Risk Management with Non-Convex and Non-Monotone Preferences

by

Yunran Wei

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, 2019

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Ajay Subramanian
Professor, Dept. of Risk Management and Insurance,
Georgia State University

Supervisor(s): Ruodu Wang
Associate Professor, Dept. of Statistics and Actuarial Science,
University of Waterloo

Supervisor(s): Gordon E. Willmot
Professor, Dept. of Statistics and Actuarial Science,
University of Waterloo

Internal Member: Mario Ghossoub
Assistant Professor, Dept. of Statistics and Actuarial Science,
University of Waterloo

Internal Member: Alexander Schied
Professor, Dept. of Statistics and Actuarial Science,
University of Waterloo

Internal-External Member: Qi-Ming He
Professor, Dept. of Management Sciences, University of Waterloo

Author's declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the 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.

Statement of Contribution

This thesis is based on research papers that are co-authored with my supervisors Ruodu Wang and Gordon E. Willmot. My contributions including writing some proofs in the thesis. My supervisors proposed the theorems and proofs and edited the papers for publication.

Abstract

This thesis studies two types of problems, the theory of risk functionals and the risk sharing problem. We put a special focus on a class of non-monotone law-invariant risk functionals, called the signed Choquet integrals.

The contribution can be seen as three portions.

The first portion of this thesis contains various results on signed Choquet integrals. A functional characterization via comonotonic additivity is established, along with some theoretical properties including six equivalent conditions for a signed Choquet integral to be convex. We proceed to address two practical issues currently popular in risk management, namely, robustness (continuity) issues and risk aggregation with dependence uncertainty, for signed Choquet integrals. Our results generalize in several directions those in the literature of risk functionals. From the results obtained in this chapter, we see that many existing elegant mathematical results in the theory of risk measures hold for the general class of signed Choquet integrals; thus they do not rely on the assumption of monotonicity.

In the second portion, we analyze the "convex level sets" (CxLS) property of risk functionals, which is a necessary condition for the notions of elicitability, identifiability, and backtestability, popular in the recent statistics and risk management literature. We put the CxLS property in the context of multi-dimensional risk functionals. We obtain two main analytical results in dimension one and dimension two, by characterizing the CxLS property of all one-dimensional signed Choquet integrals, and that of all two-dimensional signed Choquet integrals with a quantile component. Using these results, we proceed to show that a comonotonic-additive coherent risk measure is co-elicitable with a Value-at-Risk if and only if it is a convex combination of the mean and the corresponding Expected Shortfall. The new findings generalize several results in the recent literature and partially answer an open question on the characterization of multi-dimensional elicitability.

In the third portion, we study a risk sharing problem. Unlike classic risk sharing problems based on expected utilities or convex risk measures, quantile-based risk sharing problems exhibit two special features. First, quantile-based risk measures (such as the Value-at-Risk) are often not convex, and second, they ignore some part of the distribution of the risk. These features create technical challenges in establishing a full characterization of optimal allocations, a question left unanswered in the literature. In this paper, we address the issues on the existence and the characterization of (Pareto-)optimal allocations in quantile-based risk sharing problems. It turns out that negative dependence, mutual exclusivity in particular, plays an important role in the optimal allocations, in contrast to positive dependence appearing in classic risk sharing problems. As a by-product of our

main finding, we obtain some results on the optimization of the Value-at-Risk and the Expected Shortfall.

Acknowledgements

I would like to express my sincere gratitude to Professors Ruodu Wang and Gordon E. Willmot. Professor Wang spent a lot of time advising me and inspired me to start an academia career. He is a role model to me both in the career aspect and my life in general. Professor Willmot is always available for help and his wisdom guided me through my PhD years.

I am grateful that Professors Qi-Ming He, Mario Ghossoub, Alexander Schied and Ajay Subramanian have agreed to be in the committee and spent their valuable time reading the thesis. I in particular thank Professor Ajay Subramanian for flying to Waterloo to attend the oral defence. I benefited a lot from their advice.

I thank Ms Mary Lou Dufton and the other administrative staff for their help. I thank the University of Waterloo and the Hickman Scholar program of the Society of Actuaries for the financial support.

I thank my family and friends for their support.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Over the past few decades, measures of risk and variability are introduced to quantify various characteristics of random financial losses of a financial institution. These measures are mappings from the set of random variables to real numbers (thus, *risk functionals*). Typical examples of risk measures include the Value-at-Risk, the Expected Shortfall and various coherent or convex risk measures as introduced by Artzner et al. (1999) and Föllmer and Schied (2002), and typical examples of variability measures include the variance, the standard deviation, the mean absolute deviation, and various deviation measures as introduced by Rockafellar et al. (2006). We refer to McNeil et al. (2015) for a comprehensive treatment of the use of risk measures in modern risk management.

There has been extensive study on risk functionals. Most of them focus on monotone or convex ones due to the axiomatic definition given in the seminal work of Artzner et al. (1999). In addition, another separate stream of research targeting non-monotone risk functionals - measures of dispersion (Bickel and Lehmann (1976)) or deviation measures (Rockafellar et al. (2006)) has also received considerable attention. To put the two streams of work under the same umbrella, we study risk functionals and preferences that are not necessary monotone or convex.

In the practice of risk management, one very often assesses a risk through its distribution, which is obtained via statistical and simulation analysis. In academic terms, this means that commonly used measures of risk and variability are *law-invariant*. From the work of Kusuoka (2001) and Grechuk et al. (2009), a class of risk functionals becomes the building block of law-invariant risk measures and variability measures: the (law-invariant) *signed Choquet integrals*. Roughly speaking, each of them maps the distribution of a risk to a real number with the use of some given function h, which represents the decision maker's distorted belief.

From a risk management perspective, we focus on law-invariant functionals in this thesis. In what follows we shall omit the term "law-invariant", as all risk functionals we discuss are law-invariant.

In the following, we broadly describe the background and the main contribution of each chapter.

In Chapter 2, a signed Choquet integral $I_h : L^\infty \to \mathbb{R}$ is defined as

$$I_h(X) = \int_{-\infty}^{0} (h(\mathbb{P}(X \geqslant x)) - h(1))\, \mathrm{d}x + \int_{0}^{\infty} h(\mathbb{P}(X \geqslant x))\, \mathrm{d}x, \tag{1.1}$$

where L^∞ is the set of bounded random variables in a probability space, and $h : [0,1] \to \mathbb{R}$ is a function of bounded variation with $h(0) = 0$. Note that a signed Choquet integral in Chapter 3 is defined as a functional mapping from the set of distributions instead of random variables. The notion of signed Choquet integrals without law-invariance originates from Choquet (1954) in the framework of *capacities*, and is further characterized and studied in decision theory by Schmeidler (1986, 1989) and extended by Cerreia-Vioglio et al. (2012, 2015) to general spaces.

There has been an extensive literature on a subclass of signed Choquet integrals, in which h is increasing and $h(1) = 1$; we simply call this class of functionals *increasing Choquet integrals*. In different contexts, such functionals I_h are referred to as L-functionals (Huber and Ronchetti (2009)) in statistics, Yaari's dual utilities (Yaari (1987)) in decision theory, distorted premium principles (Denneberg (1994) and Wang et al. (1997)) in insurance, and distortion risk measures (Kusuoka (2001) and Acerbi (2002)) in finance. In particular, the two most important risk measures used in current banking and insurance regulation, the Value-at-Risk and the Expected Shortfall, are increasing Choquet integrals. For properties and recent advances on various issues related to increasing Choquet integrals, we refer to Dhaene et al. (2012), Wang et al. (2015), Kou and Peng (2016), Delbaen et al. (2016) and Ziegel (2016).

On the other hand, there has been relatively limited research on signed Choquet integrals compared to that on increasing Choquet integrals. The major difference between an increasing Choquet integral and a signed one is that the latter, being more general, is not necessarily monotone. We are particularly interested in signed Choquet integrals for various practical and theoretical reasons. First, although a suitable risk measure should be monotone as argued by Artzner et al. (1999), this issue is irrelevant for a measure of variability. Indeed, all practical measures of variability are not monotone (for instance, variance, standard deviation, or deviation measures in Rockafellar et al. (2006) and Grechuk et al. (2009)). Therefore, instead of obtaining an increasing-Choquet-integral-based representation for law-invariant coherent risk measures as in Kusuoka (2001), one naturally arrives at a signed-Choquet-integral-based representation of deviation measures as in Grechuk et al. (2009). In other words, signed Choquet integrals are relevant as long as a measure of variability is concerned. Second, there are many preferences or risk measures used in practice which are not monotone. A prominent example is the mean-variance and the mean-standard-deviation preferences as already studied by Markowitz (1952); see also Filipović and Svindland (2008) for a study of risk sharing with non-monotone risk measures, and Furman et al. (2017) for the class of Gini Shortfall risk measures, which are not necessarily monotone. Third, in economic decision theory, signed Choquet integrals appear naturally in many rank-based decision making; see Quiggin (1982), Gilboa and Schmeidler (1989) and De Waegenaere and Wakker (2001). Fourth, from a mathematical perspective, we aim to generalize some elegant results, which are known to hold true for increasing Choquet integrals, to the broader class of signed Choquet integrals.

Chapter 2 gives the characterization of the signed Choquet integrals using a result in Schmeidler (1986), and presents some theoretical properties. Though some of the properties are known, few literature lists them completely. Moreover, we give the sufficient conditions of the robustness properties and study the problem of risk aggregation under uncertainty for homogeneous portfolio as well as heterogeneous portfolio. We demonstrate our results through a numerical illustration.

Chapter 3 characterizes risk functionals with the "convex level sets" (CxLS) property, a necessary condition to the notions of elicitability (Osband (1985)), identifiability, and backtestability, which received an increasing attention in the statistics and risk management literature (e.g. Gneiting (2011), Ziegel (2016), Fissler and Ziegel (2016), Kou and Peng (2016), Acerbi and Szekely (2017)). These concepts refer to the quality and validity of risk forecasts, and have been a prominent issue in banking regulation and model risk management (see e.g. BCBS (2016)).

In decision theory, the CxLS property is closely related to (slightly weaker than) the axiom of *betweenness*, one of the possible relaxations of the independence axiom of the von Neumann-Morgenstern expected utility theory; see, e.g. Dekel (1986) and Chew (1989). As mentioned above, the recently growing importance of the CxLS property in risk management is mainly due to its close relation with the statistical notions of elicitability, identifiability and backtestability.

In the literature of risk measures, many results on characterization of elicitable risk measures are obtained via characterizing the CxLS property; see Weber (2006), Bellini and Bignozzi (2015) and Delbaen et al. (2016) for convex risk measures, Ziegel (2016) for coherent risk measures, Kou and Peng (2016) and Wang and Ziegel (2015) for distortion risk measures, and Liu and Wang (2016) for tail risk measures. For higher-dimensional elicitability and their statistical implications, see Lambert et al. (2008), Fissler and Ziegel (2016), Nolde and Ziegel (2017) and Acerbi and Szekely (2017).

In the above literature, characterization results are obtained for one-dimensional increasing or convex risk functionals. As mentioned above, although risk measures are typically increasing functionals, many statistical quantities, such as measures of variability or shape, are not monotone with respect to the natural order on $\mathcal{M}$, and they play an important role in the statistical analysis of risks. In this chapter, we study the CxLS property of non-monotone, non-convex, and multi-dimensional functionals, with a particular focus on the class of signed Choquet integrals (multi-dimensional). One dimensional signed Choquet integrals include many commonly used risk functionals, such as risk measures and variability measures. Moreover, as discussed by Fissler and Ziegel (2016), a two-dimensional signed Choquet integral (Example 3.4.1) gives the first example of a multi-dimensional elicitable risk functional that is not connected to one-dimensional ones via a bijection. Characterization of elicitability or CxLS for multi-dimensional signed Choquet integrals is generally an open question, as mentioned by both Kou and Peng (2016) and Fissler and Ziegel (2016).

We study the CxLS property in dimension one as well as in higher dimensions with a special focus on signed Choquet integrals. In dimension one, we show that the exact choices of h are very restricted after imposing the CxLS property. In dimension two, we characterize the CxLS property of all two-dimensional signed Choquet integrals with a quantile component. Using these results, we proceed to show that a comonotonic-additive coherent risk measure is co-elicitable with a Value-at-Risk if and only if it is a convex combination of the mean and the corresponding Expected Shortfall. The new findings generalize several results in the recent literature and partially answer an open question on the characterization of multi-dimensional elicitability.

Chapter 4 discusses existence and uniqueness of the quantile-based risk sharing problem. Quantile-based risk sharing problems, as studied by Embrechts et al. (2018), have recently drawn considerable interest in the literature of risk management, due to the popularity of quantile-based risk measures such as the Value-at-Risk (VaR) and the Expected Shortfall (ES) in current banking and insurance regulation. The key feature of these risk sharing problems is that each agent's preference is modelled by a quantile-based risk measure (called an RVaR), and these risk measures are often not convex; see Section 4.2 for more details. This feature distinguishes quantile-based risk sharing problems from the classic ones based on utility functions or convex risk measures. The non-convexity of the preferences brings in substantial challenges for studying risk sharing problems, as well as interesting mathematical and economic observations. For recent results and financial implications of quantile-based risk sharing, we refer to Embrechts et al. (2018) and the references therein. Weber (2018) contains discussions on quantile-based optimal risk sharing problem in the context of Solvency II.

The existing literature on this topic focuses on finding the minimum possible aggregate risk value and giving some optimal risk allocations, whereas existence and characterization issues are left partially or completely unaddressed. Embrechts et al. (2018) obtained some Pareto-optimal allocations and Weber (2018) generalized the underlying risk measures from the RVaR family to the so-called VaR-type distortion risk measures with concave active parts (see Section 4.5.2). The case of heterogeneous beliefs is analyzed by Embrechts et al. (2019). These papers give some solutions, but do not characterize the whole family of the optimal allocations.

In Chapter 4, we provide a complete answer to the questions of the existence and the characterization of Pareto-optimal allocations in quantile-based risk sharing problems. As a by-product of our main finding, we give some technical lemmas on the optimization of the Value-at-Risk and the Expected Shortfall and briefly discuss how the approaches can be applied to VaR-type risk measures.

As noted by Embrechts et al. (2018), Pareto-optimal allocations (which we shall simply refer to as *optimal allocations*) are often equivalent to sum-optimal allocations. As we shall see from the main results, the characterization of all optimal allocations is highly non-trivial, since the quantile-based risk measures often ignore part of the distribution of the risk, creating a considerable amount of probabilistic freedom. A further complication arises when the total risk is not continuously distributed, leading to various issues with non-uniqueness of the quantile. Our results show that an optimal allocation exhibits a strong negative dependence, in sharp contrast to the classic risk sharing problems where an optimal allocation is always strongly positively dependent (see Section 4.5.1).

Chapter 4 builds on the main results of Embrechts et al. (2018) on quantile-based risk sharing. As mentioned before, techniques in this framework are different from the classic risk sharing problems with convex risk measures or expected utilities; for the latter, we refer to Barrieu et al. (2005), Acciaio (2007), Jouini et al. (2008), Filipović and Svindland (2008), Anthropelos and Kardaras (2017) and the references therein. The RVaR family of risk measures are introduced by Cont et al. (2010) featuring its robustness properties, and Li et al. (2018) and Embrechts et al. (2018) contain more discussions on its properties and financial applications. In this chapter, the term "risk sharing problem" refers to the search for Pareto-optimal allocations. For discussions on competitive equilibria, see Embrechts et al. (2018, 2019) and Chapter 5. As a first attempt to characterize the forms of optimal allocations, the risk sharing problems we consider are formulated in a static setting with homogeneous beliefs, as opposed to the more sophisticated settings of dynamic equilibrium (see e.g. Beissner and Riedel (2018)) or heterogeneous beliefs (see e.g. Embrechts et al. (2019)).

The proofs of some theorems and lemmas are in the end of each chapter. We end the thesis with conclusions and possible future work.

Due to the different contexts of each chapter, we remind the reader that VaR, ES and the signed Choquet integrals are defined separately and may have different formulations and parametrizations in each chapter.

Chapter 2

Characterization, Robustness and Aggregation of Signed Choquet Integrals

2.1 Introduction

This chapter contains various results on signed Choquet integrals and discusses robustness issues and risk aggregation under uncertainty. It appears in large part in the submitted paper Wang et al. (2018).

The main contributions that we offer are summarized below. In Section 2.2, we establish a characterization of signed Choquet integrals via comonotonic additivity based on the seminal work of Schmeidler (1986). Furthermore, various theoretical properties of signed Choquet integrals are studied, such as monotonicity, additivity, quantile representations, convexity, quasi-convexity, convex order consistency, and mixture-concavity. The characterization and properties are partially known in the literature; yet we are unaware of a good summarizing article (hopefully this chapter serves as one). In particular, few results were found for an atomless probability space.

We proceed to discuss in Sections 2.3 and 2.4 two practically relevant and currently popular problems concerning signed Choquet integrals: robustness issues and risk aggregation with dependence uncertainty. As pointed out by the recent Basel accords (see BCBS (2016)), model uncertainty and robustness become a focal point in both academic research and industry practice of risk assessment over the past few years. We refer to Embrechts et al. (2014) and Emmer et al. (2015) for a summary on these issues and their relation to the recent development in banking and insurance regulation. For more on robustness of risk measures, see Cont et al. (2010), Kou et al. (2013), Krätschmer et al. (2014) and Embrechts et al. (2015), and for more on risk aggregation with dependence uncertainty, see Embrechts et al. (2013), Bernard et al. (2014) and Cai et al. (2017). Our results generalize those of Cont et al. (2010), Embrechts et al. (2015) and Pesenti et al. (2016) on robustness of distortion risk measures and L-statistics, and those of Wang et al. (2015) and Cai et al. (2017) on extreme risk aggregation for distortion risk measures. In particular, the detailed analysis in Wang et al. (2015) used to characterize an extreme-aggregation measure cannot be applied to signed Choquet integrals, and in this chapter, we develop a completely different and more systemic approach based on some recent results on asymptotics of the set of risk aggregation.

2.2 Characterization and properties

2.2.1 Notation and definition

We first list some notation which will be used throughout. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an atomless probability space. Consistently with the literature on risk measures, we work with the space L^∞ of essentially bounded random variables in $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with L^∞-norm $||\cdot||_\infty$; this choice of common domain ensures all functionals we encounter are well-defined. A functional $\rho: L^\infty \to \mathbb{R}$ is *law-invariant* if $\rho(X) = \rho(Y)$ for any $X, Y \in L^\infty$ that have the same distribution under $\mathbb{P}$ (denoted as $X \stackrel{\mathrm{d}}{=} Y$). For all functionals discussed in this chapter, we assume law-invariance. Moreover, we denote by $\mathcal{M}$ the set of distribution functions of $X \in L^\infty$. Terms of "increasing" and "decreasing" are in the non-strict sense.

For $F \in \mathcal{M}$, we write $X \sim F$ for $X \in L^\infty$ and X has distribution F. The left-continuous generalized inverse of F (left-quantile) is denoted by

$$F^{-1}(t) = \inf\{x \in \mathbb{R} : F(x) \geqslant t\},\ t \in (0,1],\ \text{and } F^{-1}(0) = \sup\{x \in \mathbb{R} : F(x) = 0\},$$

whereas its right-continuous generalized inverse (right-quantile) is

$$F^{-1+}(t) = \sup\{x \in \mathbb{R} : F(x) \leqslant t\},\ t \in [0,1),\ \text{and } F^{-1+}(1) = F^{-1}(1).$$

For any random variable X, we use F_X to denote its distribution function. Further, write

$$\mathcal{H} = \{h : h \text{ maps } [0,1] \text{ to } \mathbb{R},\ h \text{ is of bounded variation and } h(0) = 0\}.$$

Next we present the definition of a signed Choquet integral, which originates from the seminal work of Choquet (1954) in the theory of capacities without assuming law-invariance.

Definition 2.2.1. A *signed Choquet integral* $I_h : L^\infty \to \mathbb{R}$ is defined as

$$I_h(X) = \int_{-\infty}^{0} \left(h(\mathbb{P}(X \geqslant x)) - h(1)\right) \mathrm{d}x + \int_{0}^{\infty} h(\mathbb{P}(X \geqslant x))\, \mathrm{d}x, \tag{2.1}$$

where $h \in \mathcal{H}$. The function h is called the *distortion function* of I_h.

We first note that I_h is always finite on L^∞. For $X \in L^\infty$, since $h \in \mathcal{H}$ is of bounded variation, it is measurable and bounded. We can take $M > 0$ such that $|X| \leqslant M$ and hence

$$I_h(X) = \int_{-M}^{0} \left(h(\mathbb{P}(X \geqslant x)) - h(1)\right) \mathrm{d}x + \int_{0}^{M} h(\mathbb{P}(X \geqslant x))\, \mathrm{d}x.$$

As h is bounded, we have $|I_h(X)| < \infty$.

Remark 2.2.1. I_h has an alternative formulation by replacing $\mathbb{P}(X \geqslant x)$ with $\mathbb{P}(X > x)$ in (2.1). For $X \in L^\infty$, the functions $h(\mathbb{P}(X \geqslant x))$ and $h(\mathbb{P}(X > x))$ are equal almost everywhere for $x \in \mathbb{R}$, and therefore

$$I_h(X) = \int_{-\infty}^{0} \left(h(\mathbb{P}(X > x)) - h(1)\right) \mathrm{d}x + \int_{0}^{\infty} h(\mathbb{P}(X > x))\, \mathrm{d}x. \tag{2.2}$$

In different places we shall use either of (2.1) and (2.2), whichever is more convenient.

From (2.1), it is clear that $I_{ah_1+bh_2} = aI_{h_1} + bI_{h_2}$ for $h_1, h_2 \in \mathcal{H}$ and $a, b \in \mathbb{R}$. In particular, for any $h \in \mathcal{H}$, we have $I_h = I_{h_+} - I_{h_-}$, where $h_+ \in \mathcal{H}$ and $h_- \in \mathcal{H}$ are increasing functions such that $h = h_+ - h_-$ via the Jordan decomposition. This decomposition will be used repeatedly in this chapter, as often results are available in the literature for Choquet integrals with an increasing distortion function.

Before we proceed with characterizing signed Choquet integrals, we present some more terminology used throughout the chapter. A most relevant concept to signed Choquet integrals is comonotonicity. Random variables X and Y are said to be *comonotonic* if there exists $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ such that $\omega, \omega' \in \Omega_0$,

$$(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geqslant 0.$$

For a functional $\rho : L^\infty \to \mathbb{R}$, we say that ρ is *comonotonic-additive*, if for any comonotonic random variables $X, Y \in L^\infty$, $\rho(X + Y) = \rho(X) + \rho(Y)$; ρ is *positively homogeneous*, if for

$X \in L^\infty$ and constant $\lambda > 0$, $\rho(\lambda X) = \lambda\rho(X)$; ρ is *(uniformly) norm-continuous*, if it is (uniformly) continuous with respect to L^∞-norm; ρ is *quasi-convex* if $\rho(\lambda X + (1-\lambda)Y) \leqslant \max\{\rho(X), \rho(Y)\}$ for all $X, Y \in L^\infty$ and $\lambda \in [0, 1]$.

A random variable X is said to be smaller than a random variable Y in *convex order*, denoted by $X \leqslant_{\mathrm{cx}} Y$, if $\mathbb{E}[\phi(X)] \leqslant \mathbb{E}[\phi(Y)]$ for all convex $\phi : \mathbb{R} \to \mathbb{R}$, provided that both expectations exist. The following fact about comonotonicity and convex order is well-known (see e.g. Theorem 3.5 of Rüschendorf (2013)): for any integrable random variables X, Y, X^c and Y^c such that $X \stackrel{\mathrm{d}}{=} X^c$, $Y \stackrel{\mathrm{d}}{=} Y^c$, and X^c and Y^c are comonotonic, one has $X + Y \leqslant_{\mathrm{cx}} X^c + Y^c$. We say that a functional $\rho : L^\infty \to \mathbb{R}$ is *convex order consistent* if $\rho(X) \leqslant \rho(Y)$ for all random variables $X, Y \in L^\infty$ satisfying $X \leqslant_{\mathrm{cx}} Y$.

2.2.2 Characterization

In the following, we establish a functional characterization for signed Choquet integrals. As far as we are aware of, this characterization is not known to the literature without assuming monotonicity. We shall first show that a law-invariant, comonotonic-additive and uniformly norm-continuous functional from L^∞ to $\mathbb{R}$ is necessarily a signed Choquet integral, based on a remarkable result of Schmeidler (1986), which we list as Theorem 2.5.1 in the appendix for completeness. The converse is also true, but it will be verified later as we establish some further properties of the signed Choquet integrals.

Theorem 2.2.1. *A functional $I : L^\infty \to \mathbb{R}$ is law-invariant, comonotonic-additive and uniformly norm-continuous if and only if I is a signed Choquet integral.*

Proof. (i) "$\Rightarrow$": By Theorem 2.5.1 (Proposition 2 of Schmeidler (1986)), a comonotonic-additive and norm-continuous functional I has a representation

$$I(X) = \int_{-\infty}^{0} (v(X \geqslant x) - v(\Omega))\, \mathrm{d}x + \int_0^\infty v(X \geqslant x)\, \mathrm{d}x, \quad X \in L^\infty, \tag{2.3}$$

where the set function $v : \mathcal{F} \to \mathbb{R}$ is given by $v(E) = I(\mathbf{1}_E)$, $E \in \mathcal{F}$. Note that I is law-invariant, which means $I(\mathbf{1}_E) = h(\mathbb{P}(E))$ for some function $h : [0, 1] \to \mathbb{R}$. Hence $v(E) = h(\mathbb{P}(E))$ for $E \in \mathcal{F}$, and (2.3) can be rewritten as

$$I(X) = \int_{-\infty}^{0} (h(\mathbb{P}(X \geqslant x)) - h(1))\, \mathrm{d}x + \int_0^\infty h(\mathbb{P}(X \geqslant x))\, \mathrm{d}x. \tag{2.4}$$

Next we verify $h \in \mathcal{H}$, so that I is indeed a signed Choquet integral. Noting that comonotonic additivity gives $I(0) + I(0) = I(0) = h(0)$, we have $h(0) = 0$. It remains to verify that h is of bounded variation. Let U be a uniform random variable on $[0,1]$. First, notice that $h(t) = I(\mathbf{1}_{\{U<t\}}) < \infty$ for $t \in [0,1]$. Thus h is finite. As I is uniformly norm-continuous, for a fixed $\varepsilon > 0$, there exists a $\delta > 0$ such that $|I(X) - I(Y)| < \varepsilon$, whenever $\|X - Y\|_\infty < \delta$. Let $\mathcal{P} = \{t_0, \dots, t_n\}$ be an arbitrary partition of $[0,1]$, where $0 = t_0 < \cdots < t_n = 1$. In the summation $\sum_{i=1}^n |h(t_i) - h(t_{i-1})|$, there are exactly n terms of $h(x)$, $x \in \mathcal{P}$ with a positive sign, and n terms of $h(y)$, $y \in \mathcal{P}$ with a negative sign. Therefore, we can write two increasing sequences $\{x_1, \dots, x_n\} \subset \mathcal{P}$ and $\{y_1, \dots, y_n\} \subset \mathcal{P}$ such that

$$\sum_{i=1}^n |h(t_i) - h(t_{i-1})| = \sum_{i=1}^n h(x_i) - \sum_{i=1}^n h(y_i).$$

Since positive and negative terms in the summation $\sum_{i=1}^n |h(t_i) - h(t_{i-1})|$ appear in pairs, we have $x_i, y_i \in [t_{i-1}, t_i]$, $i = 1, \dots, n$.

Next, let $f : [0,1] \to \mathbb{R}$ and $g : [0,1] \to \mathbb{R}$ be given by $f(t) = \sum_{i=1}^n \mathbf{1}_{\{t>1-x_i\}}$ and $g(t) = \sum_{i=1}^n \mathbf{1}_{\{t>1-y_i\}}$. Let $X = \delta f(U)$ and $Y = \delta g(U)$. Clearly, $\|X - Y\|_\infty \leqslant \delta$ because $x_i, y_i \in [t_{i-1}, t_i]$ for each $i = 1, \dots, n$. It is straightforward to calculate

$$I(X) = \int_0^\infty h(\mathbb{P}(\delta f(U) > x))\,\mathrm{d}x = \delta \int_0^\infty h(\mathbb{P}(f(U) > y))\,\mathrm{d}y = \delta \sum_{i=1}^n h(x_i),$$

and similarly $I(Y) = \delta \sum_{i=1}^n h(y_i)$. Noting that $\|X - Y\|_\infty < \delta$, we have

$$|I(X) - I(Y)| = \delta \left| \sum_{i=1}^n h(x_i) - \sum_{i=1}^n h(y_i) \right| < \varepsilon.$$

It follows that

$$\sum_{i=1}^n |h(t_i) - h(t_{i-1})| < \frac{\varepsilon}{\delta} < \infty,$$

and this holds for an arbitrary partition $\mathcal{P} = \{t_0, \dots, t_n\}$. Thus h has bounded variation.

(ii) "$\Leftarrow$": Law-invariance is obvious. The uniform norm-continuity of a signed Choquet integral is verified by Lemma 2.3.1 in Section 2.3. Comonotonic additivity is implied by Lemma 2.2.4 below; see Remark 2.2.4. □

Remark 2.2.2. Theorem 4.2 of Murofushi et al. (1994) characterizes signed Choquet integrals that are not necessarily law-invariant. Comparing the above result with our Theorem

2.2.1, our result suggests that this extra law-invariance condition implies the existence of a function h such that $I = I_h$, and h has bounded variation on $[0,1]$. The corresponding condition in Murofushi et al. (1994) is that the set function μ, which is $h \circ \mathbb{P}$ in our work, has bounded variation on $(\Omega, \mathcal{F})$. In fact, we can verify from the definition of total variation in Murofushi et al. (1994) that the total variation of h on $[0,1]$ is equal to the total variation of $h \circ \mathbb{P}$ on $(\Omega, \mathcal{F})$.

We can also compare Theorem 2.2.1 with Theorem 22 of Cerreia-Vioglio et al. (2015) which characterize signed Choquet integrals without law-invariance on general spaces. In the latter result, a property of functional bounded variation is imposed, instead of the uniform norm-continuity in Theorem 2.2.1. Generally, uniform norm-continuity is not sufficient for functional bounded variation used in Cerreia-Vioglio et al. (2012, 2015). The assumption of law-invariance provides extra regularity and continuity for the underlying functional, due to a huge dimension reduction resulting from mapping random variables to their distributions. This phenomenon is well documented in the risk management literature, see e.g. Jouini et al. (2006) for the case of convex risk measures on L^∞ and more recently, Gao et al. (2017) and Gao and Xanthos (2017) for the case of convex risk measures on Orlicz hearts. Generally speaking, assuming the same set of other properties, law-invariant functionals have better regularity conditions than non-law-invariant ones.

Remark 2.2.3. From the proof of Theorem 2.2.1, we see that, if uniform norm-continuity of I is weakened to norm-continuity, a representation of the form (2.4) holds with a function h not necessarily of bounded variation (thus, not a signed Choquet integral according to our definition). Indeed, for a positively homogeneous functional, uniform continuity is equivalent to Lipschitz continuity; see also Lemma 2.3.1.

2.2.3 Basic properties

In this and the next few sections, we give several basic properties of signed Choquet integrals which will be useful in Sections 2.3 and 2.4. These properties are partially known in the literature (see e.g. De Waegenaere and Wakker (2001) and Acerbi (2002) for special cases), and can be derived from classic properties of increasing Choquet integrals; for the sake of completeness we provide short self-contained proofs in the appendix.

Lemma 2.2.2. *For $h_1, h_2 \in \mathcal{H}$, if $h_1(1) = h_2(1)$, then*

$$h_1 \leqslant h_2 \text{ on } [0,1] \quad \Leftrightarrow \quad I_{h_1} \leqslant I_{h_2} \text{ on } L^\infty.$$

In particular, $h_1 = h_2$ holds if and only if $I_{h_1} = I_{h_2}$ on L^∞.

For $h \in \mathcal{H}$, I_h is said to be *increasing* (or *decreasing*) if, for all random variables $X, Y \in L^\infty$, $X \leqslant Y$ almost surely implies $I_h(X) \leqslant I_h(Y)$ (or $I_h(X) \geqslant I_h(Y)$, respectively).

Lemma 2.2.3. *For $h \in \mathcal{H}$,*

(i) I_h is increasing (respectively decreasing) if and only if h is increasing (respectively decreasing);

(ii) for $X \in L^\infty$ and $c \in \mathbb{R}$, $I_h(X + c) = I_h(X) + ch(1)$;

(iii) for $X \in L^\infty$ and $\lambda > 0$, $I_h(\lambda X) = \lambda I_h(X)$;

(iv) for $X \in L^\infty$, $I_h(-X) = I_{\hat{h}}(X)$, where $\hat{h} : [0, 1] \to \mathbb{R}$ is given by $\hat{h}(x) = h(1-x)-h(1)$.

In the context of non-law-invariant comonotonic-additive functionals, similar results to Lemma 2.2.3 (i)-(iii) can be found in Proposition 4.11 of Marinacci and Montrucchio (2004).

2.2.4 Quantile representation

In this section, we present an important property of signed Choquet integrals, namely, the quantile representation. This result will be referred to repeatedly in this chapter. In particular, it is used to show the following properties of a signed Choquet integral: comonotonic additivity (Theorem 2.2.1 above), convex order consistency (Theorem 2.2.5 below), continuity with respect to weak convergence (Theorem 2.3.2 below), and extreme-aggregation for heterogeneous portfolios (Theorem 2.4.4 below). In the following Lemma, the first two conditions (i) and (ii) for a quantile representation are known for increasing Choquet integrals (see e.g. Denneberg (1994) and Theorems 4 and 6 of Dhaene et al. (2012)). Although (i) and (ii) can be obtained from corresponding results on increasing Choquet integrals via a Jordan decomposition, we give an independent short proof here.

Lemma 2.2.4. *For $h \in \mathcal{H}$ and $X \in L^\infty$,*

(i) if h is right-continuous, then $I_h(X) = \int_0^1 F_X^{-1+}(1-p)\,\mathrm{d}h(p)$;

(ii) if h is left-continuous, then $I_h(X) = \int_0^1 F_X^{-1}(1-p)\,\mathrm{d}h(p)$;

(iii) if F_X^{-1} is continuous, then $I_h(X) = \int_0^1 F_X^{-1}(1-p)\,\mathrm{d}h(p)$.

Proof. (i) Without loss of generality, we may assume $X \geqslant 0$, and the general case can be easily obtained via Lemma 2.2.3. Noting that h is right-continuous, $h(\mathbb{P}(X > x)) = \int_0^{\mathbb{P}(X>x)} \mathrm{d}h(p)$. Since h is of bounded variation, one can apply Fubini's theorem to the following Lebesgue-Stieltjes integral,

$$I_h(X) = \int_0^\infty \int_0^{\mathbb{P}(X>x)} \mathrm{d}h(p)\,\mathrm{d}x = \int_0^1 \int_0^{F_X^{-1+}(1-p)} \mathrm{d}x\,\mathrm{d}h(p) = \int_0^1 F_X^{-1+}(1-p)\,\mathrm{d}h(p),$$

where the second equality is due to $p \leqslant \mathbb{P}(X > x) \Leftrightarrow x \leqslant F_X^{-1+}(1-p)$.

(ii) Note that $\hat{h}$ in part (iii) of Lemma 2.2.3 is right-continuous, and $F_X^{-1}(p) = -F_{-X}^{-1+}(1-p)$. Applying part (iii) of Lemma 2.2.3, we obtain

$$\begin{aligned} I_h(X) = I_{\hat{h}}(-X) = \int_0^1 F_{-X}^{-1+}(1-p)\,\mathrm{d}\hat{h}(p) &= -\int_0^1 F_X^{-1}(p)\,\mathrm{d}h(1-p) \\ &= \int_0^1 F_X^{-1}(1-p)\,\mathrm{d}h(p). \end{aligned}$$

(iii) As h can be replaced by its Jordan decomposition $h = h_+ - h_-$, it suffices to show the representation for h increasing. First note that $\int_0^1 F_X^{-1}(1-p)\,\mathrm{d}h(p)$ is finite, and through integration-by-parts,

$$\int_0^1 F_X^{-1}(1-p)\,\mathrm{d}h(p) = F_X^{-1}(0)h(1) - \int_0^1 h(p)\,\mathrm{d}F_X^{-1}(1-p).$$

For $p \in [0,1]$, we have

$$p \in [\mathbb{P}(X > F_X^{-1}(1-p)), \mathbb{P}(X \geqslant F_X^{-1}(1-p))].$$

Define $g_1^* : \mathbb{R} \to \mathbb{R}$ by

$$g_1^*(x) = \sup\{h(y) \in \mathbb{R} : y \in [\mathbb{P}(X > x), \mathbb{P}(X \geqslant x)]\}.$$

For $p \in [0,1]$,

$$h(p) \leqslant g_1^*(F_X^{-1}(1-p)) = \sup\{h(y) \in \mathbb{R} : y \in [\mathbb{P}(X > F_X^{-1}(1-p)), \mathbb{P}(X \geqslant F_X^{-1}(1-p))]\},$$

and therefore,

$$\begin{aligned} \int_0^1 h(p)\,\mathrm{d}F_X^{-1}(1-p) \leqslant \int_0^1 g_1^*(F_X^{-1}(1-p))\,\mathrm{d}F_X^{-1}(1-p) &= \int_{F_X^{-1}(1)}^{F_X^{-1}(0)} g_1^*(t)\,\mathrm{d}t \\ &= \int_{F_X^{-1}(1)}^{F_X^{-1}(0)} h(\mathbb{P}(X \geqslant t))\,\mathrm{d}t. \end{aligned}$$

Via a symmetric argument through replacing g_1^* by $g_2^* : \mathbb{R} \to \mathbb{R}$,

$$g_2^*(x) = \inf\{h(y) \in \mathbb{R} : y \in [\mathbb{P}(X > x), \mathbb{P}(X \geqslant x)]\},$$

we obtain

$$\int_{F_X^{-1}(1)}^{F_X^{-1}(0)} h(\mathbb{P}(X > x))\,\mathrm{d}x \leqslant \int_0^1 h(p)\,\mathrm{d}F_X^{-1}(1-p).$$

Note that

$$\int_{F_X^{-1}(1)}^{F_X^{-1}(0)} h(\mathbb{P}(X > x))\,\mathrm{d}x = \int_{F_X^{-1}(1)}^{F_X^{-1}(0)} h(\mathbb{P}(X \geqslant x))\,\mathrm{d}x,$$

and therefore we have

$$\int_0^1 h(p)\,\mathrm{d}F_X^{-1}(1-p) = \int_{F_X^{-1}(1)}^{F_X^{-1}(0)} h(\mathbb{P}(X \geqslant x))\,\mathrm{d}x.$$

Finally, for $X \in L^\infty$,

$$\begin{aligned} I_h(X) &= \int_{F_X^{-1}(0)}^{F_X^{-1}(1)} h(\mathbb{P}(X \geqslant x))\,\mathrm{d}x - \int_{F_X^{-1}(0)}^{0} h(1)\,\mathrm{d}x \\ &= -\int_0^1 h(p)\,\mathrm{d}F_X^{-1}(1-p) + F_X^{-1}(0)h(1) \\ &= \int_0^1 F_X^{-1}(1-p)\,\mathrm{d}h(p). \end{aligned}$$

This completes the proof. □

Remark 2.2.4. Part (i) of Lemma 2.2.4 implies comonotonic additivity of a signed Choquet integral I_h. First, we decompose $h = h_l + h_r$, where h_l and h_r are left-continuous and right-continuous, respectively. This is always possible as h has countably many points of discontinuity. Then, it follows from Lemma 2.2.4 that I_{h_l} and I_{h_r} are both comonotonic-additive, as the left- and right-quantiles are comonotonic-additive (a well-known fact; see e.g. Proposition 7.20 of McNeil et al. (2015) for the case of left-quantiles).

2.2.5 Convexity, convex order consistency, and mixture-concavity

Next we show that convex order consistency of a signed Choquet integral is equivalent to its distortion function being concave. For increasing Choquet integrals, this result is established by Yaari (1987).

Theorem 2.2.5. *For random variables $X, Y \in L^\infty$, $X \leqslant_{\mathrm{cx}} Y$ if and only if $I_h(X) \leqslant I_h(Y)$ for all concave functions $h \in \mathcal{H}$.*

Proof. (i) " $\Rightarrow$": Given $X, Y \in L^\infty$ with distributions F and G respectively, let

$$a = \operatorname{ess\,inf}\{X\} \wedge \operatorname{ess\,inf}\{Y\},\ \ b = \operatorname{ess\,sup}\{X\} \vee \operatorname{ess\,sup}\{Y\}$$

and $f = 1 - F,\ g = 1 - G$. If $X \leqslant_{\mathrm{cx}} Y$, by Equation (3.A.7) of Shaked and Shanthikumar (2007),

$$\mathbb{E}[X] = \mathbb{E}[Y] \text{ and } \int_x^\infty \overline{F}(u)\,\mathrm{d}u \leqslant \int_x^\infty \overline{G}(u)\,\mathrm{d}u \text{ for all } x,$$

which is

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^b g(t)\,\mathrm{d}t \quad \text{and} \quad \int_a^x f(t)\,\mathrm{d}t \geqslant \int_a^x g(t)\,\mathrm{d}t \quad \text{for} \quad a \leqslant x \leqslant b.$$

A concave function h defined on $[0,1]$ is necessarily continuous on $(0,1)$. Define $\phi : [0,1] \to \mathbb{R}$ by

$$\phi(x) = \begin{cases} h(x) & \text{for } 0 < x < 1; \\ \lim\limits_{x \downarrow 0} h(x) & \text{for } x = 0; \\ \lim\limits_{x \uparrow 1} h(x) & \text{for } x = 1. \end{cases}$$

Since h has bounded variation, it can be written as the difference of two increasing functions. As the bounded monotone functions have finite limits, $\lim_{x\downarrow 0} h(x)$ and $\lim_{x\uparrow 1} h(x)$ are well defined. Note that ϕ is a continuous concave function, $\phi = h$ on $(0,1)$ and $\phi \geqslant h$ on $[0,1]$. By the classic Hardy-Littlewood-Pólya inequality (listed as Theorem 2.5.2 for the sake of completeness),

$$\int_a^b \phi(f(x))\,\mathrm{d}x \leqslant \int_a^b \phi(g(x))\,\mathrm{d}x.$$

By Equation (3.A.12) of Shaked and Shanthikumar (2007), $a = \operatorname{ess\,inf}\{Y\}$ and $b = \operatorname{ess\,sup}\{Y\}$, and therefore $h(g(x)) = \phi(g(x))$ for $x \in (a,b)$. Moreover, $h(f(x)) = h(g(x))$ for $x > b$ or $x < a$. Utilizing the above observations, we have

$$\begin{aligned} I_h(X) - I_h(Y) &= \int_a^b \left(h(f(x)) - h(g(x))\right)\,\mathrm{d}x \\ &= \int_a^b \left(h(f(x)) - \phi(g(x))\right)\,\mathrm{d}x \\ &\leqslant \int_a^b \left(\phi(f(x)) - \phi(g(x))\right)\,\mathrm{d}x \leqslant 0. \end{aligned}$$

(ii) " $\Leftarrow$": For all $p \in [0,1]$, $t \in [0,1]$, let $h(t) = -\mathbf{1}_{\{t \geqslant 1-p\}}(t-1+p)$, and then h is concave and in $\mathcal{H}$. For fixed p, by Lemma 2.2.4,

$$I_h(X) = -\int_{1-p}^{1} F_X^{-1}(1-t)\,\mathrm{d}t = -\int_0^p F_X^{-1}(u)\,\mathrm{d}u.$$

Thus for all $p \in [0,1]$, $I_h(X) \leqslant I_h(Y)$ results in

$$\int_0^p F^{-1}(t)\,\mathrm{d}t \geqslant \int_0^p G^{-1}(t)\,\mathrm{d}t,$$

which implies $X \leqslant_{\mathrm{cx}} Y$ by Theorem 3.A.5 of Shaked and Shanthikumar (2007). $\square$

Remark 2.2.5. The forward implication of Theorem 2.2.5 can also be deduced by noticing that $\nu = h \circ \mathbb{P}$ defines a submodular game (see Marinacci and Montrucchio (2004)) whenever h is concave. Then an application of Corollary 4.2 of Marinacci and Montrucchio (2004) and Theorem 4.1 of Dana (2005) establishes the claim.

