Risk Aversion in Regulatory Capital Principles

Tiantian Mao* Ruodu Wang†

October 15, 2019

Abstract

We incorporate a notion of risk aversion favoring prudent decisions from financial institutions into regulatory capital calculation principles. In the context of Basel III, IV as well as Solvency II, regulatory capital calculation is carried out through the tools of monetary risk measures. The notion of risk aversion that we focus on has four equivalent formulations: through consistency with second-order stochastic dominance, or with conditional expectations, or with portfolio diversification, and finally through expected social impact. The class of monetary risk measures representing this notion of risk aversion is referred to as consistent risk measures. We characterize the class of consistent risk measures by establishing an Expected Shortfall-based representation, and as a by-product, we obtain new results on the representation of convex risk measures. We present several examples where consistent risk measures naturally appear. Using the obtained representation results, we study risk sharing and optimal investment problems and find several new analytical solutions.

Key-words: regulatory capital, risk measures, risk aversion, risk sharing, stochastic dominance.

1 Introduction

1.1 Risk aversion in regulatory capital calculation

In modern frameworks for financial regulation such as Basel III, IV as well as Solvency II, financial institutions are regulated to maintain a certain level of capital to prepare for potential future losses. In this paper, we take the perspective of a regulator who designs a *regulatory capital principle* to calculate the amount of capital required for financial institutions bearing risks. Such a principle is described by a *risk measure*: For a risk (random loss/profit) X borne by a financial institution over a fixed time period, a risk measure ρ assigns a number $\rho(X)$ quantifying today the future (end of time period) riskiness of X; a precise

*Department of Statistics and Finance, University of Science and Technology of China. Email: tmao@ustc.edu.cn.

†Department of Statistics and Actuarial Science, University of Waterloo. Email: wang@uwaterloo.ca.

definition is in Section 2.1. When $\rho(X)$ interprets into the amount of regulatory capital required for bearing a risk X, ρ is referred to as a *regulatory risk measure*.[1]

The two most popular regulatory risk measures in use throughout the financial industry are the Value-at-Risk (VaR) and the Expected Shortfall (ES). For a random loss/profit X, VaR at a confidence level $\alpha \in [0,1]$ is defined as

$$q_\alpha(X) = \inf\{x : \mathbb{P}(X \leqslant x) \geqslant \alpha\},\ \alpha \in (0,1]; \quad q_0(X) = \inf\{x : \mathbb{P}(X \leqslant x) > 0\}, \tag{1.1}$$

and ES at level $\alpha \in [0,1]$ is defined as

$$\psi_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 q_t(X)\mathrm{d}t,\ \alpha \in [0,1); \quad \psi_1(X) = q_1(X). \tag{1.2}$$

For a summary on the use of risk measures in risk management, see McNeil et al. (2015); Bénéplanc and Rochet (2011) contains a more economically oriented discussion.

In Bénéplanc and Rochet (2011, Section 1.1), the goals of risk management for a corporation are listed as four *decisions* varying from how much risk to take, or to retain/insure, or how big a capital buffer needs to be, to liquidity considerations. A financial regulator needs to balance these corporate goals and decisions by a prudent framework which allows for a healthy and transparent economic environment for business activities to thrive, and at the same time, protects the society at large from the danger of systemic risk (especially in the case of banking) or secures a high level of protection for policyholders in the case of the insurance industry. Especially concerning the former (banking regulation), in the wake of Basel II and as a consequence of the 2007 - 2009 crisis, global regulatory systems (as well as the regulatory risk measures in use) were found wanting; see for instance Acharya (2009). In our opinion, the purely analytical ansatz of regulatory risk measures needed a widening taking decision theory into account. This precisely is the focus of our paper.

