risk units respectively such that $n \neq m$. We use X to represent the risk random vector of the company A and Y to represent the risk random vector of the company B. In this way, X has a dimension of n, and Y has a dimension of m. If those two companies merge, we may not simply merge n risk units of company A into m risk units of the company B. In this case, the subadditivity cannot be applied directly to the random vectors X and Y. One of the possible solutions is to define the subadditivity on the aggregate risk measure of company A and B.

4.2.2 Distortion risk measures

To start, we review the definition of distortion risk measure and some related transformation of distortion risk measure. There are a few different variant in the definition of distortion risk measure such as Dhaene, Kukush, et al. (2012) and Cai et al. (2020). Here, we stay with Cai et al. (2020).

Definition 4.2.2. Given a random variable X, and $h \in \mathcal{H}$, the functional

$$\rho_h(X) := \int_0^{+\infty} (1 - h(\mathbb{P}(X \le x)))dx - \int_{-\infty}^0 h(\mathbb{P}(X \le x))dx \tag{4.2.1}$$

$$= \int_0^{+\infty} h^*(\mathbb{P}(X > x))dx - \int_{-\infty}^0 (1 - h^*(\mathbb{P}(X > x)))dx. \tag{4.2.2}$$

is called a distortion risk measure, where h^* is the dual distortion function of h.

In the next proposition, we can see that under some certain conditions, distortion risk measure can be written as the expectation of a function with respect to a new probability measure.

Proposition 4.2.1. *If h is right continuous, (4.2.1) can be also written as*

$$\rho_h(X) = \int_{\mathbb{R}} x dh(F(x)) = \int_{[0,1]} VaR_\alpha(X) dh(\alpha). \tag{4.2.3}$$

Proof. For details, see Lemma 2.1 of Cai et al. (2020). □

4.2.3 Univariate model

After we go over some basic concepts that we are going to use, we move to the preference function involved risk model. Before we talk about multivariate generalized shortfalls induced by cumulative prospect theory (CPT), we go through the background of CPT and related concept in univariate cases.

Expected loss/utility based on cumulative prospect theory (CPT)

The idea of expected loss/utility based on CPT was first brought up by Tversky and Kahneman (1992). Later, a few researchers came up with several applications based on CPT including Kaluszka and Krzeszowiec (2012a), Kaluszka and Krzeszowiec (2012b), and Mao and Cai (2018). The definition given by two researchers is in different forms. In Kaluszka and Krzeszowiec (2012a) and Kaluszka and Krzeszowiec (2012b), expected loss/utility based on CPT is defined as

$$\rho_g\left(u_1\left(X^+\right)\right) - \rho_g\left(u_2\left((-X)^+\right)\right) \tag{4.2.4}$$

where $u_1, u_2 \in \mathcal{U}$, $h_1, h_2 \in \mathcal{H}$ and $\rho_g(u_i(X^+))$ is defined as the form of (4.2.1), for $i = 1, 2$.. The concept is used in insurance premium calculation. In Mao and Cai (2018), Expected loss based on CPT is defined as

$$H_{v,h_1,h_2}(X) = \int_0^\infty v(x)dh_1(F_X(x)) + \int_{-\infty}^0 v(x)dh_2(F_X(x)), \tag{4.2.5}$$

where F is the distribution function of X. In this form, CPT is viewed as the expectation of preference function v with respect to probability measure $h_1 \circ F$ for the positive part and with respect to probability measure $h_2 \circ F$ for negative part. Both definitions are generally not equivalent. In this chapter, we will stay with the definition in Mao and Cai (2018).

In Mao and Cai (2018), the function v is restricted to be continuous increasing function on both $\mathbb{R}_+$ and $\mathbb{R}_-$. In this chapter, as a generalization, we would like to ease the restriction on v to make it continuous only and study the properties of the model on new restrictions.

Generalized shortfall induced by CPT

For the following content, we use the definition of of Mao and Cai (2018) for generalized shortfall induced by CPT.

Definition 4.2.3. For $v \in \mathcal{U}$, and $h_1, h_2 \in \mathcal{H}$, the risk set is defined as

$$\mathcal{X}^v_{h_1,h_2} = \{X \in \mathcal{L}^0(\mathbb{R}) : H_{v,h_1,h_2}(X - x) < \infty, \ \ \forall x \in \mathbb{R}\}, \tag{4.2.6}$$

Then, define $\rho_{v,h_1,h_2} : \mathcal{X}_v^{h_1,h_2} \to \mathbb{R}$ and

$$\rho_{v,h_1,h_2}(X) = \inf\{x \in \mathbb{R} : H_{v,h_1,h_2}(X - x) \leq 0\}. \tag{4.2.7}$$

Such a risk measure is called a generalized shortfall induced by CPT.

Some properties of CPT include that $\mathcal{L}^\infty \subseteq \mathcal{X}^v_{h_1,h_2}$ and that $\mathcal{X}^v_{h_1,h_2}$ is closed under translations. Our multivariate risk measure will be extended from univariate generalized shortfall induced by CPT.

4.3 Multivariate extension of univariate generalized shortfalls induced by CPT

One of the challenge, when a univariate distortion risk measure is extended to multivariate distortion risk measure, is how do we properly apply the distortion function. One approach used by Estany et al. (2018) is to apply the distortion function g to joint survival function $S(x_1, \ldots, x_d)$. Then the distortion risk measure for multivariate nonnegative risks is defined as

$$\rho(\mathbf{X}) = \int_0^\infty \ldots \int_0^\infty g(S(x_1, \ldots, x_d))dx_1 \ldots dx_d. \tag{4.3.1}$$

However, one of the obvious limitations is that this can be only applied to nonnegative multivariate risks. In this chapter, we use a different approach similar to Armenti et al. (2018). In this approach. we first apply an aggregation function l which is called loss function in Armenti et al. (2018). The loss function l transforms a random vector $\mathbf{X}$ to a univariate random variable $Y = l(\mathbf{X})$. In this way, we can apply the distortion function just like in the univariate case.

4.3.1 Risk set and set of acceptable monetary allocation

Similar to the univariate case, we first define the risk set. Also, we define the set of acceptable monetary allocation. Before we go any deeper, we propose some conditions on loss function l which are useful to guarantee the existence of risk measure and desirable properties of the obtained risk measure.

Assumption 4.3.1. Let $l : \mathbb{R}^d \to \mathbb{R}$ be strictly increasing, proper convex function where strictly increasing means that if $\mathbf{x} < \mathbf{y}$ for each $x_i < y_i$, then $l(\mathbf{x}) < l(\mathbf{y})$ and strictly decreasing means that if $\mathbf{x} < \mathbf{y}$ for each $x_i < y_i$, then $l(\mathbf{x}) > l(\mathbf{y})$ for all $1 \leq i \leq d$,. Furthermore, assume $\inf l < 0$, $l(0) = 0$ and $l(\mathbf{x}) \geq \sum_{i=1}^d x_i$. Let $v \in \mathcal{C}$, and $h_1, h_2 \in \mathcal{H}$.

In this assumption, the properness of l is used in Lemma 4.3.17 to show the properness of H. The convexity of l is used in Proposition 4.3.3 to show the convexity of function H, Proposition 4.3.18 to show the unbias of function l and Theorem 4.3.22 to solve the solution of the objective function. Furthermore, the assumpton that $l(0) = 0$ is used in Theorem 4.3.19 to show the existence of solution. Finally, the assumption that $l(\mathbf{x}) \geq \sum_{i=1}^d x_i$ is also used in Theorem 4.3.19 to show the existence of solution. Here, we provide an example of possible selection of l.

Example 4.3.1 (Armenti et al. (2018), (6.1))**.** Let $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$. Consider $l(\mathbf{x}) = \sum_{i=1}^d x_i^+ - \frac{1}{2}\sum_{i=1}^d x_i^-$. We show that this l satisfies Assumption 4.3.1.
First, we show that l is strictly increasing. As $x = x^+ - x^-$, if $x < y$, then we need to have either $x^+ < y^+$ and $-x^- \leq -y^-$, or $x^+ \leq y^+$ and $-x^- < -y^-$. In either case, $x^+ - \frac{1}{2}x^- < y^+ - \frac{1}{2}y^-$. Now, if $\mathbf{x} < \mathbf{y}$,

$$l(\mathbf{x}) = \sum_{i=1}^d x_i^+ - \frac{1}{2}\sum_{i=1}^d x_i^- = \sum_{i=1}^d \left(x_i^+ - \frac{1}{2}x_i^-\right) < \sum_{i=1}^d \left(y_i^+ - \frac{1}{2}y_i^-\right) = l(\mathbf{y}). \tag{4.3.2}$$

Next, it is obvious that $l(0) = 0$. Then, Finally, as $x_i \to -\infty$ for $i = 1, \ldots, d$, $l(\mathbf{x}) = \sum_{i=1}^d x_i^+ - \frac{1}{2}\sum_{i=1}^d x_i^- = -\frac{1}{2}\sum_{i=1}^d x_i^- \to -\infty$. Therefore, $\inf l < 0$.
Finally, we see that

$$l(\mathbf{x}) = \sum_{i=1}^d x_i^+ - \frac{1}{2}\sum_{i=1}^d x_i^- = \sum_{i=1}^d x_i^+ - \sum_{i=1}^d x_i^- + \frac{1}{2}\sum_{i=1}^d x_i^- = \sum_{i=1}^d x_i + \frac{1}{2}\sum_{i=1}^d x_i^- \geq \sum_{i=1}^d x_i. \tag{4.3.3}$$

Therefore, $l(\mathbf{x}) \geq \sum_{i=1}^d x_i$.

Next, we introduce the notations $H_{v,h_1}(X)$ and $\mathcal{X}_{h_1,h_2}^{d,v,l}$ which will be used as decision criterion for our risk measure. Let $X \in \mathbb{R}$, let

$$H_{v,h}(X)) = \int_{\mathbb{R}} v(x) dh\left(F_X(x)\right). \tag{4.3.4}$$

The risk set is defined as

$$\mathcal{X}_{h_1,h_2}^{d,v,l} = \{\mathbf{X} \in \mathcal{L}^0(\mathbb{R}^d) : H_{v,h_1,h_2}(l(\mathbf{X} - \mathbf{x})) < \infty, \;\; \forall \mathbf{x} \in \mathbb{R}^d\}. \tag{4.3.5}$$

As we mentioned before, since we use the loss function l to transform $\mathbf{X}$ from the dimension of d to univariate random variables, we can directly apply the distortion function to the new random variable. Therefore, we can define the objective function same as univariate generalized shortfall induced by CPT in Mao and Cai (2018).

Remark 4.3.1. *In the above setup, technically, we can combine the aggregation function l with the adjustment function v to form a single function. We separate them mainly for two purpose. First, since our model can be also viewed as an extension of Armenti et al. (2018)'s model, we would like to be consistent with the structure in Armenti et al. (2018) in order to make further comparison. Also, the separation of two functions can make the model easier to be understood and implemented. While the aggregation function l is usually based on some subjective aggregation rules, the adjustment function v totally depends on the stakeholders' risk appetite. As the reasons above, we decide to leave the aggregation function l and adjustment function v as two separate components in the model.*

In the next proposition, we provide a sufficient condition for the risk set to be convex, and we start with a lemma that would be used in the proof of the proposition.

Lemma 4.3.2. *Let* $X \in \mathcal{X}^{1,v,l}_{h_1,h_2}$, $v_+(x) = v(x)$ *and* $v_-(x) = v(-x)$, *then*

$$H_{v,h_1,h_2}(X) = H_{v_+,h_1}\left(X^+\right) + H_{v_-,h_2^*}\left(X^-\right). \tag{4.3.6}$$

Proof. We start with right hand side of the equation.

$$\begin{aligned}
& H_{v_+,h_1}\left(X^+\right) + H_{v_-,h_2^*}\left(X^-\right) \\
=& \int_{\mathbb{R}} v_+(x) dh_1\left(F_{X^+}(x)\right) + \int_{\mathbb{R}} v_-(x) dh_2^*\left(F_{X^-}(x)\right) \text{ from (4.3.4)} \\
=& \int_{\mathbb{R}} v_+(x) dh_1(P(X^+ \le x)) + \int_{\mathbb{R}} v_-(x) dh_2^*(P(X^- \le x)) \\
=& \int_0^\infty v_+(x) dh_1(P(X^+ \le x)) + \int_0^\infty v_-(x) dh_2^*(P((-X)^+ \le x)) \\
=& \int_0^\infty v_+(x) dh_1(P(X \le x)) + \int_0^\infty v_-(x) dh_2^*(P(-X \le x)) \\
=& \int_0^\infty v_+(x) dh_1(P(X \le x)) + \int_0^\infty v_-(x) dh_2^*(P(X \ge -x)) \\
=& \int_0^\infty v_+(x) dh_1(P(X \le x)) - \int_0^\infty v_-(x) dh(P(X < -x)) \\
=& \int_0^\infty v_+(x) dh_1(P(X \le x)) - \int_0^{-\infty} v_-(-x) dh(P(X < x)) \\
=& \int_0^\infty v_+(x) dh_1(P(X \le x)) + \int_{-\infty}^0 v_-(-x) dh(P(X < x)) \\
=& \int_0^\infty v(x) dh_1(P(X \le x)) + \int_{-\infty}^0 v(x) dh(P(X < x)) \\
=& \int_0^\infty v(x) dh_1(P(X \le x)) + \int_{-\infty}^0 v(x) dh(P(X \le x))
\end{aligned} \tag{4.3.7}$$

where the last line is by the definition of Lebesgue–Stieltjes (L-S) integral. For details, please see Merkle et al. (2014). □