At this point, we are ready to establish six equivalent conditions characterizing the convexity of a signed Choquet integral. For a law-invariant functional ρ on L^∞, define $\tilde{\rho} : \mathcal{M} \to \mathbb{R}$ by $\tilde{\rho}(F) = \rho(X)$ where $X \sim F$, and we say that ρ is *concave on mixtures* if $\tilde{\rho}$ is concave.

Theorem 2.2.6. *For $h \in \mathcal{H}$, the following are equivalent: (i) h is concave; (ii) I_h is convex order consistent; (iii) I_h is subadditive; (iv) I_h is convex; (v) I_h is quasi-convex; (vi) I_h is concave on mixtures.*

Proof. We complete the proof in the order $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (v) \Rightarrow (vi) \Rightarrow (i)$.

$(i) \Rightarrow (ii)$: Guaranteed by Theorem 2.2.5.

$(ii) \Rightarrow (iii)$: By Theorem 2.2.1, I_h is law-invariant and comonotonic-additive. We take random variables X, Y and comonotonic random variables X^c, Y^c whose distribution functions are identical to X, Y, respectively. Then $X + Y \leqslant_{\mathrm{cx}} X^c + Y^c$ as mentioned in Section 2.2.1. Thus

$$I_h(X+Y) \leqslant I_h(X^c + Y^c) = I_h(X^c) + I_h(Y^c) = I_h(X) + I_h(Y),$$

and I_h is subadditive.

$(iii) \Rightarrow (iv)$: As I_h is positively homogeneous, subadditivity is equivalent to convexity.

$(iv) \Rightarrow (v)$: Convexity is stronger than quasi-convexity by definition.

$(v) \Rightarrow (vi)$: Take any $x, y \in [0,1]$, $x \leqslant y$. Define random variables X, Y, Z by

$$\mathbb{P}(X=0)=1-y,\ \mathbb{P}(X=1/2)=y-x,\ \mathbb{P}(X=1)=x,$$

and the joint distribution function of Y and Z is given by

$$\mathbb{P}(Y=0, Z=0)=1-y,\ \mathbb{P}(Y=1, Z=1)=x,$$

$$\mathbb{P}(Y=1, Z=0)=\mathbb{P}(Y=0, Z=1)=\frac{y-x}{2}.$$

Clearly $X \stackrel{\mathrm{d}}{=} \frac{1}{2}Y+\frac{1}{2}Z$ and $Y \stackrel{\mathrm{d}}{=} Z$. Since I_h is quasi-convex and law-invariant, we have

$$I_h(X)=I_h\left(\frac{1}{2}Y+\frac{1}{2}Z\right) \leqslant \max\{I_h(Y), I_h(Z)\}=I_h(Y).$$

Note that

$$\mathbb{P}(X \geqslant t)=\begin{cases} 1 & t \leqslant 0; \\ y & 0<t \leqslant \frac{1}{2}; \\ x & \frac{1}{2}<t \leqslant 1; \\ 0 & t>1, \end{cases} \qquad \text{and} \qquad \mathbb{P}(Y \geqslant t)=\begin{cases} 1 & t \leqslant 0; \\ \frac{1}{2}x+\frac{1}{2}y & 0<t \leqslant 1; \\ 0 & t>1. \end{cases}$$

As

$$I_h(X)=\int_0^{\frac{1}{2}} h(y)\,\mathrm{d}t+\int_{\frac{1}{2}}^1 h(x)\,\mathrm{d}t=\frac{1}{2}h(x)+\frac{1}{2}h(y),$$

and

$$I_h(Y)=\int_{-\infty}^0 (h(1)-h(1))\,\mathrm{d}t+\int_0^1 h\left(\frac{1}{2}x+\frac{1}{2}y\right)\,\mathrm{d}t+\int_1^\infty h(0)\,\mathrm{d}t=h\left(\frac{1}{2}x+\frac{1}{2}y\right),$$

$I_h(X) \leqslant I_h(Y)$ leads to $\frac{1}{2}h(x)+\frac{1}{2}h(y) \leqslant h\left(\frac{1}{2}x+\frac{1}{2}y\right)$; thus h is mid-point concave. By the Sierpinski Theorem (see page 12 of Donoghue (1969)), a mid-point concave and Lebesgue measurable function is a concave function. Therefore h is concave; (i) holds. With concavity of h, (vi) is straightforward from the definition of Choquet integral in (2.1).

$(vi) \Rightarrow (i)$: For $p, q, \lambda \in [0,1]$, let F be a Bernoulli distribution with mean p and G be a Bernoulli distribution with mean q. Then $\lambda F+(1-\lambda)G$ is the Bernoulli distribution with mean $\lambda p+(1-\lambda)q$. It follows from simple calculation that

$$\lambda h(p)+(1-\lambda)h(q)=\lambda \tilde{I}_h(F)+(1-\lambda)\tilde{I}_h(G) \leqslant \tilde{I}_h(\lambda F+(1-\lambda)G)=h(\lambda p+(1-\lambda)q),$$

and thus h is concave. $\square$

Remark 2.2.6. Concavity on mixtures (mixture-concavity) is a natural property for risk functionals, especially measures of variability, as it assigns a higher risk value to a mixture of two distributions with equal risk value; see e.g. Acciaio and Svindland (2013). This property is satisfied by classic variability measures, such as the variance, the standard deviation and the Gini deviation; see Section 2.2.6 below. Although being equivalent for signed Choquet integrals, mixture-concavity is essentially different from convexity for general functionals, in terms of both mathematical and economic interpretations. For instance, taking a supremum over convex signed Choquet integrals preserves convexity and may lose mixture-concavity, whereas taking an infimum over convex signed Choquet integrals preserves mixture-concavity and may lose convexity.

2.2.6 Some examples

Example 2.2.1. We first present some examples of signed Choquet integrals used as measures of distributional variability. Note that all distortion functions below are concave but not monotone.

(i) The range:

$$\mathrm{Range}(X) = \operatorname{ess\,sup}(X) - \operatorname{ess\,inf}(X), \quad X \in L^\infty.$$

The range is a signed Choquet integral with a concave distortion function h given by $h(t) = \mathbf{1}_{\{0<t<1\}},\ t \in [0,1]$. This fact can be checked by straightforward calculation:

$$I_h(X) = \int_{\operatorname{ess\,inf}(X)}^{0} \mathrm{d}x + \int_0^{\operatorname{ess\,sup}(X)} \mathrm{d}x = \operatorname{ess\,sup}(X) - \operatorname{ess\,inf}(X) = \mathrm{Range}(X).$$

(ii) The mean median difference:

$$\mathrm{MD}(X) = \min_{x\in\mathbb{R}} \mathbb{E}[|X-x|] = \mathbb{E}\left[\left|X - F_X^{-1}\left(\frac{1}{2}\right)\right|\right], \quad X \in L^\infty.$$

The mean median difference is a signed Choquet integral with a concave distortion function h given by $h(t) = \min\{t, 1-t\},\ t \in [0,1]$. This can be checked using Lemma

2.2.4:

$$\begin{aligned}
I_h(X) &= \int_{\frac{1}{2}}^{1} F_X^{-1}(u)\,\mathrm{d}u - \int_0^{\frac{1}{2}} F_X^{-1}(u)\,\mathrm{d}u \\
&= \frac{1}{2}F_X^{-1}\left(\frac{1}{2}\right) - \int_0^{\frac{1}{2}} F_X^{-1}(u)\,\mathrm{d}u + \int_{\frac{1}{2}}^{1} F_X^{-1}(u)\,\mathrm{d}u - \frac{1}{2}F_X^{-1}\left(\frac{1}{2}\right) \\
&= \int_0^1 \left| F_X^{-1}(u) - F_X^{-1}\left(\frac{1}{2}\right)\right| \mathrm{d}u = \mathrm{MD}(X).
\end{aligned}$$

(iii) The Gini deviation:

$$\mathrm{Gini}(X) = \frac{1}{2}\mathbb{E}[|X_1 - X_2|], \quad X \in L^\infty,\ X_1, X_2, X \text{ are iid.}$$

The Gini deviation is a signed Choquet integral with a concave distortion function h given by $h(t) = t - t^2,\ t \in [0,1]$. This is due to its alternative form (see e.g. Denneberg (1990))

$$\mathrm{Gini}(X) = \int_0^1 F_X^{-1}(t)(2t-1)\,\mathrm{d}t.$$

Example 2.2.2. Next we present some examples of signed Choquet integrals used as measures of risk. The first two popular risk measures used in regulation are increasing signed Choquet integrals. The last one does not necessarily have an increasing distortion function.

(i) The Value-at-Risk (VaR) for $p \in (0,1)$:

$$\mathrm{VaR}_p(X) = \inf\{x : \mathbb{P}(X \leqslant x) \geqslant p\}, \quad X \in L^\infty.$$

VaR_p for $p \in (0,1)$ is a signed Choquet integral with distortion function h given by $h(t) = \mathbf{1}_{\{t > 1-p\}},\ t \in [0,1]$. This can be directly checked via Lemma 2.2.4:

$$I_h(X) = F_X^{-1}(p) = \mathrm{VaR}_p(X).$$

(ii) The Expected Shortfall (ES) for $p \in (0,1)$:

$$\mathrm{ES}_p(X) = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_t(X)\,\mathrm{d}t, \quad X \in L^\infty.$$

ES_p for $p \in (0,1)$ is a signed Choquet integral with distortion function h given by $h(t) = \min\{\frac{t}{1-p}, 1\},\ t \in [0,1]$. This can be directly checked via Lemma 2.2.4:

$$I_h(X) = \frac{1}{1-p}\int_p^1 F_X^{-1}(t)\,\mathrm{d}t = \mathrm{ES}_p(X).$$

(iii) The Gini Shortfall (GS) for $p \in [0,1)$ and $\lambda \geqslant 0$:

$$\mathrm{GS}_p^\lambda(X) = \mathrm{ES}_p(X) + \lambda \mathrm{TGini}_p(X), \quad X \in L^\infty,$$

where TGini_p is the tail-Gini functional,

$$\mathrm{TGini}_p(X) = \frac{2}{(1-p)^2}\int_p^1 F_X^{-1}(t)(2t-(1+p))\,\mathrm{d}t, \quad X \in L^\infty.$$

By Theorem 4.1 of Furman et al. (2017), GS_p^λ for $p \in [0,1)$ and $\lambda \geqslant 0$ is a signed Choquet integral with distortion function h given by

$$h(t) = \frac{1}{(1-p)^2}\left((1-p)t + 4\lambda t\left(1 - \frac{t}{2} - \frac{1+p}{2}\right)\right)\mathbf{1}_{\{t\leqslant 1-p\}} + \mathbf{1}_{\{t>1-p\}},\ t \in [0,1].$$

A Gini Shortfall is an increasing Choquet integral if and only if $\lambda \in [0, \frac{1}{2}]$.

Example 2.2.3. Below we look at the standard deviation, which is not a signed Choquet integral, but a supremum over some signed Choquet integrals. Let

$$\tilde{\mathcal{H}} = \left\{h \in \mathcal{H} : h(1) = 0,\ \int_0^1 (h'(t))^2\,\mathrm{d}t \leqslant 1,\ h \text{ is concave}\right\}.$$

The standard deviation, defined as

$$\sigma(X) = \sqrt{\mathbb{E}[X^2] - (\mathbb{E}[X])^2}, \quad X \in L^\infty,$$

has the following representation

$$\sigma(X) = \sup_{h \in \tilde{\mathcal{H}}} I_h(X), \quad X \in L^\infty, \tag{2.5}$$

and hence it is the supremum over a class of signed Choquet integrals.

To show this, let $\mathcal{Z} = \{Z \in L^\infty : \mathbb{E}[Z] = 0,\ \mathbb{E}[Z^2] \leqslant 1\}$. It is clear that, for $X \in L^\infty$,

$$\sigma(X) = \frac{\mathbb{E}[X(X-\mathbb{E}[X])]}{\sigma(X)} = \mathbb{E}\left[X\frac{X-\mathbb{E}[X]}{\sigma(X)}\right] \leqslant \sup_{Z\in\mathcal{Z}} \mathbb{E}[XZ],$$

and for any $Z \in \mathcal{Z}$, we have $\mathbb{E}[XZ] = \mathrm{cov}(X, Z) \leqslant \sigma(Z)\sigma(X) \leqslant \sigma(X)$. Therefore, by the Fréchet-Hoeffding inequality (see e.g. Lemma 4.60 of Föllmer and Schied (2016)),

$$\sigma(X) = \sup_{Z\in\mathcal{Z}} \mathbb{E}[XZ] = \sup_{Z\in\mathcal{Z}} \int_0^1 F_Z^{-1}(u) F_X^{-1}(u)\,\mathrm{d}u.$$

Via the relation $h(t) = \int_0^t F_Z^{-1}(1-u)\,\mathrm{d}u,\ t \in [0,1]$, we establish a one-to-one mapping from the distributions of $Z \in \mathcal{Z}$ to functions in $\tilde{\mathcal{H}}$. Therefore, using Lemma 2.2.4,

$$\sigma(X) = \sup_{Z\in\mathcal{Z}} \int_0^1 F_Z^{-1}(u) F_X^{-1}(u)\,\mathrm{d}u = \sup_{h\in\tilde{\mathcal{H}}} \int_0^1 F_X^{-1}(1-u)\,\mathrm{d}h(u) = \sup_{h\in\tilde{\mathcal{H}}} I_h(X), \quad X \in L^\infty.$$

Note that $h \in \tilde{\mathcal{H}}$ is concave. As a consequence, each $I_h,\ h \in \tilde{\mathcal{H}}$ is subadditive, convex and consistent with the convex order (see Theorem 2.2.6), and so is the standard deviation σ by noting that these properties are preserved when taking a supremum.

Indeed, all law-invariant deviation measures in the sense of Rockafellar et al. (2006) admit a signed Choquet integral representation similar to (2.5); this result is established in Grechuk et al. (2009).

Example 2.2.4. We look at two further examples of measures of distributional variability based on risk measures in Example 2.2.2. They will be revisited in Sections 2.3 and 2.4.

(i) The inter-quantile range (IQR) for $p \in (1/2, 1)$:

$$\mathrm{IQR}_p(X) = \mathrm{VaR}_p(X) - \mathrm{VaR}_{1-p}(X), \quad X \in L^\infty.$$

The inter-quantile range is a commonly used measure of dispersion in statistics, and the typical choice of p is 0.75, yielding the difference between the first quarter and the third quarter quantiles. IQR_p for $p \in (1/2, 1)$ is a signed Choquet integral with distortion function h given by $h(t) = \mathbf{1}_{\{1-p<t\leqslant p\}}$, $t \in [0,1]$; see Figure 2.1. Unlike the other measures of variability in Example 2.2.1, the distortion function h of IQR_p is not concave, and hence IQR_p is not convex or convex-order consistent by Theorem 2.2.6. For $X \in L^\infty$ with a continuous quantile at $1-p$, noting that $F_X^{+}(1-p) = -F_{-X}(p)$ by Lemma 2.2.3 (iv), we can alternatively write

$$\mathrm{IQR}_p(X) = \mathrm{VaR}_p(X) + \mathrm{VaR}_p(-X). \tag{2.6}$$

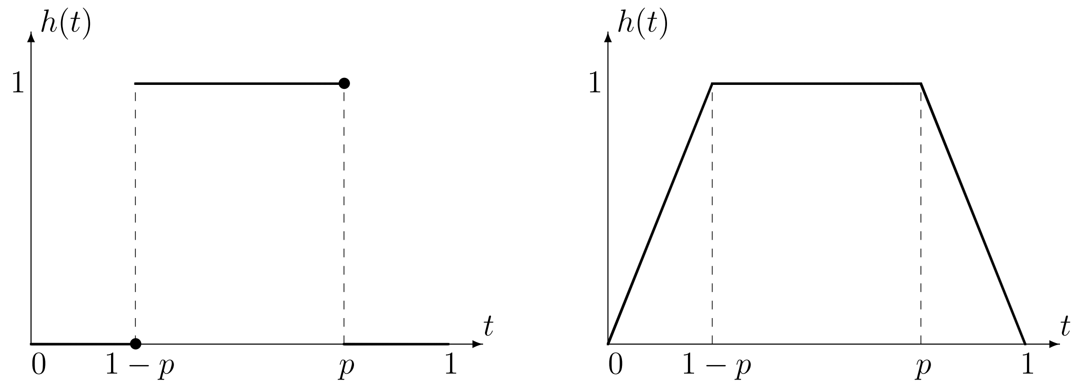

Figure 2.1: Distortion functions of IQR_p (left) and IER_p (right)

(ii) The inter-ES range (IER) for $p \in (1/2, 1)$:

$$\mathrm{IER}_p(X) = \frac{1}{1-p}\left(\int_p^1 \mathrm{VaR}_t(X)\,\mathrm{d}t - \int_0^{1-p} \mathrm{VaR}_t(X)\,\mathrm{d}t\right), \quad X \in L^\infty.$$

Similarly to (2.6), we can write, without assuming a continuous quantile,

$$\mathrm{IER}_p(X) = \mathrm{ES}_p(X) + \mathrm{ES}_p(-X), \quad X \in L^\infty.$$

IER_p for $p \in (1/2, 1)$ is a signed Choquet integral with distortion function h given by $h(t) = \min\{\frac{t}{1-p}, 1\} + \min\{\frac{p-t}{1-p}, 0\}$, $t \in [0, 1]$; see Figure 2.1. In sharp contrast to IQR_p, IER_p has a concave distortion function and hence it is convex and convex-order consistent.

2.3 Continuity

In this section, we discuss some issues related to continuity of signed Choquet integrals. We first demonstrate the simple fact that a signed Choquet integral is Lipschitz-continuous with respect to L^∞-norm. This result completes the proof of Theorem 2.2.1 above, and will be used later in Section 2.4 to study risk aggregation. This result can be derived (with a small effort) from Proposition 4.11 of Marinacci and Montrucchio (2004) on the continuity of signed Choquet integrals without law-invariance. A simple self-contained proof is put in the appendix.

Lemma 2.3.1. *For $h \in \mathcal{H}$ and $X, Y \in L^\infty$,*

$$|I_h(X) - I_h(Y)| \leqslant \mathrm{TV}_h \|X - Y\|_\infty, \tag{2.7}$$

where TV_h is the total variation of h on $[0,1]$.

Next we study continuity with respect to convergence in distribution (equivalently, weak convergence in the set of distributions $\mathcal{M}$). In general, a signed Choquet integral is not necessarily continuous with respect to convergence in distribution in L^∞, a well-known property of L-statistics; see Cont et al. (2010) for a discussion on increasing Choquet integrals (termed distortion risk measures) in risk management.

In risk management practice, convergence in distribution is the most common type of convergence, due to the statistical nature of data analysis and simulation studies. This issue is closely related to the notion of *qualitative robustness* of statistical functionals as pioneered by Hampel (1971). It would then be of interest to study under what extra conditions a risk functional can be *robust*, thus continuous with respect to convergence in distribution. This direction of research is explored by Embrechts et al. (2015), Pesenti et al. (2016) and Krätschmer et al. (2017).

The following uniform integrability condition turns out to be relevant. A set $\mathcal{D} \subset L^\infty$ is *h-uniformly integrable* for $h \in \mathcal{H}$, if

$$\lim_{k \downarrow 0} \sup_{X \in \mathcal{D}} \int_0^k \left| F_X^{-1}(1-t) \right| \mathrm{d}h(t) = 0, \tag{2.8}$$

and

$$\lim_{k \uparrow 1} \sup_{X \in \mathcal{D}} \int_k^1 \left| F_X^{-1}(1-t) \right| \mathrm{d}h(t) = 0. \tag{2.9}$$

Note that if $h \in \mathcal{H}$ is linear and non-constant in some neighborhoods of 0 and 1, then h-uniform integrability reduces to the usual uniform integrability.

Theorem 2.3.2. *For $h \in \mathcal{H}$ and $X, X_1, X_2, \dots \in L^\infty$, assume that $X_n \to X$ in distribution as $n \to \infty$ and $\{X, X_1, X_2, \dots\}$ is h-uniformly integrable. If (i) h is continuous, or (ii) X has a continuous inverse distribution function, then $I_h(X_n) \to I_h(X)$ as $n \to \infty$.*

Proof. For $n \in \mathbb{N}$, let F_n and F be the distribution functions of X_n and X, respectively.

We first assume (i). By Lemma 2.2.4, we have

$$I_h(X_n) = \int_0^1 F_n^{-1}(1-p)\,\mathrm{d}h(p) \quad \text{and} \quad I_h(X) = \int_0^1 F^{-1}(1-p)\,\mathrm{d}h(p). \tag{2.10}$$

As h can be replaced by its Jordan decomposition $h = h_+ - h_-$, it suffices to show the statement for an increasing and continuous h. The increasing function h induces a finite Borel measure μ on $[0,1]$ via $\mu([0,x]) = h(x)$, $x \in [0,1]$. Since $F_n^{-1} \to F^{-1}$ as $n \to \infty$ almost everywhere on $\mathbb{R}$ and h is continuous, the convergence is also μ-almost surely. Moveover, the h-uniform integrability of $\{X_i\}_{i\in\mathbb{N}}$ implies that $\{F_n^{-1}\}_{n\in\mathbb{N}}$ is uniformly integrable with respect to the measure μ. Therefore, using Vitali's Convergence Theorem (Rudin (1987, p. 133)), we have $I_h(X_n) \to I_h(X)$ as $n \to \infty$.

Next we assume (ii). In this case the convergence $F_n^{-1} \to F^{-1}$ is point-wise on $(0,1)$. Suppose for the moment that h is left-continuous. By Lemma 2.2.4, (2.10) holds. Similarly to the case above, we assume that h is increasing, and it induces a finite Borel measure μ on $[0,1]$ via $\mu([0,x)) = h(x)$, $x \in [0,1]$. Note that the h-uniform integrability of $\{X, X_1, X_2, \dots\}$ implies that, if $\mu(\{0\}) > 0$, then $F_n^{-1}(1) \to 0$ and $F^{-1}(1) = 0$. Analogously, if $\mu(\{1\}) > 0$, then $F_n^{-1}(0) \to 0$ and $F^{-1}(0) = 0$. Combining the above facts, $F_n^{-1} \to F^{-1}$ μ-almost surely as $n \to \infty$. Using Vitali's Convergence Theorem, we have $I_h(X_n) \to I_h(X)$ as $n \to \infty$.

If h is right-continuous, define the Borel measure μ on $[0,1]$ via $\mu([0,x]) = h(x)$, $x \in [0,1]$ and use the representation in Lemma 2.2.4 (i). The conclusion follows analogously.

Finally, for a general h, we decompose $h = h_l + h_r$, where h_l and h_r are left-continuous and right-continuous, respectively. Then we have

$$|I_h(X_n) - I_h(X)| \leqslant |I_{h_l}(X_n) - I_{h_l}(X)| + |I_{h_r}(X_n) - I_{h_r}(X)| \to 0, \quad \text{as } n \to \infty.$$

The proof is complete. ☐

Next we present a condition on h which implies the h-uniform integrability for all random variables. We say that $h \in \mathcal{H}$ is *flat in neighborhoods of* 0 *and* 1 if it satisfies the following condition: if there exists some $\delta > 0$ such that for all $0 < \varepsilon < \delta$, $h(\varepsilon) = h(0)$ and $h(1-\varepsilon) = h(1)$. In this case, clearly, any set of random variables is h-uniformly integrable. This condition is satisfied if, for instance, I_h is a finite linear combination of some quantile functionals.

Corollary 2.3.3. *For $h \in \mathcal{H}$ and $X, X_1, X_2, \dots \in L^\infty$, assume that $X_n \to X$ in distribution as $n \to \infty$ and h is flat in neighborhoods of 0 and 1. If (i) h is continuous, or (ii) X has a continuous inverse distribution function, then $I_h(X_n) \to I_h(X)$ as $n \to \infty$.*

Robustness properties of the two popular classes of risk measures VaR and ES (defined in Section 2.2.6) are well-studied in the literature. With respect to convergence in distribution, it is known that VaR_p is continuous at random variables with a continuous quantile function, and ES_p is continuous at random variables among a uniformly integrable set. These are special cases of Theorem 2.3.2 and Corollary 2.3.3.

Theorem 2.3.2 and Corollary 2.3.3 generalize Theorem 1 of Cont et al. (2010) for distortion risk measures, Theorem 2.5 of Embrechts et al. (2015) on robustness in the set of risk aggregation, and Theorem 3.5 of Pesenti et al. (2016) for finite-valued convex risk measures. Moreover, different from the settings of Krätschmer et al. (2014, 2017) and Pesenti et al. (2016), our results do not rely on any convexity assumptions.

Example 2.3.1 (Continuity of measures of distributional variability)**.** As mentioned above, continuity of risk measures with respect to convergence in distribution are well studied in the recent risk management literature. Below, we apply Theorem 2.3.2 and Corollary 2.3.3 to the measures of variability in Examples 2.2.1, 2.2.3 and 2.2.4.

(i) The range is generally not continuous with respect to convergence in distribution. Note that for the distortion function h of the range, given by $h(t) = \mathbf{1}_{\{0<t<1\}},\ t \in [0,1]$, the h-uniform integrality condition in (2.8) and (2.9) implies $X = 0$ a.s. for $X \in \mathcal{D}$, which is very restrictive.

(ii) The mean median difference has a continuous distortion function h given by $h(t) = \min\{t, 1-t\}$, $t \in [0,1]$. Since h is linear in neighbourhoods of 0 and 1, the h-uniform integrability is equivalent to the usual uniform integrability. Hence, by Theorem 2.3.2, the mean median difference is continuous with respect to convergence in distribution over any uniformly integrable set.

(iii) The Gini deviation has a continuous distortion function h given by $h(t) = t - t^2,\ t \in [0,1]$. For this distortion function h, as it has non-zero (one-sided) derivatives at 0 and 1, the h-uniform integrability is equivalent to the usual uniform integrability. Therefore, by Theorem 2.3.2, the Gini deviation is also continuous with respect to convergence in distribution over any uniformly integrable set.

(iv) The inter-quantile range for $p \in (1/2, 1)$ has a distortion function h given by $h(t) = \mathbf{1}_{\{1-p<t\leqslant p\}},\ t \in [0,1]$. Note that h is flat in neighborhoods of 0 and 1, but it is not continuous. Hence, by Corollary 2.3.3, the inter-quantile range is continuous with respect to convergence in distribution over any set of random variables with continuous quantile functions.

(v) The inter-ES range for $p \in (1/2, 1)$ has a distortion function h given by $h(t) = \min\{\frac{t}{1-p}, 1\} + \min\{\frac{p-t}{1-p}, 0\},\ t \in [0,1]$. Since h is linear in neighbourhoods of 0 and

1, the h-uniform integrability is equivalent to the usual uniform integrability. Hence, by Theorem 2.3.2, the inter-ES range is continuous with respect to convergence in distribution over any uniformly integrable set.

(vi) The standard deviation is continuous with respect to convergence in distribution over any uniformly square-integrable set. One can show this statement by applying Vitali's Convergence Theorem to the first and second moments. Theorem 2.3.2 does not directly lead to this statement. Nevertheless, in Example 2.2.3 we have seen $\sigma(X) = \sup_{h\in\tilde{\mathcal{H}}} I_h(X),\ X \in L^\infty$. By Hölder's inequality, uniform square-integrability implies h-uniform integrability for each $h \in \tilde{\mathcal{H}}$. Hence, Theorem 2.3.2 implies that I_h is continuous with respect to convergence in distribution over any uniformly square-integrable set.

Remark 2.3.1. The continuity results of signed Choquet integrals in Theorem 2.3.2 also hold on a set larger than L^∞, as long as I_h is well-defined on the corresponding set. In this chapter, due to the limitation of space, we focus on random variables in L^∞. For robustness properties of risk functionals defined on Orlicz hearts, we refer to Krätschmer et al. (2014, 2017).

2.4 Risk aggregation under uncertainty

In the literature of risk management, risk aggregation concerns quantities related to the sum $S = X_1 + \cdots + X_n$ (e.g. the distribution or a risk measure of S) of a risk vector $(X_1, \ldots, X_n)$ representing random losses from a certain portfolio. A currently popular direction of research is *risk aggregation with dependence uncertainty*, where for each $i = 1, \ldots, n$, the marginal distribution F_i of X_i, is known while the joint distribution of $(X_1, ..., X_n)$ remains unspecified. We refer to Embrechts et al. (2013, 2014) and Wang et al. (2013) for the case of the risk measure VaR (defined in Section 2.2.6), Bernard et al. (2017a,b) for some recent development, and Section 8.4 of McNeil et al. (2015) for a general discussion. As the precise distribution of S is unknown, one typically studies the worst-case value of the aggregate risk evaluated by a risk measure ρ, that is,

$$\sup\{\rho(X_1 + \cdots + X_n) : X_i \sim F_i,\ i = 1, ..., n\}. \tag{2.11}$$

Another important quantity related to portfolio diversification is the worst-case diversification ratio, defined as

$$\sup\left\{\frac{\rho(X_1 + \cdots + X_n)}{\rho(X_1) + \cdots + \rho(X_n)} : X_i \sim F_i,\ i = 1, ..., n\right\}. \tag{2.12}$$

If the functional ρ is not convex, the quantities in (2.11)-(2.12) are generally difficult to analytically compute. In this section we investigate them for signed Choquet integrals.

2.4.1 Homogeneous portfolios and the extreme-aggregation measure

To investigate the asymptotic behaviour of the values in (2.11)-(2.12) for homogeneous portfolios, Wang et al. (2015) introduced the extreme-aggregation measure as follows. Denote the set of possible sums of n F-distributed random variables by $\mathcal{S}_n(F) = \{X_1 + \cdots + X_n : X_i \sim F,\ i = 1, ..., n\},\ n \in \mathbb{N}$.

Definition 2.4.1. The *extreme-aggregation measure* Γ_ρ *induced by a law-invariant functional* $\rho : L^\infty \to \mathbb{R}$ is defined as

$$\Gamma_\rho : L^\infty \to (-\infty, \infty],\ \Gamma_\rho(X) = \limsup_{n\to\infty} \left\{ \frac{1}{n} \sup\{\rho(S) : S \in \mathcal{S}_n(F_X)\} \right\}.$$

Γ_ρ provides a limit of (2.11)-(2.12) for homogenous portfolios. The VaR-ES relation $\Gamma_{\mathrm{VaR}_p} = \mathrm{ES}_p$ for $p \in (0, 1)$ is shown in Wang and Wang (2015) via direct construction; some first special cases of this relation are established by Puccetti and Rüschendorf (2014). Generalizations to inhomogeneous portfolios are given in Embrechts et al. (2015) (VaR and ES) and Cai et al. (2017) (distortion risk measures and convex risk measures). For a distortion risk measure (equivalently, an increasing Choquet integral) ρ, Wang et al. (2015) obtained an explicit expression for Γ_ρ, which is the smallest subadditive distortion risk measure dominating ρ. The proof used in Wang et al. (2015) is based on analyzing the precise form of h, which requires a lot of delicate analysis and random variable construction. Below we give a much more concise proof, generalizing the characterization of Γ_ρ to signed Choquet integrals.

To present our main result, for $h \in \mathcal{H}$, define its *concave envelope*

$$h^*(t) = \inf\{g(t) : g \text{ is a concave function on } [0, 1] \text{ and } g \geqslant h\}, \quad t \in [0, 1]. \tag{2.13}$$

Note that calculating h^* for a given $h \in \mathcal{H}$ is equivalent to finding the convex hull of the set $\{(x, y) \in [0, 1] \times \mathbb{R} : h(x) \geqslant y\}$.

It is clear that h^* is concave as it is an infimum of concave functions. Further, $h^*(0) = h(0) = 0$ and $h^*(1) = h(1)$; to see this, as $h \in \mathcal{H}$ is bounded, we can define a concave function $g:[0,1]\to\mathbb{R}$ as

$$g(t)=\begin{cases} 0 & t=0\\ \sup_{t\in[0,1]} h(t) & 0<t<1\\ h(1) & t=1\end{cases}$$

Then one has $0=h(0)\leqslant h^*(0)\leqslant g(0)=0$ and $h(1)\leqslant h^*(1)\leqslant g(1)=h(1)$.

Our main result is the following theorem, which generalizes Theorem 3.2 of Wang et al. (2015) for increasing Choquet integrals.

Theorem 2.4.1. *For $h\in\mathcal{H}$, the extreme-aggregation measure induced by I_h is I_{h^*}, and it is the smallest law-invariant convex functional on L^∞ dominating I_h.*

The key to our proof of Theorem 2.4.1 is to show that the law-invariant functional Γ_{I_h} is comonotonic-additive and uniformly norm-continuous, and from there we can rely on Theorem 2.2.1 to justify that it is a signed Choquet integral. We first demonstrate some useful facts built on several results in Mao and Wang (2015). Denote for a distribution function $F\in\mathcal{M}$,

$$\mathcal{B}_n(F)=\left\{\frac{1}{n}(X_1+\cdots+X_n):X_i\sim F,\ i=1,...,n\right\},$$

and

$$\mathcal{C}(F)=\left\{X:X\leqslant_{\mathrm{cx}}Y,\ \text{where } Y\sim F\right\}.$$

Lemma 2.4.2. *For $h\in\mathcal{H}$, the following statements hold.*

(i) For $F\in\mathcal{M}$,

$$\limsup_{n\to\infty}\left\{\sup\{I_h(X):X\in\mathcal{B}_n(F)\}\right\}=\sup\{I_h(X):X\in\mathcal{C}(F)\}. \tag{2.14}$$

(ii) The functional on L^∞, $X\mapsto\sup_{Y\in\mathcal{C}(F_X)}I_h(Y)$ is comonotonic-additive and convex order consistent.

Proof. (i) Lemma 3.4 of Mao and Wang (2015) (which can be seen as a special case of Lemma 1 of O'Cinneide (1991)) states $\mathcal{B}_n(F)\subset\mathcal{C}(F)$ for $n\in\mathbb{N}$. Therefore

$$\limsup_{n\to\infty}\left\{\sup\{I_h(T):T\in\mathcal{B}_n(F)\}\right\}\leqslant\sup\{I_h(T):T\in\mathcal{C}(F)\}. \tag{2.15}$$

On the other hand, by Proposition 3.6 of Mao and Wang (2015), $\overline{\limsup_{n\to\infty} \mathcal{B}_n(F)}^* = \mathcal{C}(F)$, where $\overline{B}^*$ is the L^∞-closure of a set B. It follows that, for each $Y \in \mathcal{C}(F)$, $\varepsilon > 0$ and $m \in \mathbb{N}$, there exists $k \in \mathbb{N}$ and $X_k \in \mathcal{B}_k(F)$ such that $||X_k - Y||_\infty < \varepsilon$. Hence, by Lemma 2.3.1,

$$\sup\{I_h(X) : X \in \mathcal{B}_k(F)\} \geqslant \sup\{I_h(X) : X \in \mathcal{C}(F)\} - \varepsilon \mathrm{TV}_h.$$

Therefore,

$$\limsup_{n\to\infty} \{\sup\{I_h(X) : X \in \mathcal{B}_n(F)\}\} \geqslant \sup\{I_h(X) : X \in \mathcal{C}(F)\} - \varepsilon \mathrm{TV}_h.$$

As ε is arbitrary, we obtain

$$\limsup_{n\to\infty} \{\sup\{I_h(X) : X \in \mathcal{B}_n(F)\}\} \geqslant \sup\{I_h(X) : X \in \mathcal{C}(F)\}. \tag{2.16}$$

Combining (2.15)-(2.16), we obtain (2.14).

(ii) Corollary 4.3 of Mao and Wang (2015) states that the functional $X \mapsto \sup_{Y\in\mathcal{C}(F_X)} \rho(Y)$ is comonotonic-additive and convex order consistent if ρ is comonotonic-additive, which is the case if $\rho = I_h$. □

Proof of Theorem 2.4.1. For any $X \in L^\infty$, by positive homogeneity of I_h and the definition of $\mathcal{B}_n(F_X)$, we have

$$\begin{aligned}\Gamma_{I_h}(X) &= \limsup_{n\to\infty} \left\{ \sup \left\{ I_h\left(\frac{1}{n}S\right) : S \in \mathcal{S}_n(F_X) \right\} \right\} \\ &= \limsup_{n\to\infty} \{\sup\{I_h(T) : T \in \mathcal{B}_n(F_X)\}\}.\end{aligned}$$

Applying Lemma 2.4.2 (i), it is

$$\Gamma_{I_h}(X) = \sup\{I_h(T) : T \in \mathcal{C}(F_X)\}, \quad X \in L^\infty.$$

By Lemma 2.4.2 (ii), Γ_{I_h} is comonotonic-additive and convex order consistent.

Next we verify that Γ_{I_h} is uniformly norm-continuous. Fix $n \in \mathbb{N}$. For any $S \in \mathcal{S}_n(F_X)$, write $S = X_1 + \cdots + X_n$, where $X_i \sim F_X,\ i = 1, \ldots, n$. Let $U_1, \ldots, U_n$ be uniform random variables on $[0, 1]$ such that $F_X^{-1}(U_i) = X_i$ almost surely, $i = 1, \ldots, n$. The existence of such $U_1, \dots, U_n$ is given by, for instance, Lemma A.32 of Föllmer and Schied (2016). Let $Z = F_Y^{-1}(U_1) + \cdots + F_Y^{-1}(U_n)$. Clearly, $Z \in \mathcal{S}_n(F_Y)$. By Lemma 2.3.1,

$$\begin{aligned}
&I_h(S) - \sup\{I_h(T) : T \in \mathcal{S}_n(F_Y)\} \\
&\leqslant I_h(S) - I_h(Z) \\
&\leqslant \mathrm{TV}_h \|S - Z\|_\infty \\
&= \mathrm{TV}_h \left\| \left(F_X^{-1}(U_1) + \cdots + F_X^{-1}(U_n)\right) - \left(F_Y^{-1}(U_1) + \cdots + F_Y^{-1}(U_n)\right) \right\|_\infty \\
&\leqslant n \mathrm{TV}_h \|X - Y\|_\infty,
\end{aligned}$$

where the last inequality is due to the well-known fact that $\|F_X^{-1}(U_1) - F_Y^{-1}(U_1)\|_\infty \leqslant \|X - Y\|_\infty$ (see e.g. Lemma 8.2 of Bickel and Freedman (1981)). It follows from taking a supremum over $S \in \mathcal{S}_n(F_X)$ that

$$\frac{1}{n} \sup\{I_h(S) : S \in \mathcal{S}_n(F_X)\} \leqslant \frac{1}{n} \sup\{I_h(T) : T \in \mathcal{S}_n(F_Y)\} + \mathrm{TV}_h \|X - Y\|_\infty.$$

Therefore,

$$\Gamma_{I_h}(X) \leqslant \Gamma_{I_h}(Y) + \mathrm{TV}_h \|X - Y\|_\infty.$$

By symmetry, we have $|\Gamma_{I_h}(X) - \Gamma_{I_h}(Y)| \leqslant \mathrm{TV}_h \|X - Y\|_\infty$; thus Γ_{I_h} is uniformly norm-continuous.

At this point, we know that the law-invariant functional Γ_{I_h} is norm-continuous, comonotonic-additive and convex order consistent. By Theorem 2.2.1, there exists $g \in \mathcal{H}$ such that Γ_{I_h} is identified with a signed Choquet integral $I_g = \Gamma_{I_h}$. Note that

$$I_g(X) = \sup_{T \in \mathcal{C}(F_X)} I_h(T) \geqslant I_h(X),$$

and therefore $g \geqslant h$ by Lemma 2.2.2. Since Γ_{I_h} is convex order consistent, g is concave by Theorem 2.2.6. From the definition of h^*, $h^* \leqslant g$, and this implies $I_{h^*} \leqslant I_g$ by Lemma 2.2.2 again.

On the other hand, $h^* \geqslant h$, and hence $I_{h^*} \geqslant I_h$. Noting that h^* is concave and thus I_{h^*} is also convex order consistent, we have

$$I_{h^*}(X) = \sup_{T \in \mathcal{C}(F_X)} I_{h^*}(T) \geqslant \sup_{T \in \mathcal{C}(F_X)} I_h(T) = I_g(X)$$

Therefore, we conclude that $I_{h^*} = I_g = \Gamma_{I_h}$.