The standpoint of a regulator and that of a financial institution are essentially different. The former, through its governmental mandate, is responsible to taxpayers and policyholders in the society whereas the latter is logically mainly responsible to its shareholders. Although a financial institution may not necessarily take decisions favoured by the regulator, the latter can influence decisions of the financial institution, through a regulatory risk measure which assigns higher regulatory capital to a more socially dangerous risk. More precisely, suppose that ρ is the regulatory risk measure chosen by the regulator, and X and Y are any two risks the financial institution has to choose between. If $\rho(X) < \rho(Y)$, that is, the regulatory capital of X is smaller than that of Y, then the financial institution has an incentive to choose X over Y, and by doing so it reduces its regulatory capital. As such, ρ serves as an objective disutility functional for financial

[1] Risk measures as a tool for capital calculation was the original motivation in Artzner et al. (1999). Although regulatory capital is the primary interpretation of risk measures in this paper, the mathematical results are not limited to such an interpretation.

institutions to make decisions, and the regulator should thereby design a risk measure which encourages well-understood prudent decisions over (less understood) risky ones. Therefore, a regulatory risk measure serves dual purposes. First, by imposing an appropriate risk measure, a financial institution is encouraged to take responsible financial decisions and it is penalized through regulatory capital charges for taking risky decisions; second, the risk measure determines an appropriate level of regulatory capital for financial institutions to protect the society from possible consequences of insolvency. The latter requires more capital for risks with higher potential damage to society.

To fulfill both the purposes above, a regulatory risk measure should reflect *desirable preferences over risks*, or, in other words, *notions of risk aversion*.[2] As such, the main question that we address is the following:

What are appropriate notions of risk aversion for regulatory risk measures, and how do we characterize a risk measure that is consistent with such notions of risk aversion?

Including the classic notion of *second-order stochastically dominance* (SSD), four closely related formulations of risk aversion with different interpretations are considered in this paper, some from the perspective of decision principles for financial institution and some from the perspective of expected impact on the society. For *monetary risk measures*, the four formulations of risk aversion are shown to be mathematically equivalent. A monetary risk measure which respects the above risk aversion (in particular, consistent with SSD) is referred to as a *consistent risk measure*. The paper is dedicated to an axiomatic study of consistent risk measures, its properties and relations to other risk measures, and its applications.

1.2 Contribution and structure of the paper

In Section 2, consistent risk measures are formally defined, and four formulations of risk aversion are shown to be equivalent in this setting. For risk measures used in regulatory practice, this consistency distinguishes VaR (non-consistent) from ES (consistent).[3]

In Section 3, we characterize the class of consistent risk measures by showing that a monetary risk measure ρ on the set of bounded random variables $\mathcal{X}$ is consistent if and only if it has a representation

$$\rho(X) = \min_{g\in\mathcal{G}} \sup_{\alpha\in[0,1]} \{\psi_\alpha(X) - g(\alpha)\}, \ \ X \in \mathcal{X},$$

[2]Here, we mainly speak of risk aversion from the regulator's side. More precisely, regulators would naturally like to see more risk aversion from the risk managers. It might be unrealistic to expect individual risk managers to be risk-averse without any regulatory enforcement.

[3]It is well known that VaR and ES are also distinguished by convexity.

where ψ_α is defined in (1.2) and the *adjustment set* $\mathcal{G}$ is a set of measurable functions mapping $[0,1]$ to $\mathbb{R}$. In a general sense, a risk-averse regulator or risk manager is essentially using a collection of Expected Shortfalls up to some adjustments. The large class of consistent risk measures contains, but is not limited to, law-invariant coherent as well as convex risk measures in the mathematical finance literature. With the results in this paper, we arrive at a grand summary of representations for various classes of law-invariant risk measures in Section 3.4.