Proposition 4.3.3. *The risk set* $\mathcal{X}^{d,v,l}_{h_1,h_2}$ *is convex if*
(a) v *is convex and strictly increasing on* $\mathbb{R}_+$ *and convex and strictly decreasing on* $\mathbb{R}_-$ *and* h_1, $h_2 \in \mathcal{H}$, h_1 *is convex and* h_2 *is concave; or*
(b) v *is strictly increasing and convex on both* $\mathbb{R}_+$ *and* $\mathbb{R}_-$, h_1, h_2 *are convex, and there exists a constant* d *such that*

$$d \ge \sup_{p \in (0,1)} \frac{(h_2)'_-(p)}{(h_1)'_+(p)}. \tag{4.3.8}$$

Proof. This proof is modified from Proposition A.1 of Mao and Cai (2018).
(a) From Lemma 4.3.2, we know that

$$H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) = H_{v_+,h_1}\left((l(\mathbf{X}-\mathbf{x}))^+\right) + H_{v_-,h_2^*}\left((l(\mathbf{X}-\mathbf{x}))^-\right), \tag{4.3.9}$$

where $h_2^*(p) = 1 - h_2(1-p)$. Next, let $\mathbf{X}$ and $\mathbf{Y} \in \mathcal{X}_{h_1,h_2}^{d,v,l}$, and $\lambda \in (0,1)$. Set $\mathbf{Z} = \lambda\mathbf{X} + (1-\lambda)\mathbf{Y}$. As v_+ is convex, for any $\mathbf{x} \in \mathbb{R}^d$,

$$\begin{aligned} v_+\left((l(\mathbf{Z}-\mathbf{x}))^+\right) =& v_+\left((l(\lambda(\mathbf{X}-\mathbf{x}) + (1-\lambda)(\mathbf{Y}-\mathbf{x})))^+\right) \\ \leq& v_+\left((\lambda l(\mathbf{X}-\mathbf{x}) + (1-\lambda)l(\mathbf{Y}-\mathbf{x}))^+\right) \text{ by convexity of } l \text{ and } v_+ \text{ is increasing,} \\ \leq& v_+\left(\lambda(l(\mathbf{X}-\mathbf{x}))^+ + (1-\lambda)(l(\mathbf{Y}-\mathbf{x}))^+\right) \text{ by convexity of } (\cdot)^+, \\ \leq& \lambda v_+\left(l(\mathbf{X}-\mathbf{x}))^+\right) + (1-\lambda)v_+\left(l(\mathbf{Y}-\mathbf{x}))^+\right) \text{ by convexity of } v_+. \end{aligned} \tag{4.3.10}$$

Next, by property of distortion risk measure, $H_{v_+,h_1}(X) = \rho_{h_1^*}(v_+(X))$ for any random variable X if h_1 is left continuous (see Dhaene, Kukush, et al. (2012)), where ρ_h is the distortion risk measure defined in 4.2.1. Next,

$$H_{v_+,h_1}\left((l(\mathbf{Z}-\mathbf{x}))^+\right) \leq H_{v_+,h_1}\left(\lambda v_+\left(l(\mathbf{X}-\mathbf{x}))^+\right) + (1-\lambda)v_+\left(l(\mathbf{Y}-\mathbf{x}))^+\right)\right) \tag{4.3.11}$$

$$\leq \lambda H_{v_+,h_1}\left((l(\mathbf{X}-\mathbf{x}))^+\right) + (1-\lambda)H_{v_+,h_1}\left((l(\mathbf{Y}-\mathbf{x}))^+\right) \tag{4.3.12}$$

$$< \infty$$

where (4.3.11) is because $h_1 \circ F_{l(\mathbf{X})}$ defines a probability measure and (4.3.12) is from Hong et al. (1987), we know that rank-dependent utility $H_{u,h}(X)$ is convex if and only if u is convex and h is convex.

For $H_{v_-,h_2^*}((l(\mathbf{X}-\mathbf{x}))^-)$, since v is decreasing on $\mathbb{R}^-$, v_- is increasing on $\mathbb{R}^-$. With similary approach, we can show that $H_{v_-,h_2^*}((l(\mathbf{Z}-\mathbf{x}))^-)$ is also finite for any $\mathbf{x} \in \mathbb{R}^d$.

(b) First, we notice that since l is convex, for $\mathbf{Z} = \lambda\mathbf{X} + (1-\lambda)\mathbf{Y}$, we have

$$\begin{aligned} l(\mathbf{Z}) =& l(\lambda\mathbf{X} + (1-\lambda)\mathbf{Y}) \\ \leq& \lambda l(\mathbf{X}) + (1-\lambda)l(\mathbf{Y}) \end{aligned} \tag{4.3.13}$$

Therefore, $H_{v,h_1,h_2}(l(\mathbf{Z})) \leq H_{v,h_1,h_2}(\lambda l(\mathbf{X}) + (1-\lambda)l(\mathbf{Y}))$. The rest of the proof follows the proof of Theorem 2.7 (iii)→(i) in the Appendix of Mao and Cai (2018) by letting $X = l(\mathbf{X})$ and $Y = l(\mathbf{Y})$. □

With our basic setup defined, now, we are going to define the concept of acceptance set and acceptable monetary allocation set.

Definition 4.3.1. We call a risk random vector $\mathbf{X}$ is acceptable if it satisfies

$$H_{v,h_1,h_2}(l(\mathbf{X})) \leq c. \tag{4.3.14}$$

The set

$$\mathcal{A} := \{\mathbf{X} \in \mathcal{X}_{h_1,h_2}^{d,v,l} : H_{v,h_1,h_2}(l(\mathbf{X})) \leq c\} \tag{4.3.15}$$

is called acceptance set of our risk measure.

In this definition, we set an acceptable risk level which is represented by a constant c. We use objective function H with preference function v, distortion function h_1, h_2 and loss function l together to measure the total risk level of risk random vector $\mathbf{X}$. If the measured risk level is below c, that means the risk is under control and we can accept this risk position. Otherwise, the risk level is too high and we should manage the risk to make it acceptable or avoid it.

Remark 4.3.4. *The idea of acceptable monetary allocation is very similar to the idea of acceptance set what was first brought up in Artzner et al. (1999). In Artzner et al. (1999), the author defines a coherent risk measure based on the properties of the acceptance set. After that, this idea is also used by many other researchers. In Föllmer and Schied (2002), the authors use this idea to define a convex risk measure. In Feinstein et al. (2017) and Biagini et al. (2019), both authors use acceptance set to define risk measures for systemic risk.*

Next, using the idea of acceptance set and following the idea of Armenti et al. (2018), we define the set of acceptable monetary allocation.

Definition 4.3.2. A monetary allocation $\mathbf{x} \in \mathbb{R}$ is acceptable for $\mathbf{X}$ if

$$H_{v,h_1,h_2}(l(\mathbf{X} - \mathbf{x})) \leq c \tag{4.3.16}$$

for some constant c. The set of accepatable monetary allocation $A(\mathbf{X})$ is

$$A(\mathbf{X}) := \{\mathbf{x} \in \mathbb{R}^d : \mathbf{X} - \mathbf{x} \in \mathcal{A}\} = \{\mathbf{x} \in \mathbb{R}^d : H_{v,h_1,h_2}(l(\mathbf{X} - \mathbf{x})) \leq c\}. \tag{4.3.17}$$

Remark 4.3.5. *The idea for acceptable monetary allocation is that if a risk random vector* $\mathbf{X}$ *is not acceptable, we should prepare the risk capital* $\mathbf{x}$*. After subtracting the risk capital from the random vector, the new uncertain risk position should be acceptable. The amount of capital that makes the risk position acceptable is called acceptable monetary allocation. Naturally, the acceptable monetary allocation is not unique as if a capital* $\mathbf{x}$ *is acceptable, then any other capital* $\mathbf{y} \geq \mathbf{x}$ *should also be acceptable. We notice that the acceptable monetary allocation itself can be used as a risk measure for capital allocation. However, as a set of values, the nonuniqueness may create confusions when it is applied to risk management. Therefore, based on the set of acceptable monetary allocation, we define our* *multivariate generalized shortfalls induced by CPT with corresponding risk allocation in Section 4.3.3.*

Remark 4.3.6. *Generally, the acceptance criterion can be set to be any level $c \in \mathbb{R}$ such that $H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) \leq c$. In Mao and Cai (2018), c is set to be 0. However, in this chapter we go back to use general value c as the value of c together with selection of v, l, h_1 and h_2 will decide the existence and uniqueness of acceptable monetary allocation and our risk measure.*

Before we talk about the properties of acceptable monetary allocation set, we provide an example of possible selection of v, h_1, and h_2 to give a direct impression of the possible application of our risk measure.

Example 4.3.2. If we set $v(x) = x$, $h_1(x) = x$, and $h_2(x) = x$, then our acceptance criterion is reduced to

$$\begin{aligned} H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) =& \int_0^\infty sd(P(l(\mathbf{X}-\mathbf{x}) \leq s) + \int_{-\infty}^0 sdP(l(\mathbf{X}-\mathbf{x}) \leq s) \\ =& \int_{-\infty}^\infty sd(F_{l(\mathbf{X}-\mathbf{x})}(s)) \\ =& E(l(\mathbf{X}-\mathbf{x})) \end{aligned} \tag{4.3.18}$$

This acceptance criterion is used in Armenti et al. (2018). With this acceptance criterion, the risk measure defined later in Definition 4.3.3 is reduced to Multivariate shortfall risk defined in Armenti et al. (2018).

Remark 4.3.7. *Here, we would like to provide an interpretation for this acceptable money allocation. First, the acceptable money allocation decision criterion H is decided by the sum of two parts: the undercapitalization part X_+ and the overcapitalization X_-. Both risks are reweighted with preference function v and the distribution is distorted by the function h_1 and h_2 for the positive part and negative part respectively. If an acceptable monetary allocation $\mathbf{x}$ is applied, the risk level that is measured by our decision criterion functions H should be below c.*

4.3.2 Properties of acceptable monetary allocation set

In this section, we first discuss the properties of acceptable monetary allocation set under certain restrictions applied to the preference function, distortion functions and value c which would make this risk measure desirable and then consider the conditions that are sufficient for the existence of solutions. We follow this order as the existence of solutions depends on the properties of preference functions, distortion functions and c. We would only provide the existence conditions for some cases. We introduce the following propositions by the order of complexity of the restrictions on v, h_1 and h_2. We start with translation invariance property.

Proposition 4.3.8 (Translation invariance)**.** *For* $\mathbf{X}$, $\mathbf{Y} \in \mathcal{X}^{d,v,l}_{h_1,h_2}$, $A(\mathbf{X}+\mathbf{m}) = A(\mathbf{X}) + \mathbf{m}$ *for any* $\mathbf{m} \in \mathbb{R}^d$*;*

Proof. By (4.3.17), $A(\mathbf{X}+\mathbf{m}) = \{\mathbf{y} \in \mathbb{R}^d : H_{v,h_1,h_2}(l(\mathbf{X}+\mathbf{m}-\mathbf{y})) \leq c\}$. Let $\mathbf{x} = \mathbf{y} - \mathbf{m}$, we have $A(\mathbf{X}+\mathbf{m}) = \{\mathbf{x}+\mathbf{m} \in \mathbb{R}^d : H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) \leq c\} = A(\mathbf{X}) + \mathbf{m}$. ☐

The translation invariance property has a similar interpretation as the univariate risk measure's. If the risk random vector is increased by a constant amount for each component, then the capital needed to make the position acceptable for each risk component is also increased by the same constant amount. In the next proposition, we are going to show sufficient conditions for the permutation invariance property of the acceptable monetary allocation set.

Proposition 4.3.9. *Let* $\mathbf{X}$, $\mathbf{Y} \in \mathcal{X}^{d,v,l}_{h_1,h_2}$. *Let* π *be a permutation function and* π^{-1} *be the inverse function of* π. *We define* $\pi A(\mathbf{X}) = \{\mathbf{y} \in \mathbb{R}^d : \exists \mathbf{x} \in A, \mathbf{y} = \pi(\mathbf{x})\}$ *and we say a function* l *is permutation invariance if* $l(\mathbf{X}) = l(\pi(\mathbf{X}))$. *If* l *is permutation invariant, then* $A(\pi(\mathbf{X})) = \pi A(\mathbf{X})$ *for every permutation* π.

Proof. $A(\pi(\mathbf{X})) = \{\mathbf{x} \in \mathbb{R}^d : H_{v,h_1,h_2}(l(\pi(\mathbf{X}) - \mathbf{x})) \leq c\}$. If l is permutation invariant, then $H_{v,h_1,h_2}(l(\pi(\mathbf{X}) - \mathbf{x})) = H_{v,h_1,h_2}(l(\pi(\mathbf{X} - \pi^{-\mathbf{1}}(\mathbf{x})))) = H_{v,h_1,h_2}(l(\mathbf{X} - \pi^{-\mathbf{1}}(\mathbf{x})))$. Therefore, $\mathbf{x} \in A(\pi(\mathbf{X}))$ iff $\pi^{-1}(\mathbf{x}) \in A(\mathbf{X})$. As $\pi(\pi^{-1}(\mathbf{x})) = \mathbf{x}$, $\pi^{-1}(\mathbf{x}) \in A(\mathbf{X})$ implies $\mathbf{x} \in \pi A(\mathbf{X})$. Finally, we have $\mathbf{x} \in A(\pi(\mathbf{X}))$ iff $\mathbf{x} \in \pi A(\mathbf{X})$. ☐

This proposition provides us with a condition to guarantee the permutation invariance property of the acceptable monetary allocation set. This property is necessary as only changing the order of risk component should not change the risk measure for each risk component. This can prevent the incentive of reducing the risk for certain risk component by changing the order of the risk component. In the next proposition, we show sufficient conditions to make the acceptable monetary allocation set convex which is used later to show the existence and uniqueness of the risk measure.