Finally we show that I_{h^*} is the smallest law-invariant convex functional on L^∞ dominating I_h. Suppose that $I : L^\infty \to \mathbb{R}$ is a law-invariant convex functional and $I \geqslant I_h$. For any $n \in \mathbb{N}$ and $X \in L^\infty$,

$$\begin{aligned}\sup\{I_h(T) : T \in \mathcal{B}_n(F_X)\} &\leqslant \sup\{I(T) : T \in \mathcal{B}_n(F_X)\} \\ &\leqslant \sup\left\{\frac{1}{n}I(X_1) + \cdots + \frac{1}{n}I(X_n) : X_i \sim F_X, 1 \leqslant i \leqslant n\right\} \\ &= \frac{1}{n}nI(X) = I(X).\end{aligned}$$

By taking a limit on both sides of the above equation, we conclude that $I_{h^*} \leqslant I$. Thus, I_{h^*} is the smallest law-invariant convex functional dominating I_h. □

Example 2.4.1. Theorem 2.4.1 implies two well-known facts in the literature of risk measures on the relation between VaR_p and ES_p for $p \in (0, 1)$: First, the worst-case aggregation of VaR_p is asymptotically equivalent to that of ES_p (Corollary 3.7 of Wang and Wang (2015)). Second, ES_p is the smallest law-invariant convex risk measure dominating VaR_p (Theorem 9 of Kusuoka (2001); see also Theorem 4.67 of Föllmer and Schied (2016)). Theorem 2.4.1 generalizes these results to all signed Choquet integrals, and our approach is different from those in the literature.

Example 2.4.2 (The inter-quantile range and inter-ES range)**.** In Examples 2.2.4 and 2.3.1 we have already seen many differences between the two measures of variability IQR_p and IER_p in terms of convexity, convex-order consistency, and continuity. Next, we will see, by applying Theorem 2.4.1, an interesting connection between the two signed Choquet integrals. Recall that for $p \in (1/2, 1)$, IQR_p has a (non-concave) distortion function h given by $h(t) = \mathbf{1}_{\{1-p<t\leqslant p\}}$, $t \in [0, 1]$. It is straightforward that the smallest concave function h^* dominating h is given by $h^*(t) = \min\{\frac{t}{1-p}, 1\} + \min\{\frac{p-t}{1-p}, 0\}$, $t \in [0, 1]$, which is the distortion function of IER_p; see Figure 2.1 for these distortion functions. Therefore, for $I_h = \mathrm{IQR}_p$, we have $I_{h^*} = \mathrm{IER}_p$. By Theorem 2.4.1, IER_p is the extreme-aggregation measure of IQR_p; in other words, the worst-case value of the aggregate risk evaluated by IQR_p is asymptotically equivalent to that evaluated by IER_p. This relationship will be further illustrated by the numerical example in Section 2.4.3.

Remark 2.4.1. Wang et al. (2015) gives a few conditions for the upper limit in the definition of Γ_ρ to be replaced by a supremum or a limit. If ρ is a positively homogeneous functional, then the upper limit can be replaced by a supremum. Furthermore, for $h \in \mathcal{H}$ and $X \in L^\infty$, the upper limit can be replaced by either a limit or a supremum, namely

$$\Gamma_{I_h}(X) = \lim_{n\to\infty}\left\{\frac{1}{n}\sup\{I_h(S) : S \in \mathcal{S}_n(F_X)\}\right\} = \sup_{n\in\mathbb{N}}\left\{\frac{1}{n}\sup\{I_h(S) : S \in \mathcal{S}_n(F_X)\}\right\}.$$

We conclude this section by showing that the result in Theorem 2.4.1 can be generalized to suprema over a set of signed Choquet integrals. Indeed, suprema over a set of signed Choquet integrals represent all continuous law-invariant coherent risk measures and deviation measures, as established in Kusuoka (2001) and Grechuk et al. (2009), respectively.

Corollary 2.4.3. *Define a functional $\rho = \sup_{h\in\mathcal{H}_0} I_h$, where $\mathcal{H}_0 \subset \mathcal{H}$. Then $\Gamma_\rho = \sup_{h\in\mathcal{H}_0} I_{h^*}$ and Γ_ρ is the smallest law-invariant convex functional on L^∞ dominating ρ.*

Proof. For $X \in L^\infty$, as ρ is positively homogeneous, we can write

$$\Gamma_\rho(X) = \sup_{n\in\mathbb{N}}\{\sup\{\rho(T) : T \in \mathcal{B}_n(F_X)\}\};$$

see Remark 2.4.1. By exchanging the order of suprema, we have

$$\Gamma_\rho(X) = \sup_{n\in\mathbb{N}} \sup_{T\in\mathcal{B}_n(F_X)} \sup_{h\in\mathcal{H}_0} I_h(T) = \sup_{h\in\mathcal{H}_0} \sup_{n\in\mathbb{N}} \sup_{T\in\mathcal{B}_n(F_X)} I_h(T) = \sup_{h\in\mathcal{H}_0} \Gamma_{I_h}(X) = \sup_{h\in\mathcal{H}_0} I_{h^*}(X),$$

where the last equality is due to Theorem 2.4.1. The last statement of the corollary can be obtained via an argument analogous to the last part of the proof of Theorem 2.4.1. □

2.4.2 Heterogeneous portfolios

For a sequence of distribution functions $\{F_i\}_{i\in\mathbb{N}} \subset \mathcal{M}$, denote the set of possible sums of n random variables with respective distributions by

$$\mathcal{S}_n(F_1, \ldots, F_n) = \{X_1 + \cdots + X_n : X_i \sim F_i,\ i = 1, ..., n\}, \quad n \in \mathbb{N}.$$

To investigate risk aggregation for heterogeneous portfolios, we study an asymptotic equivalence of the following type,

$$\lim_{n\to\infty} \frac{\sup\{I_h(S) : S \in \mathcal{S}_n(F_1, \ldots, F_n)\}}{\sup\{I_{h^*}(S) : S \in \mathcal{S}_n(F_1, \ldots, F_n)\}} = 1. \tag{2.17}$$

The asymptotic equivalence (2.17) is established in Theorem 3.5 of Cai et al. (2017) for increasing Choquet integrals under some regularity conditions. Interpreting (2.17), to evaluate large portfolios with dependence uncertainty via a non-convex functional I_h, one can replace I_h by a convex functional I_{h^*}, the extreme-aggregation measure induced by I_h, which is much easier to calculate due to its comonotonic-additivity and subadditivity. It is clear that if $F_1 = F_2 = \cdots = F_X$, then (2.17) reads as

$$\lim_{n\to\infty} \frac{\sup\{I_h(S) : S \in \mathcal{S}_n(F_X)\}}{nI_{h^*}(X)} = 1,$$

which is precisely Theorem 2.4.1 if $I_{h^*}(X) \neq 0$. For the same relation to hold for heterogeneous portfolios, one needs some regularity conditions on $\{F_i\}_{i\in\mathbb{N}}$ and $h \in \mathcal{H}$.

Condition C1 (non-vanishing). $\lim_{n\to\infty} |\sum_{i=1}^n I_{h^*}(X_i)| = \infty$, where $X_i \sim F_i$, $i \in \mathbb{N}$.

Condition C2 (bounded ranges). $\sup_{i\in\mathbb{N}}\{F_i^{-1}(1) - F_i^{-1}(0)\} < \infty$.

Sections 2.2 and 3.3 of Cai et al. (2017) contain counter-examples where (2.17) fails to hold without some regularity conditions. Next we present the asymptotic equivalence for signed Choquet integrals with a continuous distortion function h under Conditions C1-C2.

Theorem 2.4.4. *For a continuous $h \in \mathcal{H}$ and $\{F_i\}_{i\in\mathbb{N}} \subset \mathcal{M}$ satisfying Conditions C1-C2, we have*

$$\lim_{n\to\infty} \frac{\sup\{I_h(S) : S \in \mathcal{S}_n(F_1, \dots, F_n)\}}{\sup\{I_{h^*}(S) : S \in \mathcal{S}_n(F_1, \dots, F_n)\}} = 1. \tag{2.18}$$

Proof. Our proof is similar to that of Theorem 3.5 of Cai et al. (2017), although the conditions in the latter result are different from Conditions C1-C2. Since h is continuous, we directly work with the quantile representation in Lemma 2.2.4 (ii). By Lemma 5.1 of Brighi and Chipot (1994), there exist disjoint open intervals (a_k, b_k), $k \in K \subset \mathbb{N}$ on which $h \neq h^*$, and h^* is linear on each of $[a_k, b_k]$, $k \in K$. Define $A_k = (1 - b_k, 1 - a_k)$, $k \in K$. Let U, V be independent U[0, 1] random variables, and

$$S_n^c = F_1^{-1}(U) + \cdots + F_n^{-1}(U),$$

and

$$R_n = \begin{cases} F_1^{-1}(U) + \cdots + F_n^{-1}(U), & \text{if } U \notin \cup_{k\in K} A_k, \\ \mathbb{E}\left[F_1^{-1}(U) + \cdots + F_n^{-1}(U) \mid U \in A_k\right], & \text{if } U \in A_k,\ k \in K. \end{cases}$$

Clearly, $F_i^{-1}(U) \sim F_i$, $i = 1, \dots, n$, and hence $S_n^c \in \mathcal{S}_n(F_1, \dots, F_n)$. Since

$$\mathbb{E}\left[F_i^{-1}(U) \mid U \in A_k\right] = \frac{\int_{(a_k,b_k)} F_i^{-1}(1-t)\mathrm{d}t}{b_k - a_k} \quad \text{and} \quad F_{S_n^c}^{-1}(t) = \sum_{i=1}^n F_i^{-1}(t) \quad \text{for } t \in (0,1),$$

we have

$$\begin{aligned} &\int_{(a_k,b_k)} F_{S_n^c}^{-1}(1-t)\mathrm{d}h^*(t) - \int_{(a_k,b_k)} F_{R_n}^{-1}(1-t)\mathrm{d}h^*(t) \\ &= \frac{h^*(b_k) - h^*(a_k)}{b_k - a_k} \sum_{i=1}^n \int_{(a_k,b_k)} F_i^{-1}(1-t)\mathrm{d}t - \sum_{i=1}^n \frac{\int_{(a_k,b_k)} F_i^{-1}(1-t)\mathrm{d}t}{b_k - a_k} \int_{(a_k,b_k)} \mathrm{d}h^*(t) = 0. \end{aligned}$$

It follows that

$$I_{h^*}(S_n^c) - I_{h^*}(R_n) = \int_0^1 F_{S_n^c}^{-1}(1-t)\mathrm{d}h^*(t) - \int_0^1 F_{R_n}^{-1}(1-t)\mathrm{d}h^*(t)$$
$$= \sum_{k\in K}\left[\int_{a_k}^{b_k} F_{S_n^c}^{-1}(1-t)\mathrm{d}h^*(t) - \int_{a_k}^{b_k} F_{R_n}^{-1}(1-t)\mathrm{d}h^*(t)\right] = 0. \tag{2.19}$$

By Corollary A.3 of Embrechts et al. (2015), for each k, we can find random variables $Y_{1k},\dots,Y_{nk}$, independent of U (guaranteed by the existence of V), such that Y_{ik} is identically distributed as $F_i^{-1}(U)|U\in A_k$, $i=1,\dots,n$, and

$$\left|Y_{1k}+\cdots+Y_{nk} - \mathbb{E}\left[F_1^{-1}(U)+\cdots+F_n^{-1}(U) \mid U\in A_k\right]\right|$$
$$\leqslant \max_{i=1,\dots,n}\{F_i^{-1}(1-a_k) - F_i^{-1}(1-b_k)\}.$$

Let $X_i^* = F_i^{-1}(U)1_{\{U\notin\cup_{k\in K}A_k\}} + \sum_{k\in K} Y_{ik}1_{\{U\in A_k\}}$, $i=1,\dots,n$. It is easy to check that $X_i^*\sim F_i$, $i=1,\dots,n$. Denote by $S_n^* = X_1^*+\cdots+X_n^*$. Clearly, $S_n^*\in\mathcal{S}_n(F_1,\dots,F_n)$ and by definition

$$|R_n - S_n^*| \leqslant \sup_{k\in K}\max_{i=1,\dots,n}\{F_i^{-1}(1-a_k) - F_i^{-1}(1-b_k)\} \leqslant M,$$

where $M=\sup_{i\in\mathbb{N}}\{F_i^{-1}(1)-F_i^{-1}(0)\}$ and $M<\infty$ by Condition C2. Therefore, by Lemma 2.3.1, we have

$$|I_h(R_n) - I_h(S_n^*)| \leqslant \mathrm{TV}_h \times M. \tag{2.20}$$

Integration by parts yields

$$I_{h^*}(R_n) - I_h(R_n) = \int_0^1 F_{R_n}^{-1}(1-t)\mathrm{d}h^*(t) - \int_0^1 F_{R_n}^{-1}(1-t)\mathrm{d}h(t)$$
$$= \int_0^1 (h(t)-h^*(t))\,\mathrm{d}F_{R_n}^{-1}(1-t)$$
$$= \sum_{k\in K}\int_{(a_k,b_k)} (h(t)-h^*(t))\,\mathrm{d}F_{R_n}^{-1}(1-t) = 0, \tag{2.21}$$

where the last equality follows as $F_{R_n}^{-1}(1-t)$ is constant for t in each (a_k,b_k). Combining (2.19)-(2.21), we have

$$|I_{h^*}(S_n^c) - I_h(S_n^*)| = |(I_{h^*}(S_n^c) - I_{h^*}(R_n)) + (I_{h^*}(R_n) - I_h(R_n)) + (I_h(R_n) - I_h(S_n^*))|$$
$$= |0+0+(I_h(R_n) - I_h(S_n^*))| = \mathrm{TV}_h\times M. \tag{2.22}$$

Since $h \leqslant h^*$, we have

$$\sup\{I_h(S) : S \in \mathcal{S}_n(F_1, \dots, F_n)\} \leqslant \sup\{I_{h^*}(S) : S \in \mathcal{S}_n(F_1, \dots, F_n)\} = I_{h^*}(S_n^c)$$

and hence (2.22) implies $|I_{h^*}(S_n^c) - \sup\{I_h(S) : S \in \mathcal{S}_n(F_1, \dots, F_n)\}| \leqslant \mathrm{TV}_h \times M$. By Condition C1, $\lim_{n\to\infty} |I_{h^*}(S_n^c)| = \infty$. Therefore, as $n \to \infty$,

$$\left| \frac{\sup\{I_h(S) : S \in \mathcal{S}_n(F_1, \dots, F_n)\}}{\sup\{I_{h^*}(S) : S \in \mathcal{S}_n(F_1, \dots, F_n)\}} - 1 \right| \leqslant \frac{\mathrm{TV}_h \times M}{|I_{h^*}(S_n^c)|} \to 0.$$

The desired result follows. □

Remark 2.4.2. We can compare Theorem 2.4.4 with Theorem 3.5 of Cai et al. (2017), and there are several major differences on the assumptions. First, the latter result is about increasing Choquet integrals. Second, random variables are non-negative in Cai et al. (2017) because they focus on random losses and risk measures. For signed Choquet integrals, non-negativity seems irrelevant, and we assume instead that the random variables have a bounded sequence of ranges. Third, our Condition C1 is weaker than their Condition A1, and our Condition C2 is stronger than their Condition A2. Fourth, we assume h to be continuous for technical convenience.

2.4.3 Numerical illustration

In this section, we present numerical examples of risk aggregation under dependence uncertainty for the inter-quantile range and the inter-ES range. As explained in Example 2.4.2, IQR_p and IER_p are asymptotically equivalent in terms of the worst-case risk aggregation under dependence uncertainty; this also holds for inhomogeneous portfolios as implied by Theorem 2.4.4. Although we work with bounded random variables throughout the chapter to establish the theoretical results, the numerical examples in this section are built for unbounded risks to be more realistic for risk management practice. As we shall see below, the results of Theorem 2.4.4 are numerically valid for unbounded risks although they do not satisfy Condition C2.

We consider the following three representative models studied in Embrechts et al. (2015). The portfolios in Models (A) and (B) are inhomogeneous whereas the portfolio in Model (C) is homogeneous and very heavy-tailed.

(A) (Mixed portfolio) $F_i = \text{Pareto}(2 + 0.1i)$, $i = 1, \dots, 5$; $F_i = \text{Exp}(i - 5)$, $i = 6, \dots, 10$; $F_i = \text{Log-Normal}(0, (0.1(i - 10))^2)$, $i = 11, \dots, 20$.

(B) (Light-tailed portfolio) $F_i = \text{Exp}(i),\ i = 1, \ldots, 5$; $F_i = \text{Weibull}(i-5, 1/2)$, $i = 6, \ldots, 10$; $F_i = F_{i-10},\ i = 11, \ldots, 20$.

(C) (Very heavy-tailed portfolio) $F_i = \text{Pareto}(1.5)$, $i = 1, \cdots, 50$.

As the common choice of p for the inter-quantile range is 0.75, we compare the values of $\text{IQR}_{0.75}$ and $\text{IER}_{0.75}$ in each of the above models. We look at the influence on the number of risks in the portfolio ($n = 5, 10, 20$ for Models (A) and (B) and $n = 5, 10, 20, 50$ for Model (C)), and on different dependence structures. We report the following quantities for the sum of random variables $X_i \sim F_i$, $i = 1, \ldots, n$.

(i) $\text{IQR}_{0.75}(S_n^{\perp})$: $S_n^{\perp} = \sum_{i=1}^n X_i$ and we assume $X_1, \ldots, X_n$ are independent.

(ii) $\text{IER}_{0.75}(S_n^{\perp})$: $S_n^{\perp}$ is same as in (i).

(iii) $\text{IQR}_{0.75}(S_n^c)$: $S_n^c = \sum_{i=1}^n X_i$ and we assume $X_1, \ldots, X_n$ are comonotonic. By comonotonic-additivity, $\text{IQR}_{0.75}(S_n^c) = \sum_{i=1}^n \text{IQR}_{0.75}(X_i)$.

(iv) $\overline{\text{IQR}}_{0.75}(S_n)$: the worst-case value of $\text{IQR}_{0.75}(S)$ over $S \in \mathcal{S}_n(F_1, \ldots, F_n)$.

(v) $\overline{\text{IER}}_{0.75}(S_n)$: the worst-case value of $\text{IER}_{0.75}(S)$ over $S \in \mathcal{S}_n(F_1, \ldots, F_n)$. By comonotonic-additivity and subadditivity, $\overline{\text{IER}}_{0.75}(S_n) = \text{IER}_{0.75}(S_n^c)$ $= \sum_{i=1}^n \text{IER}_{0.75}(X_i)$.

(vi) $\frac{\overline{\text{IER}}_{0.75}(S_n)}{\overline{\text{IQR}}_{0.75}(S_n)}$: the ratio of the worst-case value of $\text{IER}_{0.75}(S)$ to that of $\text{IQR}_{0.75}(S)$.

The calculation for the independence model in (i) and (ii) is carried out through a Monte-Carlo simulation with sample size $N = 10^6$, and the marginal values in (iii) and (v) are carried out by analytical formulas. The numerical calculation of the worst-case value of $\text{IQR}_{0.75}$ in (iv) is carried out through the Rearrangement Algorithm (RA) of Embrechts et al. (2013) with tail discretization parameter $N = 10^6$ (`R` package: `QRM`)[1]. The numerical results are reported in Tables 2.1-2.2.

From Tables 2.1-2.2, we make the following observations.

[1]Although the RA is designed for the worst-case risk aggregation of VaR_p, it also works for IQR_p since $\text{VaR}_p(S)$ and $-\text{VaR}_{1-p}(S)$ can be simultaneously maximized over $S \in \mathcal{S}_n(F_1, \ldots, F_n)$; this is because the worst-case scenario for quantiles concerns only tail events; see e.g. Theorem 4.6 of Bernard et al. (2014).

Table 2.1: Numerical values for two inhomogeneous portfolio models

	Model (A)			Model (B)		
	$n=5$	$n=10$	$n=20$	$n=5$	$n=10$	$n=20$
$\mathrm{IQR}_{0.75}(S_n^\perp)$	2.7108	3.2664	5.3565	1.4649	1.6836	2.4712
$\mathrm{IER}_{0.75}(S_n^\perp)$	6.7298	7.5964	11.4492	2.9024	3.2967	4.7444
$\mathrm{IQR}_{0.75}(S_n^c)$	3.4939	6.0024	13.7364	2.5085	3.9262	7.8523
$\overline{\mathrm{IQR}}_{0.75}(S_n)$	9.7960	15.3198	33.3608	4.9976	7.8054	15.7044
$\overline{\mathrm{IER}}_{0.75}(S_n)$	11.0144	16.1504	33.8089	5.1360	7.8541	15.7082
$\frac{\overline{\mathrm{IER}}_{0.75}(S_n)}{\overline{\mathrm{IQR}}_{0.75}(S_n)}$	1.1243	1.0542	1.0134	1.0277	1.0062	1.0002

Table 2.2: Numerical values for a very heavy-tailed portfolio model

	Model (C)			
	$n=5$	$n=10$	$n=20$	$n=50$
$\mathrm{IQR}_{0.75}(S_n^\perp)$	6.4091	11.6276	20.4428	41.3704
$\mathrm{IER}_{0.75}(S_n^\perp)$	21.9388	37.3421	62.6127	121.5008
$\mathrm{IQR}_{0.75}(S_n^c)$	6.5421	13.0843	26.1686	65.4214
$\overline{\mathrm{IQR}}_{0.75}(S_n)$	22.6459	51.7782	111.7977	296.4094
$\overline{\mathrm{IER}}_{0.75}(S_n)$	32.3112	64.6225	129.2450	323.1125
$\frac{\overline{\mathrm{IER}}_{0.75}(S_n)}{\overline{\mathrm{IQR}}_{0.75}(S_n)}$	1.4268	1.2481	1.1561	1.0901

(i) In all models, $\overline{\mathrm{IQR}}_{0.75}(S_n)$ is much larger than $\mathrm{IQR}_{0.75}(S_n^c)$ and $\mathrm{IER}_{0.75}(S_n^{\perp})$. This suggests that neither independence or comonotonicity serves as a conservative benchmark when studying risk aggregation with dependence uncertainty for I_h with a non-concave h such as $I_h = \mathrm{IQR}_p$.

(ii) The ratio of $\overline{\mathrm{IER}}_{0.75}(S_n)$ to $\overline{\mathrm{IQR}}_{0.75}(S_n)$ goes to 1 as n grows for all models (for bounded risks this is shown in Theorem 2.4.4). The convergence is very fast for the light-tailed model (B) and relatively slow for the heavy-tailed model (C).

(iii) The values of $\overline{\mathrm{IER}}_{0.75}(S_n)$ and $\overline{\mathrm{IQR}}_{0.75}(S_n)$ are very close for the light-tailed model (B) even for small n such as $n = 5$.

(iv) The difference between $\overline{\mathrm{IQR}}_{0.75}(S_n)$ and $\mathrm{IQR}_{0.75}(S_n^c)$ is more pronounced for the heavy-tailed model (C), compared to the light-tailed model (B).

2.5 Appendix

2.5.1 Two classic results

Here we list two classic results used in this paper for the sake of completeness. The first result is used in the proof of Theorem 2.2.1. The original choice of notation in Schmeidler (1986) is kept here.

Theorem 2.5.1 (Proposition 2 of Schmeidler (1986))**.** *Let Σ denote a nonempty algebra of subsets of a set S, let B denote the set of all bounded, real-valued, Σ-measurable functions on S. Suppose that $I : B \to \mathbb{R}$ is comonotonic additive and continuous with respect to supremum norm in B. Then, for any E in Σ, defining $v(E) = I(\mathbf{1}_E)$ on Σ, we have for all $a \in B$,*

$$I(a) = \int_{-\infty}^{0} (v(a \geqslant \alpha) - v(S))\,\mathrm{d}\alpha + \int_{0}^{\infty} v(a \geqslant \alpha)\,\mathrm{d}\alpha.$$

The second result is the classic Hardy-Littlewood-Pólya inequality (see e.g. page 22 of Olkin and Marshall (2016)), which is used in the proof of Theorem 2.2.5.

Theorem 2.5.2 (Hardy-Littlewood-Pólya)**.** *Let f and g be two decreasing integrable functions on $[a, b]$, taking values in $[0, 1]$. Then*

$$\int_a^b \phi(f(x))\,\mathrm{d}x \leqslant \int_a^b \phi(g(x))\,\mathrm{d}x$$

holds for all continuous concave function ϕ (for which both functions $\phi \circ f$ and $\phi \circ g$ are integrable) if and only if

$$\int_a^x f(t)\,\mathrm{d}t \geqslant \int_a^x g(t)\,\mathrm{d}t \qquad \text{for } a \leqslant x \leqslant b$$

and

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^b g(t)\,\mathrm{d}t.$$

2.5.2 Proofs in Section 2.2

Proof of Lemma 2.2.2. (i) " $\Rightarrow$": This is trivial from the definition of signed Choquet integrals.

(ii) " $\Leftarrow$": Fix $p \in [0,1]$, we take a Bernoulli random variable X such that

$$\mathbb{P}(X=0)=p,\ \mathbb{P}(X=1)=1-p.$$

It follows that

$$I_{h_i}(X) = h_i(1-p) + \int_1^\infty h_i(0)\,\mathrm{d}x = h_i(1-p), \quad i=1,2.$$

As $I_{h_1} \leqslant I_{h_2}$ and p is arbitrary, we conclude $h_1 \leqslant h_2$. □

Proof of Lemma 2.2.3. (i) " $\Rightarrow$": We only show the case when I_h is increasing. Take $U \sim \mathrm{U}[0,1]$, for any $t_1, t_2 \in [0,1]$, $t_1 \leqslant t_2$, we let $X = \mathbf{1}_{\{U \leqslant t_1\}}$ and $Y = \mathbf{1}_{\{U \leqslant t_2\}}$. $X \leqslant Y$ implies $h(t_1) = I_h(X) \leqslant I_h(Y) = h(t_2)$. This shows that h is increasing.

" $\Leftarrow$": We only show the case when h is increasing. For random variables $X, Y \in L^\infty$, $x \in \mathbb{R}$, if $X \leqslant Y$, then $\mathbb{P}(X \geqslant x) \leqslant \mathbb{P}(Y \geqslant x)$. If h is increasing, then

$$h(\mathbb{P}(X \geqslant x)) \leqslant h(\mathbb{P}(Y \geqslant x)),$$

which implies $I_h(X) \leqslant I_h(Y)$. Hence I_h is increasing.

(ii) By straightforward calculation,

$$\begin{aligned} I_h(c) &= \int_{-\infty}^0 \left(h(\mathbb{P}(c \geqslant x)) - h(1)\right)\,\mathrm{d}x + \int_0^\infty h(\mathbb{P}(c \geqslant x))\,\mathrm{d}x \\ &= \int_{c\wedge 0}^0 \left(-h(1)\right)\,\mathrm{d}x + \int_0^{0\vee c} h(1)\,\mathrm{d}x = ch(1). \end{aligned}$$

Since I_h is comonotonic-additive, and X and c are comonotonic, $I_h(X+c) = I_h(X) + I_h(c) = I_h(X) + ch(1)$.

(iii) By (2.1),

$$I_h(\lambda X) = \int_{-\infty}^{0} (h(\mathbb{P}(\lambda X \geqslant x)) - h(1))\, \mathrm{d}x + \int_{0}^{\infty} h(\mathbb{P}(\lambda X \geqslant x))\, \mathrm{d}x$$
$$= \lambda \int_{-\infty}^{0} (h(\mathbb{P}(X \geqslant y)) - h(1))\, \mathrm{d}y + \lambda \int_{0}^{\infty} h(\mathbb{P}(X \geqslant y))\, \mathrm{d}y = \lambda I_h(X).$$

(iv) By (2.1) and (2.2),

$$I_h(-X) = \int_{-\infty}^{0} (h(\mathbb{P}(X < -x)) - h(1))\, \mathrm{d}x + \int_{0}^{\infty} h(\mathbb{P}(X < -x))\, \mathrm{d}x$$
$$= \int_{-\infty}^{0} \hat{h}(\mathbb{P}(X \geqslant -x))\, \mathrm{d}x + \int_{0}^{\infty} \left(\hat{h}(\mathbb{P}(X \geqslant -x)) - \hat{h}(1)\right) \mathrm{d}x$$
$$= \int_{0}^{\infty} \hat{h}(\mathbb{P}(X \geqslant x))\, \mathrm{d}x + \int_{-\infty}^{0} \left(\hat{h}(\mathbb{P}(X \geqslant x)) - \hat{h}(1)\right) \mathrm{d}x = I_{\hat{h}}(X). \qquad \square$$

2.5.3 Proofs in Section 2.3

Proof of Lemma 2.3.1. Replace h by its Jordan decomposition $h = h_+ - h_-$, where h_+, h_- are increasing. We have $\mathrm{TV}_{h_+} + \mathrm{TV}_{h_-} = \mathrm{TV}_h$, and

$$|I_h(X) - I_h(Y)| = |I_{h_+}(X) - I_{h_-}(X) - I_{h_+}(Y) + I_{h_-}(Y)| \leqslant |I_{h_+}(X) - I_{h_+}(Y)| + |I_{h_-}(X) - I_{h_-}(Y)|.$$

Therefore, it suffices to show (2.7) for h increasing. From Lemma 2.2.3, we have

$$I_h(Y) \leqslant I_h(X + ||X - Y||_\infty) = I_h(X) + h(1)||X - Y||_\infty = I_h(X) + \mathrm{TV}_h ||X - Y||_\infty.$$

Therefore, by symmetry, (2.7) holds. $\square$

Chapter 3

Risk Functionals with Convex Level Sets

3.1 Introduction

In this chapter, we focus on the CxLS property of (possibly multi-dimensional) risk functionals. This chapter appears in large part in the submitted paper Wang and Wei (2018b). To make the chapter concentrated, we will present the formal definitions of elicitability, identifiability and backtestability, their risk management implications, and their relation with the CxLS property in Section 3.6. Below, we give the definition of the CxLS property, the main object of this chapter. Denote by $\mathcal{M}_0$ the set of distributions (i.e. Borel probability measures on $\mathbb{R}$).

Definition 3.1.1. (The CxLS property) For $\mathcal{M} \subset \mathcal{M}_0$, we say a functional $\rho : \mathcal{M} \to \mathbb{R}^d$ has *convex level sets (CxLS)* on $\mathcal{M}$ if the level set $\{F \in \mathcal{M} : \rho(F) = \gamma\}$ is convex for each $\gamma \in \mathbb{R}^d$.

In other words, the CxLS property of ρ means that, if $F, G \in \mathcal{M}$ satisfy $\rho(F) = \rho(G)$, then $\rho(\lambda F + (1-\lambda)G) = \rho(F)$ for all $\lambda \in [0, 1]$, assuming $\lambda F + (1-\lambda)G \in \mathcal{M}$ (guaranteed if $\mathcal{M}$ is convex). Although Definition 3.1.1 does not require $\mathcal{M}$ itself to be convex, common choices of $\mathcal{M}$ are convex sets, such as the set of distributions with bounded support, or the set of distributions with positive densities.

To interpret the CxLS property, it means that if two risk models are assessed as equally risky, then a mixture of the two models should remain at the same risk level.

This chapter contains two main contributions to the theory of risk functionals with the CxLS property, and hence to the theory of elicitability. First, in Section 3.3, we characterize all one-dimensional signed Choquet integrals with the CxLS property (Theorem 3.3.1). This result generalizes the existing characterization in Kou and Peng (2016) on distortion risk measures, a class of increasing Choquet integrals. It turns out that the only signed Choquet integrals that have CxLS are the monotone ones with CxLS multiplied by a constant. This result requires some new techniques which we provide in a few lemmas, as non-monotonicity of signed Choquet integrals creates great challenges, which cannot be addressed by the existing methods in the literature. Second, in Sections 3.4-3.5, we extend our discussion to the multi-dimensional setting. Because of the special role of the quantile functionals in the theory of elicitabiltiy and backtestability, we characterize the CxLS property of all two-dimensional signed Choquet integrals, whose one component is a quantile (Theorem 3.5.2). To our knowledge, this result is the first CxLS characterization in the literature beyond the one-dimensional case, and it partially answers the open question of Kou and Peng (2016) and Fissler and Ziegel (2016). To establish this result, we provide some general results on multi-dimensional CxLS property in Section 3.4. Since CxLS is a necessary condition of elicitability, identifiability and backtestability, our characterization identifies candidates for risk functionals with these statistical properties. In particular, based on the result in dimension two, we show that the only coherent risk measure co-elicitable with a Value-at-Risk is a combination of the mean and the corresponding Expected Shortfall (Theorem 3.6.5). To better illustrate the concept of the CxLS property, we present a list of commonly used functionals with or without CxLS in Section 3.2, and in Section 3.6 we give an overview on the relationship among the statistical concepts of elicitability, identifiability, backtestability and the CxLS property. The proofs of the technical results are put in the Appendix.

Notation

For $q \in [1, \infty)$, let $\mathcal{M}_q$ be the set of distributions with finite q-th moments. Denote by $\mathcal{M}_\infty$ the set of distributions of bounded random variables, $\mathcal{M}_{\text{con}} = \{F \in \mathcal{M}_\infty : F \text{ has a density}\}$, and $\mathcal{M}_{\text{dis}} = \{F \in \mathcal{M}_\infty : F \text{ is discrete}\}$. For the ease of presentation, we identify distributions in $\mathcal{M}_0$ with the corresponding cumulative distribution functions, that is, for $F \in \mathcal{M}_0$ and $x \in \mathbb{R}$, $F(x)$ is understood as $F((-\infty, x])$. For $F \in \mathcal{M}_0$, define the left-continuous generalized inverse (left-quantile) as

$$F^{-1}(t) = \inf\{x \in \mathbb{R} : F(x) \geqslant t\}, \quad t \in (0, 1],$$

and in addition let $F^{-1}(0) = \sup\{x \in \mathbb{R} : F(x) = 0\}$. For $x \in \mathbb{R}$, δ_x denotes the point mass at x. Throughout this article, we stick to the following convention. In a result, if we do not specify $\mathcal{M}$, then the statements hold for any set $\mathcal{M} \subset \mathcal{M}_0$ such that the risk functional at consideration is finite on $\mathcal{M}$.

3.2 Examples of risk functionals and their CxLS properties

We first present some common examples of one-dimensional risk functionals with or without CxLS. These examples will help to understand the main concepts in this chapter, and they will be referred to repeatedly throughout. We start with a few interesting common quantities that have CxLS. They are, in fact, increasing Choquet integrals; see Section 3.3. A full characterization of all signed Choquet integrals with CxLS will be given in Theorem 3.3.1 below.

Example 3.2.1. (i) The expectation:

$$\mathbb{E}[F] = \int_{-\infty}^{\infty} x \, \mathrm{d}F(x), \quad F \in \mathcal{M}_1.$$

Note that in this chapter we define $\mathbb{E}$ on the set of distributions $\mathcal{M}_1$ instead of the set of integrable random variables.

(ii) The left-quantile, or the Value-at-Risk (VaR): For $p \in (0, 1]$, define

$$\mathrm{VaR}_p(F) = \inf\{x \in \mathbb{R} : F(x) \geqslant p\}, \quad F \in \mathcal{M}_0.$$

In addition, let $\operatorname{ess\,sup} = \mathrm{VaR}_1$.

(iii) The right-quantile: For $p \in [0, 1)$, define

$$\mathrm{VaR}_p^+(F) = \inf\{x \in \mathbb{R} : F(x) > p\}, \quad F \in \mathcal{M}_0.$$

In addition, let $\operatorname{ess\,inf} = \mathrm{VaR}_0^+$.

(iv) The mixed-quantile: For $p \in (0, 1)$ and $c \in [0, 1]$, define

$$\mathrm{VaR}_p^c = c\mathrm{VaR}_p^+ + (1 - c)\mathrm{VaR}_p.$$

In addition, we include the cases $p = 0$ and $p = 1$ by letting $\mathrm{VaR}_1^c = \mathrm{VaR}_1 = \operatorname{ess\,sup}$ and $\mathrm{VaR}_0^c = \mathrm{VaR}_0^+ = \operatorname{ess\,inf}$ for all $c \in [0, 1]$.

(v) The mid-point of range:

$$\text{Mid-range}(F) = \frac{1}{2}\operatorname{ess\,sup}(F) + \frac{1}{2}\operatorname{ess\,inf}(F), \quad F \in \mathcal{M}_\infty.$$

Next, let $f : \mathbb{R} \to \mathbb{R}$ be a measurable function and $\rho : \mathcal{M} \to \mathbb{R}$, $\rho(F) = \int f \,\mathrm{d}F$. In this case, ρ has CxLS since it is linear in $F \in \mathcal{M}$. Common examples of such functionals include von Neumann-Morgenstern expected utilities, expected loss functionals, and moments.

Example 3.2.2. (i) A von Neumann-Morgenstern expected utility functional has the form $F \mapsto \int_{-\infty}^{\infty} u(x)\,\mathrm{d}F(x)$ for some increasing utility function $u : \mathbb{R} \to [-\infty, \infty)$.

(ii) Second moment:

$$F \mapsto \int_{-\infty}^{\infty} x^2 \,\mathrm{d}F(x), \quad F \in \mathcal{M}_2.$$

(iii) The entropic risk measure for $\gamma \in (0, \infty)$:

$$e_\gamma(F) = \frac{1}{\gamma}\log\left(\int_{-\infty}^{\infty} e^{-\gamma x}\,\mathrm{d}F(x)\right), \quad F \in \mathcal{M}_\infty.$$

(iv) Excess loss function: for some $k \in \mathbb{R}$,

$$F \mapsto \int_{-\infty}^{\infty} (x-k)_+ \,\mathrm{d}F(x), \quad F \in \mathcal{M}_1.$$

A lot of results on CxLS of risk measures appear in the recent literature. For definitions of these risk measures, see McNeil et al. (2015). A coherent risk measure does not have CxLS on $\mathcal{M}_\infty$ unless it is an expectile (Corollary 4.6 of Ziegel (2016)). A convex risk measure has CxLS on $\mathcal{M}_\infty$ if and only if it is a shortfall risk measure (Theorem 3.10 of Delbaen et al. (2016); this includes Examples 3.2.2 (iii) and 3.2.3 (ii)). In particular, the Expected Shortfall (ES) does not have CxLS, as noted by Gneiting (2011).

Example 3.2.3. (i) The Expected Shortfall (ES) for $p \in (0, 1)$, defined as

$$\mathrm{ES}_p(F) = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_t(F)\,\mathrm{d}t, \quad F \in \mathcal{M}_1,$$

is a coherent risk measure, and it does not have CxLS. The following famous ES-VaR relation of Rockafellar and Uryasev (2002) will be used repeatedly in this chapter.

$$\begin{cases} [\mathrm{VaR}_p(F), \mathrm{VaR}_p^+(F)] = \underset{x\in\mathbb{R}}{\arg\min}\left\{x + \dfrac{1}{1-p}\displaystyle\int_{-\infty}^{\infty}(y-x)_+\,\mathrm{d}F(y)\right\}; \\ \mathrm{ES}_p(F) = \underset{x\in\mathbb{R}}{\min}\left\{x + \dfrac{1}{1-p}\displaystyle\int_{-\infty}^{\infty}(y-x)_+\,\mathrm{d}F(y)\right\}. \end{cases} \tag{3.1}$$

(ii) The expectile (see e.g. Newey and Powell (1987)) for $p \in (0,1)$, defined as

$$e_p(F) = \underset{x\in\mathbb{R}}{\arg\min} \left\{ p \int_x^{\infty} (y-x)^2 \,\mathrm{d}F(y) + (1-p) \int_{-\infty}^{x} (y-x)^2 \,\mathrm{d}F(y) \right\}, \quad F \in \mathcal{M}_2,$$

is a coherent risk measure, and it has CxLS since it is a shortfall risk measure.

A deviation measure, such as the variance or the standard deviation, is a functional D that satisfies $D(\delta_c) = 0$ for a constant c and $D(\mu) > 0$ for $\mu \in \mathcal{M}$ not a point mass. For a general theory on deviation measures and measures of variability, see Bickel and Lehmann (1976), Rockafellar et al. (2006) and Furman et al. (2017). Deviation measures generally cannot have CxLS, as easily seen from the following argument. For any $x, y \in \mathbb{R}$, $x \neq y$, by definition, $\mathcal{D}(\lambda\delta_x + (1-\lambda)\delta_y) > 0 = \mathcal{D}(\delta_x) = \mathcal{D}(\delta_y)$ for $\lambda \in (0,1)$. This implies that $\mathcal{D}$ does not have CxLS on $\mathcal{M}$. We summarize this observation in the following proposition. Nevertheless, later in Section 3.4 we shall see that deviation measures may have CxLS when they are paired with other risk functionals.

Proposition 3.2.1. *Let $\mathcal{M}$ be a set that contains all point masses and two-point distributions. A deviation measure does not have CxLS on $\mathcal{M}$.*

Example 3.2.4. As shown in Proposition 3.2.1 above, the following deviation measures do not have CxLS on their domains.

(i) The variance $\mathrm{Var}(F) = \int (x - \mathbb{E}[F])^2 \,\mathrm{d}F(x)$, $F \in \mathcal{M}_2$.

(ii) The standard deviation $\mathrm{SD}(F) = \sqrt{\mathrm{Var}(F)}$, $F \in \mathcal{M}_2$.

(iii) The mean median deviation (MD):

$$\mathrm{MD}(F) = \min_{x\in\mathbb{R}} \int_{-\infty}^{\infty} |y-x| \,\mathrm{d}F(y) = \int_{-\infty}^{\infty} \left(\left| y - F^{-1}\left(\frac{1}{2}\right) \right| \right) \mathrm{d}F(y), \quad F \in \mathcal{M}_1.$$

(iv) The Gini deviation:

$$\mathrm{Gini}(F) = \int_0^1 F^{-1}(t)(2t-1)\,\mathrm{d}t, \quad F \in \mathcal{M}_1.$$

(v) The range:

$$\mathrm{Range}(F) = \operatorname{ess\,sup}(F) - \operatorname{ess\,inf}(F), \quad F \in \mathcal{M}_\infty.$$

Finally, we give a few other notable functionals with or without CxLS.

Example 3.2.5. (i) Let $\mathcal{M}$ be the set of all the continuous distributions with a unique mode. Then the mode functional $\mathrm{T}(F) = \max_{x\in\mathbb{R}} \frac{\mathrm{d}}{\mathrm{d}x}F(x)$ has CxLS on $\mathcal{M}$ by definition.