Two applications highlighting consistent risk measures as a powerful tool are the risk sharing problem and the optimal investment problem. In Section 4, it is shown that for a total risk X shared by n agents using consistent risk measures $\rho_1, \dots, \rho_n$ with adjustment sets $\mathcal{G}_1, \dots, \mathcal{G}_n$, respectively, an allocation $(X_1, \dots, X_n)$ of X is Pareto-optimal if and only if

$$\sum_{i=1}^{n} \rho_i(X_i) = \min_{g \in \mathcal{G}_1 + \dots + \mathcal{G}_n} \sup_{\alpha \in [0,1]} \{\psi_\alpha(X) - g(\alpha)\}.$$

Throughout the paper, the sum of two sets $\mathcal{G}$ and $\mathcal{H}$ is defined as $\mathcal{G} + \mathcal{H} = \{g + h : g \in \mathcal{G},\ h \in \mathcal{H}\}$. We continue in Section 5 by solving optimal investment problems for consistent risk measures. It turns out that, due to its ES-based representation, many optimal investment problems can be solved analytically, even if the consistent risk measure itself is not convex, a surprisingly nice feature.

Some relevant discussions on consistent risk measures are provided in Section 6. Technical proofs are put in Appendix A. In the main text, we focus our study on bounded risks; the generalization of our results to unbounded risks is provided in Appendix B.

2 Putting risk aversion into risk measures

2.1 Monetary risk measures and acceptance sets

Throughout this paper, we work with an atomless probability space $(\Omega, \mathcal{A}, \mathbb{P})$. A *risk measure* ρ is a functional $\rho : \mathcal{X} \to \mathbb{R}$, where $\mathcal{X}$ is the space of all bounded random variables in $(\Omega, \mathcal{A}, \mathbb{P})$.[4] The extension of $\mathcal{X}$ to more general spaces (such as the space of integrable random variables) is discussed in Appendix B. Throughout the paper, positive values of risks in $\mathcal{X}$ represent (discounted) losses and negative values represent surpluses. We write $X \overset{\mathrm{d}}{=} Y$ if two random variables X and Y have the same law.

A *monetary risk measure* is a risk measure ρ on $\mathcal{X}$ satisfying the following two properties:

(M) Monotonicity: $\rho(X) \leqslant \rho(Y)$ for $X, Y \in \mathcal{X}$, $X \leqslant Y$ almost surely (a.s.);

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

[4] In this paper, almost surely equal random variables (under $\mathbb{P}$) are treated as identical. Technically speaking, a bounded random variable in this paper is an essentially bounded random variable under $\mathbb{P}$.

The property (M) is self-explanatory and is common in the literature on decision theory and risk measures. (TI) reflects the interpretation that $\rho(X)$ is the least amount of capital required to be injected to X so that $X - \rho(X)$ is an acceptable position, and hence injecting a cash flow of m units would reduce the capital requirement by an amount of m. VaR and ES, defined via (1.1) and (1.2), satisfy (M) and (TI).[5] In decision theory, (TI) appears as a special case of the *risk independence* condition in Jia and Dyer (1996) or the *constant absolute risk aversion* in Zank (2001). (M) and (TI) are widely assumed in the research of regulatory risk measures; the reader is referred to Föllmer and Schied (2011, Chapter 4) and Delbaen (2012) for a comprehensive treatment.

For any monetary risk measure ρ, its *acceptance set* is defined as $\mathcal{A}_\rho = \{X \in \mathcal{X} : \rho(X) \leqslant 0\}$, which represents all risks acceptable (without requiring further capital) to a regulator using ρ to calculate regulatory capital.

2.2 Risk aversion for decisions of financial institutions

We focus on three notions of risk aversion with different interpretations; they are later shown to be equivalent. The first notion of risk aversion is described via second-order stochastic dominance (SSD). We first recall the definition of SSD and the related concept of mean-preserving spread (MPS).[6]

Definition 2.1. For $X, Y \in \mathcal{X}$, we say that X is *second-order stochastically dominated by* Y, denoted as $X \prec_{\rm sd} Y$, if $\mathbb{E}[f(X)] \leqslant \mathbb{E}[f(Y)]$ for all increasing convex functions $f : \mathbb{R} \to \mathbb{R}$ provided that both expectations exist. If, in addition, $\mathbb{E}[X] = \mathbb{E}[Y]$, then we say that Y is a *mean-preserving spread* (MPS) of X, denoted as $X \prec_{\rm mps} Y$.

The consistency with respect to SSD for risk measures is defined below.