Proposition 4.3.10. *For* $\mathbf{X}$, $\mathbf{Y} \in \mathcal{X}^{d,v,l}_{h_1,h_2}$, *if* v *is strictly increasing and convex on both* $\mathbb{R}_+$ *and* $\mathbb{R}_-$, h_1, h_2 *are convex, and there exists a constant* d *such that*

$$d \geq \sup_{p\in(0,1)} \frac{(h_2)'_-(p)}{(h_1)'_+(p)}$$

then, it holds:
(i) $A(\mathbf{X})$ *is convex;*
(ii) Define $A(\mathbf{X})+(1-\alpha)A(\mathbf{Y}) := \{\mathbf{z}|\exists \mathbf{x} \in A(X), \exists \mathbf{y} \in A(Y),\ \text{such that}\ \mathbf{z} = \alpha\mathbf{x}+(1-\alpha)\mathbf{y}\}$. *Then,* $A(\alpha\mathbf{X} + (1-\alpha)\mathbf{Y}) \supseteq \alpha A(\mathbf{X}) + (1-\alpha)A(\mathbf{Y})$, *for any* $\alpha \in (0,1)$.

Proof. (i) Since $l(\mathbf{X}-\mathbf{z}) \leq \lambda l(\mathbf{X}-\mathbf{x}) + (1-\lambda)l(\mathbf{X}-\mathbf{y})$, similar to Proposition 4.3.3 (b), we have $H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{z})) \leq \lambda H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) + (1-\lambda)H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{y})) \leq c$ which implies the set $A(\mathbf{X})$ is convex.
(ii)We need to show that if $\mathbf{x}$ satisfies $H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) \leq c$, and $\mathbf{y}$ satisfies $H_{v,h_1,h_2}(l(\mathbf{Y}-\mathbf{y})) \leq c$, then $\alpha\mathbf{x}+(1-\alpha)\mathbf{y}$ satisfies $H_{v,h_1,h_2}(l(\alpha\mathbf{X}+(1-\alpha)\mathbf{Y}-(\alpha\mathbf{x}+(1-\alpha)\mathbf{y}))) \leq c$. This can be easily shown with the same method as Proposition 4.3.3. □

In this proposition, we provide a condition that can make this set convex. In the next proposition, we provide the condition that makes the acceptable monetary allocation set monotonic.

Proposition 4.3.11. *For* $\mathbf{X}$, $\mathbf{Y} \in \mathcal{X}^{d,v,l}_{h_1,h_2}$, *if*
v *is increasing on both* $\mathbb{R}_+$ *and* $\mathbb{R}_-$, h_1, $h_2 \in \mathcal{H}$,
then $A(\mathbf{X}) \supseteq A(\mathbf{Y})$ *whenever* $\mathbf{X} \leq \mathbf{Y}$.

Proof. This is equivalent to show that if $\mathbf{X} \leq \mathbf{Y}$, then $H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) \leq H_{v,h_1,h_2}(l(\mathbf{Y}-\mathbf{x}))$. Let $\mathbf{X} \leq \mathbf{Y}$. Then, due to the increasing of v_+ on $\mathbb{R}_+$ and l is increasing, we have $v_+((l(\mathbf{X}-\mathbf{x}))^+) \leq v_+((l(\mathbf{Y}-\mathbf{x}))^+)$. As $h_1 \circ F_{l(\mathbf{X})}$ defines a probability measure, we have $H_{v_+,h_1}((l(\mathbf{X}-\mathbf{x}))^+) \leq H_{v_+,h_1}((l(\mathbf{Y}-\mathbf{x}))^+)$. For $H_{v_-,h_2}((l(\mathbf{X}-\mathbf{x}))^-)$, as both v and l are increasing, we have v_- is decreasing and $l(\mathbf{X}-\mathbf{x}) \leq l(\mathbf{Y}-\mathbf{x})$. Then, $(l(\mathbf{X}-\mathbf{x}))^- \geq (l(\mathbf{Y}-\mathbf{x}))^-$ and $v_-(l(\mathbf{X}-\mathbf{x})^-) \leq v_-(l(\mathbf{Y}-\mathbf{x})^-)$ as v^- is decreasing. Finally, $H_{v_-,h_2}((l(\mathbf{X}-\mathbf{x}))^-) \leq H_{v_-,h_2}((l(\mathbf{Y}-\mathbf{x}))^-)$ and by 4.3.9, $H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) \leq H_{v,h_1,h_2}(l(\mathbf{Y}-\mathbf{x}))$. □

Mathematically, this proposition shows that if a risk random vector $\mathbf{X}$ is less than $\mathbf{Y}$ componentwise almost surely, then the acceptable monetary allocation for $\mathbf{Y}$ is also acceptable for $\mathbf{X}$. Intuitively, this is also reasonable since acceptable capital allocation for larger risk should be acceptable for smaller risk.

Remark 4.3.12. *Proposition 4.3.10, 4.3.11 provide sufficient conditions for the acceptable monetary allocation set to satisfy the above properties. We observe that in Example 4.3.2, v satisfies conditions in all of above propositions. For details, please see Armenti et al. (2018).*

In the next proposition, we provide conditions that can make the acceptable monetary allocation set positive homogeneous.

Proposition 4.3.13. *For $h_1, h_2 \in \mathcal{H}$, if the distortion function h_1 is continuous at $q_1 = \inf\{p : h_1(p) = 1\}$ and h_2 is continuous at $q_2 = \sup\{p : h_2(p) = 0\}$ with $q_1 > q_2$, and l is positive homogeneous, then $A(c\mathbf{X}) = cA(\mathbf{X})$ for every $c > 0$ if and only if there exist $\beta > 0$, and arbitrary constant a_1, a_2 such that*

$$v(x) = \begin{cases} a_1 x^\beta, & x \geq 0, \\ a_2(-x)^\beta, & x < 0. \end{cases} \tag{4.3.19}$$

Proof. As l is positive homogeneous, $l(\lambda(\mathbf{X} - \mathbf{x})) = \lambda l(\mathbf{X} - \mathbf{x})$ for $\lambda > 0$. The rest of the proof follows the proof of Proposition 2.9 of Mao and Cai (2018) by replacing $X - x$ with $l(\mathbf{X} - \mathbf{x})$. Furthermore, as we remove of the restriction of $v(x)$ is strictly increasing, the resulted a_1 a_2 does not have to be positive as the result in Mao and Cai (2018). This is because from the proof of Proposition 2.9 of Mao and Cai (2018), it is shown that the solution of $v(x)$ is that $v(x) = v_1(x)$ for $x \geq 0$ and $v(x) = -v_2(-x)$ for $x < 0$. Then by solving the equation of $v_1(x)$ and $v_2(x)$, we have the general solution is that $v_1(x) = a_1 x^\beta$ and $v_2(x) = a_2 x^\beta$. Now, from the continuity of $v(x)$, we need to have $v_1(0) = -v_2(-0)$. Therefore, we have $\beta > 0$. As we remove the restriction of strictly increasing of $v(x)$, a_1 and a_2 can be arbitrary constants. □

This property is the same as the univariate positive homogeneity's. If the risk is magnified, the acceptable capital is also magnified by the same factor.

After we discuss the properties of acceptable monetary allocation set, we move to the sufficient conditions for the existence of solutions if the preference functions and distortion functions are defined as in Proposition 4.3.10.

Proposition 4.3.14. *For $\mathbf{X}$, $\mathbf{Y} \in \mathcal{X}^{d,v,l}_{h_1,h_2}$, if $c \in (\inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X} - \mathbf{x})), \infty)$, then $\emptyset \neq A(\mathbf{X}) \neq \mathbb{R}^d$. Furthermore, if v is strictly increasing on both $\mathbb{R}_+$ and $\mathbb{R}_-$, then*
$\inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X} - \mathbf{x})) \leq 0$.

Proof. $H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$ can also be written as

$$H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) = H_{v,h_1}\left((l(\mathbf{X}-\mathbf{x}))^+\right) + H_{v,h_2}\left(-(l(\mathbf{X}-\mathbf{x}))^-\right), \tag{4.3.20}$$

This is obvious if $\inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) = -\infty$. Therefore, we only need to prove the case that $\inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) > -\infty$.
Let $\mathbf{x}^* = \arg\inf_{\mathbf{x}\in\mathbb{R}^d} l(\mathbf{x})$. Let $\mathbf{x}_n \to \infty$ for each x_i where $1 \le i \le d$. Then, as $\inf l < 0$,

$$\begin{aligned} H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}_n)) =& H_{v,h_1}(0) + H_{v,h_2}\left(-(l(\mathbf{X}-\mathbf{x}_n))^-\right) \\ \le& H_{v,h_1}(0) + H_{v,h_2}(0) \\ \le& 0, \end{aligned} \tag{4.3.21}$$

Therefore, $\inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) \le 0$. To see $A(\mathbf{X}) \neq \mathbb{R}^d$, similar to (a), we let $\mathbf{x_n} \to -\infty$ for each x_i with $1 \le i \le d$. Then $H_{v,h_1}((l(\mathbf{X}-\mathbf{x}))^+) \to \infty$ and $H_{v,h_2}(-(l(\mathbf{X}-\mathbf{x}))^-) \to 0$ as l is strictly increasing on S_+ and v_+ is strictly increasing on $\mathbb{R}_+$, which implies $H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}_n)) > c$. □

This proposition is useful to show the existence of risk measure. Also, the second part $A(\mathbf{X}) \neq \mathbb{R}^d$ shows that not all capital allocation methods are acceptable.

4.3.3 Properties of multivariate generalized shortfalls induced by CPT

In this section, we define the risk measure multivariate generalized shortfalls induced by CPT and risk allocation and then following by its properties. After that, we discuss conditions that can guarantee the existence and uniqueness of risk allocation.

Definition 4.3.3. Multivariate generalized shortfall induced by CPT is defined as

$$\rho^{d,l,g}_{h_1,h_2}(\mathbf{X}) = \inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x} \in A(\mathbf{X})\right\} = \inf\left\{\sum_{k=1}^{d} x_k, \mathbf{x} \in \mathbb{R}^d : H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) \le c\right\}. \tag{4.3.22}$$

This risk measure can be viewed as the minimum total capital needed of a company or a system with acceptable monetary allocation to each risk component. Next, we define the concept of risk allocation based on acceptable monetary allocation set. We use same definition as Armenti et al. (2018) for risk allocation.

Definition 4.3.4. A risk allocation is an acceptable monetary allocation $\mathbf{x} = (x_1 \ldots, x_d) \in A(\mathbf{X})$ such that $R(\mathbf{X}) := \rho^{d,l,g}_{h_1,h_2}(\mathbf{X}) = \sum_{i=1}^d x_i$. When a risk allocation is uniquely determined, we denote it by $RA(\mathbf{X})$.

Before we talk about the properties of multivariate generalized shortfalls induced by CPT, we would like to draw the connection between multivariate generalized shortfalls induced by CPT and univariate generalized shortfall induced by CPT (Mao and Cai (2018)). Univariate generalized shortfall induced by CPT (Mao and Cai (2018)) can be seen as a special case of our multivariate generalized shortfall induced by CPT. We show that in the next example.

Example 4.3.3. Let $l(\mathbf{x}) = \sum_{i=1}^d x_i$ and $d = 1$, the decision criterion becomes

$$H_{v,h_1,h_2}(l(\mathbf{X} - \mathbf{x})) = H_{v,h_1,h_2}(X - x), \tag{4.3.23}$$

and our risk measure is reduced to

$$\rho^{1,l,g}_{h_1,h_2}(X) = \inf \left\{ x \in \mathbb{R}^1 : H_{v,h_1,h_2}(X - x)) \leq c \right\} \tag{4.3.24}$$

which is generalized shortfall induced by CPT in Mao and Cai (2018). Here, we notice that the loss function l satisfies Assumption 4.3.1.

In next theorem, we show that under certain conditions, multivariate generalized shortfalls induced by CPT is convex, monotone, translation invariant, and positive homogeneous. Those properties play an essential role in univariate risk measure and also make our multivariate risk measure a desireble one.

Theorem 4.3.15. *For* $\mathbf{X},\ \mathbf{Y} \in \mathcal{X}^{d,v,l}_{h_1,h_2}$, *and* $c \in (\inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X} - \mathbf{x})), \infty)$, *if* v *is strictly increasing and convex on both* $\mathbb{R}_+$ *and* $\mathbb{R}_-$, h_1, h_2 *are convex, and there exists a constant* d *such that*

$$d \geq \sup_{p\in(0,1)} \frac{(h_2)'_-(p)}{(h_1)'_+(p)}$$

and l *is positive homogeneous strictly increasing on* $\mathbb{R}$,
then $\rho^{d,l,g}_{h_1,h_2}(\mathbf{X})$ *is real valued, convex, monotone, and translation invariant where translation invariance is defined as* $\rho(\mathbf{X} + \mathbf{m}) = \rho(\mathbf{X}) + \sum_{i=1}^d m_i$ *for* $\mathbf{m} = (m_1, \ldots, m_d) \in \mathbb{R}^d$.
In addition, if the distortion function h_1 *is continuous at* $q_1 = \inf\{p : h_1(p) = 1\}$ *and* h_2 *is continuous at* $q_2 = \sup\{p : h_2(p) = 0\}$ *with* $q_1 > q_2$, *and there exist* β, a_1, a_2 *such that*

$$v(x) = \begin{cases} a_1 x^\beta, & x \geq 0, \\ a_2(-x)^\beta, & x < 0. \end{cases} \tag{4.3.25}$$

then $\rho^{d,l,g}_{h_1,h_2}(\mathbf{X})$ *is positive homogeneous.*

Proof. The proof will follow the properties of $A(\mathbf{X})$. The convexity can be derived from Proposition 4.3.10. As $A(\alpha\mathbf{X}+(1-\alpha)\mathbf{Y}) \supseteq \alpha A(\mathbf{X})+(1-\alpha)A(\mathbf{Y})$,

$$\begin{aligned}\rho^{d,l,g}_{h_1,h_2}(\alpha\mathbf{X}+(1-\alpha)\mathbf{Y}) =& \inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x}\in A(\alpha\mathbf{X}+(1-\alpha)\mathbf{Y})\right\}\\ \leq& \inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x}\in \alpha A(\mathbf{X})+(1-\alpha)A(\mathbf{Y})\right\}\\ =&\alpha\inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x}\in A(\mathbf{X})\right\}+(1-\alpha)\inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x}\in A(\mathbf{Y})\right\}\\ =&\alpha\rho^{d,l,g}_{h_1,h_2}(\mathbf{X})+(1-\alpha)\rho^{d,l,g}_{h_1,h_2}(\mathbf{Y}). \end{aligned} \tag{4.3.26}$$