(ii) The skewness functional SK is defined as $\mathrm{SK}(F) = \int \left(\frac{x-\mu}{\sigma}\right)^3 \mathrm{d}F(x)$ for $F \in M_3$ with $\mathrm{SD}(F) > 0$, where $\mu = \mathbb{E}[F]$ and $\sigma = \mathrm{SD}(F)$. By definition, we can calculate $\mathrm{SK}\left(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_0\right) = \mathrm{SK}\left(\frac{1}{2}\delta_0 + \frac{1}{2}\delta_2\right) = 0$, and

$$\mathrm{SK}\left(\frac{2}{3}\left(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_0\right) + \frac{1}{3}\left(\frac{1}{2}\delta_0 + \frac{1}{2}\delta_2\right)\right) = \mathrm{SK}\left(\frac{1}{3}\delta_{-1} + \frac{1}{2}\delta_0 + \frac{1}{6}\delta_2\right) = 1 \neq 0.$$

The skewness functional does not have CxLS on its domain.

3.3 Main characterization result in dimension one

3.3.1 Signed Choquet integrals

We first define signed Choquet integrals, a popular class of functionals in risk management and statistics and the main object of this section. Let

$$\mathcal{H} = \{h : h \text{ maps } [0,1] \text{ to } \mathbb{R},\, h \text{ is of bounded variation and } h(0) = 0\}.$$

Definition 3.3.1. A *signed Choquet integral* $I_h : \mathcal{M} \to \mathbb{R}$ is defined as

$$I_h(F) = \int_{-\infty}^{0} \left(h(1 - F(x))) - h(1)\right) \mathrm{d}x + \int_{0}^{\infty} h(1 - F(x))\, \mathrm{d}x,$$

where $h \in \mathcal{H}$ and $\mathcal{M}$ is a convex subset of $\mathcal{M}_0$ such that I_h is well-defined. The function h is called the *distortion function* of I_h. An increasing (resp. decreasing) Choquet integral is a signed Choquet integral with an increasing (resp. decreasing) distortion function.

Signed Choquet integrals are studied extensively in the literature; for its axiomatic characterization and economic interpretation, we refer to Yaari (1987), De Waegenaere and Wakker (2001), Kou and Peng (2016) and Wang et al. (2018). A functional on $\mathcal{M}$ is a signed Choquet integral if and only if it is comonotonic-additive and satisfies a continuity condition (Theorem 2.1 of Wang et al. (2018)); see Proposition 3.5.1 below for a precise statement of this characterization in multi-dimension. Many of examples of signed Choquet integrals are listed in Section 3.2. In particular, increasing Choquet integrals I_h with $h(1) = 1$ are also known as distortion risk measures, which include the mean, the quantiles, and the Expected Shortfalls, and many measures of variability belong to the class of Signed Choquet integrals. We give the distortion functions of these examples below.

(i) The expectation in Example 3.2.1 has distortion function $h(t) = t,\ t \in [0, 1]$.

(ii) VaR_p in Example 3.2.1 for $p \in (0, 1]$ has distortion function $h(t) = \mathbf{1}_{\{t>1-p\}},\ t \in [0, 1]$.

(iii) VaR_p^+ in Example 3.2.1 for $p \in [0, 1)$ has distortion function $h(t) = \mathbf{1}_{\{t\geqslant 1-p\}},\ t \in [0, 1]$.

(iv) The mid-point range in Example 3.2.1 has distortion function $h(t) = \frac{1}{2}\mathbf{1}_{\{0<t<1\}} + \mathbf{1}_{\{t=1\}},\ t \in [0, 1]$.

(v) ES_p in Example 3.2.3 for $p \in (0, 1)$ has distortion function $h(t) = \min\{\frac{t}{1-p}, 1\},\ t \in [0, 1]$.

(vi) MD in Example 3.2.4 has distortion function $h(t) = \min\{t, 1 - t\},\ t \in [0, 1]$.

(vii) The Gini deviation in Example 3.2.4 has distortion function $h(t) = t - t^2,\ t \in [0, 1]$.

(viii) The range in Example 3.2.4 has distortion function $h(t) = \mathbf{1}_{\{0<t<1\}},\ t \in [0, 1]$.

In the above examples, (i)-(iv) have CxLS and (v)-(viii) do not; (i)-(v) are increasing Choquet integrals and (vi)-(viii) are not.

3.3.2 Characterization of signed Choquet integrals with CxLS

As we have seen in Section 3.2, some signed Choquet integrals have CxLS whereas some others do not. The main result in this section characterizes all signed Choquet integrals with CxLS. It turns out that the following three forms of h in the subsets $\mathcal{H}_1^*, \mathcal{H}_2^*$ and $\mathcal{H}_3^*$ of $\mathcal{H}$ are important for the CxLS property.

(i) $h \in \mathcal{H}_1^*$: For some $c \in [0, 1]$, $h(t) = ch(1)\mathbf{1}_{\{0<t<1\}} + h(1)\mathbf{1}_{\{t=1\}}$, $t \in [0, 1]$. In this case,

$$I_h(F) = h(1)(c\operatorname{ess\,sup}(F) + (1 - c)\operatorname{ess\,inf}(F)),\ F \in \mathcal{M}. \tag{3.2}$$

(ii) $h \in \mathcal{H}_2^*$: $h(t) = th(1)$, $t \in [0, 1]$. In this case,

$$I_h(F) = h(1)\mathbb{E}[F],\ F \in \mathcal{M}. \tag{3.3}$$

(iii) $h \in \mathcal{H}_3^*$: For some $\alpha \in (0,1)$ and $c \in [0,1]$, $h(t) = ch(1)\mathbf{1}_{\{t=\alpha\}} + h(1)\mathbf{1}_{\{t>\alpha\}}$, $t \in [0,1]$. In this case,

$$I_h(F) = h(1)\mathrm{VaR}^c_{1-\alpha}(F),\ F \in \mathcal{M}. \tag{3.4}$$

We also denote by $\mathcal{H}^* = \bigcup_{i=1}^3 \mathcal{H}_i^*$.

For increasing Choquet integrals I_h with $h(1) = 1$, Kou and Peng (2016) show that only the above three cases are possible for I_h to have CxLS. Without monotonicity, the class of signed Choquet integrals is much larger than the class of increasing ones. Nevertheless, in the next theorem, we show that the only possible signed Choquet integrals with CxLS are still the ones listed above.

Theorem 3.3.1. *Let $\mathcal{M}$ be a convex set that contains all three-point distributions. A signed Choquet integral I_h has CxLS on $\mathcal{M}$ if and only if $h \in \mathcal{H}^*$.*

Comparing Theorem 3.3.1 with the characterization of increasing Choquet integrals with CxLS in Kou and Peng (2016), we see that removing monotonicity does not lead to many more choices of functionals with CxLS. More precisely, all functionals in (3.2)-(3.4) have either an increasing or decreasing distortion function, and they are monotone with respect to stochastic order (e.g. Lemma 2.3 of Wang et al. (2018)).

Corollary 3.3.2. *If I_h has CxLS on $\mathcal{M}_{\mathrm{dis}}$, then it is monotone (i.e. increasing or decreasing).*

Corollary 3.3.2 is a surprising result, as the CxLS property by definition is not related to monotonicity. For instance, for any measurable function $f : \mathbb{R} \to \mathbb{R}$, the mapping $\rho : F \mapsto \int f(x)\,\mathrm{d}F(x)$ has CxLS; whether f is monotone (i.e. whether ρ is monotone with respect to stochastic order) is irrelevant to the CxLS property. Among the class of signed Choquet integrals, however, the CxLS property surprisingly implies monotonicity.

On the other hand, the proofs for signed Choquet integrals are much more involved than the case of increasing ones, due to the lack of monotonicity. Two new technical lemmas are needed for a proof of Theorem 3.3.1, which are put in the Appendix.

3.3.3 The choice of $\mathcal{M}$

In risk management practice, one may be only interested in backtestability or elicitability over the set of continuous distribution models. In the following we will show that, when constrained on the set of continuous distributions, the only possible choices of h to allow for the CxLS property of I_h are still the three cases in Theorem 3.3.1.

We define the metric w on $\mathcal{M}_\infty$ by $w(F,G) = \sup_{t\in[0,1]} |F^{-1}(t) - G^{-1}(t)|$, known as the Wasserstein-L^∞ metric. This metric will be the mathematical tool to bridge between continuous and discrete models. For a sequence $\{F_n\}_{n\in\mathbb{N}} \subset \mathcal{M}_\infty$, we write $F_n \overset{\text{w}}{\to} F$ if the sequence $\{F_n\}_{n\in\mathbb{N}}$ converges to F in the metric w. We first provide some technical properties of the w-convergence. They will be useful in bridging the gap between continuous and discrete distributions, as well as in the characterization of multi-dimensional signed Choquet integrals in Section 3.5.

Proposition 3.3.3. *(i) Suppose that $F_n, G_n \in \mathcal{M}_\infty$, $n \in \mathbb{N}$, $F_n \overset{\text{w}}{\to} F \in \mathcal{M}_\infty$ and $G_n \overset{\text{w}}{\to} G \in \mathcal{M}_\infty$. Then $\lambda F_n + (1-\lambda)G_n \overset{\text{w}}{\to} \lambda F + (1-\lambda)G$ for all $\lambda \in [0,1]$.*

(ii) For $h \in \mathcal{H}$, I_h is uniformly continuous with respect to w on $\mathcal{M}_\infty$.

With the help of Proposition 3.3.3, we can show that CxLS on $\mathcal{M}_{\text{con}}$ implies CxLS on $\mathcal{M}_{\text{dis}}$ for I_h, and as a consequence of Theorem 3.3.1, we know $h \in \mathcal{H}^*$.

Proposition 3.3.4. *For $h \in \mathcal{H}$, if I_h has CxLS on $\mathcal{M}_{\text{con}}$, then I_h has CxLS on $\mathcal{M}_{\text{dis}}$, and, as a consequence, $h \in \mathcal{H}^*$.*

3.4 CxLS in multi-dimension

3.4.1 Some general properties

In this section, we provide some simple results for CxLS in multi-dimension, which will be found useful later to show our main characterization result in dimension two, as well as the CxLS in some examples. These results are easy to verify, and self-contained proofs are provided in the Appendix for completeness. First, a d-dimensional functional has CxLS if it is an injective function of other one-dimensional functionals with CxLS. Since one-dimensional functionals with CxLS are well studied, this result gives a convenient way to construct simple functionals with CxLS in multi-dimension.

Proposition 3.4.1. *A functional $\rho : \mathcal{M} \to \mathbb{R}^d$ has CxLS on $\mathcal{M}$ if it is an injective function of some $\mathbb{R}$-valued functionals with CxLS on $\mathcal{M}$.*

As a direct consequence of Proposition 3.4.1, a functional $\rho : \mathcal{M} \to \mathbb{R}^d$ has CxLS on $\mathcal{M}$ if each of its component has CxLS on $\mathcal{M}$.

The next proposition establishes a link between a 2-dimensional functional with CxLS and its components, assuming one of the component already has CxLS.

Proposition 3.4.2. *Let ρ_1 and ρ_2 be the functionals from $\mathcal{M}$ to $\mathbb{R}$ such that ρ_2 has CxLS on $\mathcal{M}$. The pair of functionals (ρ_1, ρ_2) has CxLS on $\mathcal{M}$ if and only if ρ_1 has CxLS on $\mathcal{M}(r)$ for all $r \in \mathbb{R}$, where $\mathcal{M}(r) = \{F \in \mathcal{M} : \rho_2(F) = r\}$.*

Proposition 3.4.2 is very useful to prove or disprove the CxLS property of two-dimensional functionals. For instance, it would justify the CxLS of $(\mathrm{VaR}_p, \mathrm{ES}_p)$ in Example 3.4.1 below, and it is used repeatedly in the proofs of Theorems 3.5.2 and 3.6.5.

3.4.2 Examples in multi-dimension

In this section, we list some multi-dimensional functionals with CxLS. Since any functional that have all components with CxLS automatically has CxLS, we only give examples where at least one of the components do not have CxLS.

Example 3.4.1 (Two-dimensional examples)**.**

(i) $(\mathrm{VaR}_p, \mathrm{ES}_p)$ for $p \in (0, 1)$ on $\mathcal{M}_1$: We know VaR_p has CxLS by Theorem 3.3.1. On the set $\mathcal{M}(r) = \{F \in \mathcal{M}_1 : \mathrm{VaR}_p(F) = r\}$, $r \in \mathbb{R}$, using the ES-VaR relation (3.1), $\mathrm{ES}_p(F) = \frac{1}{1-p}\int_r^\infty (x - r)\,\mathrm{d}F(x) + r$. Note that ES_p is linear in F for $F \in \mathcal{M}(r)$. Therefore, by Proposition 3.4.2, we know that $(\mathrm{VaR}_p, \mathrm{ES}_p)$ has CxLS.

(ii) $(\mathrm{Median}, \mathrm{MD})$ on $\mathcal{M}_1$: The median (i.e. $\mathrm{VaR}_{1/2}$) has the CxLS property. From the definition of MD and similar steps as in (i) above, the pair $(\mathrm{Median}, \mathrm{MD})$ has CxLS.

(iii) $(\text{Mid-range}, \mathrm{Range})$ has CxLS on $\mathcal{M}_\infty$ by Proposition 3.4.1, since the pair is a bijection from $(\operatorname{ess\,inf}, \operatorname{ess\,sup})$, which has components with CxLS.

(iv) $(\mathbb{E}, \mathrm{Var})$ has CxLS on $\mathcal{M}_2$ by Proposition 3.4.1, since the pair is a bijection from the pair of the first two moments, which has components with CxLS. Consequently, the pair $(\mathbb{E}, \mathrm{SD})$ has CxLS as well.

Example 3.4.2 (Three-dimensional examples)**.**

(i) The Range-Value-at-Risk (RVaR; see Embrechts et al. (2018)) is a signed Choquet integral with distortion function $h(t) = \frac{t-1+q}{q-p}\mathbf{1}_{\{1-q\leqslant t\leqslant 1-p\}}$, $t \in [0,1]$, for some $0 < p < q < 1$. An RVaR does not have CxLS as its distortion function does not belong to the cases in Theorem 3.3.1. By definition, we can write

$$\mathrm{RVaR}_{p,q}(F) = \frac{1}{q-p}((1-p)\mathrm{ES}_p(F) - (1-q)\mathrm{ES}_q(F)),\ \ F \in \mathcal{M},$$

and therefore the triplet $(\mathrm{VaR}_p, \mathrm{VaR}_q, \mathrm{RVaR}_{p,q})$ has CxLS by a similar argument in (i) of Example 3.4.1.

(ii) Recall that the skewness functional SK in (ii) of Example 3.2.5 does not have CxLS. The triplet $(\mathbb{E}, \mathrm{SD}, \mathrm{SK})$ has CxLS on $\mathcal{M} = \{F \in \mathcal{M}_3 : \mathrm{SD}(F) > 0\}$, by using Proposition 3.4.1, since the triplet is a bijection from the triplet of the first three moments on $\mathcal{M}$, which has components with CxLS.

3.5 Main characterization result in dimension two

3.5.1 Multi-dimensional signed Choquet integrals

In this section we investigate the CxLS property in higher dimension. It is natural to define the signed Choquet integral in dimension $d \geqslant 2$ as follows.

Definition 3.5.1. Let $\mathbf{h} = (h_1, \cdots, h_n)$, where each $h_i \in \mathcal{H}$, $i = 1, \ldots, d$. A *signed Choquet integral* $I_{\mathbf{h}} : \mathcal{M} \to \mathbb{R}^d$ is defined as

$$I_{\mathbf{h}}(F) = \left(I_{h_1}(F), \cdots, I_{h_d}(F)\right), \tag{3.5}$$

where each I_{h_i}, $i = 1, \ldots, d$ is the one-dimensional signed Choquet integral defined in Definition 3.3.1.

A functional $\rho : \mathcal{M} \to \mathbb{R}$ is *law-invariant* if $\rho(F) = \rho(G)$ for any $F, G \in \mathcal{M}$. Signed Choquet integrals in multi-dimension shares a similar characterization via *comonotonic-additivity* (see Section 2.2.1 for the definition of comonotonic-additivity).

Recall that we use $\mathcal{M}_\infty$ to denote the set of distributions of bounded random variables. By applying Theorem 2.1 of Wang et al. (2018) to each component of I, and noting that the w-continuity of ρ is equivalent to L^∞-continuity of $\hat{\rho}$, we obtain the following characterization of multi-dimensional signed Choquet integrals.

Proposition 3.5.1. *A functional $I : \mathcal{M}_\infty \to \mathbb{R}^d$ is comonotonic-additive and uniformly continuous with respect to w if and only if I is a d-dimensional signed Choquet integral.*

Clearly, for $h_1, h_2 \in \mathcal{H}^*$, the 2-dimensional signed Choquet integral (I_{h_1}, I_{h_2}) has CxLS, due to Proposition 3.4.1. More interestingly, (I_{h_1}, I_{h_2}) may have CxLS even if it is not a injection from one-dimensional signed Choquet integrals with CxLS. A famous example is $(\mathrm{VaR}_p, \mathrm{ES}_p)$ for $p \in (0,1)$ as in Example 3.4.1, as shown by Fissler and Ziegel (2016) and Acerbi and Szekely (2017).

Characterization of multi-dimensional signed Choquet integrals with CxLS seems to be an extremely challenging task, which is left as an open question by Fissler and Ziegel (2016) and Kou and Peng (2016). Since VaR_p is a canonical candidate for a one-dimensional signed Choquet integral with CxLS, below we explore for which $h \in \mathcal{H}$, (I_h, VaR_p) has CxLS.

3.5.2 Characterizing a signed Choquet integral and a VaR with CxLS

Below we will characterize all pairs (I_h, VaR_p) with CxLS. To the best of our knowledge, this is the first result in the literature on characterizing CxLS in multi-dimension. By Proposition 3.4.1, (I_h, VaR_p) has CxLS if and only if $(I_h + a\mathrm{VaR}_p, \mathrm{VaR}_p)$ has CxLS for one (or all) $a \in \mathbb{R}$. Therefore, adding a constant times VaR_p to I_h does not change the CxLS property of (I_h, VaR_p). This explains why the term $a\mathrm{VaR}_p$ appears in all cases in the following theorem.

Theorem 3.5.2. *For $p \in (0,1)$ and $h \in \mathcal{H}$, (I_h, VaR_p) has CxLS on $\mathcal{M}_{\mathrm{dis}}$ if and only if I_h is one of the following cases*

(i) $I_h = a\mathrm{VaR}_p + I_{h^*}$ *for some* $a \in \mathbb{R}$ *and* $h^* \in \mathcal{H}^*$*;*

(ii) $I_h = a\mathrm{VaR}_p + b\mathbb{E} + c\mathrm{ES}_p$ *for some constants* $a, b, c \in \mathbb{R}$*;*

(iii) $I_h = a\mathrm{VaR}_p + b\mathrm{VaR}_p^+ + c\,\mathrm{ess\,sup}$ *for some constants* $a, b, c \in \mathbb{R}$ *with* $bc > 0$*;*

(iv) $I_h = a\mathrm{VaR}_p + b\mathrm{VaR}_p^+ + c\,\mathrm{ess\,inf}$ *for some constants* $a, b, c \in \mathbb{R}$ *with* $bc < 0$*.*

Theorem 3.5.2 reveals a characterization of all signed Choquet integrals that has CxLS jointly with VaR_p, $p \in (0,1)$. In particular, for risk management practice, one usually does not distinguish VaR_p and VaR_p^+. If we treat VaR_p and VaR_p^+ as identical, then cases (iii) and (iv) can be combined into case (i). Therefore, we are left with the following two cases.

(i) $I_h = a\mathrm{VaR}_p + I_{h^*}$ for some $a \in \mathbb{R}$ and $h^* \in \mathcal{H}^*$;

(ii) $I_h = a\mathrm{VaR}_p + b\mathbb{E} + c\mathrm{ES}_p$ for some constants $a, b, c \in \mathbb{R}$.

From (i) and (ii), if (I_h, VaR_p) is not a bijection from a pair of signed Choquet integrals with CxLS, then I_h has to be a linear combination of VaR_p, $\mathbb{E}$ and ES_p. Regarding the CxLS property, the class of VaR_p for $p \in (0,1)$ plays a unique role among the class of distortion risk measures (Kou and Peng (2016)) and among positively homogeneous tail risk measures (Liu and Wang (2016)). The above observation shows that ES_p is also very special regarding CxLS. In particular, as the CxLS property is necessary for elicitability (see Section 3.6), we will see in Theorem 3.6.5 that a convex combination of ES_p and $\mathbb{E}$ is the only type of comonotonic-additive coherent risk measure that is co-elicitable with VaR_p.

It may be worth noting that the roles of VaR_p and VaR_p^+ are symmetric. To get functionals I_h such that (I_h, VaR_p^+) has CxLS, one simply switches the positions in the pairs $(\mathrm{VaR}_p, \mathrm{VaR}_p^+)$ and $(\operatorname{ess\,inf}, \operatorname{ess\,sup})$ in Theorem 3.5.2. This statement is due to the following relation. For any distribution $F \in \mathcal{M}_0$, let $\bar{F}$ be given by $\bar{F}(A) = F(-A)$, where $-A = \{-x : x \in A\}$, $A \in \mathcal{B}(\mathbb{R})$.

Proposition 3.5.3. *For $p \in (0,1)$ and $h \in \mathcal{H}$, (I_h, VaR_p) has CxLS on $\mathcal{M}$ if and only if $(I_{\bar{h}}, \mathrm{VaR}_{1-p}^+)$ has CxLS on $\bar{\mathcal{M}}$, where $\bar{h} \in \mathcal{H}$ is given by $\bar{h}(t) = h(1-t) - h(1)$, $t \in [0,1]$ and $\bar{\mathcal{M}} = \{\bar{F} : F \in \mathcal{M}\}$.*

3.6 Backtestability, elicitability and identifiability

This section gives formal definitions of elicitability, backtestability and identifiability as studied by Osband (1985), Gneiting (2011) and Acerbi and Szekely (2017), and discusses their relation with the CxLS property. As the main focus of this chapter is the CxLS property, this section collects some major relevant facts for the interested reader with self-contained proofs, and we refer to Fissler and Ziegel (2016), Kou and Peng (2016) and Acerbi and Szekely (2017) for excellent summaries and detailed discussions on the implications in statistical inference, risk management, and banking regulation. In addition to the known facts, we obtain a new characterization result (Theorem 3.6.5) on co-elicitability of coherent risk measures with VaR_p.

Elicitability refers to the existence of a scoring function for the forecasted value of a risk functional and realized value of future observations, so that the mean of the scoring function attains its minimum value if and only if the value of the risk functional is truly forecasted; see Gneiting (2011) for elicitability in a decision-theoretic framework. Comparative backtest is discussed in Nolde and Ziegel (2017) as an alternative to the traditional backtest, for which elicitability is a necessary condition.

Definition 3.6.1 (Elicitability)**.** The functional $\rho : \mathcal{M} \to \mathbb{R}^d$ is $\mathcal{M}$-*elicitable* if there exists a *strictly consistent scoring function* $S : \mathbb{R}^{d+1} \to \mathbb{R}$ such that for all $F \in \mathcal{M}$,

$$\rho(F) = \underset{\mathbf{x} \in \mathbb{R}^d}{\arg\min} \int_{-\infty}^{\infty} S(\mathbf{x}, y) \, \mathrm{d}F(y). \tag{3.6}$$

We also say that $\rho_1 : \mathcal{M} \to \mathbb{R}$ is *co-elicitable* with $\rho_2 : \mathcal{M} \to \mathbb{R}$ on $\mathcal{M}$ if (ρ_1, ρ_2) is $\mathcal{M}$-elicitable.

In the literature, elicitability is often defined for set-valued risk functionals (e.g. generally, quantiles are interval-valued), as the minimizer to the scoring function is not necessarily a singleton. In this chapter, as we look at risk functionals mapping $\mathcal{M}$ to $\mathbb{R}^d$, we use the slightly narrower definition on $\mathbb{R}^d$-valued functionals. This choice of definition does not affect our discussion.

Next, identifiability refers to the existence of an identification function for the forecasted value of a risk functional and realized value of future observations. The mean of the identification function is zero if and only if the value of the risk functional is truly forecasted. Therefore, the realized average value of the identification function can identify whether a risk forecast is accurate.

Definition 3.6.2 (Identifiability)**.** A functional $\rho : \mathcal{M} \to \mathbb{R}^d$ is said to be $\mathcal{M}$-*identifiable*, if there exists an *identification function* $I : \mathbb{R}^{d+1} \to \mathbb{R}$ such that for all $F \in \mathcal{M}$,

$$\int_{-\infty}^{\infty} I(\mathbf{x}, y) \, \mathrm{d}F(y) = 0 \quad \text{if and only if} \quad \mathbf{x} = \rho(F).$$

Finally, we define backtestability as in Acerbi and Szekely (2017). Backtestability refers to the existence of a backtest function, whose mean reports positive value if the risk functional is under-forecasted, and negative value if the risk functional is over-forecasted. Thus, the realized average value of this backtest functional can distinguish between under- and over-estimation in risk forecasts. Due to the lack of a natural order in $\mathbb{R}^d$, one cannot speak of under-estimation or over-estimation for $\mathbb{R}^d$-valued risk functionals. Therefore, the notion of backtestability is suitable for dimension one only (a related notion for multi-dimensional functionals is *ridge backtests*; see Acerbi and Szekely (2017)).

Definition 3.6.3 (Backtestability)**.** A functional $\rho : \mathcal{M} \to \mathbb{R}$ is said to be $\mathcal{M}$-*backtestable*, if there exists a *backtest function* $Z : \mathbb{R}^2 \to \mathbb{R}$ such that for all $F \in \mathcal{M}$,

$$\int_{-\infty}^{\infty} Z(x, y)\,\mathrm{d}F(y) = 0 \quad \text{if and only if} \quad x = \rho(F),$$

and $\int Z(t, \cdot)\,\mathrm{d}F$ is strictly increasing in the prediction $t \in \mathbb{R}$.

The three notions introduced above are model-free in the sense that the statements holds for all $F \in \mathcal{M}$, that is, in order to compare scores, to identify forecasts, or to quantify backtests, one does not need to know the underlying distribution that generates the random observations.

In what follows, we illustrate the relationship among the above three concepts and the CxLS property. First, in dimension one, identifiability follows directly from backtestability, and backtestability is generally stronger than elicitability. In any dimension, both elicitability and identifiability imply the CxLS property. Finally, for one-dimensional signed Choquet integrals, CxLS is sufficient for backtestability except for the case of $h \in \mathcal{H}_1^*$.

$$\begin{tikzcd}
 & \text{Identifiability} \arrow[dr, Rightarrow, "\text{(Prop. 3.6.1)}"] & \\
\text{Backtestability} \arrow[ur, Rightarrow, "\text{(by definition)}"] \arrow[dr, Rightarrow, "\text{(Prop. 3.6.2)}"'] & & \text{CxLS} \arrow[ll, Rightarrow, "(I_h,\ \text{Prop. 3.6.4})"'] \\
 & \text{Elicitability} \arrow[ur, Rightarrow, "\text{(Prop. 3.6.1)}"'] &
\end{tikzcd}$$

The above statements will be verified (with some conditions) below.

We shall first see that elicitability and backtestability both imply the CxLS property. Suppose that $F, G \in \mathcal{M}$ satisfy $\rho(F) = \rho(G) = \mathbf{x} \in \mathbb{R}^d$. If ρ is $\mathcal{M}$-elicitable, let S be its scoring function in (3.6). As $\mathbf{x}$ is a minimizer for both $\int_{-\infty}^{\infty} S(\cdot, y)\,\mathrm{d}F(y)$ and $\int_{-\infty}^{\infty} S(\cdot, y)\,\mathrm{d}G(y)$, it must also be a minimizer for $\int_{-\infty}^{\infty} S(\cdot, y)\,\mathrm{d}(\lambda F + (1-\lambda)G)(y)$. By definition of elicitability, $\mathbf{x} = \rho(\lambda F + (1-\lambda)G)$, and ρ has CxLS. Similarly, if ρ is $\mathcal{M}$-identifiable, let $I : \mathbb{R}^{d+1} \to \mathbb{R}$ is its identification function. As $\mathbf{x}$ satisfies $\int_{-\infty}^{\infty} I(\mathbf{x}, y)\,\mathrm{d}F(y) = 0$ and $\int_{-\infty}^{\infty} I(\mathbf{x}, y)\,\mathrm{d}G(y) = 0$, we know $\int_{-\infty}^{\infty} I(\mathbf{x}, y)\,\mathrm{d}(\lambda F + (1-\lambda)G) = 0$ for all $\lambda \in [0, 1]$. By the definition of identifiability, we have $\mathbf{x} = \rho(\lambda F + (1-\lambda)G)$. We summarize these simple arguments in the following proposition.

Proposition 3.6.1. *For $\mathcal{M} \subset \mathcal{M}_0$, if $\rho : \mathcal{M} \to \mathbb{R}^d$ is $\mathcal{M}$-elicitable or $\mathcal{M}$-identifiable, then ρ has CxLS on $\mathcal{M}$.*

The next proposition verifies that one-dimensional backtestability implies elicitability, as shown by Acerbi and Szekely (2017).

Proposition 3.6.2. *If $\rho : \mathcal{M} \to \mathbb{R}$ is $\mathcal{M}$-backtestable with backtest function Z, then ρ is $\mathcal{M}'$-elicitable, where $\mathcal{M}' = \{F \in \mathcal{M} : \int_{-\infty}^{\infty} \int_0^z |Z(x, y)| \, \mathrm{d}x \, \mathrm{d}F(y) < \infty$ for all $z \in \mathbb{R}\}$.*

By Theorem 3.3.1, if a signed Choquet integral I_h has CxLS, then $h \in \mathcal{H}^*$, belonging to one of the three cases (3.2)-(3.4). Hence, it suffices to analyze backtestability in these cases. Note that by Theorem 3.3.1, a signed Choquet integral I_h with CxLS is either a mean, a mix-quantile, or a convex combination of $\operatorname{ess\,sup}$ and $\operatorname{ess\,inf}$, multiplied by a constant equal to $h(1)$. The next proposition verifies that the constant multiplier does not affect backtestability as long as it is not zero.

Proposition 3.6.3. *Suppose that $\rho : \mathcal{M} \to \mathbb{R}$ is backtestable with backtest function Z, then for $c \neq 0$, the functional $c\rho$ is backtestable with backtest function $Z^*(x, y) = cZ(x/c, y)$, $x, y \in \mathbb{R}$.*

For the signed Choquet integral I_h, $h \in \mathcal{H}^*$, if $h(1) = 0$, then $I_h(F) = 0$ for all $F \in \mathcal{M}$. This trivial functional is backtestable with the backtest function $Z(x, y) = x$ for all $x, y \in \mathbb{R}$. If $h(1) \neq 0$, by Proposition 3.6.3, the backtestability of I_h reduces to that of increasing Choquet integrals studied in Acerbi and Szekely (2017). We list them here for completeness.

Proposition 3.6.4. *For $h \in \mathcal{H}^*$ with $h(1) \neq 0$,*

(i) if $h \in \mathcal{H}_1^*$, I_h is not $\mathcal{M}_{\mathrm{dis}}$-backtestable;

(ii) if $h \in \mathcal{H}_2^*$, I_h is $\mathcal{M}_1$-backtestable;

(iii) if $h \in \mathcal{H}_3^*$, I_h is $\mathcal{M}_0^*$-backtestable, where $\mathcal{M}_0^* = \{F \in \mathcal{M}_0 : F^{-1}$ is continuous$\}$.

Remark 3.6.1. For a given $p \in (0, 1)$, VaR_p (or VaR_p^c as in case (iii) of Proposition 3.6.4) is $\mathcal{M}_0^*(p)$-backtestable (and $\mathcal{M}$-elicitable), where $\mathcal{M}_0^*(p) = \{F \in \mathcal{M}_0 : F^{-1}$ is continuous at $p\}$. The choice of $\mathcal{M} = \mathcal{M}_0^*(p)$ is the biggest such that VaR_p is $\mathcal{M}$-backtestable (or $\mathcal{M}$-elicitable).

We conclude this chapter by a characterization theorem on spectral risk measures that are co-elicitable with a VaR. Let $\mathcal{X}$ be the set of bounded random variables. According to Artzner et al. (1999), a functional $\hat{\rho}: \mathcal{X} \to \mathbb{R}$ is set to be a *coherent risk measure* if it is increasing, cash-additive, convex, and positively homogeneous. Translating this definition into our setting, we say that the functional $\rho: \mathcal{M}_\infty \to \mathbb{R}$ is a coherent risk measure, if $\hat{\rho}$ is a coherent risk measure in the sense of Artzner et al. (1999), where $\hat{\rho}: \mathcal{X} \to \mathbb{R}$ is given by $\hat{\rho}(X) = \rho(F)$ and F is the distribution of X.

We focus on comonotonic-additive coherent risk measures, a popular class of one-dimensional signed Choquet integrals. Elicitability of comonotonic-additive risk measures is studied by, for instance, Ziegel (2016), Kou and Peng (2016) and Fissler and Ziegel (2016). Using the characterization result of Kusuoka (2001), a functional $\rho: \mathcal{M}_\infty \to \mathbb{R}$ is a comonotonic-additive coherent risk measure if and only if it can be written as $\rho = \int_0^1 \mathrm{ES}_p \,\mathrm{d}\mu(p)$ for a Borel probability measure μ on $[0,1]$, or equivalently (see e.g. Theorem 2.8 of Wang et al. (2018)), $\rho = I_h$ for a concave and increasing $h \in \mathcal{H}$ satisfying $h(1) = 1$.

Since VaR_p is elicitable only on $\mathcal{M}_0^*(p)$, we consider co-elicitability on $\mathcal{M}_\infty^*(p) = \{F \in \mathcal{M}_\infty : F^{-1}$ is continuous at $p\}$. Among the forms of risk measures identified by Theorem 3.5.2, it is easy to see that the only choice of coherent risk measures is $\rho = a\mathrm{ES}_p + (1-a)\mathbb{E}$ for some $a \in [0,1]$. Therefore, naturally we would expect that the above form of ρ is the only coherent risk measure that is co-elicitable with VaR_p, although some detailed analysis needs to be carried out to translate from the CxLS property on the set $\mathcal{M}_\infty^*(p)$ to that on the set $\mathcal{M}_{\mathrm{dis}}$, in order to utilize Theorem 3.5.2.

Theorem 3.6.5. *For a comonotonic-additive coherent risk measure $\rho: \mathcal{M}_\infty \to \mathbb{R}$ and $p \in (0,1)$, ρ is co-elicitable with VaR_p on $\mathcal{M}_\infty^*(p)$ if and only if it is a convex combination of $\mathbb{E}$ and ES_p.*

To make sense of the co-elicitability of $(\mathrm{VaR}_p, a\mathrm{ES}_p + (1-a)\mathbb{E})$ in Theorem 3.6.5, we obtain its scoring function in the following proposition.

Proposition 3.6.6. *For $p \in (0,1)$ and $a \in (0,1]$, let $\rho = a\mathrm{ES}_p + (1-a)\mathbb{E}$. The functional (VaR_p, ρ) is $\mathcal{M}_\infty^*(p)$-elicitable with the strictly consistent scoring function*

$$S(x_1, x_2, y) = g(x_2) + g'(x_2)\left(a\left(x_1 + \frac{1}{1-p}(y - x_1)_+\right) + (1-a)y - x_2\right), \qquad x_1, x_2, y \in \mathbb{R},$$

where g is any differentiable, strictly increasing and strictly concave function on $\mathbb{R}$.

3.7 Proofs of the main results

3.7.1 Proofs in Section 3.3

Proof of Theorem 3.3.1. In order to prove Theorem 3.3.1, we present two technical lemmas.

Lemma 3.7.1. *If $h \in \mathcal{H}$ satisfies, for all $t, q \in [0,1]$, $0 \leqslant h(t) \leqslant h(1) = 1$ and*

$$h(t) = h(t)h(1-q+qt) + (1-h(t))h(qt), \tag{3.7}$$

then $h \in \mathcal{H}^*$.

Proof of the lemma. We first get some smoothness of h using the same trick as in Wang and Ziegel (2015). Integrating both sides of (3.7) over $q \in [0,1]$, we obtain for $t \in (0,1)$,

$$\begin{aligned} h(t) &= h(t)\int_0^1 h(1-(1-t)q)\,\mathrm{d}q + (1-h(t))\int_0^1 h(tq)\,\mathrm{d}q \\ &= \frac{h(t)}{1-t}\int_t^1 h(x)\,\mathrm{d}x + \frac{1-h(t)}{t}\int_0^t h(x)\,\mathrm{d}x \\ &= \frac{h(t)}{1-t}(g(1)-g(t)) + \frac{1-h(t)}{t}g(t), \end{aligned}$$

where $g(t) = \int_0^t h(x)\,\mathrm{d}x$. Rearranging terms, we have

$$h(t)\Big(t - \frac{t}{1-t}(g(1)-g(t)) + g(t)\Big) = g(t). \tag{3.8}$$

Note that the function g is continuous on $(0,1)$. For $t \in (0,1)$, if $g(t) \neq 0$, then (3.8) implies that h is continuous at t. If $g(t) = 0$ and h is not continuous at t, then $t - \frac{t}{1-t}(g(1)-g(t)) + g(t) = 0$, which implies $t = 1-g(1)$. To summarize, either h has a jump at $t = 1-g(1)$ and continuous elsewhere, or h is continuous on $(0,1)$. This fact implies that g is continuously differentiable on $(0,1)$ except for at the point $t = 1-g(1)$. Using the above relation (3.8) again, we know that h is continuously differentiable on $(0,1)$ except at the point $t = 1-g(1)$.

Differentiating both sides of (3.7) with respect to q, we get

$$0 = \frac{\mathrm{d}}{\mathrm{d}q}(h(t)h(1-q+qt) + (1-h(t))h(qt)).$$

By the product rule,

$$\frac{\mathrm{d}}{\mathrm{d}q}(h(t)h(1-q+qt)+(1-h(t))h(qt)) = h(t)h'(1-q+qt)(t-1)+(1-h(t))h'(qt)t, \quad (3.9)$$

assuming the derivatives in the right-hand-side of (3.9) exist. Plugging in $q = 1$ in (3.9) and rearranging terms, we have

$$h(t)h'(t) = h'(t)t \quad \text{ for all } t \in (0, 1) \setminus \{1 - g(1)\}.$$

As a consequence,

$$h(t) = t \text{ or } h'(t) = 0 \quad \text{for all } t \in (0, 1) \setminus \{1 - g(1)\}. \quad (3.10)$$

Pick any point $t_0 \in (0,1) \setminus \{1-g(1)\}$, and assume that $h'(t_0) \neq 0$ and $h'(t_0) \neq 1$. Using (3.10), we have $h(t_0) = t_0$. Since $h'(t_0) \neq 1$, there exists a neighbourhood $(t_0 - \varepsilon, t_0 + \varepsilon)$ such that $h(t) \neq t$ for all $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$ and $t \neq t_0$. Using (3.10) again, we know that $h'(t) = 0$ for all $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$ and $t \neq t_0$. The continuous differentiability of h at t_0 then implies $h'(t_0) = 0$, a contradiction. Therefore, we conclude

$$h'(t) = 1 \text{ or } h'(t) = 0 \quad \text{for all } t \in (0, 1) \setminus \{1 - g(1)\}. \quad (3.11)$$

First, suppose that h is continuously differentiable on $(0, 1)$. In this case, h' cannot switch between 0 and 1. Therefore, we have, either $h'(t) = 0$ on $(0, 1)$ or $h'(t) = 1$ on $(0, 1)$. This means either $h(t) = c$ on $(0, 1)$ for some constant $c \in [0, 1]$, or $h(t) = t$ on $(0, 1)$. In other words, $h \in \mathcal{H}_1^*$ or $h \in H_2^*$.

Next, suppose that h is not continuously differentiable at $t_0 = 1 - g(1)$. Note that this implies $g(t_0) = 0$, and hence $h(t) = 0$ a.e on $(0, t_0)$. Further, since $1 - t_0 = g(1) = \int_0^1 h(t)\,\mathrm{d}t = \int_{t_0}^1 h(t)\,\mathrm{d}t \leqslant 1-t_0$, we know that $h(t) = 1$ a.e. on $(t_0, 1)$. Since h is continuously differentiable on $(0, t_0)$ and $(t_0, 1)$, we know that $h(t) = 0$ on $(0, t_0)$ and $h(t) = 1$ on $(t_0, 1)$. Thus, $h \in \mathcal{H}_3^*$. □

The next lemma gives a sufficient condition for $h(t)$ to have the same sign as $h(1)$. Since h is not necessarily monotone for a signed Choquet integral, it is an important step to verify that $h(t)$ has the same sign as $h(1)$ in order to utilize Lemma 3.7.1.

Lemma 3.7.2. *Fix $h \in \mathcal{H}$ and $t \in (0, 1]$, and suppose $h(t) \neq 0$ and $h(1) \neq 0$. For $x, y \in \mathbb{R}$, where x and y satisfy $0 < x < y$ and $y = \left(1 - \frac{h(1)}{h(t)}\right) x + \frac{h(1)}{h(t)}$, if h satisfies*

$$h(1) = I_h(q((1 - t)\delta_x + t\delta_y) + (1 - q)\delta_1), \quad (3.12)$$

for all $q \in [0, 1]$, then $x < 1 \leqslant y$ and $\frac{h(1)}{h(t)} \geqslant 1$.

Proof of the lemma. Without loss of generality, we assume $h(1) = 1$. Because $y - x = \frac{1}{h(t)}(1-x)$ and $y - 1 = \left(1 - \frac{1}{h(t)}\right)(x-1)$, by the fact that $x < y$, there exist three cases.

(a) $x > 1, y > 1$ and $h(t) < 0$: Equation (3.12) reduces to

$$h(q)h(t) = h(tq). \tag{3.13}$$

(b) $x < 1, y < 1$ and $h(t) > 1$: Equation (3.12) reduces to

$$h(t) = h(1 - q(1-t)) + h(1-q)(h(t) - 1). \tag{3.14}$$

Then

$$h(t) = \frac{h(1 - q(1-t)) - h(1-q)}{1 - h(1-q)}.$$

(c) $x < 1, y \geqslant 1$ and $0 < h(t) \leqslant 1$: Equation (3.12) reduces to

$$h(t) = h(t)h(1 - q(1-t)) + (1 - h(t))h(tq). \tag{3.15}$$

We show that cases (a) and (b) above are actually not possible. In other words, a function $h \in \mathcal{H}$ satisfying $h(1) = 1$ and (a)-(c) takes values in $[0, 1]$.