(SC) SSD-consistency: $\rho(X) \leqslant \rho(Y)$ if $X \prec_{\rm sd} Y$, $X, Y \in \mathcal{X}$.

(SC) is often termed *strong risk aversion* in economic decision theory. By using a risk measure satisfying (SC), a financial institution makes decisions consistent with the common notion of risk aversion, and in particular, favours a risk with small variability over one with a large variability.

A closely related property is the consistency with respect to mean-preserving spread. It is not difficult to show that for monetary risk measures, (SC) is equivalent to the following property (MC).

[5] (TI) is also called *cash-additivity*. The justification for (TI) may fail if there is randomness in the interest rate; see El Karoui and Ravanelli (2009) and Cerreia-Vioglio et al. (2011) for *cash-subadditive* risk measures which also have a decision theoretic foundation.

[6] Second-order stochastic dominance is also known as *increasing convex order* in probability theory and *stop-loss order* in actuarial science. Mean-preserving spread is also known as *convex order* in probability theory.

(MC) MPS-consistency: $\rho(X) \leqslant \rho(Y)$ if $X \prec_{\mathrm{mps}} Y$, $X, Y \in \mathcal{X}$.

The second notion of risk aversion that we consider is described via conditional expectations. In the context of coherent risk measures (see Section 3.2), it is under the name *dilatation monotonicity* in Leitner (2004) and Cherny and Grigoriev (2007).

(DM) Dilatation monotonicity: $\rho(X) \leqslant \rho(Y)$ if $(X, Y) \in \mathcal{X}^2$ such that $X = \mathbb{E}[Y|X]$.

The property (DM) also describes a preference on the variability of a risk. If (X, Y) is a martingale, then X is more "predictable" than Y, leading to less variability; as such, it is natural to require that $\rho(X) \leqslant \rho(Y)$ for any risk measure ρ that penalizes more risky behaviour. (DM) can also be interpreted as a preference about the maturity of financial investments with zero risk premium. For instance, let $\{X_t : t \in [0, T]\}$ be the discounted price process of a security in an arbitrage-free financial market where $\{X_t : t \in [0, T]\}$ is a martingale under $\mathbb{P}$. Then (DM) implies that $\rho(X_s) \leqslant \rho(X_t)$ for $0 \leqslant s \leqslant t \leqslant T$, i.e. the time-$t$ price (far future) is less preferable than the time-s price (near future) of the same security.

The third notion of risk aversion that we consider is related to portfolio diversification. We first recall the definition of comonotonicity.

Definition 2.2. A random vector $(X_1, \dots, X_n)$ is called *comonotonic* if there exists a random variable Z and non-decreasing functions $f_1, \dots, f_n$, such that $X_i = f_i(Z)$ almost surely for $i = 1, \dots, n$.

See Puccetti and Wang (2015) for more on comonotonicity and related properties. Comonotonic risks represent risks that are not diversified. For instance, a portfolio consisting of a call option and a long position of the underlying asset is not diversified at all (the call option price and the asset price are comonotonic in standard models), and hence is not favourable. The following property encourages diversification of risks in this sense.

(DC) Diversification consistency: $\rho(X + Y) \leqslant \rho(X^c + Y^c)$ if $X, Y, X^c, Y^c \in \mathcal{X}$, $X \overset{\mathrm{d}}{=} X^c$, $Y \overset{\mathrm{d}}{=} Y^c$, and (X^c, Y^c) is comonotonic.

(DC) is a natural requirement: putting two comonotonic (absolutely not diversified) risks X^c and Y^c in a portfolio results in a larger capital requirement compared to a portfolio consisting of risks X and Y, equal to X^c and Y^c in distribution, respectively, but not comonotonic. To the best of our knowledge, (DC) is not explicitly formulated in the previous literature as a property of risk measures.