The second last equality is because that if $\mathbf{z}\in\alpha A(\mathbf{X})+(1-\alpha)A(\mathbf{Y})$, from definition, there exists an $\mathbf{x}\in A(\mathbf{X})$, and $\mathbf{y}\in A(\mathbf{Y})$, such that $\mathbf{z}=\alpha\mathbf{x}+(1-\alpha)\mathbf{y}$. Then, let $\{\mathbf{z}_n\}|_{n=1}^{\infty}\subseteq A(\mathbf{X})+(1-\alpha)A(\mathbf{Y})$ be a sequence of vectors such that $\inf_{A(\mathbf{X})+(1-\alpha)A(\mathbf{Y})}\sum_{k=1}^{d} z_k = \lim_{n=1}^{\infty}\sum_{i=1}^{d} z_{ni}$ where $\mathbf{z}_n=(z_{n1},\dots,z_{nd})$. Then, $\inf_{A(\mathbf{X})+(1-\alpha)A(\mathbf{Y})}\sum_{k=1}^{d} z_k=\lim_{n=1}^{\infty}\sum_{i=1}^{d}\alpha x_{ni}+(1-\alpha)y_{ni}$, where $\mathbf{x}_n=(x_{n1},\dots,x_{nd})$ and $\mathbf{y}_n=(y_{n1},\dots,y_{nd})$. At the same time, $\{\mathbf{x}_n\}|_{n=1}^{\infty}\subseteq A(\mathbf{X})$ and $\{\mathbf{y}_n\}|_{n=1}^{\infty}\subseteq A(\mathbf{Y})$. Therefore, $\lim_{n=1}^{\infty}\sum_{i=1}^{d} x_{ni}\leq\inf_{A(\mathbf{X})}\sum_{i=1}^{d} x_i$ and $\lim_{n=1}^{\infty}\sum_{i=1}^{d} y_{ni}\leq\inf_{A(\mathbf{Y})}\sum_{i=1}^{d} y_i$. Therefore,

$$\begin{aligned}&\inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x}\in \alpha A(\mathbf{X})+(1-\alpha)A(\mathbf{Y})\right\}\\ \geq&\alpha\inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x}\in A(\mathbf{X})\right\}+(1-\alpha)\inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x}\in A(\mathbf{Y})\right\}.\end{aligned} \tag{4.3.27}$$

On the other hand, we can find a sequence of vectors $\{\mathbf{x}_n\}|_{n=1}^{\infty}\subseteq A(\mathbf{X})$ and $\{\mathbf{y}_n\}|_{n=1}^{\infty}\subseteq A(\mathbf{Y})$ such that $\inf_{A(\mathbf{X})}\sum_{k=1}^{d} x_k=\lim_{n=1}^{\infty}\sum_{i=1}^{d} x_{ni}$ and $\inf_{A(\mathbf{Y})}\sum_{k=1}^{d} y_k=\lim_{n=1}^{\infty}\sum_{i=1}^{d} y_{ni}$. Then, by definition of $A(\mathbf{X})+(1-\alpha)A(\mathbf{Y})$, for every $\mathbf{x}_n=(x_{n1},\dots,x_{nd})$ and $\mathbf{y}_n=(y_{n1},\dots,y_{nd})$, there exists a $\mathbf{z}_n\in A(\mathbf{X})+(1-\alpha)A(\mathbf{Y})$ such that $\mathbf{z}_n=\alpha\mathbf{x}_n+(1-\alpha)\mathbf{y}_n$. Then, $\inf_{A(\mathbf{X})+(1-\alpha)A(\mathbf{Y})}\sum_{k=1}^{d} z_k\leq\sum_{i=1}^{d} z_{ni}=\sum_{i=1}^{d}\alpha x_{ni}+(1-\alpha)y_{ni}$ where $\mathbf{z}_n=(z_{n1},\dots,z_{nd})$. Finally, as $n\to\infty$, we have

$$\begin{aligned}&\inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x}\in \alpha A(\mathbf{X})+(1-\alpha)A(\mathbf{Y})\right\}\\ \leq&\alpha\inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x}\in A(\mathbf{X})\right\}+(1-\alpha)\inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x}\in A(\mathbf{Y})\right\}.\end{aligned} \tag{4.3.28}$$

By combining, (4.3.27) and (4.3.28), we have the second last equality.

Monotonicity can be derived from Proposition 4.3.11. As $A(\mathbf{X}) \supseteq A(\mathbf{Y})$ we have

$$\begin{aligned}\rho_{h_1,h_2}^{d,l,g}(\mathbf{X}) =& \inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x} \in A(\mathbf{X})\right\} \\ \leq& \inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x} \in A(\mathbf{Y})\right\} \\ =& \rho_{h_1,h_2}^{d,l,g}(\mathbf{Y}).\end{aligned} \tag{4.3.29}$$

Translation invariance can be derived from Proposition 4.3.8. As $A(\mathbf{X}+\mathbf{m}) = A(\mathbf{X})+\mathbf{m}$, we have

$$\begin{aligned}\rho_{h_1,h_2}^{d,l,g}(\mathbf{X}+\mathbf{m}) =& \inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x} \in A(\mathbf{X}+\mathbf{m})\right\} \\ =& \inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x} \in A(\mathbf{X})+\mathbf{m}\right\} \\ =& \inf\left\{\sum_{k=1}^{d} x_k + m_k : \mathbf{x} \in A(\mathbf{X})\right\} \\ =& \rho_{h_1,h_2}^{d,l,g}(\mathbf{X}) + \sum_{i=1}^{d} m_i.\end{aligned} \tag{4.3.30}$$

Positive homogeneity can be derived from Proposition 4.3.13. As $A(c\mathbf{X}) = cA(\mathbf{X})$, we have

$$\begin{aligned}\rho_{h_1,h_2}^{d,l,g}(c\mathbf{X}) =& \inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x} \in A(c\mathbf{X})\right\} \\ =& \inf\left\{\sum_{k=1}^{d} x_k : \mathbf{x} \in cA(\mathbf{X})\right\} \\ =& \inf\left\{c\sum_{k=1}^{d} x_k : \mathbf{x} \in A(\mathbf{X})\right\} \\ =& c\rho_{h_1,h_2}^{d,l,g}(\mathbf{X}).\end{aligned} \tag{4.3.31}$$

□

Another important problem that we need to consider is that if this risk allocation is attainable and unique. In other words, we need to make sure this risk allocation is in set $A(\mathbf{X})$. In this case, we first need $A(\mathbf{X})$ to be closed. In the next proposition, we provide conditions that can make $A(\mathbf{X})$ a closed set.

Proposition 4.3.16. *For* $\mathbf{X} \in \mathcal{X}_{h_1,h_2}^{d,v,l}$, *and* $c \in (\inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})), \infty)$, *if* l *is continuous, then* $A(\mathbf{X})$ *is closed. Furthermore, if* $\inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$ *is attainable, for*

$c = \inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$, $A(\mathbf{X})$ *is closed.*

Proof. We want to show that for any $\mathbf{x}_n \in A(\mathbf{X})$ for $n \in 1, 2, \cdots$, and $\mathbf{x}_n \to \mathbf{x}$, then $\mathbf{x} \in A(\mathbf{X})$. First, we want to show that $H_{v,h_1}((l(\mathbf{X}-\mathbf{x}))^+) < \infty$ and $H_{v_-,h_2^*}((l(\mathbf{X}-\mathbf{x}))^-) < \infty$. Since we have $\mathbf{x}_n \in A(\mathbf{X})$, this implies $H_{v,h_1}((l(\mathbf{X}-\mathbf{x}_n))^+) < \infty$ and $H_{v_-,h_2^*}((l(\mathbf{X}-\mathbf{x}_n))^-) < \infty$. By translation invariance of $A(\mathbf{X})$ (Proposition 4.3.8), we have $\mathbf{x} \in A(\mathbf{X}+\mathbf{x}-\mathbf{x}_n)$ which implies $H_{v,h_1}((l(\mathbf{X}-\mathbf{x}))^+) < \infty$ and $H_{v_-,h_2^*}((l(\mathbf{X}-\mathbf{x}))^-) < \infty$.
Next, we have $\mathbf{X}-\mathbf{x}_n \to \mathbf{X}-\mathbf{x}$. Then, by continuity of l, we have $l(\mathbf{X}-\mathbf{x}_n) \to l(\mathbf{X}-\mathbf{x})$. Furthermore, we have $(l(\mathbf{X}-\mathbf{x}_n))^+ \to (l(\mathbf{X}-\mathbf{x}))^+$ and $(l(\mathbf{X}-\mathbf{x}_n))^- \to (l(\mathbf{X}-\mathbf{x}))^-$. As $H_{v,h_1}((l(\mathbf{X}-\mathbf{x}))^+) < \infty$ and $H_{v_-,h_2^*}((l(\mathbf{X}-\mathbf{x}))^-) < \infty$, by dominated convergence theorem, we have $H_{v,h_1}((l(\mathbf{X}-\mathbf{x}_n))^+) \to H_{v,h_1}((l(\mathbf{X}-\mathbf{x}))^+)$ and $H_{v_-,h_2^*}((l(\mathbf{X}-\mathbf{x}_n))^-) \to H_{v_-,h_2^*}((l(\mathbf{X}-\mathbf{x}))^-)$. As $H_{v,h_1h_2}(l(\mathbf{X}-\mathbf{x})) = H_{v,h_1}((l(\mathbf{X}-\mathbf{x}))^+) + H_{v_-,h_2^*}((l(\mathbf{X}-\mathbf{x}))^-)$, we have $H_{v,h_1h_2}(l(\mathbf{X}-\mathbf{x}_n)) \to H_{v,h_1h_2}(l(\mathbf{X}-\mathbf{x}))$.
Now, since $\mathbf{x}_n \in A(\mathbf{X})$, if $H(l(\mathbf{X}-\mathbf{x}_n)) = c$ for some $n \in \mathbb{N}$, we have $H(l(\mathbf{X}-\mathbf{x}_m)) = c$ $\forall m \geq n$. In this case, $H(l(\mathbf{X}-\mathbf{x})) = c$, and we conclude $\mathbf{x} \in A(\mathbf{X})$. If $H(l(\mathbf{X}-\mathbf{x}_n)) < c$, then, as $H(l(\mathbf{X}-\mathbf{x}_n)) \to H(l(\mathbf{X}-\mathbf{x}))$, we have $H(l(\mathbf{X}-\mathbf{x})) < c$, then we also have $\mathbf{x} \in A(\mathbf{X})$ which complete our proof. □

Next, we also need the objective function $H^{\mathbf{X},l}_{v,h_1,h_2}(\mathbf{x}) = H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$ to be a proper closed convex function which will be used in Theorem 4.3.19 and we state the definition of closed function here.

Definition 4.3.5 (Boyd and Vandenberghe (2004), A.3.3)**.** A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be closed if, for each $\alpha \in \mathbb{R}$, the sublevel set

$$\{\mathbf{x} \in \mathrm{dom} f | f(\mathbf{x}) \leq \alpha\} \tag{4.3.32}$$

is closed.

In the following lemma, we provide a sufficient condition to make H a proper closed convex function.

Lemma 4.3.17. *Let $\mathbf{X} \in \mathcal{X}^{d,v,l}_{h_1,h_2}$, and l be a continuous function. If v is strictly increasing and convex on both $\mathbb{R}_+$ and $\mathbb{R}_-$, h_1, $h_2 \in \mathcal{H}$, h_1, h_2 are convex, and there exists a constant d_2 such that*

$$d_2 \geq \sup_{p\in(0,1)} \frac{(h_2)'_-(p)}{(h_1)'_+(p)},$$

then $H^{\mathbf{X},l}_{v,h_1,h_2}(\mathbf{x}) = H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$ is a proper closed convex function.

Proof. Convex property can be seen from Proposition 4.3.3 and Proposition 4.3.10. We prove the closed property with Definition 4.3.5 and Lemma 4.3.17. Let $\alpha \in (\inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})), \infty)$

Then, from Lemma 4.3.17, the sublevel set $\{\mathbf{x} \in \text{dom} H^{\mathbf{X},l}_{v,h_1,h_2}(\mathbf{x}) | H^{\mathbf{X},l}_{v,h_1,h_2}(\mathbf{x}) \leq \alpha\} = A(\mathbf{X})$ is closed. If $\alpha = \inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$, then if $\inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$ is attainable, by Lemma 4.3.17 again, the sublevel set is closed. If $\inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$ is not attainable, the sublevel set is empty, thus closed. If $\alpha \in (-\infty, \inf_{\mathbf{x}\in\mathbb{R}^d} H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})))$, the sublevel set is again, empty, thus closed.

Now, we need to show that $H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$ is proper. It is obvious that $H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) < \infty$ for some $\mathbf{x}$. We only need to show that $H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) > -\infty$ for all $\mathbf{x}$. We have

$$
\begin{aligned}
H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}_n)) =& H_{v,h_1}\left((l(\mathbf{X}-\mathbf{x}_n))^+\right) + H_{v,h_2}\left(-(l(\mathbf{X}-\mathbf{x}_n))^-\right) \\
\geq& H_{v,h_1}(0) + H_{v,h_2}\left(-(l(\mathbf{X}-\mathbf{x}_n))^-\right) \\
>& -\infty,
\end{aligned}
\tag{4.3.33}
$$

where the last inequality is from l is proper. $\square$

Now, we define the concept of an unbiased function which will also be used to show the existence of the solution.

Definition 4.3.6. We call a loss function l unbiased if, for every zero-sum allocation $\mathbf{u} = (u_1, \ldots, u_d)$, $l(\lambda\mathbf{u}) = 0$ for any $\lambda > 0$ implies that $l(-\lambda\mathbf{u}) = 0$ for any $\lambda > 0$ where zero-sum allocation means that $\sum_{i=1}^d u_i = 0$.