We first show that such a function h is non-negative. Note that (c) implies

$$\text{for any } t \in [0,1], \text{ if } h(t) = 0, \text{ then } h(s) = 0 \text{ for all } s \in [0,t]. \tag{3.16}$$

Suppose that there exists $t \in [0,1]$ such that $h(t) < 0$. If $h(\sqrt{t}) < 0$, by (a), we have $h(t) = h(\sqrt{t})h(\sqrt{t}) > 0$, which is a contradiction. Hence, $h(\sqrt{t}) \geqslant 0$. Note that (3.13) also holds if $h(t) = 0$, due to (3.16). By (a), we know that $h(t)h(\sqrt{t}) = h(t\sqrt{t}) \leqslant 0$. Using (a) again,

$$0 \geqslant h(t)h(\sqrt{t})h(\sqrt{t}) = h(t\sqrt{t})h(\sqrt{t}) = h(t^2).$$

On the other hand, (a) also gives

$$h(t^2) = h(t)h(t) > 0,$$

a clear contradiction. Therefore, $h(t) \geqslant 0$ for all $t \in [0,1]$.

Next we show $h(t) \leqslant 1$ for all $t \in [0,1]$. We note the following two useful facts. First, for any $t \in [0,1]$ such that $h(t) \in (0,1]$, by (c), we have

$$1 = h(1 - q(1-t)) + \frac{1 - h(t)}{h(t)}h(tq).$$

Using the fact that h is non-negative, we have $h(1-q(1-t)) \leqslant 1$ for all $q \in [0,1]$. Therefore, we conclude the following statement:

$$\text{For any } t \in [0,1] \text{ such that } h(t) \in (0,1],\ h(s) \in [0,1] \text{ for all } s \in [t,1]. \tag{3.17}$$

Second, for any $t \in [0,1]$ such that $h(t) > 1$, by taking $q = 1-t$, (3.14) gives

$$h(t) = h(1-(1-t)^2) + h(t)(h(t)-1).$$

Rearranging terms, we have $(h(t)-1)^2 = 1-h(1-(1-t)^2)$. This implies $h(1-(1-t)^2) < 1$. Therefore, we have:

$$\text{For any } t \in [0,1] \text{ such that } h(t) > 1,\ h(1-(1-t)^2) < 1. \tag{3.18}$$

Suppose that there exists $t \in (0,1)$ such that $h(t) > 1$. Let $s = 1-\sqrt{1-t}$. Clearly, $s < t$. If $h(s) > 1$, then by (3.18), we have $h(1-(1-s)^2) = h(t) < 1$, a contradiction. If $h(s) \in (0,1]$, then by (3.17), $h(t) \in [0,1]$, another contradiction. Hence, $h(s) = 0$.

Plugging $q = 1-s = \sqrt{1-t}$ in (3.14),

$$h(t) = h(1-(1-t)(1-s)) + h(s)(h(t)-1)) = h(1-(1-t)\sqrt{1-t}) > 1. \tag{3.19}$$

In particular,

$$h(t) = h(1-(1-t)\sqrt{1-t}) > 1, \tag{3.20}$$

and setting $w = 1-(1-t)\sqrt{1-t}$ in (3.19), we get

$$h(1-(1-w)\sqrt{1-t}) = h(1-(1-t)^2) > 1.$$

This is a contraction to (3.18). Combining all cases, there does not exist $t \in (0,1)$ such that $h(t) > 1$. Together with the fact that h is non-negative, we come to the conclusion that $h(t) \in [0,1]$ for all $t \in [0,1]$. Thus only case (c) is possible. □

Proof of Theorem 3.3.1 continued. It is easy to verify that the three classes of functionals in (3.2)-(3.4) have CxLS, and hence $h \in \mathcal{H}^*$ is sufficient for the CxLS property. Below we show the necessity of $h \in \mathcal{H}^*$. Suppose $h(t_0) = 0$ for some fixed $t_0 \in [0,1]$. Observe that $I_h(\delta_0) = 0$ and $I_h((1-t_0)\delta_0 + t_0\delta_1) = h(t_0) = 0$. Since I_h has CxLS, for any fixed $q \in [0,1]$,

$$I_h\left((1-q)\delta_0 + q(1-t_0)\delta_0 + qt_0\delta_1\right) = h(t_0 q) = I_h(\delta_0) = 0. \tag{3.21}$$

It follows that if $h(1) = 0$, then $h(t) = 0$ on $[0,1]$. This is included in each of cases (i)-(iii). In the following, $h(t) \neq 0$ for any $t \in (0,1]$, and we can assume $h(1) = 1$ without loss of generality, since the set $\mathcal{H}^*$ is invariant under a constant multiplier. For $0 < x < y$ and any fixed $t \in (0,1]$, we have

$$I_h((1-t)\delta_x + t\delta_y) = x + h(t)(y-x).$$

Note that $I_h(\delta_1) = h(1) = 1$. In the following we choose $y = \left(1 - \frac{1}{h(t)}\right) x + \frac{1}{h(t)}$, so that $I_h((1-t)\delta_x + t\delta_y) = 1$. Since I_h has CxLS, for all $q \in [0,1]$,

$$1 = I_h(q((1-t)\delta_x + t\delta_y) + (1-q)\delta_1). \tag{3.22}$$

By Lemma 3.7.2, we have $x < 1 \leqslant y$ and $h(t) \in (0,1]$. Hence, (3.22) reduces to

$$h(t) = h(t)h(1-q+qt) + (1-h(t))h(qt), \tag{3.23}$$

for all $t \in (0,1]$ with $h(t) \neq 0$ and $q \in [0,1]$. Note that (3.22) holds for $t = 0$ and if $h(t) = 0$, then (3.23) holds automatically by (3.21). Therefore, (3.23) holds for all $t, q \in [0,1]$. This is precisely (3.7). By applying Lemma 3.7.1, we obtain $h \in \mathcal{H}^*$. $\square$

Proof of Proposition 3.3.3. For both statements, we note that $F_n \xrightarrow{\text{w}} F$ if and only if $F_n^{-1}(U) \to F^{-1}(U)$ in L^∞ for any $\mathrm{U}[0,1]$ random variable U.

(i) Let $A \in \mathcal{F}$ with $\mathbb{P}(A) = \lambda$ and $U \sim \mathrm{U}[0,1]$ be independent. It is easy to see that the random variable $\mathbf{1}_A F_n^{-1}(U) + \mathbf{1}_{A^c} G_n^{-1}(U)$ has the distribution $\lambda F_n + (1-\lambda)G_n$, $n \in \mathbb{N}$. Moreover, by the w-convergence of $\{F_n\}_{n\in\mathbb{N}}$ and $\{G_n\}_{n\in\mathbb{N}}$, $\mathbf{1}_A F_n^{-1}(U) + \mathbf{1}_{A^c} G_n^{-1}(U) \to \mathbf{1}_A F^{-1}(U) + \mathbf{1}_{A^c} G^{-1}(U)$ in L^∞. Therefore, $\lambda F_n + (1-\lambda)G_n \xrightarrow{\text{w}} \lambda F + (1-\lambda)G$.

(ii) The conclusion follows from the fact that a Signed Choquet integral as a functional on L^∞ is uniformly continuous with respect to L^∞-norm (Theorem 2.1 of Wang et al. (2018)). $\square$

Proof of Proposition 3.3.4. (i) Assume $h(1) = 0$. Let $F_n = \mathrm{U}[0, 1/n]$ and $G_n = \mathrm{U}[1, 1 + 1/n]$, $n \in \mathbb{N}$. Note that $F_n \xrightarrow{\text{w}} \delta_0$ and $G_n \xrightarrow{\text{w}} \delta_1$. By Proposition 3.3.3, we have $\lambda F_n + (1-\lambda)G_n \xrightarrow{\text{w}} \mathrm{Bernoulli}(\lambda)$ for $\lambda \in (0,1)$. Also note that from the translation invariance of I_h (e.g. Lemma 2.4 of Wang et al. (2018)), $I_h(G_n) = I_h(F_n) + h(1) = I_h(F_n)$. The CxLS on $\mathcal{M}_{\text{con}}$ implies $I_h(\lambda F_n + (1-\lambda)G_n) = I_h(F_n)$. Therefore, by Proposition 3.3.3,

$$I_h(\mathrm{Bernoulli}(\lambda)) = \lim_{n\to\infty} I_h(\lambda F_n + (1-\lambda)G_n) = \lim_{n\to\infty} I_h(F_n) = I_h(\delta_0) = 0.$$

Also note that $I_h(\mathrm{Bernoulli}(\lambda)) = h(\lambda)$. This gives $h(t) = 0$ for $t \in [0,1]$. Hence $I_h(F) = 0$ for all $F \in \mathcal{M}_0$, and it has CxLS on any set of distributions.

(ii) Assume $h(1) \neq 0$. Take $F, G \in \mathcal{M}_{\text{dis}}$ such that $I_h(F) = I_h(G)$. Write $c = h(1)$ and $b = I_h(F)$. Take two sequences of distributions $\{F_n\}_{n\in\mathbb{N}} \subset \mathcal{M}_{\text{con}}$ and $\{G_n\}_{n\in\mathbb{N}} \subset \mathcal{M}_{\text{con}}$ such that $F_n \xrightarrow{\text{w}} F$ and $G_n \xrightarrow{\text{w}} G$. Such sequences are easy to construct by, e.g., replacing each point mass with a uniform on a small interval of length $1/n$. Let U be a U$[0,1]$ random variable. For $n \in \mathbb{N}$, let F_n^* be the distribution of $F_n^{-1}(U) + \frac{1}{c}(b - I_h(F_n))$. We have $F_n^* \xrightarrow{\text{w}} F$, because $F_n \xrightarrow{\text{w}} F$ and $I_h(F_n) \to b$ by Proposition 3.3.3. Moreover, by the translation invariance of I_h again, $I_h(F_n^*) = I_h(F_n) + b - I_h(F_n) = b$. Similarly, let G_n^* be the distribution of $G_n^{-1}(U) + \frac{1}{c}(b - I_h(G_n))$, and we have $G_n^* \xrightarrow{\text{w}} G$ and $I_h(G_n^*) = b$. By Proposition 3.3.3, we have $\lambda F_n^* + (1-\lambda)G_n^* \xrightarrow{\text{w}} \lambda F + (1-\lambda)G$. Finally, noting that $F_n^*, G_n^* \in \mathcal{M}_{\text{con}}$, the CxLS on $\mathcal{M}_{\text{con}}$ implies that $I_h(\lambda F_n^* + (1-\lambda)G_n^*) = I_h(F_n^*) = b$. Therefore, by Proposition 3.3.3 again,

$$I_h(\lambda F + (1-\lambda)G) = \lim_{n\to\infty} I_h(\lambda F_n^* + (1-\lambda)G_n^*) = b = I_h(F) = I_h(G).$$

Hence, I_h has CxLS on $\mathcal{M}_{\text{dis}}$. $\square$

3.7.2 Proofs in Section 3.4

Proof of Proposition 3.4.1. Fix an $m \in \mathbb{N}$. The functional ρ can be written as $\rho(\cdot) = h(f_1(\cdot), \cdots, f_m(\cdot))$, where $h : \mathbb{R}^m \to \mathbb{R}^d$ is an injective function and $(f_i)_{1\leqslant i\leqslant m} : \mathcal{M} \to \mathbb{R}$ has CxLS. For any $F, G \in \mathcal{M}$ that satisfy $\rho(F) = \rho(G)$, we have $(f_1(F), \cdots, f_m(F)) = (f_1(G), \cdots, f_m(G))$ by the fact that h is an injection. Because each f_i has CxLS, then $(f_1(\lambda F + (1-\lambda)G), \cdots, f_m(\lambda F + (1-\lambda)G)) = (f_1(F), \cdots, f_m(F))$ for any $\lambda \in [0,1]$. So $\rho(\lambda F + (1-\lambda)G) = h(f_1(\lambda F + (1-\lambda)G), \cdots, f_m(\lambda F + (1-\lambda)G)) = h(f_1(F), \cdots, f_m(F)) = \rho(F) = \rho(G)$ for any $\lambda \in [0,1]$. $\square$

Proof of Proposition 3.4.2. We only need to show the "if" direction. For any $F, G \in \mathcal{M}$, if $(\rho_1(F), \rho_2(F)) = (\rho_1(G), \rho_2(G)) = (r_1, r_2)$, then $\rho_2(\lambda F + (1-\lambda)G) = r_2$ for any $\lambda \in [0,1]$. Since ρ_1 has CxLS on $\mathcal{M}(r_2)$, $\rho_1(\lambda F + (1-\lambda)G) = r_2$ for any $\lambda \in [0,1]$, which means ρ_1 has CxLS on $\mathcal{M}$. $\square$

3.7.3 Proofs in Section 3.5

Proof of Theorem 3.5.2. In order to prove Theorem 3.5.2, we need the following technical lemma, which connects the problem on dimension two with the result in dimension one.

Lemma 3.7.3. *For $p \in (0,1)$ and $h \in \mathcal{H}$, if (I_h, VaR_p) has CxLS on $\mathcal{M}_{\mathrm{dis}}$, then*

$$h(t) = h_1\left(\frac{t}{1-p}\right)\mathbf{1}_{\{t \leqslant 1-p\}} + \left(h_2\left(\frac{1-t}{p}\right) + h(1)\right)\mathbf{1}_{\{t>1-p\}}, \tag{3.24}$$

for some h_1 and $h_2 \in \mathcal{H}^*$.

Proof of the lemma. Define $h_1, h_2 : [0,1] \to \mathbb{R}$ by $h_1(t) = h(t(1-p))$, $t \in [0,1]$ and $h_2(t) = h(1-tp) - h(1)$, $t \in [0,1)$. Further, let $h_2(1) = \lim_{t\uparrow 1} h_2(t)$, so that h_2 is continuous at $t = 1$. Clearly, $h_1, h_2 \in \mathcal{H}$ and (3.24) holds. We shall show $h_1, h_2 \in \mathcal{H}^*$ below.

Let $\mathcal{M}_+ = \{F \in \mathcal{M}_{\mathrm{dis}} : F((-\infty, 0)) = 0\}$ which is the set of distributions supported on $[0,\infty)$. Further, let $\mathcal{M}^* = \{p\delta_0 + (1-p)F : F \in \mathcal{M}_+\}$. Note that $\mathrm{VaR}_p(F) = 0$ for $F \in \mathcal{M}^*$ by definition. Therefore, by Proposition 3.4.2, I_h has CxLS on $\mathcal{M}^*$. By definition of I_h, for $F \in \mathcal{M}_+$,

$$I_h(p\delta_0 + (1-p)F) = \int_0^\infty h(1-(p+(1-p)F(x)))\,\mathrm{d}x = \int_0^\infty h_1(1-F(x))\,\mathrm{d}x = I_{h_1}(F).$$

Note that in the above equation we treat the measure F as the corresponding cumulative distribution function. Since I_h has CxLS on $\mathcal{M}^*$, and there is a linear mapping between $\mathcal{M}^*$ and $\mathcal{M}_+$, we know that I_{h_1} has CxLS on $\mathcal{M}_+$.

For $G \in \mathcal{M}_{\mathrm{dis}}$, we can write $G = T_xF$ for some $x \in \mathbb{R}$ and $F \in \mathcal{M}_+$, where T_x is the operator of left-shift by $x \in \mathbb{R}$, that is, $T_xF(y) = F(y+x)$ for $y \in \mathbb{R}$. By definition of a signed Choquet integral (or, see Lemma 2.4 of Wang et al. (2018)), we have $I_{h_1}(G) = I_{h_1}(T_xF) = I_{h_1}(F) - xh_1(1)$. Therefore, if $G_1, G_2 \in \mathcal{M}_{\mathrm{dis}}$ satisfy $I_{h_1}(G_1) = I_{h_2}(G_2)$, then for some $x \in \mathbb{R}$ such that $T_{-x}G_1, T_{-x}G_2 \in \mathcal{M}_+$, we know $I_{h_1}(T_{-x}G_1) = I_{h_2}(T_{-x}G_2)$. Since I_{h_1} has CxLS on $\mathcal{M}_+$,

$$\begin{aligned} I_{h_1}(\lambda G_1 + (1-\lambda)G_2) &= I_{h_1}(\lambda T_{-x}G_1 + (1-\lambda)T_{-x}G_2) + xh_1(1) \\ &= I_{h_1}(T_{-x}G_1) + xh_1(1) = I_{h_1}(G_1). \end{aligned}$$

As a consequence, I_{h_1} has CxLS on $\mathcal{M}_{\mathrm{dis}}$. By Theorem 3.3.1, we know $h_1 \in \mathcal{H}^*$.

The statement for h_2 is somehow more complicated as it is not symmetric to the case of h_1. Fix $q \in (0,p)$, and let $\mathcal{M}_- = \{F \in \mathcal{M}_{\mathrm{dis}} : F((-\infty, 0]) = 1\}$, $g_q(t) = h(1-tq) - h(1)$, $t \in [0,1]$, and $\mathcal{M}_q^* = \{(1-q)\delta_0 + qF : F \in \mathcal{M}_-\}$. Note that $\mathrm{VaR}_p(F) = 0$ for $F \in \mathcal{M}_q^*$ by definition. Therefore, by Proposition 3.4.2, I_h has CxLS on $\mathcal{M}_q^*$. Let $\bar{g}_q(t) = g_q(1-t) -$

$g_q(1)$, $t \in [0,1]$. Note that $\bar{g}_q \in \mathcal{H}$ and $\bar{g}_q(1) = -g_q(1)$. By definition of I_h, for $F \in \mathcal{M}_-$,

$$\begin{aligned} I_h((1-q)\delta_0 + qF) &= \int_{-\infty}^{0} (h(1-qF(x)) - h(1))\,\mathrm{d}x \\ &= \int_{-\infty}^{0} g_q\,(F(x))\,\mathrm{d}x = \int_{-\infty}^{0} (\bar{g}_q\,(1-F(x)) - \bar{g}_q(1))\,\mathrm{d}x = I_{\bar{g}_q}(F). \end{aligned}$$

Since I_h has CxLS on $\mathcal{M}_q^*$, and there is a linear mapping between $\mathcal{M}_q^*$ and $\mathcal{M}_-$, we know that I_{g_q} has CxLS on $\mathcal{M}_-$. Following the similar arguments for h_1, we obtain $\bar{g}_q \in \mathcal{H}^*$. Checking the three forms of functions in $\mathcal{H}^*$, we know that $g_q \in \mathcal{H}^*$. Note that q is arbitrarily chosen in $(0,p)$, and $h_2(t) = g_q(pt/q) = g_{tp}(1)$ for $t \leqslant q/p$. If $g_q(1) = 0$ for all $q \in (0,1)$, then h_2 is zero on $(0,1)$, thus $h_2 \in \mathcal{H}_1^*$. Next, assume that there exists $q_0 \in (0,p)$ such that $g_{q_0}(1) = c \neq 0$. There are four cases to analyze. If $g_{q_0} \in \mathcal{H}_1^*$ and $g_{q_0}(1-) = 0$, i.e. h_2 is zero on $(0, q_0/p)$, then, by letting q vary in (q_0,p), constrained by $g_q \in \mathcal{H}^*$, h_2 must be equal to a constant d on $(q_0/p, 1)$ with $dc > 0$ and $|d| \geqslant |c|$, thus $h_2 \in \mathcal{H}_3^*$. If $g_{q_0} \in \mathcal{H}_1^*$ and $g_{q_0}(1-) \neq 0$, i.e. h_2 is a non-zero constant on $(0, q_0/p)$, then, by letting q vary in (q_0,p), constrained by $g_q \in \mathcal{H}^*$, h_2 must be equal to c on $(q_0/p, 1)$, thus $h_2 \in \mathcal{H}_1^*$. If $g_{q_0} \in \mathcal{H}_2^*$, i.e. h_2 is linear on $(0, q_0/p)$, then, by letting q vary in (q_0,p), constrained by $g_q \in \mathcal{H}^*$, h_2 must be linear on $(0,1)$, thus $h_2 \in \mathcal{H}_2^*$. If $g_{q_0} \in \mathcal{H}_3^*$, i.e. there is a jump of h_2 in $(0, q_0/p)$, then, by letting q vary in (q_0,p), constrained by $g_q \in \mathcal{H}^*$, h_2 must be equal to c on $(q_0/p, 1)$, thus $h_2 \in \mathcal{H}_3^*$. In all cases, $h_2 \in \mathcal{H}^*$. □

Proof of Theorem 3.5.2 continued. For $r \in \mathbb{R}$, denote by $\mathcal{M}(r) = \{F \in \mathcal{M}_{\mathrm{dis}} : \mathrm{VaR}_p(F) = r\}$. We first verify (I_h, VaR_p) in cases (i)-(iv) indeed has CxLS.

(i) Both VaR_p and I_{h^*} have CxLS by Theorem 3.3.1. Hence, (I_h, VaR_p) has CxLS as justified by Proposition 3.4.1.

(ii) For $F \in \mathcal{M}(r)$, using the ES-VaR formula of Rockafellar and Uryasev (2002),

$$I_h(F) = ar + b\int_{-\infty}^{\infty} x\,\mathrm{d}F(x) + c\left(r + \frac{1}{1-p}\int_{r}^{\infty}(x-r)\,\mathrm{d}F(x)\right).$$

Hence, I_h is linear for $F \in \mathcal{M}(r)$. So I_h has CxLS on $\mathcal{M}(r)$. By Proposition 3.4.2, we know that (I_h, VaR_p) has CxLS.

(iii) For $F, G \in \mathcal{M}(r)$, if $I_h(F) = I_h(G)$, then $b\mathrm{VaR}_p^+(F) + c\,\mathrm{ess\,sup}(F) = b\mathrm{VaR}_p^+(G) + c\,\mathrm{ess\,sup}(G)$. Without loss of generality, assume $\mathrm{VaR}_p^+(F) \geqslant \mathrm{VaR}_p^+(G)$, which implies $\mathrm{ess\,sup}(F) \leqslant \mathrm{ess\,sup}(G)$ since $bc > 0$. Note that for $\lambda \in (0,1)$, since $\mathrm{VaR}_p(F) =$

$\mathrm{VaR}_p(G) = \mathrm{VaR}_p(\lambda F + (1-\lambda G)) = r$, we have $\mathrm{VaR}_p^+(\lambda F + (1-\lambda)G) = \mathrm{VaR}_p^+(G)$. Therefore, for $\lambda \in (0,1)$.

$$b\mathrm{VaR}_p^+(\lambda F + (1-\lambda)G) + c \operatorname{ess\,sup}(\lambda F + (1-\lambda)G) = b\mathrm{VaR}_p^+(G) + c \operatorname{ess\,sup}(G).$$

Hence, I_h has CxLS on $\mathcal{M}(r)$. By Proposition 3.4.2, we know that (I_h, VaR_p) has CxLS.

(iv) For $F, G \in \mathcal{M}(r)$, if $I_h(F) = I_h(G)$, then $b\mathrm{VaR}_p^+(F) + c \operatorname{ess\,inf}(F) = b\mathrm{VaR}_p^+(G) + c \operatorname{ess\,inf}(G)$. Without loss of generality, assume $\mathrm{VaR}_p^+(F) \geqslant \mathrm{VaR}_p^+(G)$, which implies $\operatorname{ess\,inf}(F) \geqslant \operatorname{ess\,inf}(G)$ since $bc < 0$. Note that for $\lambda \in (0,1)$, since $\mathrm{VaR}_p(F) = \mathrm{VaR}_p(G) = \mathrm{VaR}_p(\lambda F + (1-\lambda G)) = r$, we have $\mathrm{VaR}_p^+(\lambda F + (1-\lambda)G) = \mathrm{VaR}_p^+(G)$. Therefore, for $\lambda \in (0,1)$.

$$b\mathrm{VaR}_p^+(\lambda F + (1-\lambda)G) + c \operatorname{ess\,inf}(\lambda F + (1-\lambda)G) = b\mathrm{VaR}_p^+(G) + c \operatorname{ess\,inf}(G).$$

Hence, I_h has CxLS on $\mathcal{M}(r)$. By Proposition 3.4.2, we know that (I_h, VaR_p) has CxLS.

Next, suppose that (I_h, VaR_p) has CxLS, and we show that it has to be one of the cases (i)-(iv). To simply notation, for $p \in (0,1)$ and $c \in [0,1]$, let

$$\mathrm{ES}_p^-(F) = \frac{1}{p}\int_0^p \mathrm{VaR}_t(F)\,\mathrm{d}t, \quad F \in \mathcal{M}_\infty,$$

and

$$Q_p^c(F) = c \operatorname{ess\,sup}(F) + (1-c)\mathrm{VaR}_p^+(F), \quad F \in \mathcal{M}_\infty.$$

Note that $Q_p^1 = \operatorname{ess\,sup}$ and $Q_p^0 = \mathrm{VaR}_p^+$.

By Lemma 3.7.3, h satisfies (3.24), that is, for some h_1 and $h_2 \in \mathcal{H}^*$,

$$h(t) = h_1\left(\frac{t}{1-p}\right)\mathbf{1}_{\{t\leqslant 1-p\}} + \left(h_2\left(\frac{1-t}{p}\right) + h(1)\right)\mathbf{1}_{\{t>1-p\}} = g_1(t) + g_2(t),$$

where $g_1(t) = h_1(\frac{t}{1-p})\mathbf{1}_{\{t\leqslant 1-p\}}$ and $g_2(t) = (h_2(\frac{1-t}{p}) + h(1))\mathbf{1}_{\{t>1-p\}}$, $t \in [0,1]$. Analyzing all possible forms of I_{g_1} and I_{g_2}, we have

$$I_h = I_{g_1} + I_{g_2} = a\mathrm{VaR}_p + a_1 I + a_2 J,$$

where $I \in \{\mathrm{VaR}_\alpha^{c_1}, \mathrm{ES}_p^-\}$ and $J \in \{\mathrm{ES}_p, \mathrm{VaR}_\beta^{c_2}, Q_p^c\}$, $a, a_1, a_2 \in \mathbb{R}$, $\alpha \in [0,p)$, $\beta \in (p,1]$ and $c_1, c_2, c \in [0,1]$. Since (I_h, VaR_p) has CxLS if and only if $(I_h - a\mathrm{VaR}_p, \mathrm{VaR}_p)$ has

CxLS, we can freely set $a = 0$. Hence, we can write $I_h = a_1 I + a_2 J$. Without loss of generality we assume $a_1 \geqslant 0$; otherwise we can replace I_h by $-I_h$. There are a few cases to analyze. Below, we use the fact that ES_p^- is a linear combination of ES_p and $\mathbb{E}$ via $p\mathrm{ES}_p^- + (1-p)\mathrm{ES}_p = \mathbb{E}$.

(a) $a_1 = 0$. The case $J \in \{\mathrm{VaR}_\beta^{c_2}, \mathrm{VaR}_p^+, \operatorname{ess\,sup}\}$ is included in case (i); the case $J = \mathrm{ES}_p$ is included in case (ii), and the case $J = Q_p^c$ for $c \in (0,1)$ is included in case (iii).

(b) $a_2 = 0$. The case $I = \mathrm{VaR}_\alpha^{c_1}$ is included in case (i) and $I = \mathrm{ES}_p^-$ is included in case (ii).

(c) $a_1 > 0, a_2 > 0$. We claim that if I_h has CxLS on $\mathcal{M}(0)$, then either $(I, J) = (\operatorname{ess\,inf}, \operatorname{ess\,sup})$, included in case (i), or $(I, J) = (\mathrm{ES}_p^-, \mathrm{ES}_p)$, included in case (ii). Below we show our assertion.

First, suppose that $I = \mathrm{VaR}_\alpha^{c_1}$. For $\varepsilon \in (\alpha, p)$, let $F = \varepsilon\delta_{-a_2} + (p-\varepsilon)\delta_0 + (1-p)\delta_{a_1}$ and $G = \delta_0$. We can easily calculate $\mathrm{VaR}_p(F) = \mathrm{VaR}_p(G) = 0$, $I(F) = -a_2$, $J(F) = a_1$ and $I(G) = J(G) = 0$. Therefore, $I_h(F) = I_h(G) = 0$. For $\lambda \in [0,1]$,

$$\lambda F + (1-\lambda)G = \lambda\varepsilon\delta_{-a_2} + (1 - \lambda + \lambda(p-\varepsilon))\delta_0 + \lambda(1-p)\delta_{a_1}.$$

If I_h has CxLS on $\mathcal{M}(0)$, then $I_h(\lambda F + (1-\lambda)G) = 0$ for all $\lambda \in [0,1]$ and all $\varepsilon \in (\alpha, p)$. Note that the function $\lambda \mapsto I(\lambda F + (1-\lambda)G)$ has a jump at $\lambda_1 = \alpha/\varepsilon \in [0,1)$, and the function $\lambda \mapsto J(\lambda F + (1-\lambda)G)$ either has no jump ($J = \mathrm{ES}_p$), a jump at $\lambda_2 = 1$ ($J = Q_p^c$ for $c \neq 0$), a jump at $\lambda_2 = 0$ ($J = \operatorname{ess\,sup} = Q_p^0$), or a jump at $\lambda_2 = (1-\beta)/(1-p)$ ($J = \mathrm{VaR}_\beta^c$). Note that $\lambda_1 \neq \lambda_2$ for almost every $\varepsilon \in (\alpha, p)$, except for the case $(\alpha, c) = (0, 0)$. Hence, except for $(I, J) = (\operatorname{ess\,inf}, \operatorname{ess\,sup})$, the function $\lambda \mapsto I_h(\lambda F + (1-\lambda)G)$ does not take a constant value on $[0,1]$, and I_h cannot have CxLS on $\mathcal{M}(0)$.

Next, suppose that $I = \mathrm{ES}_p^-$. For $\varepsilon \in (0, p)$, let $F = \varepsilon\delta_{-a_3} + (p-\varepsilon)\delta_0 + (1-p)\delta_{a_1}$ and $G = \delta_0$, where $a_3 = pa_2/\varepsilon$. We can easily calculate $\mathrm{VaR}_p(F) = \mathrm{VaR}_p(G) = 0$, $I(F) = -\varepsilon a_3/p = -a_2$, $J(F) = a_1$ and $I(G) = J(G) = 0$. Therefore, $I_h(F) = I_h(G) = 0$. For $\lambda \in [0,1]$,

$$\lambda F + (1-\lambda)G = \lambda\varepsilon\delta_{-a_3} + (1 - \lambda + \lambda(p-\varepsilon))\delta_0 + \lambda(1-p)\delta_{a_1}.$$

If I_h has CxLS on $\mathcal{M}(0)$, then $I_h(\lambda F + (1-\lambda)G) = 0$ for all $\lambda \in [0,1]$. Note that the function $\lambda \mapsto I(\lambda F + (1-\lambda)G)$ has no jump whereas the function $\lambda \mapsto J(\lambda G + (1-\lambda)F)$ has a jump on $[0,1]$ except for $J = \mathrm{ES}_p$. Hence, except for $J = \mathrm{ES}_p$, the function $\lambda \mapsto I_h(\lambda F + (1-\lambda)G)$ does not take a constant value on $[0,1]$, and I_h cannot have CxLS on $\mathcal{M}(0)$.

(d) $a_1 > 0, a_2 < 0$. We claim that if I_h has CxLS on $\mathcal{M}(0)$, then either $(I,J) = (\operatorname{ess\,inf}, \mathrm{VaR}_p^+)$, included in case (iv), or $(I,J) = (\mathrm{ES}_p^-, \mathrm{ES}_p)$, included in case (ii). Below we show our assertion.

First, suppose that $I = \mathrm{VaR}_\alpha^{c_1}$. For $\varepsilon \in (\alpha, p)$, let $F = \varepsilon\delta_{a_2} + (1-\varepsilon)\delta_0$ and $G = p\delta_0 + (1-p)\delta_{a_1}$. We can easily calculate $\mathrm{VaR}_p(F) = \mathrm{VaR}_p(G) = 0$, $I(F) = a_2$, $J(F) = 0$, $I(G) = 0$ and $J(G) = a_1$. Therefore, $I_h(F) = I_h(G) = a_1a_2$. Note that, for $\lambda \in [0,1]$,

$$\lambda F + (1-\lambda)G = \lambda\varepsilon\delta_{a_2} + (\lambda(1-\varepsilon) + (1-\lambda)p)\delta_0 + (1-\lambda)(1-p)\delta_{a_1}.$$

If I_h has CxLS on $\mathcal{M}(0)$, then $I_h(\lambda F + (1-\lambda)G) = a_1a_2$ for all $\lambda \in [0,1]$ and all $\varepsilon \in (\alpha, p)$. Note that the function $\lambda \mapsto I(\lambda F + (1-\lambda)G)$ has a jump at $\lambda_1 = \alpha/\varepsilon \in [0,1)$, and the function $\lambda \mapsto J(\lambda G + (1-\lambda)F)$ either has no jump ($J = \mathrm{ES}_p$), a jump at $\lambda_2 = 1$ ($J = Q_p^c$ for $c \neq 1$), a jump at $\lambda_2 = 0$ ($J = \mathrm{VaR}_p^+ = Q_p^1$), or a jump at $\lambda_2 = 1 - (1-\beta)/(1-p)$ ($J = \mathrm{VaR}_\beta^c$). Note that $\lambda_1 \neq \lambda_2$ for almost every $\varepsilon \in (\alpha, p)$, except for the case $(\alpha, c) = (0,1)$. Hence, except for $(I,J) = (\operatorname{ess\,inf}, \mathrm{VaR}_p^+)$, the function $\lambda \mapsto I_h(\lambda F + (1-\lambda)G)$ does not take a constant value on $[0,1]$, and I_h cannot have CxLS on $\mathcal{M}(0)$.

Next, suppose that $I = \mathrm{ES}_p^-$. For $\varepsilon \in (0,p)$, let $F = \varepsilon\delta_{a_3} + (1-\varepsilon)\delta_0$ and $G = p\delta_0 + (1-p)\delta_{a_1}$, where $a_3 = pa_2/\varepsilon$. We can easily calculate $\mathrm{VaR}_p(F) = \mathrm{VaR}_p(G) = 0$, $I(F) = \varepsilon a_3/p = a_2$, $J(F) = 0$, $I(G) = 0$ and $J(G) = a_1$. Therefore, $I_h(F) = I_h(G) = a_1a_2$. For $\lambda \in [0,1]$,

$$\lambda F + (1-\lambda)G = \lambda\varepsilon\delta_{a_3} + (\lambda(1-\varepsilon) + (1-\lambda)p)\delta_0 + (1-\lambda)(1-p)\delta_{a_1}.$$

If I_h has CxLS on $\mathcal{M}(0)$, then $I_h(\lambda F + (1-\lambda)G) = 0$ for all $\lambda \in [0,1]$. Note that the function $\lambda \mapsto I(\lambda F + (1-\lambda)G)$ has no jump whereas the function $\lambda \mapsto J(\lambda G + (1-\lambda)F)$ has a jump except for $J = \mathrm{ES}_p$. Hence, except for $J = \mathrm{ES}_p$, the function $\lambda \mapsto I_h(\lambda F + (1-\lambda)G)$ does not take a constant value on $[0,1]$, and I_h cannot have CxLS on $\mathcal{M}(0)$.

To summarize our findings in (a)-(d), all (I_h, VaR_p) with CxLS are included in cases (i)-(iv). $\square$

Proof of Proposition 3.5.3. By definition, it is easy to verify that $I_h(F) = I_{\bar{h}}(\bar{F})$ and $\mathrm{VaR}_p(F) = -\mathrm{VaR}_{1-p}^+(\bar{F})$ for all $F \in \mathcal{M}$ (or, see Lemma 2.5 of Wang et al. (2018)). Therefore, $F, G \in \mathcal{M}$ satisfy $(I_h(F), \mathrm{VaR}_p(F)) = (I_h(G), \mathrm{VaR}_p(G))$ if and only if $\bar{F}, \bar{G} \in \bar{\mathcal{M}}$ satisfy $(I_{\bar{h}}(\bar{F}), \mathrm{VaR}_{1-p}^+(\bar{F})) = (I_{\bar{h}}(\bar{G}), \mathrm{VaR}_{1-p}^+(\bar{G}))$. Hence, the CxLS property of (I_h, VaR_p) on $\mathcal{M}$ and that of $(I_{\bar{h}}, \mathrm{VaR}_{1-p}^+)$ on $\bar{\mathcal{M}}$ are equivalent. $\square$

3.7.4 Proofs in Section 3.6

Proof of Proposition 3.6.2. We define $S : \mathbb{R}^2 \to \mathbb{R}$ by letting $S(z, y) = \int_0^z Z(x, y)\,\mathrm{d}x$. Observe that for any $F \in \mathcal{M}$, Fubini's Theorem gives

$$\int_{-\infty}^{\infty} S(z, y)\,\mathrm{d}F(y) = \int_{-\infty}^{\infty}\int_0^z Z(x, y)\,\mathrm{d}x\,\mathrm{d}F(y) = \int_0^z \int_{-\infty}^{\infty} Z(x, y)\,\mathrm{d}F(y)\,\mathrm{d}x.$$

Then we differentiate the above equation with respect to z,

$$\frac{\mathrm{d}}{\mathrm{d}z}\left(\int_0^z \int_{-\infty}^{\infty} Z(x, y)\,\mathrm{d}F(y)\,\mathrm{d}x\right) = \int_{-\infty}^{\infty} Z(z, y)\,\mathrm{d}F(y).$$

Since Z is a backtest function of ρ,

$$\int_{-\infty}^{\infty} Z(z, y)\,\mathrm{d}F(y) = 0 \text{ if and only if } \rho(F) = z,$$

and the following two inequalities hold,

$$\int_{-\infty}^{\infty} Z(z, y)\,\mathrm{d}F(y) < \int_{-\infty}^{\infty} Z(\rho(F), y)\,\mathrm{d}F(y) = 0 \text{ for } z < \rho(F),$$

and

$$\int_{-\infty}^{\infty} Z(t, y)\,\mathrm{d}F(y) > \int_{-\infty}^{\infty} Z(\rho(F), y)\,\mathrm{d}F(y) = 0 \text{ for } z > \rho(F).$$

Thus, $\int_{-\infty}^{\infty} S(z, y)\,\mathrm{d}F(y)$ achieves the global minimum at and only at $\rho(F) = z$. Hence S is strictly consistent for ρ and ρ is $\mathcal{M}$-elicitable. $\square$

Proof of Proposition 3.6.3. For $F \in \mathcal{M}$, let $z = c\rho(F)$, we have

$$\int_{-\infty}^{\infty} Z^*(z, y)\,\mathrm{d}F(y) = \int_{-\infty}^{\infty} cZ(z/c, y)\,\mathrm{d}F(y) = \int_{-\infty}^{\infty} cZ(\rho(F), y)\,\mathrm{d}F(y) = 0.$$

For $z_1 < z_2$, if $c > 0$, we have $z_1/c < z_2/c$, and by the fact that Z is a backtest function for ρ, we have

$$\int_{-\infty}^{\infty} cZ(z_1/c, y)\,\mathrm{d}F(y) < \int_{-\infty}^{\infty} cZ(z_2/c, y)\,\mathrm{d}F(y).$$

If $c < 0$, we have $z_1/c > z_2/c$, and in this case,

$$\int_{-\infty}^{\infty} cZ(z_1/c, y)\,\mathrm{d}F(y) < \int_{-\infty}^{\infty} cZ(z_2/c, y)\,\mathrm{d}F(y).$$

In both cases, the function $\int Z^*(z, y)\,\mathrm{d}F(y)$ is strictly increasing in z. Therefore, Z^* is a backtest function for $c\rho$. $\square$

Proof of Proposition 3.6.4. Without loss of generality, we assume $h(1) = 1$. If $h \in \mathcal{H}_1^*$, $I_h = c \operatorname{ess\,sup} + (1-c) \operatorname{ess\,inf}$. We show that I_h is not backtestable based on an example in Kou and Peng (2016), used to show that I_h is not elicitable (their definition is slightly different from ours). Suppose I_h is backtestable and Z is a backtest function. For any $u, v \in \mathbb{R}$, $u < v$, we have $I_h(\delta_u) = u$, and by definition of the backtest function,

$$Z(u, v) < Z(v, v) = 0 = Z(u, u) < Z(v, u). \tag{3.25}$$

For $v > u$ and $p \in (0, 1)$, let $G = p\delta_u + (1-p)\delta_v$, then $I_h(G) = cv + (1-c)u$. If $c = 0$, $I(G) = u$, and by (3.25),

$$0 = \int_{-\infty}^{\infty} Z(u, y)\, \mathrm{d}G(y) = pZ(u, u) + (1-p)Z(u, v) < 0,$$

a contradiction. If $c \in (0, 1]$, by (3.25), as $u < I_h(G)$,

$$0 = pZ(cv + (1-c)u, u) + (1-p)Z(cv + (1-c)u, v) > pZ(u, u) + (1-p)Z(u, v).$$

Letting $p \to 1$, we have

$$0 = Z(cv + (1-c)u, u) > Z(u, u),$$

a contradiction. In either case, I_h is not backtestable.

If $h \in \mathcal{H}_2^*$, one can easily check that it is backtestable with backtest function $Z(x, y) = x - y$, $x, y \in \mathbb{R}$. If $h \in \mathcal{H}_3^*$, then for some $\alpha \in (0, 1)$, $I_h = \mathrm{VaR}_{1-\alpha}$ on $\mathcal{M}_{\mathrm{con}}^*$. One can easily check that it is backtestable with backtest function $Z(x, y) = \alpha \mathbf{1}_{\{y>x\}} + (1-\alpha)\mathbf{1}_{\{y<x\}}$, $x, y \in \mathbb{R}$. The above two backtest functions can be found in Table 3 of Acerbi and Szekely (2017). Finally, using Proposition 3.6.3 we get the backtest functions for the signed Choquet integral I_h. ☐

Proof of Theorem 3.6.5. In order to show Theorem 3.6.5, we need to use the following lemma, as well as Proposition 3.6.6.

Lemma 3.7.4. *For a comonotonic-additive coherent risk measure $\rho : \mathcal{M}_\infty \to \mathbb{R}$ and $p \in (0, 1)$, if (ρ, VaR_p) has CxLS on $\mathcal{M}_\infty^*(p)$, then it has CxLS on $\mathcal{M}_{\mathrm{dis}}$.*

Proof of the lemma. Write $\rho = I_h$ where $h \in \mathcal{H}$ is concave and increasing with $h(1) = 1$. We first assume h is continuous. If h is not continuous, it can only have a jump at 0, and we will comment on that case at the end of the proof.