2.3 Risk aversion for expected impact on the society

The main purpose of regulatory capital is to reduce the negative impact on the economy of the society when there is loss from financial institutions. Let X be the one-period risk taken by a financial institution and C is the amount of risk capital reserved for this risk. In a simplified way, loss to the society occurs when the financial institution is insolvent, and the amount of loss is $(X-C)_+ := \max\{X-C, 0\}$. Typically, there are multiple financial institutions in an economy with risks $X_1, \dots, X_n$ and respective capitals $C_1, \dots, C_n$. A regulator may then be interested in the total expected loss to the society,

$$\mathbb{E}\left[\sum_{i=1}^{n}(X_i - C_i)_+\right] = \sum_{i=1}^{n}\mathbb{E}\left[(X_i - C_i)_+\right]. \tag{2.1}$$

Thus, regulating towards a small total expected loss in (2.1) reduces to regulating towards small individual expected losses of each institution.

(i) *Loss from a single institution*: Suppose that X and Y are two risks that a financial institution has to decide between, and they satisfy $\mathbb{E}[(X-K)_+] \leqslant \mathbb{E}[(Y-K)_+]$ for all $K \in \mathbb{R}$. That is, no matter what level of capital K this company holds, taking the risk Y always results in a larger expected loss than taking the risk X. If C is an adequate amount of regulatory capital for a company taking risk X, then the regulator has to require equal or more capital for taking the risk Y, in order to maintain the same standard of solvency and to control impact on the society.

(ii) *Fairness among institutions*: Suppose that $\mathbb{E}[(X_1-K)_+] \geqslant \mathbb{E}[(X_2-K)_+]$ for all $K \in \mathbb{R}$, that is, the first institution always has a worse expected impact on the society compared to the second institution, if they were required to hold the same amount of capital. In view of fairness, it is then natural to require more regulatory capital for the first institution.[7]

From both considerations (i) and (ii) and their social implications, it is important for the regulator to employ

(EI) Consistency with expected impact on the society: for $X, Y \in \mathcal{X}$, $\rho(X) \leqslant \rho(Y)$ if $\mathbb{E}[(X-K)_+] \leqslant \mathbb{E}[(Y-K)_+]$ for all $K \in \mathbb{R}$.

Mathematically, (EI) is equivalent to (SC) (for instance, Shaked and Shanthikumar (2007, Theorem 4.A.3) and Müller and Stoyan (2002, Theorem 1.5.7)); yet it further motivates us to consider the desirability of (SC) for a regulator. Second order stochastic dominance can also be characterized in terms of ψ_α, $\alpha \in [0,1]$, that is, for $X, Y \in \mathcal{X}$,

$$X \prec_{\text{sd}} Y \;\Leftrightarrow\; \psi_\alpha(X) \leqslant \psi_\alpha(Y) \text{ for all } \alpha \in [0,1]. \tag{2.2}$$

See Theorem 4.A.3 of Shaked and Shanthikumar (2007).

[7]This consideration finds its similarity with the arguments for *systemic expected shortfall* in Acharya (2009), where individual contribution to the total loss of the society is measured as regulatory capital. In this paper, we focus on regulatory risk measures for individual companies at the micro level, albeit some arguments are consistent with the literature of systemic risk.

2.4 Consistent risk measures

One may argue that each of (SC), (MC), (DM), (DC) and (EI) results in a natural property of a risk measure that reflects a notion of risk aversion in financial regulation. Fortunately, for monetary risk measures on $\mathcal{X}$, the above properties are equivalent, and each of them implies the following property (LD).

(LD) Law-invariance: $\rho(X) = \rho(Y)$ if $X, Y \in \mathcal{X}$ and $X \stackrel{\mathrm{d}}{=} Y$.

Theorem 2.1. *For a monetary risk measure ρ on $\mathcal{X}$, the properties (SC), (MC), (DM), (DC) and (EI) are all equivalent. In turn, they imply that ρ satisfies (LD).*

Some of the above equivalences can be easily seen from classic results in the literature of stochastic orders, and others need a technical proof (see Appendix A.1). As the above properties are equivalent for monetary risk measures on $\mathcal{X}$, we concentrate on (SC) for this paper, and define the class of (SSD-)consistent risk measures.