Here, we would like to provide sum examples of biased and unbiased functions.

Example 4.3.4 (Armenti et al. (2018) ,Example 3.3)**.** Consider the loss function

$$
l(x,y) = \begin{cases} x + y + \frac{(x+y)^+}{1-y} - 1 & \text{if } y < 1 \\ \infty & \text{otherwise.} \end{cases}
\tag{4.3.34}
$$

If we let $x = 0.5$ and $y = 0$, and $\lambda = 1$, we have $l(0.5, 0) = 0$. However, $l(-\lambda(x, y)) = l(-0.5, 0) = -1.5 \neq 0$. Therefore, the function is biased.

Before we discuss the example of unbiased functions, we state a lemma. From the lemma, we draw a connection between homogeneous and permutation invariant function and an unbiased loss function.

Lemma 4.3.18. *If loss function l is positive homogeneous and permutation invariant, then the function l is unbiased loss function.*

Proof. Since $\sum_{i=1}^{d} u_i = 0$, we have $u_i = \sum_{j=1,j\neq i}^{d} u_j$. Now, if $l(\lambda \mathbf{u}) = 0$, as l is convex and permutation invariante, we have

$$\frac{l(-\lambda \mathbf{u})}{d!-1} = l\left(\frac{-\lambda \mathbf{u}}{d!-1}\right) = l\left(\frac{\lambda \sum_{i=1,\pi_i(\mathbf{u})\neq \mathbf{u}}^{d!} \pi_i(\mathbf{u})}{d!-1}\right) \leq \frac{\sum_{i=1,\pi_i(\mathbf{u})\neq \mathbf{u}}^{d!} l(\lambda \pi_i(\mathbf{u}))}{d!-1} = l(\lambda \mathbf{u}) = 0, \tag{4.3.35}$$

where the first equality is from homogeneity of l and second equality is from zero-sum allocation property of $\mathbf{u}$. The last inequality is from convexity of l. On the other hand, we have

$$0 = l\left(\frac{\lambda \mathbf{u}}{d!-1}\right) = l\left(\frac{\lambda \sum_{i=1,\pi_i(\mathbf{u})\neq \mathbf{u}}^{d!} \pi_i(-\mathbf{u})}{d!-1}\right) \leq \frac{\sum_{i=1,\pi_i(\mathbf{u})\neq \mathbf{u}}^{d!} l(\lambda \pi_i(-\mathbf{u}))}{d!-1} = l(-\lambda \mathbf{u}). \tag{4.3.36}$$

It follows that $l(-\lambda \mathbf{u}) = 0$ and thus l is unbiased. □

Example 4.3.5. Again, we consider the loss function in Example 4.3.1. $l(\mathbf{x}) = \sum_{i=1}^{d} x_i^+ - \frac{1}{2}\sum_{i=1}^{d} x_i^-$. Obviously, the loss function l is positive homogeneous and permutation invariant. Therefore, by Lemma 4.3.18, the loss function l is unbiased.

Now, with our tools ready, we can move to show the existence of solution under certain conditions.

Theorem 4.3.19. *If l is a continuous, positive homogeneous, unbiased loss function, furthermore, v is strictly increasing and convex on both $\mathbb{R}_+$ and $\mathbb{R}_-$, h_1, $h_2 \in \mathcal{H}$, h_1, h_2 are convex, and there exists a constant d such that*

$$d \geq \sup_{p\in(0,1)} \frac{(h_2)'_-(p)}{(h_1)'_+(p)}.$$

Then, for every $\mathbf{X} \in \mathcal{X}_{h_1,h_2}^{d,v,l}$, risk allocations exists.

Proof. Let $\mathbf{x} \in A(\mathbf{X})$, then by Definiton 4.6.1 in Appendix,

$$0^+A(\mathbf{X}) = \{\mathbf{u} \in \mathbb{R}^d : H_{v,h_1,h_2}(l(\mathbf{X} - (\mathbf{x} + \lambda \mathbf{u}))) \leq c, \text{ for all } \lambda > 0\}. \tag{4.3.37}$$

Since $A(\mathbf{X})$ is closed from 4.3.16, by Theorem 4.6.1, if $\mathbf{u}$ satisfies $H_{v,h_1,h_2}(l(\mathbf{X}-(\mathbf{x}+\lambda \mathbf{u}))) \leq c$ for any $\mathbf{x} \in A(\mathbf{X})$, it satisfies $H_{v,h_1,h_2}(l(\mathbf{X} - (\mathbf{x} + \lambda \mathbf{u}))) \leq c$ for all $\mathbf{x} \in A(\mathbf{X})$. By Lemma 4.3.17, we know that H_{v,h_1,h_2} is a proper closed convex function. Then, by Theorem 4.6.2(2), we have $0^+A(\mathbf{X}) = 0^+H_{v,h_1,h_2}$. It turns out that

$$0^+A(\mathbf{X}) = 0^+H_{v,h_1,h_2} = \{\mathbf{u} \in \mathbb{R}^d : (H_{v,h_1,h_2}0^+)(\mathbf{y}) \leq 0\}. \tag{4.3.38}$$

Then by Theorem 4.6.2(3), we have

$$\begin{aligned}0^+A(\mathbf{X}) =&0^+H_{v,h_1,h_2}\\ =&\left\{\mathbf{u}\in\mathbb{R}^d : \sup_{\lambda>0}\frac{H_{v,h_1,h_2}(l(\mathbf{X}-(\mathbf{x}+\lambda\mathbf{u}))) - H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))}{\lambda}\leq 0\right\}\end{aligned} \tag{4.3.39}$$

Next, we define

$$f(\mathbf{x}) = \sum_{i=1}^{d} x_i + \delta(\mathbf{x}|A(\mathbf{X})) = \begin{cases}\sum_{i=1}^{d} x_i & \text{if } \mathbf{x}\in A(\mathbf{X}),\\ \infty & \text{if } \mathbf{x}\notin A(\mathbf{X}).\end{cases}$$

Then, it follows that f is increasing, convex, lower semicontinuous, proper and such that $R(\mathbf{X}) = \inf f$. Let $B := \{\mathbf{x}\in\mathbb{R}^d : f(\mathbf{x})\leq\gamma\}\neq\emptyset$. Then

$$0^+B = \{\mathbf{u}\in\mathbb{R}^d : f(\mathbf{x}+\lambda\mathbf{u}) = \sum_{i=1}^{d}(x_i+\lambda u_i) + \delta(\mathbf{x}+\lambda\mathbf{u}|A(\mathbf{X}))\leq\gamma,\ \text{for all } \lambda>0\}. \tag{4.3.40}$$

We know that f is a proper closed convex function as a proper convex function is closed if and only if it is lower semi-continuous. Then, by Theorem 4.6.2(2), we have $0^+f = 0^+B$. It turns out that

$$0^+f = 0^+B = \{\mathbf{u}\in\mathbb{R}^d : (f0^+)(\mathbf{y})\leq 0\}. \tag{4.3.41}$$

Then by Theorem 4.6.2(3), we have

$$\begin{aligned}0^+f = 0^+B &= \left\{\mathbf{u}\in\mathbb{R}^d : \sup_{\lambda>0}\frac{f(\mathbf{x}+\lambda\mathbf{u})-f(\mathbf{x})}{\lambda}\leq 0\right\}\\ &= \left\{\mathbf{u}\in\mathbb{R}^d : \sup_{\lambda>0}\frac{(\sum_{i=1}^{d}(x_i+\lambda u_i)+\delta(\mathbf{x}+\lambda\mathbf{u}|A(\mathbf{X}))) - (\sum_{i=1}^{d}x_i+\delta(\mathbf{x}|A(\mathbf{X})))}{\lambda}\leq 0\right\}\\ &= \left\{\mathbf{u}\in\mathbb{R}^d : \sup_{\lambda>0}\sum_{i=1}^{d}u_i + \frac{\delta(\mathbf{x}+\lambda\mathbf{u}|A(\mathbf{X}))) - \delta(\mathbf{x}|A(\mathbf{X}))}{\lambda}\leq 0\right\}\end{aligned} \tag{4.3.42}$$

which can be satisfied only if $\mathbf{x}+\lambda\mathbf{u}\in A(\mathbf{X})$ which implies $\mathbf{u}\in 0^+A(\mathbf{X})$. At the same time, from the definition of 0^+B, we have

$$0^+B := \{\mathbf{u}\in\mathbb{R}^d : f(\mathbf{x}+\lambda\mathbf{u})\leq\gamma,\ \text{for all } \lambda>0\}. \tag{4.3.43}$$

Then, with $\mathbf{x}+\lambda\mathbf{u}\in A(\mathbf{X})$, we must have

$$-\infty < R(\mathbf{X}) \leq \sum_{i=1}^{d} x_i + \lambda\sum_{i=1}^{d} u_i < \gamma < \infty \tag{4.3.44}$$

for all λ. Therefore, we conclude that $\sum_{i=1}^d u_i = 0$ and $0^+ f = \{\mathbf{u} \in \mathbb{R}^d : \sum_{i=1}^d u_i = 0\} \cap 0^+ A(\mathbf{X})$. By Theorem 27.1 (b) of Rockafellar (2015), the solution exists if f is constant along its directions of recession $0^+ f$. For $\mathbf{u} \in f^+ 0$, we have $\liminf_{\lambda \to \infty} f(\mathbf{x} + \lambda \mathbf{y}) < \infty$. Then, by Theorem 4.6.2(1) and (4), we know that $\mathbf{u}$ is the direction of recession of f. Therefore, we want to prove that $f(\mathbf{x} + \lambda \mathbf{y})$ is constant for all $\lambda > 0$ which by Theorem 4.6.2(5) is equivalent to show that $\mathbf{u} \in f^+ 0$ and $-\mathbf{u} \in f^+ 0$. Let $\mathbf{u} \in f^+ 0$. By 4.3.39 and Proposition 4.3.11, we have $l(\mathbf{X} - (\mathbf{x} + \lambda \mathbf{u})) \leq l(\mathbf{X} - \mathbf{x})$ for all $\mathbf{x} \in A(\mathbf{X})$ and $\lambda > 0$ which implies $l(-\lambda \mathbf{u}) \leq l(0) = 0$. Furthermore, we have $l(-\lambda \mathbf{u}) \geq -\lambda \sum_{i=1} u_i^d = 0$ from Assumption 4.3.1. Therefore, $l(-\lambda \mathbf{u}) = 0$. As l is unbiased, $l(\lambda \mathbf{u}) = 0$. From homogeneity and convexity of l, we have

$$\frac{l(\mathbf{x} + \lambda \mathbf{u})}{2} = l\left(\frac{\mathbf{x} + \lambda \mathbf{u}}{2}\right) \leq \frac{l(\mathbf{x})}{2} + \frac{l(\lambda \mathbf{u})}{2} = \frac{l(\mathbf{x})}{2} \tag{4.3.45}$$

for any $\mathbf{x} \in \mathbb{R}^d$. Therefore, $l(\mathbf{X} - \mathbf{x} + \lambda \mathbf{u}) \leq l(\mathbf{X} - \mathbf{x})$ which implies $-\mathbf{u} \in 0^+ A(\mathbf{X}) \Rightarrow -\mathbf{u} \in f^+ 0$. □

Remark 4.3.20. *Lemma 4.3.18 provides a practical way go find an unbiased function as a homogeneous and permutation invariante loss function is unbiased.*

Corollary 4.3.21. *If l is a homogeneous and permutation invariante loss function, then, for every $\mathbf{X} \in \mathcal{X}_{h_1,h_2}^{d,v,l}$, risk allocations exists.*

Proof. It follows that by Lemma 4.3.18 l is unbiased. Then, by Theorem 4.3.19, the risk allocation exists. □

After we show the existence of the solution, we need to provide a method on how the risk allocation can be calculated. In the next theorem, we show that the risk allocation can be calculated by solving a system of equation.

Theorem 4.3.22. *If all the conditions in Theorem 4.3.19 hold, and c is selected following Proposition 4.3.14. Then the risk allocation $\mathbf{x}^*$ will be the solution of*

$$\nabla f(\mathbf{x}) + \lambda \nabla H_{v,h_1,h_2}(l(\mathbf{X} - \mathbf{x})) = 0, \tag{4.3.46}$$

where $\nabla = \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_d}\right)$, and

$$H_{v,h_1,h_2}(l(\mathbf{X} - \mathbf{x})) = c, \tag{4.3.47}$$

where λ is a Lagrange multiplier. Furthermore, if the condition in Theorem 4.3.19 holds, then the solution can be further reduced to the solutions to the system of the following *equations:*

$$1 + \lambda \int_{\{\mathbf{x}: l(\mathbf{X}-\mathbf{x})>0\}} \frac{\partial}{\partial x_i} \left(v(l(\mathbf{X}-\mathbf{x}))\right) d\mu_1 + \lambda \int_{\{\mathbf{x}: l(\mathbf{X}-\mathbf{x})\leq 0\}} \frac{\partial}{\partial x_i} \left(v(l(\mathbf{X}-\mathbf{x}))\right) d\mu_2 = 0, \tag{4.3.48}$$

for $i = 1, \ldots, d,$ *and*

$$H_{v,h_1,h_2}\left(l(\mathbf{X}-\mathbf{x})\right) = c. \tag{4.3.49}$$

where $\mu_1 = h_1 \circ F_{\mathbf{x}}$ *and* $\mu_2 = h_2 \circ F_{\mathbf{x}}$.