For $r \in \mathbb{R}$, denote by $\mathcal{M}_{\mathrm{dis}}(r) = \{F \in \mathcal{M}_{\mathrm{dis}} : \mathrm{VaR}_p(F) = r\}$ and $\mathcal{M}_\infty^*(p, r) = \{F \in \mathcal{M}_\infty^*(p) : \mathrm{VaR}_p(F) = r\}$. By Proposition 3.4.2, to show that (ρ, VaR_p) has CxLS on $\mathcal{M}_{\mathrm{dis}}$, it suffices to show that ρ has CxLS on $\mathcal{M}_{\mathrm{dis}}(r)$ for all $r \in \mathbb{R}$.

Fix $r \in \mathbb{R}$ and take $F, G \in \mathcal{M}_{\mathrm{dis}}(r)$ such that $\rho(F) = \rho(G)$. We construct two sequences of distributions $\{F_n\}_{n\in\mathbb{N}} \subset \mathcal{M}^*_\infty(p)$ and $\{G_n\}_{n\in\mathbb{N}} \subset \mathcal{M}^*_\infty(p)$ as follows. Let $\varepsilon_n = (1-p)/n$, $n \in \mathbb{N}$. For $n \in \mathbb{N}$, let $\hat{F}_n^{-1}(t) = F^{-1}(t)$ for $t \in [0, p] \cup [p + \varepsilon_n, 1]$, and $\hat{F}_n^{-1}$ is linear on $[p, p + \varepsilon_n]$. Similarly, let $\hat{G}_n^{-1}(t) = G^{-1}(t)$ for $t \in [0, p] \cup [p + \varepsilon_n, 1]$, and $\hat{G}_n^{-1}$ is linear on $[p, p + \varepsilon_n]$. Note that the quantile function is always left-continuous by definition. Next, let $F_n^{-1}(t) = \min\{\hat{F}_n^{-1}(t), F^{-1}(t)\}$, $t \in [0, 1]$ and $G_n^{-1}(t) = \min\{\hat{G}_n^{-1}(t), G^{-1}(t)\}$, $t \in [0, 1]$. Since F^{-1} and G^{-1} may only have an up-side jump at p, and $\hat{F}_n$ and $\hat{G}_n$ have continuous quantiles at p, we know that F_n and G_n both have continuous quantiles at p, i.e. $\{F_n\}_{n\in\mathbb{N}} \subset \mathcal{M}^*_\infty(p, r)$ and $\{G_n\}_{n\in\mathbb{N}} \subset \mathcal{M}^*_\infty(p, r)$. It is also easy to see that $F_n \to F$ and $G_n \to G$ weakly.

Note that $F_n^{-1}(t) = F^{-1}(t)$ for t close to 0 and 1, and hence $\{F_n^{-1}\}_{n\in\mathbb{N}}$ is bounded above and below. Therefore, the uniform integrability condition in Theorem 3.2 of Wang et al. (2018) is satisfied, and we have $\rho(F_n) \to \rho(F)$ as $n \to \infty$. Similarly, $\rho(G_n) \to \rho(G)$ as $n \to \infty$.

Let B be the Bernoulli distribution with mean $c = (1 - p)/2$. Since $\rho \geqslant \mathbb{E}$, we know that $\rho(B) \geqslant c > 0$. Further, since $F_n^{-1}(t) \leqslant F^{-1}(t)$ for all $t \in [0, 1]$, we know $\rho(F_n) \leqslant \rho(F)$ since a coherent risk measure is monotone with respect to stochastic order. For $n \in \mathbb{N}$, let F_n^* be given by

$$(F_n^*)^{-1}(t) = F_n^{-1}(t) + \mathbf{1}_{\{t>1-c\}} \frac{\rho(F) - \rho(F_n)}{\rho(B)}, \quad t \in [0, 1].$$

Note that $(F_n^*)^{-1}$ is increasing and left-continuous, thus a well-defined quantile function. We can calculate, using the comonotonic-additivity of ρ, that

$$\rho(F_n^*) = \rho(F_n) + \rho(B) \frac{\rho(F) - \rho(F_n)}{\rho(B)} = \rho(F).$$

Moreover, $(F_n^*)^{-1}(p) = F_n^{-1}(p) = F^{-1}(p) = r$. Therefore, $\{F_n^*\}_{n\in\mathbb{N}} \subset \mathcal{M}^*_\infty(p, r)$. On the other hand, since $\rho(F_n) \to \rho(F)$ as $n \to \infty$, we have $F_n^* \to F$ weakly. Similarly, we can construct $\{G_n^*\}_{n\in\mathbb{N}} \subset \mathcal{M}^*_\infty(p, r)$ such that $\rho(G_n^*) = \rho(G)$, $n \in \mathbb{N}$, and $G_n^* \to G$ weakly.

It is clear that $\lambda F_n^* + (1-\lambda)G_n^* \to \lambda F + (1-\lambda)G$ weakly for $\lambda \in [0, 1]$. By noting again the uniform integrability in the sense of Theorem 3.2 of Wang et al. (2018) is satisfied by $\{\lambda F_n^* + (1-\lambda)G_n^*\}_{n\in\mathbb{N}}$, we have $\rho(\lambda F_n^* + (1-\lambda)G_n^*) \to \rho(\lambda F + (1-\lambda)G)$. Thus,

$$\rho(F) = \rho(\lambda F_n^* + (1-\lambda)G_n^*) \to \rho(\lambda F + (1-\lambda)G),$$

and hence ρ has CxLS on $\mathcal{M}_{\mathrm{dis}}(r)$. This shows that (ρ, VaR_p) has CxLS on $\mathcal{M}_{\mathrm{dis}}$.

If h has a jump at 0, then ρ can be decomposed into a convex combination of $\operatorname{ess\,sup}$ and I_g for a continuous and concave $g \in \mathcal{H}$. Since $F_n^{-1}(1) = F^{-1}(1)$ and $G_n^{-1}(1) = G^{-1}(1)$, this does not affect the arguments that $\rho(F_n) \to \rho(F)$ and $\rho(G_n) \to \rho(G)$ as $n \to \infty$, or the construction of F_n^* and G_n^*. $\square$

Proof of Proposition 3.6.6. Fix $F \in \mathcal{M}_\infty^*(p)$. For $x_1, x_2 \in \mathbb{R}$, let $H(x_1, x_2) = \int_{-\infty}^{\infty} S(x_1, x_2, y)\,\mathrm{d}F(y)$, and $G(x_1) = \int_{-\infty}^{\infty}(x_1 + \frac{1}{1-p}(y - x_1)_+)\,\mathrm{d}F(y)$. Both H and G are clearly $\mathbb{R}$-valued. Obviously,

$$H(x_1, x_2) = g(x_2) + g'(x_2)\left(aG(x_1) + (1-a)\mathbb{E}[F] - x_2\right).$$

Using the ES-VaR relation (3.1), and noting that F has a continuous quantile at p, we have $\mathrm{VaR}_p(F) = \arg\min_{x_1\in\mathbb{R}} G(x_1)$ and $\mathrm{ES}_p(F) = \min_{x_1\in\mathbb{R}} G(x_1)$. Hence, for fixed $x_2 \in \mathbb{R}$, since $af(x_2) > 0$, a minimizer of $H(x_1, x_2)$ satisfies $x_1 = \mathrm{VaR}_p(F)$. Note that

$$\begin{aligned} H(\mathrm{VaR}_p(F), x_2) &= g(x_2) + g'(x_2)(a\mathrm{ES}_p(F) + (1-a)\mathbb{E}(F) - x_2) \\ &= g(x_2) + g'(x_2)(\rho(F) - x_2). \end{aligned}$$

Since g is strictly concave, we have $g(x_2) + g'(x_2)(\rho(F) - x_2) \geqslant g(\rho(F))$ for $x_2 \in \mathbb{R}$ and the equality is attained at $x_2 = \rho(F)$. Therefore,

$$(\mathrm{VaR}_p(F), \rho(F)) = \operatorname*{arg\,min}_{(x_1,x_2)\in\mathbb{R}^2} \int_{-\infty}^{\infty} S(x_1, x_2, y)\,\mathrm{d}F(y),$$

which shows that S is a strictly consistent function for (VaR_p, ρ). $\square$

Proof of Theorem 3.6.5 continued. To show the "only-if" part, note that the elicitability of (ρ, VaR_p) implies that it has CxLS on $\mathcal{M}_\infty^*(p)$, and ρ is an increasing Choquet integral. By Lemma 3.7.4, we know that (ρ, VaR_p) has CxLS on $\mathcal{M}_{\text{dis}}^*$. Hence, we know that ρ is one of the four cases in Theorem 3.5.2. Clearly, case (ii) gives a possible coherent risk measure ρ of the form $\rho = a\mathrm{ES}_p + (1-a)\mathbb{E}$ for $a \in [0,1]$, and all other forms of ρ in Theorem 3.5.2 are not coherent. The "if" part for $a \in (0,1]$ follows from Proposition 3.6.6, and the case $a = 0$ is due to Proposition 3.4.1 since both $\mathbb{E}$ and VaR_p are $\mathcal{M}_\infty^*(p)$-elicitable. $\square$

Chapter 4

Characterizing Optimal Allocations in Quantile-based Risk Sharing

4.1 Introduction

This chapter studies the existence and uniqueness of a risk sharing problem. It appears in large part in the submitted paper Wang and Wei (2018a).

This chapter is organized in a straightforward manner. In Section 4.2 we present preliminaries on quantile-based risk sharing, as well as some existing results. In Section 4.3, we address the existence issue of optimal allocations by showing that optimal allocations exist in exactly four cases (Theorem 4.3.3). Section 4.4 contains characterization results of optimal allocations as well as some technical lemmas. In particular, Propositions 2 and 3 characterize the optimal allocations for RVaR agents based on explicit results in the cases of VaR agents (Theorem 4.4.1), ES agents (Theorem 4.4.2), and one VaR plus one ES agents (Theorem 4.4.5). Section 4.5 concludes the chapter by presenting a representative class of optimal allocations and discusses some future questions.

4.2 Preliminaries

4.2.1 Risk measures

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an atomless probability space, and $\mathcal{X}$ be the set of real, integrable random variables (i.e. random variables with finite means) defined on $(\Omega, \mathcal{F}, \mathbb{P})$. We treat almost surely equal random variables as identical in this chapter; equalities, inequalities and set inclusions should always be understood in the almost sure sense. A *risk measure* is a functional $\rho : \mathcal{X} \to [-\infty, \infty]$.

The Value-at-Risk (VaR) of $X \in \mathcal{X}$ at level $\alpha \in \mathbb{R}_+ := [0, \infty)$ is defined as the $100(1-\alpha)\%$ left quantile of X,

$$\mathrm{VaR}_\alpha(X) = \inf\{x \in [-\infty, \infty] : \ \mathbb{P}(X \leqslant x) \geqslant 1 - \alpha\}. \tag{4.1}$$

The corresponding right quantile is denoted by VaR^+_α, namely,

$$\mathrm{VaR}^+_\alpha(X) = \inf\{x \in [-\infty, \infty] : \ \mathbb{P}(X \leqslant x) > 1 - \alpha\}. \tag{4.2}$$

In (4.1)-(4.2), we use the convention $\inf\{\varnothing\} = \infty$. Note that in (4.1), for $\alpha \geqslant 1$, $\mathrm{VaR}_\alpha(X) = -\infty$ for all $X \in \mathcal{X}$. Certainly, only the case $\alpha \in [0, 1)$ is relevant in risk management; in particular, practical values of α are close to 0 in banking and insurance regulation.

For $X \in \mathcal{X}$, the Range-Value-at-Risk (RVaR) at level $(\alpha, \beta) \in \mathbb{R}^2_+$ is defined as

$$\mathrm{RVaR}_{\alpha,\beta}(X) = \begin{cases} \frac{1}{\beta}\int_\alpha^{\alpha+\beta} \mathrm{VaR}_\gamma(X)\mathrm{d}\gamma & \text{if } \beta > 0, \\ \mathrm{VaR}_\alpha(X) & \text{if } \beta = 0. \end{cases} \tag{4.3}$$

For $X \in \mathcal{X}$ and $\alpha + \beta > 1$, $\mathrm{RVaR}_{\alpha,\beta}(X) = -\infty$. Besides VaR, another special case of RVaR is the Expected Shortfall (ES), defined as

$$\mathrm{ES}_\beta(X) = \mathrm{RVaR}_{0,\beta}(X) = \frac{1}{\beta}\int_0^\beta \mathrm{VaR}_\gamma(X)\,\mathrm{d}\gamma, \quad \beta \geqslant 0.$$

Different from RVaR and VaR, an ES is subadditive.

Remark 4.2.1 (Terminological remark). There are several different conventions used in the literature of risk measures. Some papers use the convention $\mathrm{VaR}_1 = \lim_{\alpha\to 1}\mathrm{VaR}_\alpha$, which corresponds to our VaR^+_1. The convention $\mathrm{VaR}_1 = -\infty$ used in this chapter and Embrechts et al. (2018) unifies several technical results. In different contexts, ES has various alternative names, such as AVaR (Föllmer and Schied (2016)), CVaR (Rockafellar and Uryasev (2000)) and TVaR (Denuit et al. (2005)).

A useful optimization property linking VaR and ES obtained by Rockafellar and Uryasev (2000, 2002) is, for $\beta \in (0, 1)$,

$$\mathrm{ES}_\beta(X) = \min\left\{\frac{1}{\beta}\mathbb{E}[(X - x)_+] + x : x \in \mathbb{R}\right\}, \tag{4.4}$$

and

$$[\mathrm{VaR}_\beta(X), \mathrm{VaR}^+_\beta(X)] = \arg\min \left\{ \frac{1}{\beta}\mathbb{E}[(X-x)_+] + x : x \in \mathbb{R} \right\}. \tag{4.5}$$

The second parameter β in $\mathrm{RVaR}_{\alpha,\beta}$ is referred to as the *tolerance parameter*; see the discussions in Embrechts et al. (2018) after Theorem 2. RVaR was first introduced by Cont et al. (2010) featuring its robustness properties (see Embrechts et al. (2018) for more details on the family of RVaR). The RVaR family of risk measures provide a flexible and tractable framework for the study of risk sharing, including the two most practical risk measures as special cases. Following the setup of Embrechts et al. (2018), we shall focus on the RVaR family of risk measures in this chapter. As far as we know, there are very few results on risk sharing problems with other non-convex distortion risk measures; see Weber (2018) for some available results.

4.2.2 Risk sharing and inf-convolution

Similarly to Embrechts et al. (2018), we refer to a participant in the risk sharing transactions as an *agent*, which may represent an affiliate, a firm, an insured, an insurer, or an investor in various specific contexts. Let n be a positive integer which represents the number of agents. Given random variable $X \in \mathcal{X}$, we define the set of *allocations* of X as

$$\mathbb{A}_n(X) = \left\{ (X_1, \dots, X_n) \in \mathcal{X}^n : \sum_{i=1}^n X_i = X \right\}. \tag{4.6}$$

For $i = 1, \dots, n$, agent i is equipped with a risk measure $\rho_i : \mathcal{X} \to \mathbb{R}$, which is the agent's objective to minimize. In this chapter, we consider Pareto-optimal allocations defined below.

Definition 4.2.1 (Pareto-optimal allocations)**.** Fix the risk measures $\rho_1, \dots, \rho_n$ and the total risk $X \in \mathcal{X}$. An allocation $(X_1, \dots, X_n) \in \mathbb{A}_n(X)$ is *Pareto-optimal* with respect to $(\rho_1, \dots, \rho_n)$ if for any allocation $(Y_1, \dots, Y_n) \in \mathbb{A}_n(X)$, $\rho_i(Y_i) \leqslant \rho_i(X_i)$ for all $i = 1, \dots, n$ implies $\rho_i(Y_i) = \rho_i(X_i)$ for all $i = 1, \dots, n$. Throughout, we shall simply call a Pareto-optimal allocation an *optimal allocation*.

To study risk sharing problems for risk measures, define the *inf-convolution* of risk measures (see e.g. Delbaen (2012) and Rüschendorf (2013)) as

$$\mathop{\square}_{i=1}^{n} \rho_i(X) = \inf \left\{ \sum_{i=1}^n \rho_i(X_i) : (X_1, \cdots, X_n) \in \mathbb{A}_n(X) \right\}, \quad X \in \mathcal{X}. \tag{4.7}$$

Our choices of $\rho_1, \dots, \rho_n$ in this chapter do not take the value $-\infty$ on $\mathcal{X}$ and hence the infimum in (4.7) is well posed. It is well-known that for *monetary risk measures* (Artzner et al. (1999)) including the RVaR family, Pareto optimality is equivalent to optimality with respect to the sum (Proposition 1 of Embrechts et al. (2018)). More precisely, assuming that each of $\rho_i(X_i)$, $i = 1, \dots, n$ is finite, $(X_1, \dots, X_n)$ is a Pareto-optimal allocation of X if and only if

$$\sum_{i=1}^{n} \rho_i(X_i) = \mathop{\square}_{i=1}^{n} \rho_i(X). \tag{4.8}$$

In the sequel, an allocation $(X_1, \dots, X_n)$ satisfying (4.8) is called a *sum-optimal allocation*. We will omit "with respect to $(\rho_1, \dots, \rho_n)$" in most cases as long as the underlying risk measures are clear. As mentioned above, unless $\square_{i=1}^{n} \rho_i(X)$ is infinite, optimal allocations and sum-optimal ones are equivalent; see Lemma 4.3.1 and Remark 4.3.1 below for more discussions on the subtle cases $\square_{i=1}^{n} \rho_i(X) = \pm\infty$.

4.2.3 Existing results on optimal allocations

We first specify agents' preferences in the risk sharing problems in this chapter. As these preferences will be used throughout the chapter, we emphasize it in the following assumption. Throughout, for any constants $\beta_1, \dots, \beta_n \in \mathbb{R}$, write $\bigvee_{i=1}^{n} \beta_i = \max\{\beta_1, \dots, \beta_n\}$ and $\bigwedge_{i=1}^{n} \beta_i = \min\{\beta_1, \dots, \beta_n\}$.

Assumption. Unless otherwise specified, all optimal allocations are with respect to the risk measures $(\mathrm{RVaR}_{\alpha_1,\beta_1}, \dots, \mathrm{RVaR}_{\alpha_n,\beta_n})$, where $\alpha_i \in [0, 1)$, $\beta_i \in [0, 1]$ and $\alpha_i + \beta_i \leqslant 1$, $i = 1, \dots, n$.

We will always denote by $\alpha = \sum_{i=1}^{n} \alpha_i$ and $\beta = \bigvee_{i=1}^{n} \beta_i$. The above specification of parameters guarantees $\mathrm{RVaR}_{\alpha_i,\beta_i}(X) > -\infty$ for all $X \in \mathcal{X}$. If the assumption is not satisfied (i.e. $\alpha_i + \beta_i > 1$ or $\alpha_i = 1$), then $\mathrm{RVaR}_{\alpha_i,\beta_i}(X) = -\infty$ for all $X \in \mathcal{X}$, leading to a trivial case.

Below we summarize the main results from Embrechts et al. (2018) on optimal allocations. Let U_X be a uniform random variable on $[0, 1]$ such that $F^{-1}(U_X) = X$ almost surely where F is the distribution function of the random variable X and $F^{-1}(p) = \inf\{x \in \mathbb{R} : F(x) \geqslant p\}$, $p \in (0, 1)$. If X is continuously distributed, then $U_X = F(X)$. For a general random variable X, the existence of U_X is guaranteed; see, for instance, Lemma A.32 of Föllmer and Schied (2016).

Theorem 4.2.1 (Theorem 2 of Embrechts et al. (2018)). *We have*

$$\mathop{\square}_{i=1}^{n} \mathrm{RVaR}_{\alpha_i,\beta_i}(X) = \mathrm{RVaR}_{\alpha,\beta}(X), \quad X \in \mathcal{X}. \tag{4.9}$$

Moreover, if $p := \alpha + \beta < 1$, then, assuming (without loss of generality) $\beta_n = \beta$, a sum-optimal allocation $(X_1, \dots, X_n)$ of $X \in \mathcal{X}$ is given by

$$X_i = (X - m)\,\mathbf{1}_{\{1-\sum_{k=1}^{i}\alpha_k < U_X \leqslant 1-\sum_{k=1}^{i-1}\alpha_k\}}, \quad i = 1, \dots, n-1, \tag{4.10}$$

$$X_n = (X - m)\,\mathbf{1}_{\{U_X \leqslant 1-\sum_{k=1}^{n-1}\alpha_k\}} + m, \tag{4.11}$$

where $m \in (-\infty, \mathrm{VaR}_p(X)]$ is a constant.

Theorem 4.2.1 implies the following useful inequality, which is given in Theorem 1 of Embrechts et al. (2018). For all $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \geqslant 0$ and $X_1, \dots, X_n \in \mathcal{X}$, we have

$$\sum_{i=1}^{n} \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) \geqslant \mathrm{RVaR}_{\alpha,\beta}\left(\sum_{i=1}^{n} X_i\right). \tag{4.12}$$

Theorem 4.2.1 and (4.12) will be used repeatedly in this chapter. It is clear that the above results do not fully address the issue of existence, and no results on the unique forms of optimal allocations are provided. In particular, the following questions are unanswered:

(i) Theorem 4.2.1 implies that a sum-optimal allocation exists if $\alpha + \beta < 1$, and it does not exist if $\alpha + \beta > 1$. Under what conditions a Pareto-optimal allocation exists?

(ii) When an optimal allocation exists, is it possible to identify all possible optimal allocations (unique form up to certain freedom)?

This chapter is dedicated to complete answers to both questions above.

4.3 Existence of the optimal allocations

In this section, we analyze the existence of optimal allocations in a quantile-based risk sharing problem. The main results are that Pareto-optimal and sum-optimality are the equivalent for RVaR, unless $\alpha = \beta = 0$, and the existence of a Pareto-optimal allocation can be characterized in four cases (A1)-(A4) below depending on the parameters $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n$ and the total risk X.

In the sequel, for $X \in \mathcal{X}$, we say that X is bounded from below (resp. above) if $\mathrm{VaR}_1^+(X) > -\infty$ (resp. $\mathrm{VaR}_0(X) < \infty$). The following lemma clarifies the subtle difference between optimal allocations and sum-optimal ones.

Lemma 4.3.1. *For $X \in \mathcal{X}$, the following hold.*

(i) If $\mathrm{RVaR}_{\alpha,\beta}(X) = \infty$, there does not exist an optimal allocation, whereas all allocations are sum-optimal.

(ii) If $-\infty < \mathrm{RVaR}_{\alpha,\beta}(X) < \infty$, an allocation is optimal if and only if it is sum-optimal.

(iii) If $\mathrm{RVaR}_{\alpha,\beta}(X) = -\infty$, there does not exist an optimal allocation or a sum-optimal allocation.

Proof. (i) Suppose that $(X_1, \ldots, X_n)$ is an optimal allocation. As $\mathrm{RVaR}_{\alpha,\beta}(X) \leqslant \sum_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i)$, at least one of $\mathrm{RVaR}_{\alpha_i,\beta_i}(X_i)$, $i = 1, \ldots, n$ is equal to ∞. Without loss of generality, assume $\mathrm{RVaR}_{\alpha_1,\beta_1}(X_1) = \infty$. If $\mathrm{RVaR}_{\alpha_2,\beta_2}(X_2) = \infty$, then we take an allocation $(X_1 + X_2, 0, X_3 \ldots, X_n) \in \mathbb{A}_n(X)$. It is clear that $\mathrm{RVaR}_{\alpha_1,\beta_1}(X_1 + X_2) \leqslant \infty = \mathrm{RVaR}_{\alpha_1,\beta_1}(X_1)$ and $\mathrm{RVaR}_{\alpha_2,\beta_2}(0) < \infty = \mathrm{RVaR}_{\alpha_2,\beta_2}(X_2)$. Hence, $(X_1, \ldots, X_n)$ is not Pareto-optimal. If $\mathrm{RVaR}_{\alpha_2,\beta_2}(X_2) < \infty$, then we take an allocation $(X_1 + c, X_2 - c, X_3, \ldots, X_n) \in \mathbb{A}_n(X)$ for some $c > 0$. It is clear that $\mathrm{RVaR}_{\alpha_1,\beta_1}(X_1 + c) = \infty = \mathrm{RVaR}_{\alpha_1,\beta_1}(X_1)$ and $\mathrm{RVaR}_{\alpha_2,\beta_2}(X_2 - c) < \mathrm{RVaR}_{\alpha_2,\beta_2}(X_2)$. Hence, $(X_1, \ldots, X_n)$ is not Pareto-optimal.

On the other hand, as $\mathrm{RVaR}_{\alpha,\beta}(X) = \square_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(X) = \infty$, any choice of $(X_1, \ldots, X_n) \in \mathbb{A}_n(X)$ satisfies $\sum_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) = \infty$, and hence it is a sum-optimal allocation.

(ii) This is due to Proposition 1 of Embrechts et al. (2018).

(iii) Note that for our choices of parameters, $\mathrm{RVaR}_{\alpha_i,\beta_i}$ does not take the value $-\infty$. Suppose that $(Y_1, \cdots, Y_n)$ is an optimal allocation. Since $\mathrm{RVaR}_{\alpha,\beta}(X) = -\infty$, there exists $(X_1, \cdots, X_n) \in \mathbb{A}_n(X)$ such that $\sum_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) < \sum_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(Y_i)$. Then at least one $\mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) < \mathrm{RVaR}_{\alpha_i,\beta_i}(Y_i)$. Without loss of generality, assume $\mathrm{RVaR}_{\alpha_1,\beta_1}(X_1) < \mathrm{RVaR}_{\alpha_1,\beta_1}(Y_1)$. Let $c_i = \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) - \mathrm{RVaR}_{\alpha_i,\beta_i}(Y_i)$ for $i = 2, \ldots, n$. Clearly $\mathrm{RVaR}_{\alpha_i,\beta_i}(X_i - c_i) = \mathrm{RVaR}_{\alpha_i,\beta_i}(Y_i)$, $i = 2, \ldots, n$. Moreover,

$$\begin{aligned}\mathrm{RVaR}_{\alpha_1,\beta_1}\left(X_1 + \sum_{i=2}^n c_i\right) &= \sum_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) - \sum_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(Y_i) + \mathrm{RVaR}_{\alpha_1,\beta_1}(Y_1)\\ &< \mathrm{RVaR}_{\alpha_1,\beta_1}(Y_1).\end{aligned}$$

This means that $(X_1 + \sum_{i=2}^n c_i, X_2 - c_2, \cdots, X_n - c_n) \in \mathbb{A}_n(X)$ strictly dominates $(Y_1, \cdots, Y_n)$, and the latter is not Pareto-optimal.

On the other hand, since for $i = 1, \ldots, n$, $\mathrm{RVaR}_{\alpha_i,\beta_i}$ does not take the value $-\infty$, one always have, for any $(X_1, \ldots, X_n) \in \mathcal{A}_n(X)$, $\sum_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) > -\infty = \square_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(X)$. Therefore, no sum-optimal allocations exist. $\square$

Remark 4.3.1. The only possible difference between an optimal allocation and a sum-optimal one is case (i) in Lemma 4.3.1. More precisely, it corresponds to $\alpha = \beta = 0$ (implying $\alpha_1 = \cdots = \alpha_n = \beta_1 = \cdots = \beta_n = 0$) and $\mathrm{VaR}_0(X) = \infty$. As any allocation $(X_1, \ldots, X_n) \in \mathbb{A}_n(X)$ is sum-optimal in this case and no Pareto-optimal allocation exists, it is not interesting for further study.

Lemma 4.3.1 implies that, unless $\alpha = \beta = 0$ and X is unbounded from above, there is no difference between optimal allocation and sum-optimal ones. By (4.8) and Theorem 4.2.1, an allocation $(X_1, \ldots, X_n)$ is optimal if and only if

$$\sum_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) = \mathrm{RVaR}_{\alpha,\beta}(X) \quad \text{and} \quad \mathrm{RVaR}_{\alpha,\beta}(X) < \infty.$$

Below we illustrate four cases where an optimal allocation can be explicitly formulated:

(A1) $\alpha = \beta = 0$ and X is bounded from above;

(A2) $0 < \alpha + \beta < 1$;

(A3) $\alpha + \beta = 1$, $\beta > 0$ and X is bounded from below;

(A4) $\alpha + \beta = 1$, $\beta > 0$ and there exists $i \in \{1, \ldots, n\}$ such that $\alpha_i = \alpha$ and $\beta_i = \beta$.

To describe the corresponding optimal allocations under (A1)-(A4), assume, without loss of generality, $\beta_n = \beta$, i.e. β_n is the largest among $\beta_1, \ldots, \beta_n$.

Case (A1): A sum-optimal allocation is provided by (4.10) and (4.11) in Theorem 4.2.1, which is optimal by Lemma 4.3.1 (ii) due to $-\infty < \mathrm{RVaR}_{\alpha,\beta}(X) < \infty$.

Case (A2): Same as Case (A1).

Case (A3): Let $(X_1, \ldots, X_n)$ be given by (4.10) and (4.11), where $m = \mathrm{VaR}_1^+(X)$. One can check that $\mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) = 0$ for $i = 1, \ldots, n-1$ and $\mathrm{RVaR}_{\alpha_n,\beta_n}(X_n) = \mathrm{RVaR}_{\alpha,\beta}(X)$, and hence $\sum_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) = \mathrm{RVaR}_{\alpha,\beta}(X)$, i.e. $(X_1, \ldots, X_n)$ is a sum-optimal allocation.

Case (A4): Let $X_i = X$ and $X_j = 0$ for $j \neq i$. Recall that our specification of (α_i, β_i) guarantees $\alpha_i + \beta_i \leqslant 1$ and $\alpha_i < 1$; thus $\mathrm{RVaR}_{\alpha_i,\beta_i}(X) > -\infty$. We can easily see $\sum_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) = \mathrm{RVaR}_{\alpha,\beta}(X)$, and hence $(X_1, \dots, X_n)$ is a sum-optimal allocation.

In all four cases, since $\mathrm{RVaR}_{\alpha,\beta}(X) < \infty$, sum-optimal allocations are optimal.

Note that condition (A4) implies $\alpha_j = 0$ for $j \neq i$ and in fact all other agents are essentially not participating in the risk sharing transactions except for agent i, because they have very conservative risk attitude (each of them uses an ES with a smaller tolerance parameter β_j compared to β_i). Obviously, this case is very special and not very practically relevant.

Next, we shall show that (A1)-(A4) are precisely the only possible cases where an optimal allocation may exist. We first present a lemma on the sum of a VaR and an ES, which may be of independent interest.

Lemma 4.3.2. *For $\alpha \in (0,1)$ and $X, Y \in \mathcal{X}$ such that $X + Y$ is unbounded from below, we have*

$$\mathrm{VaR}_\alpha(X) + \mathrm{ES}_{1-\alpha}(Y) > \mathrm{RVaR}_{\alpha,1-\alpha}(X+Y). \tag{4.13}$$

Proof. Since one can freely replace X by $X+c$ for any constant $c \in \mathbb{R}$ in (4.13), without loss of generality we assume $\mathrm{VaR}_\alpha(X) = 0$. It suffices to show $\mathrm{ES}_{1-\alpha}(Z - X) > \mathrm{RVaR}_{\alpha,1-\alpha}(Z)$ for all $X, Z \in \mathcal{X}$ such that Z is unbounded from below and $\mathrm{VaR}_\alpha(X) = 0$. Note that $\mathrm{ES}_{1-\alpha}(Z - X) \geqslant \mathrm{ES}_{1-\alpha}(Z - X_+)$, and $\mathrm{VaR}_\alpha(X) = 0$ can be loosened to $\mathrm{VaR}_\alpha(X) \geqslant 0$, which is equivalent to $\mathbb{P}(X > 0) \leqslant \alpha$. Therefore, to prove the lemma, it suffices to show

$$\mathrm{ES}_{1-\alpha}(Z - X) > \mathrm{RVaR}_{\alpha,1-\alpha}(Z) \tag{4.14}$$

for all $X, Z \in \mathcal{X}$ such that $X \geqslant 0$, $\mathbb{P}(X > 0) \leqslant \alpha$ and Z is unbounded from below.

Fix arbitrary $X, Z \in \mathcal{X}$ satisfying the above conditions. Write $Y = Z - X$ and note that $\mathbb{P}(Z = Y) \geqslant 1 - \alpha$. As a consequence, for all $x \in \mathbb{R}$, $\mathbb{P}(Y \leqslant x) - \mathbb{P}(Z \leqslant x) \leqslant \alpha$. Using the above relation and the definition of VaR, for $\gamma \leqslant 1 - \alpha$, we have $\mathrm{VaR}_\gamma(Y) \geqslant \mathrm{VaR}_{\gamma+\alpha}(Z)$.

Also note that $\mathrm{VaR}_\gamma(Y) \geqslant \mathrm{VaR}_{1-\alpha}(Y)$. Therefore, we have

$$\begin{aligned}
\mathrm{ES}_{1-\alpha}(Y) &= \frac{1}{1-\alpha}\int_0^{1-\alpha} \mathrm{VaR}_\gamma(Y)\,\mathrm{d}\gamma \\
&\geqslant \frac{1}{1-\alpha}\int_0^{1-\alpha} (\mathrm{VaR}_{\gamma+\alpha}(Z) \vee \mathrm{VaR}_{1-\alpha}(Y))\,\mathrm{d}\gamma \\
&= \frac{1}{1-\alpha}\int_\alpha^{1} (\mathrm{VaR}_{\gamma}(Z) \vee \mathrm{VaR}_{1-\alpha}(Y))\,\mathrm{d}\gamma \\
&> \frac{1}{1-\alpha}\int_\alpha^{1} \mathrm{VaR}_{\gamma}(Z)\,\mathrm{d}\gamma = \mathrm{RVaR}_{\alpha,1-\alpha}(Z), \qquad (4.15)
\end{aligned}$$

where the last inequality is due to the fact that Z is unbounded from below and $\mathrm{VaR}_{1-\alpha}(Y)$ is a constant. Therefore, (4.14) holds and the proof is complete. □

Remark 4.3.2. The condition that $X + Y$ is unbounded from below is essential to the statement of Lemma 4.3.2. In fact, from the proof of Lemma 4.3.2 we can see that, if Z is bounded from below, then one can choose X such that $\mathrm{VaR}_{1-\alpha}(Y)$ is small enough so that the last inequality in (4.15) is an equality, leading to $\mathrm{VaR}_\alpha(X) + \mathrm{ES}_{1-\alpha}(Y) \geqslant \mathrm{RVaR}_{\alpha,1-\alpha}(X + Y)$, a special case of the inequality (4.12).

Remark 4.3.3. Obtained from the definition of RVaR, for any random variable Z, one has

$$\mathrm{RVaR}_{\alpha,1-\alpha}(Z) = -\mathrm{ES}_{1-\alpha}(-Z).$$

Therefore, Lemma 4.3.2 is equivalent to the following statement: For any $\alpha \in (0, 1)$ and $(X, Y, Z) \in \mathbb{A}_3(0)$ with Z unbounded from above,

$$\mathrm{VaR}_\alpha(X) + \mathrm{ES}_{1-\alpha}(Y) + \mathrm{ES}_{1-\alpha}(-Z) > 0.$$

With the help of Lemma 4.3.2, we are ready to give a full characterization of the existence of an optimal allocation.

Theorem 4.3.3. *For $X \in \mathcal{X}$, an optimal allocation exists if and only if one of (A1)-(A4) holds.*

Proof. As explicitly constructed above, under each of the conditions (A1)-(A4), an optimal allocation exists. We only need to show that no optimal allocation exists when none of (A1)-(A4) holds. First, note that if $\alpha + \beta > 1$, $\alpha = 1$, or $\mathrm{RVaR}_{\alpha,\beta}(X) = \infty$, no optimal allocation may exist according to Lemma 4.3.1. Hence, we only need to consider the case where $\alpha + \beta = 1$ and $\beta \in (0,1)$. As (A3) does not hold, X is unbounded from below. Furthermore, $\alpha + \beta = 1$ and (A4) does not hold, we have $\alpha_i + \beta_i < 1$ for each $i = 1, \dots, n$.

Take an arbitrary $(X_1, \dots, X_n) \in \mathbb{A}_n(X)$. For $i = 1, \dots, n$, we assert that there exists $(Y_i, Z_i) \in \mathbb{A}_2(X_i)$ such that

$$\mathrm{VaR}_{\alpha_i}(Y_i) + \mathrm{ES}_{\beta_i}(Z_i) = \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i).$$

This assertion is shown by noticing the fact that, for the risk sharing problem of two agents with VaR_{α_i} and ES_{β_i} as their preferences, an optimal allocation always exists, because condition (A2) is satisfied for this problem. Write $Y = \sum_{i=1}^n Y_i$ and $Z = \sum_{i=1}^n Z_i$. We have

$$\sum_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) = \sum_{i=1}^n \mathrm{VaR}_{\alpha_i}(Y_i) + \sum_{i=1}^n \mathrm{ES}_{\beta_i}(Z_i) \geqslant \mathrm{VaR}_{\alpha}(Y) + \mathrm{ES}_{\beta}(Z) > \mathrm{RVaR}_{\alpha,\beta}(X),$$

where the first inequality is an application of (4.12) and the second inequality is due to Lemma 4.3.2 by noting that X is unbounded from below, $Y + Z = X$, and $\alpha + \beta = 1$. Therefore, no optimal allocation exists if none of (A1)-(A4) holds. □

Note that the cases (A1)-(A3) are mutually exclusive, but (A4) may overlap with (A3). To obtain mutually exclusive cases, one can replace (A4) by

(A4') $\alpha + \beta = 1$, $\beta > 0$, X is unbounded from below, and there exists $i \in \{1, \dots, n\}$ such that $\alpha_i = \alpha$ and $\beta_i = \beta$.

With this modification, Theorem 4.3.3 reads as, for $X \in \mathcal{X}$, an optimal allocation exists if and only if precisely one of (A1)-(A3) and (A4') holds.

4.4 Characterizing optimal allocations

4.4.1 The route and notation

In this section, we characterize all optimal allocations in a quantile-based risk sharing problem. We first make an intuitive statement. Due to the fact that each risk measure in the RVaR family ignores part of the distribution, one might naturally expect that the class of optimal allocations has a lot of freedom. As we shall see in this section, this is indeed the case.

We outline the key ideas behind our main results. In order to characterize optimal allocations for RVaR agents, we note the following relationship from Theorem 2 of Embrechts et al. (2018),

$$\mathop{\square}_{i=1}^{n} \mathrm{RVaR}_{\alpha_i,\beta_i} = \mathrm{RVaR}_{\alpha,\beta} = \mathrm{VaR}_{\alpha} \square \mathrm{ES}_{\beta} = \left(\mathop{\square}_{i=1}^{n} \mathrm{VaR}_{\alpha_i}\right) \square \left(\mathop{\square}_{i=1}^{n} \mathrm{ES}_{\beta_i}\right). \tag{4.16}$$

Intuitively, a risk sharing problem for RVaR agents may be decomposed into two steps: first, allocate X to $(Y, Z) \in \mathcal{A}_2(X)$ such that $\mathrm{RVaR}_{\alpha,\beta}(X) = \mathrm{VaR}_{\alpha}(Y) + \mathrm{ES}_{\beta}(Z)$, and second, allocate Y and Z to $(Y_1, \ldots, Y_n) \in \mathcal{A}_n(Y)$ and $(Z_1, \ldots, Z_n) \in \mathcal{A}_n(Z)$ such that $\sum_{i=1}^{n} \mathrm{VaR}_{\alpha_i}(Y_i) = \mathrm{VaR}_{\alpha}(Y)$ and $\sum_{i=1}^{n} \mathrm{ES}_{\alpha_i}(Z_i) = \mathrm{ES}_{\beta}(Z)$. If all of the above allocations exist, then by letting $X_i = Y_i + Z_i$, $i = 1, \ldots, n$, we obtain an optimal allocation for the RVaR agents. Note that the above allocations are optimal with respect to the corresponding risk sharing problems, namely, the case of one VaR and one ES agent, the case of n VaR agents, and the case of n ES agents.

Following the above plan, we analyze the special case $\alpha = \beta = 0$ in Section 4.4.2, the case of n VaR agents ($\beta = 0$) in Section 4.4.3 and the case of n ES agents ($\alpha = 0$) in Section 4.4.4. In Section 4.4.5, we study the case of one VaR agent and one ES agent. Finally, in Propositions 2 and 3 in Section 4.4.6, we characterize optimal allocations for RVaR agents based on the above two-step decomposition and the results obtained in Sections 4.4.2-4.4.5.

The following notation will be useful in this section. For a set $A \in \mathcal{F}$, let $\pi_n(A)$ be the set of n-partitions of A in $\mathcal{F}^n$, namely,

$$\pi_n(A) = \left\{(A_1, \ldots, A_n) \in \mathcal{F}^n : \bigcup_{i=1}^{n} A_i = A, \text{ and } A_1, \ldots, A_n \text{ are mutually disjoint}\right\}.$$

Let $\mathbb{A}_n^+(X)$ (resp. $\mathbb{A}_n^-(X)$) be the set of non-negative (resp. non-positive) allocations of a random variable X, namely,

$$\mathbb{A}_n^+(X) = \{(X_1, \ldots, X_n) \in \mathbb{A}_n(X) : X_i \geqslant 0,\ i = 1, \ldots, n\}, \quad X \geqslant 0,$$

and

$$\mathbb{A}_n^-(X) = \{(X_1, \ldots, X_n) \in \mathbb{A}_n(X) : X_i \leqslant 0,\ i = 1, \ldots, n\}, \quad X \leqslant 0.$$

For a constant x, let $\mathbb{A}_n^c(x)$ be the set of constant allocations, namely,

$$\mathbb{A}_n^c(x) = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1 + \cdots + x_n = x\}, \quad x \in \mathbb{R}.$$

To simplify the notation, for a specified X, we always write $y_\alpha = \mathrm{VaR}_\alpha(X)$ and $y_\alpha^+ = \mathrm{VaR}_\alpha^+(X)$ for $\alpha \in [0,1]$. For given $\alpha_1, \dots, \alpha_n \geqslant 0$, if $y_\alpha \in \mathbb{R}$, let

$$\mathcal{Z}_n = \left\{ (Z_1, \dots, Z_n) \in \mathcal{X}^n : Z_i \geqslant 0,\ \mathbb{P}(Z_i > 0) \leqslant \alpha_i,\ i = 1, \dots, n \text{ and } \sum_{i=1}^n Z_i \geqslant (X - y_\alpha)_+ \right\}.$$

We can verify that $\mathcal{Z}_n$ is non-empty since $\mathbb{P}((X - y_\alpha)_+ > 0) \leqslant \alpha$.