Definition 2.3. A *consistent risk measure* is a risk measure that satisfies (TI) and (SC).

Note that (SC) implies (M) by definition and hence a consistent risk measure is monetary. The first example of consistent risk measures is the Expected Shortfall ψ_α, $\alpha \in [0,1]$ defined in (1.2). This is a classic result in the theory of stochastic orders; see for instance, Theorem 4.A.3 of Shaked and Shanthikumar (2007). On the other hand, it is well known that the Value-at-Risk q_α, $\alpha \in [0,1)$ defined in (1.1) is not consistent. The rest of this paper is dedicated to a comprehensive study of consistent risk measures.

3 Characterization of consistent risk measures

3.1 Representation of consistent risk measures via Expected Shortfalls

In this section we establish a representation of consistent risk measures based on Expected Shortfalls $\psi_\alpha,\ \alpha \in [0,1]$ defined in (1.2).

Theorem 3.1. *A risk measure ρ on $\mathcal{X}$ is a consistent risk measure if and only if it has the following representation*

$$\rho(X) = \inf_{g\in\mathcal{G}} \sup_{\alpha\in[0,1]} \left\{\psi_\alpha(X) - g(\alpha)\right\}, \quad X \in \mathcal{X}, \tag{3.1}$$

where $\mathcal{G}$ is a set of measurable functions mapping $[0,1]$ to $(-\infty,\infty]$. Moreover, $\mathcal{G}$ in (3.1) can be chosen as

$$\mathcal{G} = \left\{g_Y : [0,1] \to \mathbb{R},\ \alpha \mapsto \psi_\alpha(Y) \mid Y \in \mathcal{A}_\rho\right\}, \tag{3.2}$$

where $\mathcal{A}_\rho$ is the acceptance set of ρ, and with this choice the infimum in (3.1) is attained as a minimum.

Definition 3.1. The set $\mathcal{G}$ in (3.2) is called the *adjustment set* of a consistent risk measure ρ.

The main idea in the proof of Theorem 3.1 (in Appendix A.2) is to write the acceptance set of ρ as the union of lower sets generated by each of the acceptable positions.

A few interesting observations can be made from Theorem 3.1. First, the Expected Shortfalls ψ_α, $\alpha \in [0,1]$ are the basis of any consistent risk measure. Second, by using a consistent risk measure, a regulator or risk manager is essentially basing decisions on a collection of Expected Shortfalls ψ_α suitably adjusted by functions from a set $\mathcal{G}$. That is why we call $\mathcal{G}$ in (3.2) an adjustment set. Third, for a risk $X \in \mathcal{X}$, $\rho(X) \leqslant 0$ (that is, no further capital required) if and only if there exists $g \in \mathcal{G}$ such that $\psi_\alpha(X) \leqslant g(\alpha)$ for all $\alpha \in [0,1]$. Thus, a risk X is accepted by a regulator using ρ if and only if all of its Expected Shortfalls do not exceed one of the *pre-designed benchmarks* $g \in \mathcal{G}$. Note that $g(\alpha)$ may take the value ∞, making $\psi_\alpha(X) \leqslant g(\alpha)$ automatic. In this case, it is sufficient to check whether $\psi_\alpha(X) \leqslant g(\alpha)$ for $\alpha \in [0,1]$ such that $g(\alpha) < \infty$, for instance, $\alpha = p_1, \dots, p_n$ and $\alpha = p$ in Examples 3.1 and 3.2 below, respectively.