Proof.

$$\begin{aligned} H_{v,h_1,h_2}\left(l(\mathbf{X}-\mathbf{x})\right) &= \int_0^\infty v(s) dh_1\left(F_{l(\mathbf{X}-\mathbf{x})}(s)\right) + \int_{-\infty}^0 v(s) dh_2\left(F_{l(\mathbf{X}-\mathbf{x})}(s)\right) \\ &= \int_{\{\mathbf{x}: l(\mathbf{X}-\mathbf{x})>0\}} v(l(\mathbf{X}-\mathbf{x})) d\mu_1 + \int_{\{\mathbf{x}: l(\mathbf{X}-\mathbf{x})\leq 0\}} v(l(\mathbf{X}-\mathbf{x})) d\mu_2, \end{aligned} \tag{4.3.50}$$

where $\mu_1 = h_1 \circ F_{\mathbf{x}}$ and $\mu_2 = h_2 \circ F_{\mathbf{x}}$. Then, we have a convex optimization problem

$$\text{minimize } f(\mathbf{x})$$

$$\text{subject to } H_{v,h_1,h_2}\left(l(\mathbf{X}-\mathbf{x})\right) \leq c,$$

The associated Lagrangian L is defined as

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda(H_{v,h_1,h_2}\left(l(\mathbf{X}-\mathbf{x})\right) - c).$$

We can see that the slaters condition is hold with proper selected c. Then, the solution to this convex optimization problem, by KKT condition, is

$$\nabla f(\mathbf{x}) + \lambda \nabla H_{v,h_1,h_2}\left(l(\mathbf{X}-\mathbf{x})\right) = 0,$$

$$\Rightarrow 1 + \lambda \frac{\partial}{\partial x_i} \int_{\{\mathbf{x}: l(\mathbf{X}-\mathbf{x})>0\}} v(l(\mathbf{X}-\mathbf{x})) d\mu_1 + \lambda \frac{\partial}{\partial x_i} \int_{\{\mathbf{x}: l(\mathbf{X}-\mathbf{x})\leq 0\}} v(l(\mathbf{X}-\mathbf{x})) d\mu_2 = 0, \tag{4.3.51}$$

for $i = 1, \ldots, d$, and

$$\lambda((H_{v,h_1,h_2}\left(l(\mathbf{X}-\mathbf{x})\right) - c) = 0. \tag{4.3.52}$$

If the conditions in Theorem 4.3.19 holds, let $\delta_i = (0, \ldots, 0, \underbrace{\delta}_{\text{i-th}}, 0, \ldots, 0)$ and $e_i = (0, \ldots, 0, \underbrace{1}_{\text{i-th}}, 0, \ldots, 0)$. Let

$$w_{\delta_i}(\mathbf{X}, \mathbf{x}) = \frac{v(l(\mathbf{X}-(\mathbf{x}+\delta_i))) - v(l(\mathbf{X}-\mathbf{x}))}{\delta}. \tag{4.3.53}$$

Then, $\frac{\partial}{\partial x_i} v(l(\mathbf{X}-\mathbf{x})) = \lim_{\delta\to 0} w_{\delta_i}(\mathbf{X},\mathbf{x})$, and the derivative exists since v and l are continuous. Now, as v is convex and increasing, l is convex and increasing, $v \circ l$ is convex and increasing. Let $-w_{e_i}(\mathbf{X},\mathbf{x}) = v(l(\mathbf{X}-(\mathbf{x}+e_i))) - v(l(\mathbf{X}-\mathbf{x}))$. Then, it is integrable with respect to μ_1 and μ_2. Furthermore, we have

$$|w_{\delta_i}(\mathbf{X},\mathbf{x})| \le -w_{e_i}(\mathbf{X},\mathbf{x}) \quad \text{ for all } \mathbf{x} \in \mathbb{R}^d \text{ and } \delta > -1, \delta \neq 0. \tag{4.3.54}$$

Therefore, we can apply the change of integral and partial derivative is by Proposition 4.6.3. Finally, from equation 4.3.55 and 4.3.56, we have

$$\Rightarrow 1 + \lambda \int_{\{\mathbf{x}:l(\mathbf{X}-\mathbf{x})>0\}} \frac{\partial}{\partial x_i} \left(v(l(\mathbf{X}-\mathbf{x}))\right) d\mu_1 + \lambda \int_{\{\mathbf{x}:l(\mathbf{X}-\mathbf{x})\le 0\}} \frac{\partial}{\partial x_i} \left(v(l(\mathbf{X}-\mathbf{x}))\right) d\mu_2 = 0, \tag{4.3.55}$$

for $i = 1, \dots, d$, and

$$H_{v,h_1,h_2}\left(l(\mathbf{X}-\mathbf{x})\right) = c. \tag{4.3.56}$$

□

Generally, risk allocation is not unique. However, with the following additional restriction, the risk allocation can be unique.

Theorem 4.3.23. *If all the assumptions in Theorem 4.3.19 are satisfied, furthermore, H is strictly convex outside $\mathbb{R}^d_-$ along zero-sums allocations, then the risk allocation is unique.*

Proof. Assume the risk allocation is not unique. Then, we have two risk allocation $\mathbf{m}$ and $\mathbf{n}$ such that $\mathbf{m} \neq \mathbf{n}$ and $\rho^{d,l,g}_{h_1,h_2}(\mathbf{X}) = \sum_{i=1}^d m_i = \sum_{i=1}^d n_i$. As $\alpha \sum_{i=1}^d m_i + (1-\alpha)\sum_{i=1}^d n_i = \rho^{d,l,g}_{h_1,h_2}(\mathbf{X})$ for any $\alpha \in [0,1]$, $\alpha\mathbf{m} + (1-\alpha)\mathbf{n}$ is also a risk allocation. Furthermore, as $\sum_{i=1}^d m_i - n_i = 0$, $\mathbf{m}-\mathbf{n}$ is a zero-sum allocation. Also, $\alpha\mathbf{m}+(1-\alpha)\mathbf{n} = \mathbf{n}+\alpha(\mathbf{m}-\mathbf{n})$ is a risk allocation along zero-sum direction. Now, as $\mathbf{m}$, $\mathbf{n}$ and $\alpha\mathbf{m}+(1-\alpha)\mathbf{n}$ are all risk allocation, from Theorem 4.3.22, they satisfy $H_{v,h_1,h_2}\left(l(\mathbf{X}-\mathbf{m})\right) = c$, $H_{v,h_1,h_2}\left(l(\mathbf{X}-\mathbf{n})\right) = c$ and $H_{v,h_1,h_2}\left(l(\mathbf{X}-\alpha\mathbf{m}-(1-\alpha)\mathbf{n})\right) = c$. By convexity,

$$c = H_{v,h_1,h_2}\left(l(\mathbf{X}-\alpha\mathbf{m}-(1-\alpha)\mathbf{n})\right) \le \alpha H_{v,h_1,h_2}\left(l(\mathbf{X}-\mathbf{n})\right) + (1-\alpha) H_{v,h_1,h_2}\left(l(\mathbf{X}-\mathbf{m})\right) = c.$$

Therefore, as H is not strictly convex along $\alpha\mathbf{m}+(1-\alpha)\mathbf{n}$, we must have $\alpha\mathbf{m}+(1-\alpha)\mathbf{n} \in \mathbb{R}^d_-$ which implies $\alpha\mathbf{m}+(1-\alpha)\mathbf{n} \le 0$. By monotonicity, we have $H_{v,h_1,h_2}\left(l(\mathbf{X}-\alpha\mathbf{m}-(1-\alpha)\mathbf{n})\right) \le H_{v,h_1,h_2}\left(l(0)\right) \le 0 < c$ which is a contradiction. □

In next example, we provide a simple case of selection of l, v, h_1 and h_2 to illustrate the conditions in Theorem 4.3.22 and Theorem 4.3.23.

Example 4.3.6. (i) Let $h_1(x) = x$, $h_2(x) = x$, $v(x) = x$, and $l(\mathbf{x}) := \sum_{i=1}^{d} g_i(x_i)$, where $g_k : \mathbb{R} \to \mathbb{R}$ is univariate loss function that satisfies Assumption 4.3.1 for $k = 1, \ldots, d$. Furthermore, g_k is strictly convex on $\mathbb{R}_+$. In this way, we can see that v, h_1 and h_2 satisfy conditions in Theorem 4.3.19. Also, in this way, H satisfies the condition in Theorem 4.3.23. This objective function is actually reduced to Proposition 3.9 of Armenti et al. (2018). Therefore, the risk allocation exists and unique.
(ii) For an example that the risk allocation exists but not unique, we have same set up for v, h_1, and h_2, but let l be the loss function defined as $l(\mathbf{x}) = g(\sum_{i=1}^{d} x_i)$ where $g(x) = x^+ - \beta x^-$ and $0 \leq \beta < 1$ (Example 3.8 of Armenti et al. (2018)). We can see that l is positive homogeneous and permutation invariant. Therefore, by Lemma 4.3.18, it is unbiased. Then, by Theorem 4.3.19, the risk allocation exist. However, if $\mathbf{x}$ is a risk allocation and $\mathbf{u}$ is a zero-sum allocation, then $\sum_{i=1}^{d} x_i + u_i = \sum_{i=1}^{d} x_i$. Therefore,

$$H_{v,h_1,h_2}\left(l(\mathbf{X} - \mathbf{x} + \mathbf{u})\right) = H_{v,h_1,h_2}\left(g\left(\sum_{i=1}^{d} X_k - (x_k + u_k)\right)\right) = H_{v,h_1,h_2}\left(g\left(\sum_{i=1}^{d} X_k - x_k\right)\right) \leq c, \tag{4.3.57}$$

which means $\mathbf{x}+\mathbf{u}$ is another risk allocation. This risk allocation is not unique because l is not strictly increasing outside $\mathbb{R}^d_-$. To see that, let $\mathbf{x} \in \mathbb{R}^d$ and $\mathbf{y} \in \mathbb{R}^d$, then $\lambda\mathbf{x}+(1-\lambda)\mathbf{y} \in \mathbb{R}^d$.

$$l(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) = g\left(\sum_{i=1}^{d}(\lambda x_i + (1-\lambda)y_i)\right) = \sum_{i=1}^{d}(\lambda x_i + (1-\lambda)y_i). \tag{4.3.58}$$

On the other hand,

$$l(\mathbf{x}) = g\left(\sum_{i=1}^{d} x_i\right) = \sum_{i=1}^{d} x_i. \tag{4.3.59}$$

Therefore, $l(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) = \lambda l(\mathbf{x}) + (1-\lambda) l(\mathbf{y})$, and l is not strictly increasing outside $\mathbb{R}^d_-$.

Remark 4.3.24. *From the properties of this model, we can obtain a risk measure for the entire risk system and risk allocation to each risk unit at the same time. Therefore, it is desirable for the capital allocation problem.*

4.4 Numerical illustrations and empirical studies

In Armenti et al. (2018), researchers performed a empirical study for a default fund allocation. In this study, let $\mathrm{RA}(\mathbf{X})$ be the unique capital allocation for random vector $\mathbf{X}$

calculated with the method provided in Armenti et al. (2018), then the allocation of the default fund (DF) for unit $k = 1, \ldots, d$ is calculated as

$$A_k(\mathbf{X}) = DF \frac{\mathrm{RA}_k(\mathbf{X})}{\sum_{j=1}^{d} \mathrm{RA}_j(\mathbf{X})}. \tag{4.4.1}$$

In this study, we first provide a simple numerical illustration to show why Multivariate generalized shortfall induced by CPT is a superior risk measure comparing to the risk measure defined in Armenti et al. (2018). Then, we will use the same data as the one used in Armenti et al. (2018) which is “based on an LCH real dataset corresponding to the clearing of 74 portfolios of equity derivatives bearing on 90 underlyings” and it is available online at https://github.com/yarmenti/MSRA, to illustrate the real life application of multivariate generalized shortfalls induced by CPT. The simulation method and distributions has been discussed in Armenti et al. (2018). As the code are provided online, for comparison purpose, we modified the original code, regenerate the simulations and use the same simulation results to compare the allocation result from Armenti et al. (2018) and the allocation result from our risk measures.

4.4.1 Loss functions and distortion functions

In Example 4.3.1, we have an unbiased loss function:

$$l(\mathbf{x}) = \sum_{i=1}^{d} x_i^+ - \frac{1}{2} \sum_{i=1}^{d} x_i^-. \tag{4.4.2}$$

We would like to combine this loss function with different choice of distortion functions. In Example 3.4 of Mao and Cai (2018), researchers provide two distortion functions defined as following

1)

$$h_1(x) = \begin{cases} 0 & \text{if } 0 \le x < \alpha \\ \frac{x-\alpha}{1-\alpha} & \text{if } \alpha \le x \le 1. \end{cases} \tag{4.4.3}$$

2)

$$h_2^*(x) = \begin{cases} \frac{x}{1-\beta} & \text{if } 0 \le x < 1-\beta \\ 1 & \text{if } 1-\beta \le x \le 1. \end{cases} \tag{4.4.4}$$

We see that $h_1(x)$ is a convex function and $h_2^*(x)$ is a concave function. Therefore, we use distortion function $h_1(x)$ and $h_2(x)$ in our risk measure. In this way, we have h_1 and h_2

to be convex. In this way, $h_1^*(x)$ is distortion function for TVaR_α and $h_2^*(x)$ is distortion function for TVaR_β and furthermore, $\text{TVaR}_0 = E(X)$.

As we have shown in Theorem 4.3.19, the allocation exists and with Theorem 4.3.23, the allocation is unique. Then, we can calculate of allocation with the method provided in Theorem 4.3.22. Then, we use the allocation method in (4.4.1) to calculate the allocation for default fund explicitly and comparing the result with the one in Armenti et al. (2018). We can also adjust the level of α to see the effect on the allocation result.