4.4.2 The special case of essential supremum agents

We first consider the special case where $\alpha_1 = \cdots = \alpha_n = \beta_1 = \cdots = \beta_n = 0$, corresponding to Case (A1) in Section 4.3. In this case an optimal allocation exists if and only if $y_0 < \infty$, according to Theorem 4.3.3. This special case is obviously the simplest, and it is treated separately since its solution form is different from any of the later, more complicated, cases.

Proposition 1. *Suppose that $\alpha = \beta = 0$, $X \in \mathcal{X}$, and $y_0 < \infty$. $(X_1, \dots, X_n) \in \mathcal{X}^n$ is an optimal allocation of X if and only if*

$$X_i = Y_i + c_i,\ i = 1, \dots, n \tag{4.17}$$

for some $(Y_1, \dots, Y_n) \in \mathbb{A}_n^-(X - y_0)$ and $(c_1, \dots, c_n) \in \mathbb{A}_n^c(y_0)$.

Proof. Recall that by Theorem 4.2.1, $(X_1, \dots, X_n) \in \mathbb{A}_n(X)$ is optimal if and only if $\sum_{i=1}^n \mathrm{VaR}_0(X_i) = \mathrm{VaR}_0(X) = y_0$. It is easy to see that (4.17) defines an optimal allocation since $\sum_{i=1}^n \mathrm{VaR}_0(X_i) \leqslant \sum_{i=1}^n \mathrm{VaR}_0(c_i) = y_0$. It remains to show that any optimal allocation $(X_1, \dots, X_n)$ admits the form (4.17). Note that $\sum_{i=1}^n \mathrm{VaR}_0(X_i) = y_0 < \infty$ implies $\mathrm{VaR}_0(X_i) < \infty$ for each $i = 1, \dots, n$. Take $c_i = \mathrm{VaR}_0(X_i)$, $i = 1, \dots, n$. It is clear that $\sum_{i=1}^n (X_i - c_i) = X - y_0$ and hence (4.17) holds by taking $Y_i = X_i - c_i$, $i = 1, \dots, n$. □

4.4.3 VaR agents

We consider the case where $\beta_1 = \cdots = \beta_n = 0$ and $\alpha \in (0,1)$, that is, the objective of each agent is a VaR. In this case, by Theorem 4.3.3, an optimal allocation exists if and only if $\alpha = \sum_{i=1}^n \alpha_i$ is less than 1, i.e. (A2) holds. We introduce the following class of allocations. Let $(X_1, \dots, X_n)$ be given by

$$\begin{aligned} &X_i = Z_i + Y_i + c_i, \quad i = 1, \dots, n, \\ &\text{where } (Z_1, \dots, Z_n) \in \mathcal{Z}_n,\ (Y_1, \dots, Y_n) \in \mathbb{A}_n^-\left(X - y_\alpha - \textstyle\sum_{i=1}^n Z_i\right) \text{ and } (c_1, \dots, c_n) \in \mathbb{A}_n^c(y_\alpha). \end{aligned} \tag{4.18}$$

We assert that (4.18) gives a properly defined allocation of X by verifying a few facts:

1. As we have seen above, $\mathcal{Z}_n$ is non-empty.
2. Since $\sum_{i=1}^n Z_i \geqslant (X - y_\alpha)_+$, we have $X - y_\alpha - \sum_{i=1}^n Z_i \leqslant 0$, and hence $\mathbb{A}_n^-(X - y_\alpha - \sum_{i=1}^n Z_i)$ is non-empty.
3. It is easy to see $\sum_{i=1}^n X_i = X$ for all choices of $(Z_1, \dots, Z_n)$, $(Y_1, \dots, Y_n)$ and $(c_1, \dots, c_n)$ in (4.18).

Below we show the optimality of (4.18) and that any optimal allocation of X has the form (4.18).

Theorem 4.4.1. *Assume $\beta = 0$ and $\alpha \in (0, 1)$. For $X \in \mathcal{X}$, $(X_1, \dots, X_n) \in \mathcal{X}^n$ is an optimal allocation of X if and only if it has the form (4.18).*

Proof. We first show the "if" part. For $i = 1, \dots, n$, we have

$$\mathrm{VaR}_{\alpha_i}(X_i) = \mathrm{VaR}_{\alpha_i}(Z_i + Y_i + c_i) \leqslant \mathrm{VaR}_{\alpha_i}(Z_i + c_i) \leqslant c_i.$$

Therefore,

$$\sum_{i=1}^n \mathrm{VaR}_{\alpha_i}(X_i) \leqslant \sum_{i=1}^n c_i = y_\alpha = \mathrm{VaR}_\alpha(X).$$

Using Theorem 4.2.1, we have

$$\sum_{i=1}^n \mathrm{VaR}_{\alpha_i}(X_i) \leqslant \mathrm{VaR}_\alpha(X) = \mathop{\square}_{i=1}^n \mathrm{VaR}_{\alpha_i}(X) \leqslant \sum_{i=1}^n \mathrm{VaR}_{\alpha_i}(X_i).$$

Noting that sum-optimality is equivalent to Pareto-optimality, we conclude that $(X_1, \dots, X_n)$ is optimal.

Next we show the "only-if" part in two steps.

(i) Let $Y \in \mathcal{X}$ be such that $\mathrm{VaR}_\alpha(Y) = 0$ and $(X_1, \dots, X_n)$ be an optimal allocation of Y such that $\mathrm{VaR}_{\alpha_i}(X_i) = 0$ for $i = 1, \dots, n$. Write

$$X_i = \mathbf{1}_{\{X_i > 0\}} X_i + \mathbf{1}_{\{X_i \leqslant 0\}} X_i, \quad i = 1, \dots, n.$$

Write $Z_i = X_i \mathbf{1}_{\{X_i>0\}}$, $i = 1, \ldots, n$. Note that $\mathbb{P}(Z_i > 0) = \mathbb{P}(X_i > 0) \leqslant \alpha_i$ since $\mathrm{VaR}_{\alpha_i}(X_i) = 0$, $i = 1, \ldots, n$. We have

$$Y_+ = \left(\sum_{i=1}^n X_i\right)_+ \leqslant \left(\sum_{i=1}^n Z_i\right)_+ = \sum_{i=1}^n Z_i.$$

and $Y_i = \mathbf{1}_{\{X_i \leqslant 0\}} X_i$, $i = 1, \ldots, n$. Since $X_1 + \cdots + X_n = Y$, we have $(Y_1, \ldots, Y_n) \in \mathbb{A}_n^-(Y - \sum_{i=1}^n Z_i)$. Therefore, we have

$$X_i = Z_i + Y_i, \quad i = 1, \ldots, n$$

for some $Z_1, \ldots, Z_n$ and $Y_1, \ldots, Y_n$ satisfying $\mathbb{P}(Z_i > 0) \leqslant \alpha_i$, $i = 1, \ldots, n$, $\sum_{i=1}^n Z_i \geqslant Y_+$, and $(Y_1, \ldots, Y_n) \in \mathbb{A}_n^-(Y - \sum_{i=1}^n Z_i)$.

(ii) Let $(X_1, \ldots, X_n)$ be an optimal allocation of X. Recall the notation $y_\alpha = \mathrm{VaR}_\alpha(X)$ and we further write $x_i = \mathrm{VaR}_{\alpha_i}(X)$, $i = 1, \ldots, n$. Note that by Theorem 4.2.1,

$$\begin{aligned}
&(X_1, \ldots, X_n) \text{ is an optimal allocation of } X\\
&\Rightarrow \sum_{i=1}^n \mathrm{VaR}_{\alpha_i}(X_i) = y_\alpha\\
&\Rightarrow \sum_{i=1}^n \mathrm{VaR}_{\alpha_i}(X_i - x_i) = y_\alpha - \sum_{i=1}^n x_i = \mathrm{VaR}_\alpha(X - y_\alpha)\\
&\Rightarrow (X_1 - x_1, \ldots, X_n - x_n) \text{ is an optimal allocation of } X - y_\alpha.
\end{aligned}$$

Therefore, $(X_1 - x_1, \ldots, X_n - x_n)$ is an optimal allocation of $X - y_\alpha$. Observing $\mathrm{VaR}_\alpha(X - y_\alpha) = 0$ and $\mathrm{VaR}_{\alpha_i}(X_i - x_i) = 0$, $i = 1, \ldots, n$, by letting $Y = X - y_\alpha$ in part (i), we obtain

$$X_i - x_i = Z_i + Y_i, \quad i = 1, \ldots, n$$

where $(Z_1, \ldots, Z_n) \in \mathscr{Z}_n$ and $(Y_1, \ldots, Y_n) \in \mathbb{A}_n^-(X - y_\alpha - \sum_{i=1}^n Z_i)$. Therefore, $(X_1, \ldots, X_n)$ has the form in (4.18). □

Assuming $\beta_1 = \cdots = \beta_n = 0$, the optimal allocation (4.10)-(4.11) in Embrechts et al. (2018) is a special case of (4.18), by taking $Z_i = (X - m)\mathbf{1}_{\{1 - \sum_{k=1}^i \alpha_k < U_X \leqslant 1 - \sum_{k=1}^{i-1} \alpha_k\}}$, $i = 1, \ldots, n$, $Y_1 = \cdots = Y_{n-1} = 0$, $Y_n = (X - m)\mathbf{1}_{\{U_X \leqslant 1 - \sum_{k=1}^n \alpha_k\}} + m - y_\alpha$, $c_1 = \cdots = c_{n-1} = 0$, and $c_n = y_\alpha$.

Remark 4.4.1. If $\mathbb{P}(X > y_\alpha) = \alpha$, which is satisfied by all X with a continuous distribution function, then for any $(Z_1, \dots, Z_n) \in \mathcal{Z}_n$, we have

$$Z_i = Z\mathbf{1}_{A_i} \ \text{ a.s.}, \quad i = 1, \dots, n$$

for some $Z \geqslant (X - y_\alpha)_+$, $(A_1, \dots, A_n) \in \pi_n(\{X > y_\alpha\})$ with $\mathbb{P}(A_i) = \alpha_i$, $i = 1, \dots, n$. Note that the random vector $(Z_1, \dots, Z_n)$ is *mutually exclusive* (or *pair-wise countermonotonic*; see Section 3.2 of Puccetti and Wang (2015)), showing a strongest form of negative dependence. Theorem 4.4.1 is the first result showing that *an optimal allocation for VaR agents has to have a mutually exclusive part*, whereas in the literature (Embrechts et al. (2018, 2019)) we only know that *some* optimal allocations for VaR agents have a mutually exclusive part.

4.4.4 ES agents

Next, we consider the case where $\alpha_1 = \cdots = \alpha_n = 0$ and $\beta > 0$, that is, the objective of each agent is an ES. In this case, by Theorem 4.3.3, an optimal allocation exists for $\beta \in (0, 1]$ since (A2) holds in case $\beta \in (0, 1)$ and (A4) holds in case $\beta = 1$. We introduce the following class of allocations. Let $J = \{i \in \{1, \dots, n\} : \beta_i = \beta\}$, that is, J is the set of agents with the largest tolerance parameter. If $0 < \beta < 1$, let $(X_1, \dots, X_n)$ be given by

$$\begin{aligned}
&X_i = Z_i\mathbf{1}_{\{i\in J\}} + Y_i + c_i, \quad i = 1, \dots, n,\\
&\text{where } x \in [y_\beta, y_\beta^+],\ (Z_i)_{i\in J} \in \mathbb{A}^+_{\#J}((X - x)_+),\ (Y_1, \dots, Y_n) \in \mathbb{A}^-_n(-(x - X)_+),\\
&\qquad \text{and } (c_1, \dots, c_n) \in \mathbb{A}^c_n(x).
\end{aligned} \tag{4.19}$$

If $\beta = 1$, let $(X_1, \dots, X_n)$ be given by

$$\begin{aligned}
&X_i = Z_i\mathbf{1}_{\{i\in J\}} + c_i, \quad i = 1, \dots, n,\\
&\text{where } (Z_i)_{i\in J} \in \mathbb{A}_{\#J}(X), \text{ and } (c_1, \dots, c_n) \in \mathbb{A}^c_n(0).
\end{aligned} \tag{4.20}$$

Below we show the optimality of (4.19)-(4.20) and that any optimal allocation of X has the forms (4.19)-(4.20).

Theorem 4.4.2. *Assume $\alpha = 0$ and $\beta \in (0, 1]$. For $X \in \mathcal{X}$, $(X_1, \dots, X_n) \in \mathcal{X}^n$ is an optimal allocation of X if and only if it has the form* (4.19)-(4.20).

Proof. We first show the "if" part. Let $(X_1, \dots, X_n)$ be an optimal allocation of X. If $\beta < 1$, using the VaR-ES relation (4.4), we have

$$\begin{aligned}
\sum_{i=1}^n \mathrm{ES}_{\beta_i}(X_i) &= \sum_{i\in J} \mathrm{ES}_{\beta_i}(Z_i + Y_i) + \sum_{i\in\{1,\dots,n\}\setminus J} \mathrm{ES}_{\beta_i}(Y_i) + x \\
&\leqslant \sum_{i\in J} \mathrm{ES}_{\beta_i}(Z_i) + x \\
&= \sum_{i\in J} \min_{z_i\in\mathbb{R}} \left\{ \frac{1}{\beta}\mathbb{E}[(Z_i - z_i)_+] + z_i \right\} + x \\
&= \min_{z_i\in\mathbb{R},\ i\in J} \left\{ \frac{1}{\beta}\sum_{i\in J}\mathbb{E}[(Z_i - z_i)_+] + \sum_{i\in J} z_i \right\} + x \\
&\leqslant \frac{1}{\beta}\sum_{i\in J}\mathbb{E}[(Z_i)_+] + x = \frac{1}{\beta}\mathbb{E}[(X-x)_+] + x = \mathrm{ES}_\beta(X),
\end{aligned}$$

where the last equality is because $x \in [y_\beta, y_\beta^+]$. If $\beta = 1$, then

$$\begin{aligned}
\sum_{i=1}^n \mathrm{ES}_{\beta_i}(X_i) &= \sum_{i\in J} \mathbb{E}[Z_i + Y_i + c_i] + \sum_{i\in\{1,\dots,n\}\setminus J} \mathrm{ES}_{\beta_i}(Y_i + c_i) \\
&\leqslant \sum_{i\in J}\mathbb{E}[Z_i] + \sum_{i=1}^n c_i = \mathbb{E}[(X-x)_+] + x = \mathbb{E}[X] = \mathrm{ES}_\beta(X).
\end{aligned}$$

In both cases, $(X_1, ..., X_n)$ is optimal.

Next we show the "only-if" part. Let $(X_1, \dots, X_n)$ be an optimal allocation of X. By Theorem 4.2.1, this means

$$\sum_{i=1}^n \mathrm{ES}_{\beta_i}(X_i) = \mathrm{ES}_\beta(X). \tag{4.21}$$

(i) First we assume $\beta \in (0,1)$. Using the VaR-ES relation (4.4), there exist $x_1, \dots, x_n \in \mathbb{R}$, such that

$$\sum_{i=1}^n \mathrm{ES}_{\beta_i}(X_i) = \sum_{i=1}^n \left(\frac{1}{\beta_i}\mathbb{E}[(X_i - x_i)_+] + x_i \right).$$

It follows that

$$\begin{aligned}
\mathrm{ES}_\beta(X) &= \sum_{i=1}^n \left(\frac{1}{\beta_i}\mathbb{E}[(X_i - x_i)_+] + x_i\right) \\
&\geqslant \sum_{i=1}^n \left(\frac{1}{\beta}\mathbb{E}[(X_i - x_i)_+] + x_i\right) \\
&= \frac{1}{\beta}\mathbb{E}\left[\sum_{i=1}^n (X_i - x_i)_+\right] + \sum_{i=1}^n x_i \\
&\geqslant \frac{1}{\beta}\mathbb{E}\left[\left(\sum_{i=1}^n (X_i - x_i)\right)_+\right] + \sum_{i=1}^n x_i \\
&= \frac{1}{\beta}\mathbb{E}\left[\left(X - \sum_{i=1}^n x_i\right)_+\right] + \sum_{i=1}^n x_i \geqslant \min_{x\in\mathbb{R}}\left\{\frac{1}{\beta}\mathbb{E}\left[(X-x)_+\right] + x\right\} = \mathrm{ES}_\beta(X).
\end{aligned}$$

Therefore, the three inequalities above are all equalities, namely

$$\sum_{i=1}^n \left(\frac{1}{\beta_i}\mathbb{E}[(X_i - x_i)_+] + x_i\right) = \sum_{i=1}^n \left(\frac{1}{\beta}\mathbb{E}[(X_i - x_i)_+] + x_i\right), \tag{4.22}$$

$$\frac{1}{\beta}\mathbb{E}\left[\sum_{i=1}^n (X_i - x_i)_+\right] + \sum_{i=1}^n x_i = \frac{1}{\beta}\mathbb{E}\left[\left(X - \sum_{i=1}^n x_i\right)_+\right] + \sum_{i=1}^n x_i, \tag{4.23}$$

and

$$\frac{1}{\beta}\mathbb{E}\left[\left(X - \sum_{i=1}^n x_i\right)_+\right] + \sum_{i=1}^n x_i = \min_{x\in\mathbb{R}}\left\{\frac{1}{\beta}\mathbb{E}\left[(X-x)_+\right] + x\right\}. \tag{4.24}$$

Note that the equalities of expectations in (4.22) and (4.23) are indeed almost surely point-wise equality.

Next, write $x = \sum_{i=1}^n x_i$, $Z_i = (X_i - x_i)_+$ and $Y_i = -(x_i - X_i)_+$ for $i = 1, \dots, n$. Recall that we treat almost equal random variables as identical. By (4.22), we have $Z_i = (X_i - x_i)_+ = 0$ for each $i \notin J$. By (4.23), we have,

$$\sum_{i\in J} Z_i = \sum_{i\in J}(X_i - x_i)_+ = \sum_{i=1}^n (X_i - x_i)_+ = \left(X - \sum_{i=1}^n x_i\right)_+ = (X - x)_+.$$

Consequently, $(Z_i)_{i\in J} \in \mathbb{A}^+_{\#J}((X-x)_+)$. Since $\sum_{i=1}^n X_i = X$, we have

$$\sum_{i=1}^n Y_i = X - \sum_{i=1}^n Z_i - x = X - x - (X-x)_+ = -(x-X)_+,$$

which gives $(Y_1,\dots,Y_n) \in \mathbb{A}^-_n(-(x-X)_+)$. By (4.24) and using the VaR-ES relation (4.5), we have $x \in [y_\beta, y^+_\beta]$. Note that $X_i = (X_i - x_i)_+ - (x_i - X_i)_+ + x_i = Z_i + Y_i + x_i$ for $i = 1,\dots,n$ and $Z_i = 0$ for $i \notin J$. Therefore, $(X_1,\dots,X_n)$ has the form (4.19).

(ii) Assume $\beta = 1$. Equation (4.21) reads as

$$\sum_{i=1}^n \mathrm{ES}_{\beta_i}(X) = \mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X_i].$$

Note that for $Y \in \mathcal{X}$ and $\gamma \in [0,1)$, $\mathrm{ES}_\gamma(Y) \geqslant \mathbb{E}[Y]$ holds, and $\mathrm{ES}_\gamma(Y) = \mathbb{E}[Y]$ if and only if Y is a constant. Therefore, X_i is a constant for all $i \notin J$. This leads to the conclusion that $(X_1,\dots,X_n)$ has the form (4.20). □

Remark 4.4.2. It is trivial to observe that, if there is only one agent whose tolerance parameter is the largest, that is, $\#J = 1$, then $(Z_1,\dots,Z_n)$ is mutually exclusive. Combined with the observation in Remark 4.4.1, in both the case of ES agents and that of VaR agents, $(X_1,\dots,X_n)$ is mutually exclusive on an event that X is large. We will continue discussing this phenomenon in Section 4.5.1.

4.4.5 One VaR agent and one ES agent

We move on to consider the combined case of one VaR agent and one ES agent. For this purpose, assume $n = 2$, $\alpha_1 > 0$, $\beta_1 = \alpha_2 = 0$, and $\beta_2 > 0$. Recall that $\alpha = \alpha_1$ and $\beta = \beta_2$. According to Theorem 4.3.3, for a fixed $X \in \mathcal{X}$, there are two cases where an optimal allocation exists: either (A2) $\alpha + \beta < 1$ or (A3) $\alpha + \beta = 1$ and X is bounded from below. In both cases, $y^+_{\alpha+\beta} > -\infty$. To characterize all optimal allocations, we define the following set

$$\mathcal{A}_{\alpha,\beta} = \{A \in \mathcal{F} : \{X > y_\alpha\} \subset A,\ \mathbb{P}(A) = \alpha;\ \text{moreover, } A \subset \{X \geqslant y_\alpha\} \text{ if } y^+_{\alpha+\beta} \neq y_\alpha\}.$$

In the above notation we omit the reliance on X, which should be clear throughout this section. It is easy to see that $\mathcal{A}_{\alpha,\beta}$ is non-empty as $\mathbb{P}(X > y_\alpha) \leqslant \alpha \leqslant \mathbb{P}(X \geqslant y_\alpha)$.

A set A in $\mathcal{A}_{\alpha,\beta}$ represents an event of probability α on which X takes the largest values. It is clear that, $A = \{X > y_\alpha\}$ if $\mathbb{P}(X > y_\alpha) = \alpha$, and $\{X > y_\alpha\} \subset A \subset \{X \geqslant y_\alpha\}$

if $y^+_{\alpha+\beta} < y_\alpha$. A small complication arises when $\mathbb{P}(X > y_\alpha) \neq \alpha$ and $y^+_{\alpha+\beta} = y_\alpha$, in which case $A \setminus \{X > y_\alpha\}$ can be arbitrary as long as $\mathbb{P}(A) = \alpha$. The reason for this complication can be seen from the proof of Lemma 4.4.3, where an optimization for ES relies on the set $\mathcal{A}_{\alpha,\beta}$. For risk management practice such a special case is irrelevant; it is included in our main results for the completeness of this study.

We first present two lemmas useful in characterizing the optimal allocations, and they may be of independent interest in optimizing ES.

Lemma 4.4.3. *For any $X \in \mathcal{X}$, $\alpha > 0$ and $\beta > 0$ with $\alpha + \beta \leqslant 1$ and $y^+_{\alpha+\beta} > -\infty$, $(Y, B) \in \mathcal{X} \times \mathcal{F}$ is a solution to the problem*

$$\text{to minimize } \ \mathrm{ES}_\beta(X - Y\mathbf{1}_B) \ \text{ subject to } B \in \mathcal{F},\ \mathbb{P}(B) = \alpha \text{ and } Y \in \mathcal{X}, \tag{4.25}$$

if and only if $B \in \mathcal{A}_{\alpha,\beta}$ and $Y\mathbf{1}_B \geqslant (X - y^+_{\alpha+\beta})\mathbf{1}_B$. Moreover, the minimum of (4.25) is $\mathrm{RVaR}_{\alpha,\beta}(X)$.

Proof. We first show the "if" part. Note that by Theorem 4.2.1, for any $B \in \mathcal{F}$, $\mathbb{P}(B) = \alpha$ and $Y \in \mathcal{X}$,

$$\mathrm{RVaR}_{\alpha,\beta}(X) \leqslant \mathrm{VaR}_\alpha(Y\mathbf{1}_B) + \mathrm{ES}_\beta(X - Y\mathbf{1}_B) \leqslant \mathrm{ES}_\beta(X - Y\mathbf{1}_B).$$

Suppose $B \in \mathcal{A}_{\alpha,\beta}$ and $Y\mathbf{1}_B \geqslant (X - y^+_{\alpha+\beta})\mathbf{1}_B$. We have

$$\begin{aligned}\mathrm{ES}_\beta(X - Y\mathbf{1}_B) \leqslant \mathrm{ES}_\beta(X - (X - y^+_{\alpha+\beta})\mathbf{1}_B) &= \frac{1}{\beta}\int_0^\beta \mathrm{VaR}_\alpha(X - (X - y^+_{\alpha+\beta})\mathbf{1}_B)\,\mathrm{d}\gamma \\ &= \frac{1}{\beta}\int_\alpha^{\alpha+\beta} \mathrm{VaR}_{\gamma-\alpha}(X - (X - y^+_{\alpha+\beta})\mathbf{1}_B)\,\mathrm{d}\gamma.\end{aligned}$$

In both the case $y^+_{\alpha+\beta} < y_\alpha$ and the case $y^+_{\alpha+\beta} = y_\alpha$, we have $\mathrm{VaR}_{\gamma-\alpha}(X - (X - y^+_{\alpha+\beta})\mathbf{1}_B) \leqslant \mathrm{VaR}_\gamma(X)$ holds for $\gamma \in [\alpha, \alpha+\beta)$. Hence, $\mathrm{ES}_\beta(X - Y\mathbf{1}_B) \leqslant \mathrm{RVaR}_{\alpha,\beta}(X)$. This shows that (Y, B) satisfying $B \in \mathcal{A}_{\alpha,\beta}$ and $Y\mathbf{1}_B \geqslant (X - y^+_{\alpha+\beta})\mathbf{1}_B$ minimizes (4.25). Moreover, the corresponding minimum is $\mathrm{ES}_\beta(X - Y\mathbf{1}_B) = \mathrm{RVaR}_{\alpha,\beta}(X)$.

We next show the "only-if" direction. Suppose that (Y, B) is such that $\mathrm{ES}_\beta(X - Y\mathbf{1}_B) = \mathrm{RVaR}_{\alpha,\beta}(X)$, namely,

$$\int_\alpha^\beta \mathrm{VaR}_{\gamma-\alpha}(X - Y\mathbf{1}_B)\,\mathrm{d}\gamma = \int_\alpha^{\alpha+\beta} \mathrm{VaR}_\gamma(X)\,\mathrm{d}\gamma. \tag{4.26}$$

Observe that for $\gamma \in (\alpha, \alpha+\beta)$,

$$\mathrm{VaR}_{\gamma-\alpha}(X - Y\mathbf{1}_B) \geqslant \mathrm{VaR}_{\gamma-\alpha}(X - Y\mathbf{1}_B) + \mathrm{VaR}_{\alpha}(Y\mathbf{1}_B) \geqslant \mathrm{VaR}_{\gamma-\alpha}\square\mathrm{VaR}_{\alpha}(X) = \mathrm{VaR}_{\gamma}(X).$$

To make (4.26) hold, we need $\mathrm{VaR}_{\gamma-\alpha}(X - Y\mathbf{1}_B) = \mathrm{VaR}_{\gamma}(X)$ for $\gamma \in (\alpha, \alpha+\beta)$ a.e. By the right-continuity of the left-quantile (VaR), this requires

$$\mathrm{VaR}_{\gamma-\alpha}(X - Y\mathbf{1}_B) = \mathrm{VaR}_{\gamma}(X) \tag{4.27}$$

holds for $\gamma \in [\alpha, \alpha+\beta)$.

Suppose for the purpose of contradiction that $\mathbb{P}(Y < X - y^+_{\alpha+\beta}|B) > 0$. Then,

$$\mathbb{P}(X - Y\mathbf{1}_B > y^+_{\alpha+\beta}) > \mathbb{P}(X - (X - y^+_{\alpha+\beta})\mathbf{1}_B > y^+_{\alpha+\beta}).$$

As a consequence, there exists some $\gamma \in (\alpha, \alpha+\beta)$ such that $\mathrm{VaR}_{\gamma-\alpha}(X - Y\mathbf{1}_B) > \mathrm{VaR}_{\gamma-\alpha}(X - (X - y^+_{\alpha+\beta})\mathbf{1}_B)$. It follows that

$$\begin{aligned}\mathrm{VaR}_{\gamma-\alpha}(X - Y\mathbf{1}_B) &> \mathrm{VaR}_{\gamma-\alpha}\left(X - (X - y^+_{\alpha+\beta})\mathbf{1}_B\right)\\ &\geqslant \mathrm{VaR}_{\gamma-\alpha}\left(X - (X - y^+_{\alpha+\beta})\mathbf{1}_B\right) + \mathrm{VaR}_{\alpha}\left((X - y^+_{\alpha+\beta})\mathbf{1}_B\right) \geqslant \mathrm{VaR}_{\gamma}(X),\end{aligned}$$

contradicting (4.27). Therefore, we have $\mathbb{P}(Y < X - y^+_{\alpha+\beta}|B) = 0$, namely, $Y\mathbf{1}_B \geqslant (X - y^+_{\alpha+\beta})\mathbf{1}_B$.

Next we show $B \in \mathcal{A}_{\alpha,\beta}$. Note that we treat two sets as equal if the difference of the two sets is of measure zero. Equation (4.27) implies $\mathbb{P}(X - Y\mathbf{1}_B \leqslant \mathrm{VaR}_{\gamma}(X)) \geqslant 1 - \gamma + \alpha$. Taking $\gamma = \alpha$, we have $\mathbb{P}(X - Y\mathbf{1}_B \leqslant \mathrm{VaR}_{\alpha}(X)) = 1$. This implies $\{X > \mathrm{VaR}_{\alpha}\} \subset B$.

It remains to show $B \subset \{X \geqslant \mathrm{VaR}_{\alpha}(X)\}$ if $y^+_{\alpha+\beta} \neq y_{\alpha}$. Take $(Y^*, B^*) \in \mathcal{X} \times \mathcal{F}$ such that $B^* \in \mathcal{A}_{\alpha,\beta}$ and $Y^*\mathbf{1}_{B^*} \geqslant (X - y^+_{\alpha+\beta})\mathbf{1}_{B^*}$. From the first part of the proof, we know that (Y^*, B^*) minimizes (4.25). Since (Y, B) also minimizes (4.25), by (4.27), we know

$$\mathrm{VaR}_{\gamma-\alpha}(X - Y\mathbf{1}_B) = \mathrm{VaR}_{\gamma-\alpha}(X - Y^*\mathbf{1}_{B^*}) = \mathrm{VaR}_{\gamma}(X),$$

for $\gamma \in [\alpha, \alpha+\beta)$. Since $y^+_{\alpha+\beta} < y_{\alpha}$, the above equation implies

$$\mathbb{P}(X - Y\mathbf{1}_B \geqslant y_{\alpha}) = \mathbb{P}(X - Y^*\mathbf{1}_{B^*} \geqslant y_{\alpha}). \tag{4.28}$$

Since $Y\mathbf{1}_B \geqslant (X - y^+_{\alpha+\beta})\mathbf{1}_B$ and $Y^*\mathbf{1}_{B^*} \geqslant (X - y^+_{\alpha+\beta})\mathbf{1}_{B^*}$, $\mathbb{P}(X - Y \geqslant y_{\alpha}, B) = \mathbb{P}(X - Y^* \geqslant y_{\alpha}, B^*) = 0$. Using this relation and noting that $B^* \subset \{X \geqslant y_{\alpha}\}$, (4.28) implies

$$\begin{aligned}\mathbb{P}(X \geqslant y_{\alpha}, B^c) = \mathbb{P}(X \geqslant y_{\alpha}, (B^*)^c) = \mathbb{P}(X \geqslant y_{\alpha}) - \mathbb{P}(B^*) &= \mathbb{P}(X \geqslant y_{\alpha}) - \alpha\\ &= \mathbb{P}(X \geqslant y_{\alpha}) - \mathbb{P}(B).\end{aligned}$$

Therefore, $B \subset \{X \geqslant y_{\alpha}\}$. This shows $B \in \mathcal{A}_{\alpha,\beta}$. □

Lemma 4.4.4. *For any $X, Y \in \mathcal{X}$ with $Y \geqslant 0$ and $\beta \in (0,1)$, $\mathrm{ES}_\beta(X+Y) = \mathrm{ES}_\beta(X)$ if and only if $Y \leqslant (\mathrm{VaR}^+_\beta(X) - X)_+$.*

Proof. We first show the "if" direction. Note that

$$X + Y \leqslant X + (\mathrm{VaR}^+_\beta(X) - X)_+ = X \vee \mathrm{VaR}^+_\beta(X).$$

It is easy to see, for $\gamma \in (0, \beta)$, that

$$\mathrm{VaR}_\gamma(X \vee \mathrm{VaR}^+_\beta(X)) = \mathrm{VaR}_\gamma(X).$$

Therefore, $\mathrm{ES}_\beta(X) = \mathrm{ES}_\beta(X \vee \mathrm{VaR}^+_\beta(X)) \geqslant \mathrm{ES}_p(X+Y)$, which implies $\mathrm{ES}_\beta(X) = \mathrm{ES}_\beta(X+Y)$.

Next we show the "only-if" direction. By the fact that

$$\mathrm{ES}_\beta(X) = \frac{1}{\beta}\int_0^\beta \mathrm{VaR}^+_\gamma(X)\,\mathrm{d}\gamma = \frac{1}{\beta}\int_0^\beta \mathrm{VaR}^+_\gamma(X+Y)\,\mathrm{d}\gamma = \mathrm{ES}_\beta(X+Y),$$

and $\mathrm{VaR}^+_\gamma(X) \leqslant \mathrm{VaR}^+_\gamma(X+Y)$ for all $\gamma \in (0, \beta]$, we have $\mathrm{VaR}^+_\gamma(X) = \mathrm{VaR}^+_\gamma(X+Y)$ a.e. on $(0, \beta)$. Since $\mathrm{VaR}^+_\gamma(Z)$ is left-continuous in γ for any fixed $Z \in \mathcal{X}$, $\mathrm{VaR}^+_\gamma(X) = \mathrm{VaR}^+_\gamma(X+Y)$ holds for all $\gamma \in (0, \beta]$. Using the VaR-ES relation (4.4), it follows that

$$\mathbb{E}\left[\left(X - \mathrm{VaR}^+_\beta(X)\right)_+\right] = \mathbb{E}\left[\left(X + Y - \mathrm{VaR}^+_\beta(X)\right)_+\right]. \tag{4.29}$$

Since $Y \geqslant 0$, (4.29) means

$$\left(X - \mathrm{VaR}^+_\beta(X)\right)_+ = \left(X + Y - \mathrm{VaR}^+_\beta(X)\right)_+,$$

and therefore $Y \leqslant (\mathrm{VaR}^+_\beta(X) - X)_+$. $\square$

Now we are ready to characterize the optimal allocations in the setting of this section. Let (X_1, X_2) be given by

$$\begin{aligned} &X_1 = Y\mathbf{1}_B - Z + c, \quad X_2 = X - X_1 \\ &\text{where } B \in \mathcal{A}_{\alpha,\beta},\ Y \geqslant X - y^+_{\alpha+\beta},\ 0 \leqslant Z \leqslant (y^+_{\alpha+\beta} - X + Y\mathbf{1}_B)_+, \text{ and } c \in \mathbb{R}. \end{aligned} \tag{4.30}$$

Theorem 4.4.5. *Assume $\alpha_1 > 0$, $\beta_1 = \alpha_2 = 0$, $\beta_2 > 0$, and either (A2) or (A3) holds. For $X \in \mathcal{X}$, $(X_1, X_2) \in \mathcal{X}^2$ is an optimal allocation of X if and only if it has the form (4.30).*

Proof. Obviously, the constant c does not matter in terms of the optimality of (X_1, X_2), and we set $c = 0$ for simplicity.

We first show that (4.30) gives an optimal allocation. It is easy to verify that

$$\mathrm{VaR}_\alpha(X_1) = \mathrm{VaR}_\alpha(Y\mathbf{1}_B - Z) \leqslant \mathrm{VaR}_\alpha(Y\mathbf{1}_B) = 0.$$

Using Lemmas 4.4.3 and 4.4.4, and noting that $\mathrm{VaR}^+_\beta(X - Y\mathbf{1}_B) = y^+_{\alpha+\beta}$ as implied by (4.27), we have

$$\mathrm{ES}_\beta(X_2) = \mathrm{ES}_\beta(X - Y\mathbf{1}_B + Z) = \mathrm{ES}_\beta(X - Y\mathbf{1}_B) = \mathrm{RVaR}_{\alpha,\beta}(X).$$

Therefore, (X_1, X_2) is an optimal allocation.

Next, suppose that (X_1, X_2) is an optimal allocation. Without loss of generality, assume $\mathrm{VaR}_\alpha(X_1) = 0$, which implies $\mathbb{P}(X_1 \geqslant 0) \geqslant \alpha \geqslant \mathbb{P}(X_1 > 0)$. Therefore, there exists $B \in \mathcal{F}$ such that $\{X_1 \geqslant 0\} \subset B \subset \{X_1 > 0\}$ with $\mathbb{P}(B) = \alpha$. Write $X_1 = X_1\mathbf{1}_B - Z$ where $Z = -X_1\mathbf{1}_{B^c}$. Note that $Z = -X_1\mathbf{1}_{B^c} \geqslant 0$. Since (X_1, X_2) is an optimal allocation, we know

$$\mathrm{RVaR}_{\alpha,\beta}(X) = \mathrm{VaR}_\alpha(X_1) + \mathrm{ES}_\beta(X_2) = \mathrm{ES}_\beta(X - X_1\mathbf{1}_B + Z) \geqslant \mathrm{ES}_\beta(X - X_1\mathbf{1}_B).$$

From Lemma 4.4.3, we know that (X_1, B) minimizes (4.25). The results of Lemma 4.4.3 imply $B \in \mathcal{A}_{\alpha,\beta}$, $X_1\mathbf{1}_B \geqslant (X - y^+_{\alpha+\beta})\mathbf{1}_B$, and

$$\mathrm{RVaR}_{\alpha,\beta}(X) = \mathrm{ES}_\beta(X - X_1\mathbf{1}_B) = \mathrm{ES}_\beta(X - X_1\mathbf{1}_B + Z).$$

Using Lemma 4.4.4, we have $Z \leqslant (y^+_{\alpha+\beta} - X + X_1\mathbf{1}_B)_+$. Let $Y = X_1\mathbf{1}_B + (X - y^+_{\alpha+\beta})\mathbf{1}_{B^c}$. It is clear that $Y\mathbf{1}_B = X_1\mathbf{1}_B$, $Y \geqslant X - y^+_{\alpha+\beta}$. Therefore, $X_1 = X_1\mathbf{1}_B + X_1\mathbf{1}_{B^c} = Y\mathbf{1}_B - Z$, which has the form (4.30). $\square$

4.4.6 RVaR agents

Finally, based on the results in Sections 4.4.2-4.4.5, we are able to present some general result for in the case of RVaR agents. The main idea here is that, for each $i = 1, \dots, n$, we write $\mathrm{RVaR}_{\alpha_i,\beta_i} = \mathrm{VaR}_{\alpha_i} \square \mathrm{ES}_{\beta_i}$, and reduce the risk sharing problem to the above studied cases. We summarize this methodology in the following proposition. Since we need to translate between different cases, below we emphasize the risk measures with respect to which we speak of optimality. The case (A4') requires a special treatment which will be discussed later.

Proposition 2. *Assume (A4') does not hold.* $(X_1, \ldots, X_n)$ *is an optimal allocation of* $X \in \mathcal{X}$ *with respect to* $(\mathrm{RVaR}_{\alpha_1,\beta_1}, \ldots, \mathrm{RVaR}_{\alpha_n,\beta_n})$ *if and only if there exist an optimal allocation* (Y, Z) *of* X *with respect to* $(\mathrm{VaR}_\alpha, \mathrm{ES}_\beta)$*, an optimal allocation* $(Y_1, \ldots, Y_n)$ *of* Y *with respect to* $(\mathrm{VaR}_{\alpha_1}, \ldots, \mathrm{VaR}_{\alpha_n})$*, and an optimal allocation* $(Z_1, \ldots, Z_n)$ *of* Z *with respect to* $(\mathrm{ES}_{\beta_1}, \ldots, \mathrm{ES}_{\beta_n})$*, such that*

$$X_i = Y_i + Z_i, \quad i = 1, \ldots, n.$$

Proof. We first show the "if" part. From the construction of X, it is easy to calculate

$$\begin{aligned}
\sum_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) &= \sum_{i=1}^n \mathrm{VaR}_{\alpha_i} \square \mathrm{ES}_{\beta_i}(X_i) \\
&\leqslant \sum_{i=1}^n (\mathrm{VaR}_{\alpha_i}(Y_i) + \mathrm{ES}_{\beta_i}(Z_i)) \\
&= \sum_{i=1}^n \mathrm{VaR}_{\alpha_i}(Y_i) + \sum_{i=1}^n \mathrm{ES}_{\beta_i}(Z_i) \\
&= \mathrm{VaR}_\alpha(Y) + \mathrm{ES}_\beta(Z) = \mathrm{RVaR}_{\alpha,\beta}(X).
\end{aligned}$$

Therefore, $(X_1, \ldots, X_n)$ is an optimal allocation of X with respect to $(\mathrm{RVaR}_{\alpha_1,\beta_1}, \ldots, \mathrm{RVaR}_{\alpha_n,\beta_n})$. Next we show the "only-if" part. Suppose that $(X_1, \ldots, X_n)$ is an optimal allocation of $\mathcal{X}$ with respect to $(\mathrm{RVaR}_{\alpha_1,\beta_1}, \ldots, \mathrm{RVaR}_{\alpha_n,\beta_n})$.