4.4.2 Risk measure calculation

As our risk measure is calculated as

$$\rho_{h_1,h_2}^{d,l,g}(\mathbf{X}) = \inf\left\{\sum_{k=1}^{d} x_k, \mathbf{x} \in \mathbb{R}^d : H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) \leq c\right\}, \tag{4.4.5}$$

we set $c = 0$, and we will use

$$\begin{aligned}\bar{H}_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) = &\sum_{x_i\geq 0,\ i\in\{2,\dots,n\}} v(x_i)\left(h_1(\bar{F}_{l(\mathbf{X}-\mathbf{x})}(x_i)) - h_1(\bar{F}_{l(\mathbf{X}-\mathbf{x})}(x_{i-1}))\right)\\ &+ \sum_{x_i\leq 0\ i\in\{2\dots,n\}} v(x_i)\left(h_2(\bar{F}_{l(\mathbf{X}-\mathbf{x})}(x_i)) - h_2(\bar{F}_{l(\mathbf{X}-\mathbf{x})}(x_{i-1}))\right)\\ &+v(x_1)h_1(\bar{F}_{l(\mathbf{X}-\mathbf{x})}(x_1))\end{aligned} \tag{4.4.6}$$

to estimate

$$H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) = \int_0^{\infty} v(x)dh_1(F_{l(\mathbf{X}-\mathbf{x})}(x)) + \int_{-\infty}^{0} v(x)dh_2(F_{l(\mathbf{X}-\mathbf{x})}(x)), \tag{4.4.7}$$

where $x_1 \dots, x_n$ are simulated data points in increasing order and $\bar{F}_{l(\mathbf{X}-\mathbf{x})}(x_i)$ is the empirical cdf of $l(\mathbf{X}-\mathbf{x})$. Furthermore, $x_j \geq 0$ and $x_{j-1} < 0$. WLOG, we assume $x_1 < 0$ and $x_n > 0$. Thus, we need to show that $\bar{H}_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$ converges to $H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$.

Proposition 4.4.1. *Let $H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$ be defined as in (4.4.6) and let $\bar{H}_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$ be defined as in (4.4.7). Then, as $n \to \infty$, $\bar{H}_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) \to H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$ almost surely.*

Proof. First, by law of large number, we have $\bar{F}_{l(\mathbf{X}-\mathbf{x})}(x_i)$ converges to $F_{l(\mathbf{X}-\mathbf{x})}(x_i)$ almost surely as $n \to \infty$. Now, we look at h_1 and h_2. As h_1 and h_2 are monotone function on $[0, 1]$, by Froda's theorem, they have at most countably numbers of discontinuities. Then, let $\mathcal{S}_F$ be the set that $\bar{F}_{l(\mathbf{X}-\mathbf{x})}(x_i)$ converges to $F_{l(\mathbf{X}-\mathbf{x})}(x_i)$ and $\mathcal{S}_{h_i}$ be the set that h_i is continuous for $i = 1, 2$, then $h_i(\bar{F}_{l(\mathbf{X}-\mathbf{x})}(x))$ converges to $h_i(F_{l(\mathbf{X}-\mathbf{x})}(x))$ on $\mathcal{S}_{h_i} \cap \mathcal{S}_F$. As $P\left((\mathcal{S}_{h_i} \cap \mathcal{S}_F)^c\right) = 0$, $h_i(\bar{F}_{l(\mathbf{X}-\mathbf{x})}(x))$ converges to $h_i(F_{l(\mathbf{X}-\mathbf{x})}(x))$ almost surely. Finally, we have $\bar{H}_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x})) \to H_{v,h_1,h_2}(l(\mathbf{X}-\mathbf{x}))$ almost surely. ☐

4.4.3 Allocation result and comparison

First, we would like to use a numerical illustration to show the flexibility of our risk measure. Let (X, Y) be a bivariate discrete distribution to represent the loss distribution of a company with two business units. The probability distribution is defined in the following table:

X=x	Y=y	Pr(X=x, Y=y)
280	1000	0.05
280	200	0.45
0	0	0.45
-1000	-1000	0.05

Then, the risk measure and allocation result obtained based on Armenti et al. (2018)'s risk measure is listed in the Table 4.1.

	Risk measure	Allocation percentage
X	120.38	39.25%
Y	186.29	60.75%

Table 4.1: Allocation result and percentage weight based on Armenti et al. (2018)'s method

The risk measure and allocation result obtained based on our risk measure based on different α and β level are listed in the Table 4.2 and Table 4.3.

	Risk measure	Allocation percentage
X	280	29.61%
Y	665.72	70.39%

Table 4.2: Allocation result and percentage weight based on Multivariate generalized shortfall induced by CPT with $\alpha = 0.95$, $\beta = 0.05$

	Risk measure	Allocation percentage
X	280	27.14%
Y	751.72	72.86%

Table 4.3: Allocation result and percentage weight based on Multivariate generalized shortfall induced by CPT with $\alpha = 0.95$, $\beta = 0.5$

From the result in the tables above, we can see that based on Armenti et al. (2018)'s method, there are 50% chance that the company is going to be insolvent and under insolvent situation, unit X is very unlikely to get extra funding to become solvent as the loss 280 is more than twice of the allocation capital 120.38. In terms of the allocation percentage, as the unit Y has a much heavier tail, intuitively, we should allocate more capital to unit Y. However, under Armenti et al. (2018)'s method, the allocation weight does not emphasis enough on the heavy tail of Y.

Under multivariate generalized shortfall induced by CPT with the preference function v and distortion functions h_1, h_2 we selected, as we have two more level parameters α and β, the risk measure becomes more flexible. If $\alpha = \beta = 0$, then we can obtain the same risk measure as Armenti et al. (2018)'s method. If we set $\alpha = 0.95$ and $\beta = 0.05$, then it means that we look at the the tail, top $(1 - 0.95)\%$ loss, and ignore the top 5% profit. Then, the obtained risk measure is shown in table 4.2. With this risk measure, we can see that there are only 5% chance that unit Y will be insolvent and the extra funding it needs to reach the solvent point is much smaller. If we set the parameter $\alpha = 0.95$ and $\beta = 0.5$, then we ignore more profit to 50% level, then the risk measure for Y will increase further to 751.72. That means, in this case, the extra funding it needs to reach the solvent point is getting smaller. From the allocation weight aspect, the weight on Y is getting bigger and bigger as we increase the value of α and β.

Therefore, as a comparison result, multivariate generalized shortfall induced by CPT is superior to Armenti et al. (2018)'s risk measure and the superiority comes from the great flexibility of the preference function v and distortion function h_1, h_2. With different selection of v, h_1, and h_2, we can emphasis more on profit, loss, tail distribution and more other aspects.

In next part, we will apply multivariate generalized shortfall induced by CPT with same selection of v, h_1, and h_2 to real life data used in Armenti et al. (2018).

From simulated risk values, the allocation result for highest 10 portfolios and its percentage weight based on Armenti et al. (2018) is shown in Table 4.4:

portfolio name	$\nu = 2$		$\nu = 6$		$\nu = 50$	
PB7	15812895	16.93%	15200883	16.23%	15797036	16.79%
PB56	11324746	12.13%	10803504	11.53%	11051189	11.74%
PB59	8690535	9.31%	8998567	9.61%	8720702	9.27%
PB50	5868323	6.28%	6052060	6.46%	5725238	6.08%
PB32	4881761	5.23%	5040400	5.38%	4861512	5.17%
PB45	4258280	4.56%	4392386	4.69%	4294089	4.56%
PB41	3840451	4.11%	3977165	4.25%	3843467	4.08%
PB34	3632838	3.89%	3566306	3.81%	3752219	3.99%
PB15	3298364	3.53%	3102467	3.31%	3196824	3.40%
PB22	3211442	3.44%	3061508	3.27%	3137539	3.33%

Table 4.4: Allocation result and percentage weight based on Armenti et al. (2018)'s method

The allocation result for highest 10 portfolios and its percentage weight based on our risk measure at different α and β levels are shown in Table 4.5 and Table 4.6:

portfolio name	$\nu = 2$		$\nu = 6$		$\nu = 50$	
PB7	32972614	16.96%	32259869	16.69%	32739120	16.90%
PB56	23630313	12.16%	22770629	11.78%	23169112	11.96%
PB59	18501877	9.52%	18567673	9.61%	18357995	9.48%
PB50	12386177	6.37%	12375210	6.40%	12032855	6.21%
PB32	10348629	5.32%	10415175	5.39%	10265705	5.30%
PB45	9001322	4.63%	9121441	4.72%	9018360	4.65%
PB41	8157867	4.20%	8192979	4.24%	8098012	4.18%
PB34	7534564	3.88%	7513720	3.89%	7801296	4.03%
PB15	6846814	3.52%	6591190	3.41%	6654228	3.43%
PB22	6707251	3.45%	6445806	3.33%	6565776	3.39%

Table 4.5: Allocation result and percentage weight based on our method with $\alpha = 0.9$ $\beta = 0.1$

portfolio name	$\nu = 2$		$\nu = 6$		$\nu = 50$	
PB7	40439562	17.04%	39077500	16.66%	39166399	16.68%
PB56	28770683	12.12%	27673622	11.80%	27619335	11.76%
PB59	22532256	9.49%	22541031	9.61%	22466488	9.57%
PB50	15057948	6.34%	14907017	6.35%	14810419	6.31%
PB32	12613429	5.31%	12615961	5.38%	12580794	5.36%
PB45	10948603	4.61%	11057596	4.71%	11046438	4.70%
PB41	9951034	4.19%	9939243	4.24%	9920986	4.23%
PB34	9234160	3.89%	9123470	3.89%	9255543	3.94%
PB15	8402953	3.54%	7992394	3.41%	7988451	3.40%
PB22	8166736	3.44%	7849559	3.35%	7816522	3.33%

Table 4.6: Allocation result and percentage weight based on our method with $\alpha = 0.95$ $\beta = 0.1$

As a recap, the allocation result based on Armenti et al. (2018)'s method is equivalent to our risk measure with $\alpha = 0$ and $\beta = 0$. From the allocation result, we can see that the risk measure for each portfolio increases as α and β increase. This is because at the level of α increases, we only consider the expectation of loss based on top $(1 - \alpha)\%$. Similarly, at the level of β, we only consider the expectation of profit ignoring the top $\beta\%$. Therefore, at the level of (α, β), we are making the risk measure more conservative. The allocation weight for each portfolio also changes as our emphasis on tail changes.

4.5 Conclusions

In this chapter, we define a multivariate risk measure called multivariate generalized shortfall induced by CPT which is an extension of univariate generalized shortfall induced by CPT as defined by Mao and Cai (2018). This risk measure includes a total measure for the whole risk system and risk allocation to the risk units within the system. It can also be viewed as a generalization of the model proposed by Armenti et al. (2018). In this chapter, we have discussed the desirable properties of the risk measure and the conditions and restrictions on the distortion functions and preference function to guarantee desirable properties and the existence and uniqueness of the risk measure. We have also provided some simple examples to illustrate the possible selections of preference functions and distortion functions.

In future study, we will consider posible applications by choosing more specific distortion functions and real life data, and compare with other existing allocation methods including haircut, quantile, covariance, and CTE principles.

4.6 Appendix

In the appendix, we provide some properties for the convex function and the condition for interchangeability of integral and derivative which we need for the proofs in the main context.

4.6.1 Convex function properties

Here, we quote the summarization of definition and theorems of recession cone that is provided in Armenti et al. (2018) Appendix A which was originally stated in Rockafellar (2015).

Definition 4.6.1. For any non-empty set $C \subseteq \mathbb{R}^d$, recession cone is defined as

$$0^+C := \{\mathbf{y} \in \mathbb{R}^d : \mathbf{x} + \lambda\mathbf{y} \text{ for every } \mathbf{x} \in C \text{ and } \lambda \in \mathbb{R}_+\}$$

Definition 4.6.2. We denote by $f0^+$ the recession function of f, that is, the function with epigraph given as the recession cone of the epigraph of f, and we call

$$0^+f := \{\mathbf{y} \in \mathbb{R}^d : (f0^+)(\mathbf{y}) \leq 0\}.$$

Theorem 4.6.1. *If C is non-empty, closed and convex, and $\mathbf{y} \neq 0$,*

$$0^+C = \{\mathbf{y} \in \mathbb{R}^d : \textit{ there exists } \mathbf{x} \in C \textit{ such that } \mathbf{x} + \lambda\mathbf{y} \in C \textit{ for every } \lambda \in \mathbb{R}_+\}$$

Theorem 4.6.2. *Let f be a proper, closed and convex function on $\mathbb{R}^d$.*
1. Given $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, if $\liminf_{\lambda\to\infty} f(\mathbf{x} + \lambda\mathbf{y}) < \infty$, then $\lambda \to f(\mathbf{x} + \lambda\mathbf{y})$ is decreasing.
2. All the non-empty level sets $B := \{\mathbf{x} \in \mathbb{R}^d : f(\mathbf{x}) \leq \gamma\} \neq \emptyset$ of f have the same recession cone, namely the recession cone of f. That is:

$$0^+f = 0^+B, \textit{ for every } \gamma \in \mathbb{R} \textit{ such that } B \neq \emptyset.$$

3. $f0^+$ is a positively homogeneous, proper, closed and convex function, such that

$$(f0^+)(y) = \sup_{\lambda>0} \frac{f(\mathbf{x}+\lambda\mathbf{y}) - f(\mathbf{x})}{\lambda} = \lim_{\lambda\to\infty} \frac{f(\mathbf{x}+\lambda\mathbf{y}) - f(\mathbf{x})}{\lambda}, \mathbf{y} \in \mathbb{R}^d,$$

for every $\mathbf{x} \in dom(f)$.
4. There exists $\mathbf{x} \in dom(f)$ such that the map $\lambda \to f(\mathbf{x}+\lambda\mathbf{y})$ is decreasing, that is, $\mathbf{y}$ is a direction of recession of f, if and only if this map is decreasing for every $\mathbf{x} \in dom(f)$, which in turn is equivalent to $(f0^+)(\mathbf{y}) \leq 0$.
5. The map $\lambda \to f(\mathbf{x}+\lambda\mathbf{y})$ is constant for every $\mathbf{x} \in dom(f)$ if and only if $(f0^+)(\mathbf{y}) \leq 0$ and $(f0^+)(-\mathbf{y}) \leq 0$.