Since (A4') does not hold, it is easy to see from the existence of the optimal allocation and Theorem 4.3.3 that for each $i = 1, \ldots, n$, $\alpha_i + \beta_i < 1$ and $\mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) \neq \infty$. As a consequence, for each $i = 1, \ldots, n$, we can use Theorem 4.3.3 on X_i to conclude that there exists $(Y_i, Z_i) \in \mathbb{A}_2(X_i)$ such that $\mathrm{VaR}_{\alpha_i}(Y_i) + \mathrm{ES}_{\beta_i}(Z_i) = \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i)$. Write $Y = \sum_{i=1}^n Y_i$ and $Z = \sum_{i=1}^n Z_i$. Clearly $Y + Z = X$. It follows that

$$\begin{aligned}
\mathrm{RVaR}_{\alpha,\beta}(X) &= \sum_{i=1}^n \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) \\
&= \sum_{i=1}^n \mathrm{VaR}_{\alpha_i} \square \mathrm{ES}_{\beta_i}(X_i) \\
&= \sum_{i=1}^n (\mathrm{VaR}_{\alpha_i}(Y_i) + \mathrm{ES}_{\beta_i}(Z_i)) \\
&= \sum_{i=1}^n \mathrm{VaR}_{\alpha_i}(Y_i) + \sum_{i=1}^n \mathrm{ES}_{\beta_i}(Z_i) \geqslant \mathrm{VaR}_\alpha(Y) + \mathrm{ES}_\beta(Z) \geqslant \mathrm{RVaR}_{\alpha,\beta}(X),
\end{aligned}$$

where the two inequalities are due to Theorem 4.2.1. Noting that

$$\mathrm{RVaR}_{\alpha,\beta}(X) \geqslant \mathrm{VaR}_{\alpha}(Y) + \mathrm{ES}_{\beta}(Z) \geqslant \mathrm{RVaR}_{\alpha,\beta}(X),$$

the inequalities herein are equalities. Therefore $\sum_{i=1}^{n} \mathrm{VaR}_{\alpha_i}(Y_i) = \mathrm{VaR}_{\alpha}(Y)$, $\sum_{i=1}^{n} \mathrm{ES}_{\beta_i}(Z_i) = \mathrm{ES}_{\beta}(Z)$, and $\mathrm{VaR}_{\alpha}(Y)+\mathrm{ES}_{\beta}(Z) = \mathrm{RVaR}_{\alpha,\beta}(X)$. In other words, (Y, Z) is an optimal allocation of X with respect to $(\mathrm{VaR}_{\alpha}, \mathrm{ES}_{\beta})$, $(Y_1, \ldots, Y_n)$ is an optimal allocation of Y with respect to $(\mathrm{VaR}_{\alpha_1}, \ldots, \mathrm{VaR}_{\alpha_n})$, and $(Z_1, \ldots, Z_n)$ is an optimal allocation of Z with respect to $(\mathrm{ES}_{\beta_1}, \ldots, \mathrm{ES}_{\beta_n})$. □

The reason why case (A4') requires a special treatment can also be seen from the proof. A key step in the proof is to write $X_i = Y_i + Z_i$ where $(Y_i, Z_i) \in \mathbb{A}_2(X_i)$ satisfies $\mathrm{VaR}_{\alpha_i}(Y_i) + \mathrm{ES}_{\beta_i}(Z_i) = \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i)$. If (A4') holds, then such (Y_i, Z_i) may not exist, as shown in Lemma 4.3.2. Below we analyze the case of (A4'), which is different from all other cases. Recall that, (A4') implies that there exists $j \in \{1, \ldots, n\}$ such that $\alpha_j = \alpha$ and $\beta_j = \beta = 1 - \alpha$.

Proposition 3. *Assume that (A4') holds, and without loss of generality, $\alpha_n = \alpha$ and $\beta_n = \beta$. Then, $(X_1, \ldots, X_n)$ is an optimal allocation of $X \in \mathcal{X}$ with respect to $(\mathrm{RVaR}_{\alpha_1,\beta_1}, \ldots, \mathrm{RVaR}_{\alpha_n,\beta_n})$ if and only if $(X_1, \ldots, X_{n-1}, -X)$ is an optimal allocation of $-X_n$ with respect to $(\mathrm{ES}_{\beta_1}, \ldots, \mathrm{ES}_{\beta_n})$.*

Proof. We shall use the fact that, for all $Y \in \mathcal{X}$,

$$\mathrm{RVaR}_{\alpha,\beta}(Y) = \mathrm{RVaR}_{\alpha,1-\alpha}(Y) = -\mathrm{ES}_{1-\alpha}(-Y) = -\mathrm{ES}_{\beta_n}(-Y), \tag{4.31}$$

which is immediate from the definition of RVaR. Note that for any $(X_1, \ldots, X_n) \in \mathcal{X}^n$,

$$\sum_{i=1}^{n} \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) = \sum_{i=1}^{n-1} \mathrm{ES}_{\beta_i}(X_i) + \mathrm{RVaR}_{\alpha,1-\alpha}(X_n) = \sum_{i=1}^{n-1} \mathrm{ES}_{\beta_i}(X_i) - \mathrm{ES}_{\beta_n}(-X_n). \tag{4.32}$$

By (4.31) and (4.32), the equality

$$\sum_{i=1}^{n} \mathrm{RVaR}_{\alpha_i,\beta_i}(X_i) = \mathrm{RVaR}_{\alpha,\beta}(X) \tag{4.33}$$

is equivalent to

$$\sum_{i=1}^{n-1} \mathrm{ES}_{\beta_i}(X_i) - \mathrm{ES}_{\beta_n}(-X_n) = -\mathrm{ES}_{\beta_n}(-X).$$

Rearranging terms, it is

$$\sum_{i=1}^{n-1} \mathrm{ES}_{\beta_i}(X_i) + \mathrm{ES}_{\beta_n}(-X) = \mathrm{ES}_{\beta_n}(-X_n). \tag{4.34}$$

As (4.33) is equivalent to (4.34), the proposition holds. $\square$

Before ending this section, we remark that, although we are able to translate the general case of RVaR agents to the completely characterized cases in Sections 4.4.2-4.4.5, we were not able to write down an elegant unifying form of the optimal allocations, due to the complications raised in the two-step characterization in Proposition 2.

4.5 Discussions

4.5.1 A representative class of optimal allocations

As is seen from Section 4.4, optimal allocations may take various forms, and this is due to the fact that the RVaR family only uses partial information of the underlying distribution. Among many choices of optimal allocations, the allocation (4.10)-(4.11) obtained by Embrechts et al. (2018) is a rather simple and intuitive choice. One small disadvantage of (4.10)-(4.11) is that it is not symmetric with respect to the order of agents. Below we present a slightly more general class, which is also simple, and generalizes (4.10)-(4.11) to a symmetric form. We consider the most relevant case (A2), that is, $0 < \alpha + \beta < 1$. Define

$$\mathcal{P}_n = \left\{ (A_1, \dots, A_n) \in \mathcal{F}^n : \{X > y_\alpha\} \subset \bigcup_{i=1}^n A_i \subset \{X \geqslant y_\alpha\},\ \mathbb{P}(A_i) = \alpha_i,\ {\scriptstyle A_1, \dots, A_n \text{ are disjoint}} \right\}.$$

Recall that $J = \{i \in \{1, \dots, n\} : \beta_i = \beta\}$. Let $(X_1, \dots, X_n)$ be given by

$$\begin{aligned} &X_i = (X - y_{\alpha+\beta})\mathbf{1}_{A_i} + \frac{1}{\#J}(X - y_{\alpha+\beta})\mathbf{1}_{\{i \in J\}}\mathbf{1}_{(\bigcup_{i=1}^n A_i)^c} + c_i, \quad i = 1, \dots, n, \\ &\text{where } (A_1, \dots, A_n) \in \mathcal{P}_n, \text{ and } (c_1, \dots, c_n) \in \mathbb{A}_n^c(y_{\alpha+\beta}). \end{aligned} \tag{4.35}$$

One can easily verify that (4.35) defines a class of optimal allocations, and it includes (4.10)-(4.11) if $\#J = 1$. Comparing with the results in Section 4.4, (4.35) gives one of the the simplest forms of optimal allocations.

The economic intuition behind (4.35) is also simple: the agents first share the most dangerous outcomes, modelled by the event $\bigcup_{i=1}^n A_i$ with probability α. They divide the event into pieces so that each agent feels safe, because the probability of loss is equal to α_i for agent i, thus insensitive to the agent. Then, they share the rest of the risk among agents in J, the most tolerant agents (with the biggest β_i value). Finally, they make some side-payments $c_1, \dots, c_n$.

The dependence structure of (4.35) may be worth noting. Assume $\mathbb{P}(X > y_\alpha) = \alpha$, as satisfied by all continuously distributed X. In this case, $\bigcup_{i=1}^n A_i = \{X > y_\alpha\}$. The optimal allocation $(X_1, \dots, X_n)$ is *mutually exclusive* on the event $\{X > y_\alpha\}$, a consistent observation with the ones made in Remarks 4.4.1 and 4.4.2. Mutual exclusivity represents the strongest form of negative dependence (see e.g. Puccetti and Wang (2015)). This is in sharp contrast to the classic risk sharing problems with law-invariant and convex objective functionals, where an optimal allocations is always *comonotonic* (based on a result of Landsberger and Meilijson (1994); see Rüschendorf (2013)), representing the strongest form of positive dependence. For related discussions on this phenomenon in the context of heterogeneous beliefs, see Embrechts et al. (2019).

We remark that in order to arrive at a specific form of optimal allocations, one may need to involve a second-step optimization. Here, we give the representative allocation (4.35) only for the simplicity in its form, economic intuition and dependence structure.

4.5.2 VaR-type risk measures

In this section, we discuss how the techniques developed in this chapter can be applied to other types of risk measures. In particular, we consider the VaR-type risk measures as studied by Weber (2018). To avoid cases of $\infty - \infty$, we choose the underlying space $\mathcal{Y}$ as the set of bounded random variables as in Weber (2018). We first give the necessary definitions.

Definition 4.5.1. (i) A *distortion function* g is a left-continuous and non-decreasing function on $[0, 1]$ with $g(0) = 0$ and $g(1) = 1$. We denote by $\mathcal{G}$ the set of distortion functions.

(ii) A *distortion risk measure* ρ_g on $\mathcal{Y}$ with distortion function g is defined as Choquet integral

$$\rho_g(X) = \int X \,\mathrm{d}(g \circ \mathbb{P}) = \int_{-\infty}^0 (g \circ \mathbb{P}(X > x) - 1)\,\mathrm{d}x + \int_0^\infty g \circ \mathbb{P}(X > x)\,\mathrm{d}x, \quad X \in \mathcal{Y}. \tag{4.36}$$

(iii) For a distortion function g, the number $\alpha = \sup\{t \in [0,1] : g(t) = 0\} \in [0,1)$ is called the *parameter* of g.

(iv) A distortion risk measure is said to be *VaR-type* if the parameter α of its distortion function is positive.

(v) For a distortion function g with parameter α, the function $\hat{g}$, defined by $\hat{g}(t) = g((t+\alpha) \wedge 1)$, $t \in [0,1]$, is called the *active part* of g.

Remark 4.5.1. In the literature, the distortion risk measure ρ_g defined by (4.36) does not require g to be left-continuous. Here we consider the case of left-continuous function g as in Weber (2018). It is well known that if g is left-continuous, ρ_g can be written in a Lebesgue-Stieltjes integral form

$$\rho_g(X) = \int_0^1 \mathrm{VaR}_\gamma(X)\,\mathrm{d}g(\gamma). \tag{4.37}$$

Clearly, (4.37) includes the RVaR family by definition.

Weber (2018) studied the (sum-)optimal risk sharing problem with respect to the distortion risk measures $\rho_{g_1}, \ldots, \rho_{g_n}$, where the distortion functions have concave active parts. Clearly, for $\alpha \in (0,1)$ and $\beta \in [0, 1-\alpha)$, the risk measure $\mathrm{RVaR}_{\alpha,\beta}$ is a VaR-type distortion risk measure with parameter α, and its distortion function has a concave active part. Below, we illustrate how the technique developed in Section 4.4 can be applied to investigate optimal allocations for VaR-type risk measures. In what follows, the set of allocations and the inf-convolution are defined as in (4.6) and (4.7) with $\mathcal{X}$ replaced by $\mathcal{Y}$.

Below, we establish a connection between a VaR-type distortion risk measure and a corresponding VaR, generalizing the formula $\mathrm{RVaR}_{\alpha,\beta} = \mathrm{VaR}_\alpha \square \mathrm{ES}_\beta$ which we used repeatedly in this chapter. To make the presentation concise, for $h \in \mathcal{G}$ and $\alpha \in [0,1)$, we define $h_\alpha(t) = h((t-\alpha)_+)$, $t \in [0,1]$. Clearly, $h_0 = h$, and $h_\alpha \in \mathcal{G}$ if $h(1-\alpha) = 1$. Moreover, for $g \in \mathcal{G}$ with parameter α, we can easily get $\hat{g}_\alpha = g$.

Theorem 4.5.1. *For any $h \in \mathcal{G}$ and $\alpha \in [0,1)$, we have*

$$\mathrm{VaR}_\alpha \square \rho_h = \begin{cases} \rho_{h_\alpha} & \text{if } h(1-\alpha) = 1, \\ -\infty & \text{if } h(1-\alpha) < 1. \end{cases}$$

In particular, for any $g \in \mathcal{G}$ with parameter α, we have

$$\rho_g = \mathrm{VaR}_\alpha \square \rho_{\hat{g}}.$$

Proof. Take any $X \in \mathcal{Y}$. We first consider the case $h(1-\alpha) < 1$. Take $A \in \mathcal{F}$ with $\mathbb{P}(A) = \alpha$ and $m > 0$. Note that $\int_0^\infty h \circ \mathbb{P}(X > x)\,\mathrm{d}x < \infty$ because X is bounded. Using (4.36), we have

$$\begin{aligned}
&\rho_h\left(X\mathbf{1}_{A^c} - m\mathbf{1}_A\right) \\
&= \int_{-\infty}^{0} (h \circ \mathbb{P}(X\mathbf{1}_{A^c} - m\mathbf{1}_A > x) - 1)\,\mathrm{d}x + \int_0^\infty h \circ \mathbb{P}(X\mathbf{1}_{A^c} - m\mathbf{1}_A > x)\,\mathrm{d}x \\
&\leqslant \int_{-m}^{0} (h \circ \mathbb{P}(X\mathbf{1}_{A^c} - m\mathbf{1}_A > x) - 1)\,\mathrm{d}x + \int_0^\infty h \circ \mathbb{P}(X > x)\,\mathrm{d}x \\
&\leqslant \int_{-m}^{0} (h \circ \mathbb{P}(A^c) - 1)\,\mathrm{d}x + \int_0^\infty h \circ \mathbb{P}(X > x)\,\mathrm{d}x \\
&= \int_{-m}^{0} (h(1-\alpha) - 1)\,\mathrm{d}x + \int_0^\infty h \circ \mathbb{P}(X > x)\,\mathrm{d}x \\
&= m(h(1-\alpha) - 1) + \int_0^\infty h \circ \mathbb{P}(X > x)\,\mathrm{d}x \to -\infty \quad \text{as } m \to \infty.
\end{aligned}$$

On the other hand, $\mathrm{VaR}_\alpha((X+m)\mathbf{1}_A) \leqslant 0$ because $\mathbb{P}((X+m)\mathbf{1}_A > 0) \leqslant \mathbb{P}(A) = \alpha$. Combining the above observations, we have $\rho_h\left(X\mathbf{1}_{A^c} - m\mathbf{1}_A\right) + \mathrm{VaR}_\alpha((X+m)\mathbf{1}_A) \to -\infty$ as $m \to \infty$, and hence $\mathrm{VaR}_\alpha \square \rho_h(X) = -\infty$ for all $X \in \mathcal{Y}$.

Next, we consider the case $h(1-\alpha) = 1$. Assume $X \geqslant 0$; this is without loss of generality since both ρ_{h_α} and $\mathrm{VaR}_\alpha \square \rho_h$ satisfy the property (called cash-additivity) $\rho(X + c) = \rho(X) + c$ for any constant $c \in \mathbb{R}$. The case $\alpha = 0$ follows from the simple fact that, for all $Y \in \mathcal{Y}$,

$$\begin{aligned}
\mathrm{VaR}_0 \square \rho_h(X) \leqslant \mathrm{VaR}_0(0) + \rho_h(X) = \rho_h(X) &= \mathrm{VaR}_0(Y) + \rho_h(X - \mathrm{VaR}_0(Y)) \\
&\leqslant \mathrm{VaR}_0(Y) + \rho_h(X - Y),
\end{aligned}$$

and thus $\mathrm{VaR}_0 \square \rho_h(X) = \rho_h(X)$. In the following we assume $\alpha > 0$.

(i) We first show $\rho_{h_\alpha}(X) \geqslant \mathrm{VaR}_\alpha \Box \rho_h(X)$. Note that $\mathrm{VaR}_\alpha(X\mathbf{1}_{\{U_X>1-\alpha\}}) = 0$. We have

$$\begin{aligned}
\rho_h\left(X\mathbf{1}_{\{U_X\leqslant 1-\alpha\}}\right) + \mathrm{VaR}_\alpha\left(X\mathbf{1}_{\{U_X>1-\alpha\}}\right) &= \rho_h\left(X\mathbf{1}_{\{U_X\leqslant 1-\alpha\}}\right) \\
&= \int_0^\infty h \circ \mathbb{P}(X\mathbf{1}_{\{U_X\leqslant 1-\alpha\}} > x)\,\mathrm{d}x \\
&= \int_0^\infty h \circ \mathbb{P}(\{X > x\} \cup \{U_X \leqslant 1-\alpha\})\,\mathrm{d}x \\
&= \int_0^\infty h((\mathbb{P}(X > x) - \alpha)_+)\,\mathrm{d}x \\
&= \int_0^\infty h_\alpha(\mathbb{P}(X > x))\,\mathrm{d}x = \rho_{h_\alpha}(X).
\end{aligned}$$

By the definition of inf-convolution, we have $\rho_{h_\alpha}(X) \geqslant \mathrm{VaR}_\alpha \Box \rho_h(X)$.

(ii) Next we show $\rho_{h_\alpha}(X) \leqslant \mathrm{VaR}_\alpha \Box \rho_h(X)$. For this, it suffices to show $\rho_{h_\alpha}(X) \leqslant \rho_h(X - Y)$ for all $Y \in \mathcal{Y}$ with $\mathrm{VaR}_\alpha(Y) = 0$, again due to cash-additivity. Since $\mathrm{VaR}_\alpha(Y) = 0$ implies $\mathbb{P}(Y > 0) \leqslant \alpha$, we have

$$\mathbb{P}(X - Y > x) \geqslant (\mathbb{P}(X > x) - \mathbb{P}(Y > 0))_+ \geqslant (\mathbb{P}(X > x) - \alpha)_+, \quad x \in \mathbb{R}.$$

As a consequence,

$$\begin{aligned}
\rho_h(X - Y) &= \int_{-\infty}^0 (h \circ \mathbb{P}(X - Y > x) - 1)\,\mathrm{d}x + \int_0^\infty h \circ \mathbb{P}(X - Y > x)\,\mathrm{d}x \\
&\geqslant \int_{-\infty}^0 (h((\mathbb{P}(X > x) - \alpha)_+) - 1)\,\mathrm{d}x + \int_0^\infty h((\mathbb{P}(X > x) - \alpha)_+)\,\mathrm{d}x \\
&= \int_{-\infty}^0 (h_\alpha \circ \mathbb{P}(X > x) - 1)\,\mathrm{d}x + \int_0^\infty h_\alpha \circ \mathbb{P}(X > x)\,\mathrm{d}x = \rho_{h_\alpha}(X).
\end{aligned}$$

Therefore, we know $\rho_{h_\alpha}(X) \leqslant \mathrm{VaR}_\alpha \Box \rho_h(X)$. $\square$

As shown in the proof of Theorem 4.5.1, for $h \in \mathcal{G}$ with $h(1-\alpha) = 1$, a sum-optimal allocation (Y, Z) of $X \geqslant 0$ with respect to $(\mathrm{VaR}_\alpha, \rho_h)$ is given by $Y = X\mathbf{1}_{\{U_X>1-\alpha\}}$ and $Z = X\mathbf{1}_{\{U_X\leqslant 1-\alpha\}}$.

Theorem 4.5.1 suggests that a VaR-type distortion risk measure is simply the inf-convolution of a VaR and another distortion risk measure. Using Theorem 4.5.1, we can apply the results in Section 4.4 to study forms of optimal allocations for VaR-type distortion risk measures. For $g_1, \dots, g_n \in \mathcal{G}$ with parameters $\alpha_1, \dots, \alpha_n$, respectively, assume that

$\square_{i=1}^n \rho_{g_i}$ is finite on $\mathcal{Y}$. Write $\alpha = \sum_{i=1}^n \alpha_i$ and $\rho^* = \square_{i=1}^n \rho_{\hat{g}_i}$. Similarly to (4.16), noting that the inf-convolution is associative,

$$\mathop{\square}_{i=1}^{n} \rho_{g_i} = \mathop{\square}_{i=1}^{n} (\mathrm{VaR}_{\alpha_i} \square \rho_{\hat{g}_i}) = \left(\mathop{\square}_{i=1}^{n} \mathrm{VaR}_{\alpha_i} \right) \square \left(\mathop{\square}_{i=1}^{n} \rho_{\hat{g}_i} \right) = \mathrm{VaR}_{\alpha} \square \rho^*. \tag{4.38}$$

According to (4.38) and following the idea of Proposition 2, the problem of finding optimal allocations for the risk measures $\rho_{g_1}, \dots, \rho_{g_n}$ can be decomposed into two steps: first, allocate X to $(Y, Z) \in \mathcal{A}_2(X)$ such that $\mathrm{VaR}_\alpha \square \rho^*(X) = \mathrm{VaR}_\alpha(Y) + \rho^*(Z)$, and second, allocate Y and Z to $(Y_1, \dots, Y_n) \in \mathcal{A}_n(Y)$ and $(Z_1, \dots, Z_n) \in \mathcal{A}_n(Z)$ such that $\sum_{i=1}^n \mathrm{VaR}_{\alpha_i}(Y_i) = \mathrm{VaR}_\alpha(Y)$ and $\sum_{i=1}^n \rho_{g_i}(Z_i) = \rho^*(Z)$. If all of the above allocations exist, then by letting $X_i = Y_i + Z_i$, $i = 1, \dots, n$, we obtain an optimal allocation for the agents with risk measures $\rho_{g_1}, \dots, \rho_{g_n}$. In the above procedure, there are three optimal allocation problems:

(i) Theorem 4.4.1 gives all forms of optimal allocations $(Y_1, \dots, Y_n) \in \mathcal{A}_n(Y)$ with respect to $(\mathrm{VaR}_{\alpha_1}, \dots, \mathrm{VaR}_{\alpha_n})$.

(ii) Solution of the optimal allocations $(Z_1, \dots, Z_n) \in \mathcal{A}_n(Z)$ with respect to general choices of $\rho_{\hat{g}_1}, \dots, \rho_{\hat{g}_n}$ is not available in the literature. In the special case that each of $g_1, \dots, g_n$ has a concave active part, as studied by Weber (2018), $\rho_{\hat{g}_1}, \dots, \rho_{\hat{g}_n}$ are coherent distortion risk measures. In this case, there always exist comonotonic optimal allocations, and

$$\rho^* = \mathop{\square}_{i=1}^{n} \rho_{\hat{g}_i} = \rho_{g*}, \qquad \text{where } g^* = \textstyle\bigwedge_{i=1}^n g_i; \tag{4.39}$$

see e.g. Proposition 5 of Embrechts et al. (2018). For coherent distortion risk measures, the forms of optimal allocations $(Z_1, \dots, Z_n) \in \mathcal{A}_n(Z)$ are extensively studied in the literature; see Jouini et al. (2008) and Chapter 11 of Rüschendorf (2013).

(iii) The determination of optimal allocations $(Y, Z) \in \mathcal{A}_2(X)$ with respect to $(\mathrm{VaR}_\alpha, \rho^*)$ requires a result that is similar to Theorem 4.4.5, which depends highly on the form of ρ^*. The proof of Theorem 4.5.1 gives an optimal allocation when ρ^* is also a distortion risk measure (which is true if $g_1, \dots, g_n$ have concave active parts).

Although the above arguments do not give explicit forms of all optimal allocations for the VaR-type distortion risk measures, they offer technical tools as well as new interpretation of VaR-type risk measures and their optimal allocations. A full characterization of optimal allocations with respect to VaR-type risk measures requires future research.

Remark 4.5.2. If ρ_h is chosen as ES_β, Theorem 4.5.1 recovers the formula $\mathrm{RVaR}_{\alpha,\beta} = \mathrm{VaR}_\alpha \square \mathrm{ES}_\beta$ by checking the distortion functions of $\mathrm{RVaR}_{\alpha,\beta}$, VaR_α and ES_β in (4.37). Hence, Theorem 4.5.1 can be seen as a generalization of Theorem 2 of Embrechts et al. (2018). Theorem 4.5.1 also implies the result on $\square_{i=1}^n \rho_{g_i}$ in Theorem 11 of Weber (2018) for $g_1, \dots, g_n$ with concave active parts via (4.38) and (4.39).

Chapter 5

Conclusions and Future Work

We make some remarks and briefly discuss possible future work in this chapter.

From the results obtained in Chapter 2, we clearly see that many profound and elegant mathematical results in the theory of risk functionals remain valid for the general class of signed Choquet integrals; they do not rely on the common assumption of monotonicity. Hopefully, our results serve as a building block for future theoretical developments and applications of signed Choquet integrals.

Our discussions are confined to the space of bounded random variables L^∞, in order for signed Choquet integrals to be properly defined, and for all results to be concisely stated. Some results involve norm-continuity on the space or an operation (addition or subtraction) on several signed Choquet integrals, and hence we need to fix a suitable domain upfront. Certainly, many results can be naturally generalized to functional-specific spaces such as Λ-spaces (Lorentz (1951)) and Orlicz spaces (e.g. Rao and Ren (1991)). We leave this direction of research for future work.

In Chapter 3, we provide various results on the CxLS property of one- and multi-dimensional risk functionals. Two major characterization results are established on signed Choquet integrals I_h with CxLS and on (I_h, VaR_p) with CxLS. A particularly elegant message is that the only type of signed Choquet integral that gains CxLS when paired with VaR_p is a linear combination of ES_p and $\mathbb{E}$. Based on these results, we proceed to show that a convex combination of $\mathbb{E}$ and ES_p is the only comonotonic-additive coherent risk measure that is co-elicitable with VaR_p.

It however remains an open question to characterize all two-dimensional signed Choquet integrals (I_h, I_g) with CxLS, or, furthermore, a similar problem in higher dimension. Given the level of technical complexity displayed in the techniques used to show Theorem 3.5.2, it seems to us that a general conclusion to the above question is far from being reachable with current methods. Even if one assumes that the signed Choquet integrals are increasing as in the risk measure literature, general results in multi-dimension are not available.

Closely related to the above issue, the characterization of elicitable (or identifiable) d-dimensional signed Choquet integrals remains an open problem. As explained in Section 3.6, the issues of elicitability, identifiability and backtestability are highly relevant for risk management practice, and they all require the CxLS property as a necessary condition. Hence, our study on CxLS provides a useful tool for future studies on the statistical notions above, especially in the multi-dimensional setting. In addition to the field of risk management, elicitation is an important property in machine learning (Steinwart et al. (2014)). Many open problems related to elicitation complexity remain in the field of computer science (Frongillo and Kash (2018)).

Chapter 4 obtained the characterization of the optimal allocations in quantile-based risk sharing. The competitive equilibria discussed in Embrechts et al. (2018) is left untouched. In Embrechts et al. (2018), competitive equilibria for RVaR agents were studied under a "no-short-sell" assumption, thus an incomplete market. It is shown that an equilibrium allocation is necessarily optimal, but an optimal allocation is not necessarily an equilibrium allocation; the optimal allocation (4.10)-(4.11) is in a competitive equilibrium under some non-trivial conditions. Results in Embrechts et al. (2019) show that if the "no-short-sell" assumption is removed, thus in a complete market, competitive equilibria generally fail to exist, unless all agents are ES agents. Moreover, in the case of ES agents, optimal allocations and equilibrium allocations are equivalent. Note that the negative dependence discussed in Section 4.5.1 already alerts the non-existence of competitive equilibria in a complete market, where an equilibrium allocation should be comonotonic under mild conditions (see e.g. Boonen et al. (2018) and Xia and Zhou (2016)).

A lot of questions remain open in the settings of an incomplete market. It would be of great interest to answer the following questions in an incomplete market setting for RVaR agents:

1. What trading constraints on the incomplete market allow for a competitive equilibrium to exist?

2. Given trading constraints such as the one in Embrechts et al. (2018), what is a necessary and sufficient condition on the parameters $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n$ and the total risk X for a competitive equilibrium to exist?

3. When a competitive equilibrium exists, is it possible to identify all possible equilibria, and what conditions give uniqueness of the equilibrium price?

As is already seen from the analysis in Chapter 4, fundamental questions in quantile-based risk sharing seem much more difficult than those in utility-based risk sharing problems, where convexity or concavity is typically assumed. We anticipate great technical challenges in the above questions, and at this moment we leave them for future work.

Bibliography

Acciaio, B. (2007). Optimal risk sharing with non-monotone monetary functionals. *Finance and Stochastics*, **11**, 267–289.

Acciaio, B. and Svindland, G. (2013). Are law-invariant risk functions concave on distributions? *Dependence Modeling*, **1**, 54–64.

Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance*, **26**(7), 1505–1518.

Acerbi, C. and Szekely, B. (2017). General properties of backtestable statistics. *Preprint*. Available at SSRN: https://ssrn.com/abstract=2905109

Anthropelos, M. and Kardaras, C. (2017). Equilibrium in risk-sharing games. *Finance and Stochastics*, **21**(3), 815–865.

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

Barrieu, P. and El Karoui, N. (2005). Inf-convolution of risk measures and optimal risk transfer. *Finance and stochastics*, **9**(2), 269–298.

BCBS (2016). *Standards. Minimum Capital Requirements Market Risk. January 2016.* Basel Committee on Banking Supervision. Basel: Bank for International Settlements.

Beissner, P. and Riedel, F. (2018). Non-implementability of Arrow-Debreu equilibria by continuous trading under volatility uncertainty. *Finance and Stochastics*, **22**(3), 603–620.

Bellini, F. and Bignozzi, V. (2015). On elicitable risk measures. *Quantitative Finance*, **15**(5), 725–733.

Bernard, C., Jiang, X. and Wang, R. (2014). Risk aggregation with dependence uncertainty. *Insurance: Mathematics and Economics*, **54**, 93–108.

Bernard, C., Rüschendorf, L. and Vanduffel, S. (2017a). VaR bounds with variance constraint. *Journal of Risk and Insurance*, **84**(3), 923–959.

Bernard, C., Rüschendorf, L., Vanduffel, S. and Wang, R. (2017b). Risk bounds for factor models. *Finance and Stochastics*, **21**(3), 631–659.

Borch, K. (1962). Equilibrium in a reinsurance market. *Econometrica*, **30**, 424–444.

Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *Annals of Statistics*, 1196–1217.

Bickel, P. J. and Lehmann, E. L. (1976). Descriptive statistics for nonparametric models. III. Dispersion. *The Annals of Statistics*. **4**(6), 1139–1158.

Boonen, T., Liu, F. and Wang, R. (2018). Competitive equilibria in a comonotone market. *Preprint*. Available at SSRN: https://ssrn.com/abstract=3091424

Brighi, B. and Chipot, M. (1994): Approximated convex envelope of a function. *SIAM Journal on Numerical Analysis*, **31**, 128–148.

Cai, J., Liu, H. and Wang, R. (2017). Asymptotic equivalence of risk measures under dependence uncertainty. *Mathematical Finance*. **28**(1), 29–49.

Cerreia-Vioglio, S. and Maccheroni, F. and Marinacci, M. and Montrucchio, L. (2012). Signed integral representations of comonotonic additive functionals. *Journal of Mathematical Analysis and Applications*, **385**(2), 895–912.

Cerreia-Vioglio, S. and Maccheroni, F. and Marinacci, M. and Montrucchio, L. (2015). Choquet integration on Riesz spaces and dual comonotonicity. *Transactions of the American Mathematical Society*, **367**(12), 8521–8542.

Chateauneuf, A. and Dana, R. and Tallon, J. (2000).Optimal risk-sharing rules and equilibria with Choquet-expected-utility. *Journal of Mathematical Economics*, **34**(2), 191–214.

Chew, S.H. (1989). Axiomatic utility theories with the betweenness property. *Annals of Operations Research*, **19**, 273C-298.

Choquet, G. (1954). Theory of capacities. *Annales de l'institut Fourier*, **5**, 131–295.

Cont, R., Deguest, R. and Scandolo, G. (2010). Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance*, **10**(6), 593–606.

Dana, R.A. (2005). A representation result for concave Schur concave functions. *Mathematical Finance*, **15**(4), 613–634.

Dekel, E. (1986). An axiomatic characterization of preferences under uncertainty: weakening the independence axiom. *Journal of Econonic Theory*, **40**, 304–318.

Delbaen, F. (2012): *Monetary Utility Functions.* Osaka University Press, Osaka.

Delbaen, F., Bellini, F., Bignozzi, V. and Ziegel, J. (2016). Risk measures with the CxLS property. *Finance and Stochastics*, **20**(2), 433–453.

Denneberg, D. (1990). Premium calculation: why standard deviation should be replaced by absolute deviation. *ASTIN Bulletin*, **20**, 181–190.

Denneberg, D. (1994). *Non-additive Measure and Integral.* Springer Science & Business Media.

Denuit, M., Dhaene, J., Goovaerts, M.J., Kaas, R. (2005). *Actuarial Theory for Dependent Risks.* Wiley: Chichester, UK.

De Waegenaere, A. and Wakker, P. P. (2001). Nonmonotonic Choquet integrals. *Journal of Mathematical Economics*, **36**, 45–60.

Dhaene, J. and Kukush, A., Linders, D. and Tang, Q. (2012). Remarks on quantiles and distortion risk measures. *European Actuarial Journal*, **2**(2), 319–328.

Donoghue, W. (1969). *Distributions and Fourier Transforms.* Academic press, New York.

Embrechts, P., Liu, H. and Wang, R. (2018). Quantile-based risk sharing. *Operations Research*, **66**(4), 936–949.

Embrechts, P., Liu, H., Mao, T. and Wang, R. (2019). Quantile-based risk sharing with heterogeneous beliefs. *Mathematical Programming Series B*, 1–29.

Embrechts, P., Puccetti, G. and Rüschendorf, L. (2013). Model uncertainty and VaR aggregation. *Journal of Banking and Finance*, **37**(8), 2750–2764.

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R. and Beleraj, A. (2014). An academic response to Basel 3.5. *Risks*, **2**(1), 25–48.

Embrechts, P., Wang, B. and Wang, R. (2015). Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance and Stochastics*, **19**(4), 763–790.

Emmer, S., Kratz, M. and Tasche, D. (2015). What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk*, **18**(2), 31–60.

Filipović, D. and Svindland, G. (2008). Optimal capital and risk allocations for law- and cash-invariant convex functions. *Finance and Stochastics*, **12**, 423–439.

Fissler T. and Ziegel J. (2016). Higher order elicitability and Osbands principle. *The Annals of Statistics*, **44**(4),1680–1707.

Föllmer, H. and Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, **6**(4), 429–447.

Föllmer, H. and Schied, A. (2016). *Stochastic Finance: An Introduction in Discrete Time.* Walter de Gruyter, Berlin, Fourth Edition.

Gao, N. and Leung, D. H. and Munari, C. and Xanthos, F. (2017). Fatou Property, representations, and extensions of law-invariant risk measures on general Orlicz spaces. *Finance and Stochastics*, **22**(2), 395–415.

Frongillo, R. and Kash, I. A. (2018). Elicitation Complexity of Statistical Properties. *Preprint.* Available at https://arxiv.org/abs/1506.07212

Furman, E., Wang, R. and Zitikis, R. (2017). Gini-type measures of risk and variability: Gini shortfall, capital allocation and heavy-tailed risks. *Journal of Banking and Finance*, **83**, 70–84.

Gao, N. and Leung, D. H. and Munari, C. and Xanthos, F. (2017). Fatou Property, representations, and extensions of law-invariant risk measures on general Orlicz spaces. *Preprint.* Available at https://arxiv.org/abs/1701.05967

Gao, N. and Xanthos, F. (2017). On the C-property and w^*-representations of risk measures. *Mathematical Finance*, **28**(2), 748–754.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, **18**, 141–153.

Gneiting T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association*, **106**(494),746–762.

Grechuk, B., Molyboha, A. and Zabarankin, M. (2009). Maximum entropy principle with general deviation measures. *Mathematics of Operations Research*, **34**(2), 445–467.

Hampel, F. (1971). A general qualitative definition of robustness. *Annals of Mathematical Statistics*, **42**(6), 1887–1896.

Huber, P. J. and Ronchetti E. M. (2009). *Robust Statistics.* Second Edition, Wiley Series in Probability and Statistics. Wiley, New Jersey.

Jouini, E., Schachermayer, W. and Touzi, N. (2006). Law invariant risk measures have the Fatou property. *Advances in Mathematical Economics*, **9**, 49–71.

Jouini, E., Schachermayer, W. and Touzi, N. (2008). Optimal risk sharing for law invariant monetary utility functions. *Mathematical Finance*, **18**(2), 269–292.

Kou, S., Peng, X. and Heyde, C. C. (2013). External risk measures and Basel accords. *Mathematics of Operations Research*, **38**(3), 393–417.

Kou, S. and Peng, X. (2016). On the measurement of economic tail risk. *Operations Research*, **64**(5), 1056–1072.

Krätschmer, V., Schied, A. and Zähle, H. (2014). Comparative and quantitative robustness for law-invariant risk measures. *Finance and Stochastics*, **18**(2), 271–295.

Krätschmer, V., Schied, A. and Zähle, H. (2017). Domains of weak continuity of statistical functionals with a view toward robust statistics. *Journal of Multivariate Analysis*, **158**, 1–19.

Kusuoka, S. (2001). On law invariant coherent risk measures. *Advances in Mathematical Economics*, **3**, 83–95.

Lambert, N., Pennock, D. M. and Shoham, Y. (2008). Eliciting properties of probability distributions. *Proceedings of the 9th ACM Conference on Electronic Commerce*, 129–138.

Lambert, N. (2012). Elicitation and evaluation of statistical functionals. *Preprint*. Available at: http://www.stanford.edu/~nlambert/papers/elicitation.pdf

Landsberger, M. and Meilijson, I. (1994). Co-monotone allocations, Bickel-Lehmann dispersion and the Arrow-Pratt measure of risk aversion. *Annals of Operations Research*, **52**, 97–106.

Liu, F. and Wang, R. (2016). A theory for measures of tail risk. *Preprint*, SSRN: abstract=2841909.

Li, L., Shao, H., Wang, R. and Yang, J. (2018). Worst-case Range Value-at-Risk with partial information. *SIAM Journal on Financial Mathematics*, **9**(1), 190–218.

Lorentz, G. G. (1951). On the theory of spaces Λ. *Pacific Journal of Mathematics*, **1**(3), 411–429.

Mao, T. and Wang, R. (2015). On aggregation sets and lower-convex sets. *Journal of Multivariate Analysis*, **136**, 12–25.

Marinacci, M. and Montrucchio L. (2004). Introduction to the mathematics of ambiguity. In *Uncertainty in Economic Theory*, I. Gilboa, Ed. Routledge, New York, NY, USA. 46–107.

Markowitz, H. (1951). Portfolio selection, *Journal of Finance*, **7**(1), 77–91.

McNeil, A. J., Frey, R. and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*. Revised Edition. Princeton, NJ: Princeton University Press.

Murofushi, T., Sugeno, M. and Machida, M. (1994). Non-monotonic fuzzy measures and the Choquet integral. *Fuzzy sets and Systems*, **64**(1), 73–86.

Newey, W. and Powell, J. (1987). Asymmetric least squares estimation and testing. *Econometrica*, **55**(4), 819-C847.

Nolde, N. and Ziegel, J. (2017). Elicitability and backtesting: Perspectives for banking regulation. *The Annals of Applied Statistics*, **11**(4), 1833–1874.

O'Cinneide, C. A. (1991). Phase-type distributions and majorization. *The Annals of Applied Probability*, **1**(2), 219–227.

Olkin, I. and Marshall, A. W. (2016). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.

Osband, K. (1985).*Providing incentives for better cost forecasting.* University of California, Berkeley.

Pesenti, S. M., Millossovich, P. and Tsanakas A. (2016). Robustness regions for measures of risk aggregation. *Dependence Modeling*, **4**, 348–367.

Puccetti, G. and Rüschendorf, L. (2014). Asymptotic equivalence of conservative value-at-risk- and expected shortfall-based capital charges. *Journal of Risk*, **16**(3), 1–19.

Puccetti, G. and Wang R. (2015). Extremal dependence concepts. *Statistical Science*, **30**(4), 485–517.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, **3**(4), 323–343.

Rao, M. M. and Ren, Z. D. (1991). *Theory of Orlicz Spaces.* New York: M. Dekker.

Rockafellar, R. T., Uryasev, S. and Zabarankin, M. (2006). Generalized deviation in risk analysis. *Finance and Stochastics*, **10**, 51–74.

Rockafellar, R. T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, **2**, 21–42.

Rockafellar, R. T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, **26**(7), 1443–1471.

Rudin, W. (1987). *Real and Complex Analysis.* Tata McGraw-Hill Education.

Rüschendorf, L. (2013). *Mathematical Risk Analysis. Dependence, Risk Bounds, Optimal Allocations and Portfolios.* Springer, Heidelberg.

Schmeidler, D. (1986). Integral representation without additivity. *Proceedings of the American Mathematical Society*, **97**(2), 255–261.

Schmeidler, D. (1989). Subject probability and expected utility without additivity. *Econometrica*, **57**(3), 571–587.

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders*. Springer Series in Statistics.

Schmeidler, D. (1989). Subject probability and expected utility without additivity. *Econometrica*, **57**(3), 571–587.

Steinwart, I., Pasin, C., Williamson, R. and Zhang, S. (2014). Elicitation and identification of properties. *Conference on Learning Theory*, 482–526.

Wang, B. and Wang, R. (2015). Extreme negative dependence and risk aggregation. *Journal of Multivariate Analysis*. **136**, 12–25.

Wang, R., Bignozzi, V. and Tsakanas, A. (2015). How superadditive can a risk measure be? *SIAM Jounral on Financial Mathematics*, **6**, 776–803.

Wang, R., Peng, L. and Yang, J. (2013). Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. *Finance and Stochastics*, **17**(2), 395–417.

Wang, R. and Ziegel, J. (2015). Elicitable distortion risk measures: A concise proof. *Statistics and Probability Letters*, **100**, 172–175.

Wang, S., Young, V. R. and Panjer, H. H. (1997). Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics*, **21**(2), 173–183.

Wang, R., Wei, Y. and Willmot, G. E. (2018). Characterization, robustness and aggregation of signed Choquet integrals. *Mathematics of Operations Research*, forthcoming.

Wang, R. and Wei, Y. (2018). Characterizing optimal allocations in quantile-based risk sharing. *Preprint*. Available at SSRN: https://ssrn.com/abstract=3173503

Wang, R. and Wei, Y. (2018). Risk functionals with convex level sets. *Preprint*. Available at SSRN: https://ssrn.com/abstract=3292661

Weber, S. (2006). Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance*, **16**, 419–441.

Weber, S. (2018). Solvency II, or how to sweep the downside risk under the carpet. *Insurance: Mathematics and Economics*, **82**, 191–200.

Xia, J. and Zhou, X. (2016). Arrow-Debreu equilibria for rank-dependent utilities. *Mathematical Finance*, **26**(3), 558–588.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, **55**(1), 95–115.

Ziegel, J. (2016). Coherence and elicitability. *Mathematical Finance*, **26**, 901–918.