4.6.2 Interchange of integral and derivative

Proposition 4.6.3. *Let $I \in \mathbb{R}$ be a nontrivial open interval and let $f : \Omega \times I \to \mathbb{R}$ be a map with the following properties.*
(i) For any $x \in I$, the map $\omega \mapsto f(\omega, x)$ is in $\mathcal{L}^1(\mathbb{P})$
(ii) For almost all $\omega \in \Omega$, the map $I \to \mathbb{R}$, $x \mapsto f(\omega, x)$ is differentiable with derivative f'
(iii) There is a map $h \in \mathcal{L}^1(\mathbb{P})$, $h \geq 0$, such that $|f'(\cdot, x)| \leq h$ a.s. for all $x \in I$.
Then, for any $x \in I$, $f'(\cdot, x) \in \mathcal{L}^1(\mathbb{P})$ and the function $F : x \mapsto \int f(\omega, x)\mathbb{P}(d\omega)$ is differentiable with derivative

$$F'(x) = \int f'(\omega, x)\mathbb{P}(d\omega),$$

where $\mathcal{L}^1(\mathbb{P})$ is the set of Lebesgue integrable function with respect to measure $\mathbb{P}$.

Proof. Please see Klenke (2012) for details. □

Chapter 5

Conclusions

In Chapter 2, we introduce a new multivariate CVaR. This risk measure extends the CVaR defined by Rockafellar, Uryasev, et al. (2000) to a multivariate context with dimension d and is motivated by the multivariate geometric quantile/VaR and multivariate geometric expectile. It shares the same strategy of extending the expected loss function to the multivariate case that was first introduced by Chaudhuri (1996). After we define the risk measure, we give an interpretation of it and provide the restrictions on parameters and the conditions necessary to guarantee the existence and uniqueness of the minimum value of the expected loss function. In the following sections, we discuss the properties of the risk measure. Since it may not make sense to directly apply the risk axioms of the univariate case to our measure, we first modify some properties such as subadditivity to fit our multivariate model. Also, for a given property, for example monotonicity, that does not apply to our model or may not be suitable for it, we provide an explanation for that exception. Finally, we provide a numerical illustration to show how changes in parameters and a change of covariance between random variables would affect the risk measure, and also provide a comparison between univariate risk measures and multivariate geometric CVaR.

Chapter 3 introduces a new capital allocation principle which can be seen as an extension of the allocation principle studied in Furman and Zitikis (2008) and Cai and Wang (2020). Using this allocation principle, we can obtain the optimal total capital and the optimal capital allocation to each individual risk in the portfolio at the same time by including the component that considers the risk of the entire portfolio in our objective function. In this chapter, we discuss the conditions needed to guarantee the existence and uniqueness of the solution, and we also provide numerical illustrations and comparisons with the existing allocation principles. Our allocation principle can be viewed as the generalization of many existing allocation principles, and our model can be used to provide explanations for those allocation principles from the viewpoint of optimization problems.

In Chapter 4, we define a multivariate risk measure called multivariate generalized shortfall induced by CPT, which is an extension of univariate generalized shortfall induced by CPT introduced by Mao and Cai (2018). This risk measure includes a total measure for the whole risk system and risk allocation to the risk units in the system. It can also be viewed as a generalization of the multivariate shortfall risk allocation and systemic risk introduced by **armentFi2018multivariate**. In this chapter, we discuss the desirable properties of the risk measure and the conditions and restrictions on the distortion functions and preference function to guarantee desirable properties and the existence and uniqueness of the risk measure. We also provide some examples to illustrate the possible selections of preference functions and distortion functions. Finally, we provide a similar simulation study to those in Armenti et al. (2018) to show the superiority of our defined risk measure.

In this thesis, we propose a new approach for portfolio risk management with multivariate risk measures. With this new approach, we define three multivariate risk measures. We obtain these risk measures for the entire portfolio and each risk unit in the portfolio at the same time as optimal solutions to multivariate objective functions.

References

Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, *26*(7), 1505–1518.

Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. (2017). Measuring systemic risk. *The Review of Financial Studies*, *30*(1), 2–47.

Armenti, Y., Crépey, S., Drapeau, S., & Papapantoleon, A. (2018). Multivariate shortfall risk allocation and systemic risk. *SIAM Journal on Financial Mathematics*, *9*(1), 90–126.

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

Bartram, S. M., Brown, G. W., & Hund, J. E. (2007). Estimating systemic risk in the international financial system. *Journal of Financial Economics*, *86*(3), 835–869.

Belles-Sampera, J., Guillén, M., & Santolino, M. (2014). Beyond value-at-risk: Gluevar distortion risk measures. *Risk Analysis*, *34*(1), 121–134.

Bellini, F., & Frittelli, M. (2002). On the existence of minimax martingale measures. *Mathematical Finance*, *12*(1), 1–21.

Biagini, F., Fouque, J.-P., Frittelli, M., & Meyer-Brandis, T. (2019). A unified approach to systemic risk measures via acceptance sets. *Mathematical Finance*, *29*(1), 329–367.

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.

Brunnermeier, M. K., & Cheridito, P. (2019). Measuring and allocating systemic risk. *Risks*, *7*(2), 46.

Burgert, C., & Rüschendorf, L. (2006). Consistent risk measures for portfolio vectors. *Insurance: Mathematics and Economics*, *38*(2), 289–297.

Cai, J., & Li, H. (2005). Conditional tail expectations for multivariate phase-type distributions. *Journal of Applied Probability*, *42*(3), 810–825.

Cai, J., Li, J., & Mao, T. (2020). Distributionally robust optimization under distorted expectations. *Available at SSRN 3566708*.

Cai, J., & Mao, T. (2020). Risk measures derived from a regulator's perspective on the regulatory capital requirements for insurers. *ASTIN Bulletin*, *50*(3), 1065–1092.

Cai, J., & Wang, Y. (2020). Optimal capital allocation principles considering shortfall and surplus risks in a hierarchical corporate structure. *Submitted and under revision.*

Cai, J., Wang, Y., & Mao, T. (2017). Tail subadditivity of distortion risk measures and multivariate tail distortion risk measures. *Insurance: Mathematics and Economics*, *75*, 105–116.

Chaudhuri, P. (1996). On a geometric notion of quantiles for multivariate data. *Journal of the American Statistical Association*, *91*(434), 862–872.

Chen, C., Iyengar, G., & Moallemi, C. C. (2013). An axiomatic approach to systemic risk. *Management Science*, *59*(6), 1373–1388.

Cossette, H., Mailhot, M., Marceau, E., & Mesfioui, M. (2016). Vector-valued tail value-at-risk and capital allocation. *Methodology and Computing in Applied Probability*, *18*(3), 653–674.

Cousin, A., & Di Bernardino, E. (2013). On multivariate extensions of value-at-risk. *Journal of Multivariate Analysis*, *119*, 32–46.

Cousin, A., & Di Bernardino, E. (2014). On multivariate extensions of conditional-tail-expectation. *Insurance: Mathematics and Economics*, *55*, 272–282.

Cummins, J. D. (2000). Allocation of capital in the insurance industry. *Risk Management and Insurance Review*, *3*(1), 7–27.

Denault, M. (2001). Coherent allocation of risk capital. *Journal of Risk*, *4*, 1–34.

Dhaene, J., Goovaerts, M. J., & Kaas, R. (2003). Economic capital allocation derived from risk measures. *North American Actuarial Journal*, *7*(2), 44–56.

Dhaene, J., Henrard, L., Landsman, Z., Vandendorpe, A., & Vanduffel, S. (2008). Some results on the CTE-based capital allocation rule. *Insurance: Mathematics and Economics*, *42*(2), 855–863.

Dhaene, J., Kukush, A., Linders, D., & Tang, Q. (2012). Remarks on quantiles and distortion risk measures. *European Actuarial Journal*, *2*(2), 319–328.

Dhaene, J., Tsanakas, A., Valdez, E. A., & Vanduffel, S. (2012). Optimal capital allocation principles. *Journal of Risk and Insurance*, *79*(1), 1–28.

Embrechts, P., & Puccetti, G. (2006). Bounds for functions of multivariate risks. *Journal of Multivariate Analysis*, *97*(2), 526–547.

Estany, G. I., Sarabia, J. M., Belles-Sampera, J., Prieto, F., et al. (2018). Distortion risk measures for nonnegative multivariate risks. *Journal of Operational Risk*, *13*(2).

Feinstein, Z., Rudloff, B., & Weber, S. (2017). Measures of systemic risk. *SIAM Journal on Financial Mathematics*, *8*(1), 672–708.

Föllmer, H., & Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, *6*(4), 429–447.

Frittelli, M., & Gianin, E. R. (2002). Putting order in risk measures. *Journal of Banking & Finance*, *26*(7), 1473–1486.

Furman, E., & Zitikis, R. (2008). Weighted risk capital allocations. *Insurance: Mathematics and Economics*, *43*(2), 263–269.

Herrmann, K., Hofert, M., & Mailhot, M. (2018). Multivariate geometric expectiles. *Scandinavian Actuarial Journal*, 1–31.

Herrmann, K., Hofert, M., & Mailhot, M. (2020). Multivariate geometric tail-and range-value-at-risk. *ASTIN Bulletin: The Journal of the IAA*, *50*(1), 265–292.

Hesselager, O., & Andersson, U. (2002). Risk sharing and capital allocation. *Tryg Insurance DP*.

Hong, C. S., Karni, E., & Safra, Z. (1987). Risk aversion in the theory of expected utility with rank dependent probabilities. *Journal of Economic Theory*, *42*(2), 370–381.

Huerlimann, W. (2014). On some properties of two vector-valued VaR and CTE multivariate risk measures for Archimedean copulas. *ASTIN Bulletin: The Journal of the International Actuarial Association*, *44*(3), 613–633.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous univariate distributions*. John Wiley & Sons, Ltd.

Kalkbrener, M. (2005). An axiomatic approach to capital allocation. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, *15*(3), 425–437.

Kaluszka, M., & Krzeszowiec, M. (2012a). Mean-value principle under cumulative prospect theory. *ASTIN Bulletin: The Journal of the IAA*, *42*(1), 103–122.

Kaluszka, M., & Krzeszowiec, M. (2012b). Pricing insurance contracts under cumulative prospect theory. *Insurance: Mathematics and Economics*, *50*(1), 159–166.

Kemperman, J. (1987). The median of a finite measure on a Banach space. *Statistical data analysis based on the l1-norm and related methods* (pp. 217–230).

Klenke, A. (2012). *Probability theory: A comprehensive course*. World Publishing Corporation. https://books.google.ca/books?id=AbDvsgEACAAJ

Kromer, E., Overbeck, L., & Zilch, K. (2016). Systemic risk measures on general measurable spaces. *Mathematical Methods of Operations Research*, *84*(2), 323–357.

Landsman, Z., Makov, U., & Shushi, T. (2016). Multivariate tail conditional expectation for elliptical distributions. *Insurance: Mathematics and Economics*, *70*, 216–223.

Lee, J., & Prékopa, A. (2013). Properties and calculation of multivariate risk measures: MVaR and MCVaR. *Annals of Operations Research*, *211*(1), 225–254.

Mao, T., & Cai, J. (2018). Risk measures based on behavioural economics theory. *Finance and Stochastics*, *22*(2), 367–393.

Maume-Deschamps, V., Rullière, D., & Said, K. (2017). Multivariate extensions of expectiles risk measures. *Dependence Modeling*, *5*(1), 20–44.

Meraklı, M., & Küçükyavuz, S. (2018). Vector-valued multivariate conditional value-at-risk. *Operations Research Letters*, *46*(3), 300–305.

Merkle, M., Marinescu, D., Merkle, M. M., Monea, M., & Stroe, M. (2014). Lebesgue-Stieltjes integral and Young's inequality. *Applicable Analysis and Discrete Mathematics*, 60–72.

Myers, S. C., & Read Jr, J. A. (2001). Capital allocation for insurance companies. *Journal of Risk and Insurance*, 545–580.

Niculescu, C., & Persson, L.-E. (2006). *Convex functions and their applications: A contemporary approach*. Springer.

Noyan, N., & Rudolf, G. (2013). Optimization with multivariate conditional value-at-risk constraints. *Operations Research*, *61*(4), 990–1013.

Overbeck, L. (2000). Allocation of economic capital in loan portfolios. *Measuring risk in complex stochastic systems* (pp. 1–17). Springer.

Pflug, G. C., & Romisch, W. (2007). *Modeling, measuring and managing risk*. World Scientific.

Prékopa, A. (2012). Multivariate value at risk and related topics. *Annals of Operations Research*, *193*(1), 49–69.

Prékopa, A., & Lee, J. (2018). Risk tomography. *European Journal of Operational Research*, *265*(1), 149–168.

Rockafellar, R. T., Uryasev, S. et al. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, *2*, 21–42.

Rockafellar, R. T. (2015). *Convex analysis*. Princeton University Press.

Sherris, M. (2006). Solvency, capital allocation, and fair rate of return in insurance. *Journal of Risk and Insurance*, *73*(1), 71–96.

Shushi, T., & Yao, J. (2020). Multivariate risk measures based on conditional expectation and systemic risk for exponential dispersion model. *Insurance: Mathematics and Economics*.

Torres, R., Lillo, R. E., & Laniado, H. (2015). A directional multivariate value at risk. *Insurance: Mathematics and Economics*, *65*, 111–123.

Tsanakas, A. (2004). Dynamic capital allocation with distortion risk measures. *Insurance: Mathematics and Economics*, *35*(2), 223–243.

Tsanakas, A. (2007). Capital allocation with risk measures. *Proceedings of the 5th Actuarial and Financial Mathematics Day*, *9*, 3–17.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323.

Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance Mathematics & Economics*, *17*(1), 43–54.

Weber, S., & Weske, K. (2017). The joint impact of bankruptcy costs, fire sales and cross-holdings on systemic risk in financial networks. *Probability, Uncertainty and Quantitative Risk*, *2*(1), 9.

Zaks, Y., & Tsanakas, A. (2014). Optimal capital allocation in a hierarchical corporate structure. *Insurance: Mathematics and Economics*, *56*, 48–55.