Example 3.1 (Discrete version of the representation)**.** For some numbers $g_{i,j} \in \mathbb{R}$ and distinct numbers $p_i \in [0,1]$, $i = 1, \dots, n$, $j = 1, \dots, m$, let $\mathcal{G} = \{g_1, \dots, g_m\}$ where for $j = 1, \dots, m$,

$$g_j(p_i) = g_{i,j},\ i = 1, \dots, n \ \text{ and } \ g_j(\alpha) = \infty,\ \alpha \in [0,1] \setminus \{p_1, \dots, p_n\}.$$

A discrete version of (3.1) is obtained as

$$\rho(X) = \min_{g \in \mathcal{G}} \sup_{\alpha \in [0,1]} \{\psi_\alpha(X) - g(\alpha)\} = \min_{j=1,\dots,m} \max_{i=1,\dots,n} \{\psi_{p_i}(X) - g_{i,j}\}, \quad X \in \mathcal{X}. \tag{3.3}$$

The formula (3.3) may be used to generate simple consistent risk measures by choosing the numbers p_i and $g_{i,j}$, $i = 1, \dots, n$, $j = 1, \dots, m$.

Example 3.2 (Expected Shortfall)**.** For $p \in [0,1]$, the Expected Shortfall ψ_p itself has a representation in Theorem 3.1. The natural choice of $\mathcal{G}$ is a singleton $\mathcal{G} = \{g_p\}$ where

$$g_p(p) = 0 \text{ and } g_p(\alpha) = \infty,\ \alpha \in [0,1] \setminus \{p\}.$$

To interpret $\mathcal{G}$ as the set of pre-designed benchmarks, a regulator decides whether a risk X is acceptable by checking whether $\psi_p(X) \leqslant 0$, the only criterion imposed by $\mathcal{G}$. Note that the set $\mathcal{G}$ in (3.1) is not unique in general. For the risk measure ψ_p, its adjustment set is given by (implied by Lemma 5.1 in Section 5),

$$\mathcal{G} = \{g : [0,1] \to \mathbb{R} \mid g(p) \leqslant 0,\ g \text{ is increasing, continuous, and } \alpha \mapsto (1-\alpha)g(\alpha) \text{ is concave}\},$$

which is different from $\{g_p\}$ above.

3.2 Relation to other risk measures

In this section, we connect the consistent risk measures to the classic notions of *coherent* and *convex risk measures* in Artzner et al. (1999), Föllmer and Schied (2002) and Frittelli and Rosazza Gianin (2002). We first review some classic properties of risk measures considered in the literature. For a risk measure ρ, the following properties are relevant:

(PH) Positive homogeneity: $\rho(\lambda X) = \lambda\rho(X)$ for all $\lambda \in (0, \infty)$ and $X \in \mathcal{X}$;

(SA) Subadditivity: $\rho(X + Y) \leqslant \rho(X) + \rho(Y)$ for all $X, Y \in \mathcal{X}$;

(CX) Convexity: $\rho(\lambda X + (1 - \lambda)Y) \leqslant \lambda\rho(X) + (1 - \lambda)\rho(Y)$ for all $\lambda \in [0, 1]$ and $X, Y \in \mathcal{X}$.

Definition 3.2. A risk measure is called a *convex risk measure* if it satisfies (M), (TI) and (CX). A risk measure is called a *coherent risk measure* if it satisfies (M), (TI), (PH) and (CX).

For $\alpha \in [0, 1]$, the Value-at-Risk q_α defined in (1.1) satisfies (M), (TI), (LD) and (PH) and the Expected Shortfall ψ_α defined in (1.2) further satisfies (CX) and (SA). (CX) and (SA), which favour diversification of risks, are sometimes argued as properties respecting risk aversion. Indeed, as implied by Proposition 3.2 below, any law-invariant convex risk measure on $\mathcal{X}$ is consistent.[8]

Proposition 3.2. *A law-invariant convex risk measure on $\mathcal{X}$ is consistent, and so is the maximum, the minimum or any convex combination of law-invariant convex risk measures.*

In view of Proposition 3.2, a consistent risk measure can be interpreted naturally from two different perspectives:

(i) In connection to law-invariant convex risk measures: the convexity (CX) is weakened to (SC); all other properties are preserved.

(ii) In connection to risk-averse expected utility functions: it adds the translation-invariance (TI) but it does not assume the independence axiom (along with the continuity axiom) as in the von Neumann-Morgenstern expected utility theory.

[8]The proof of this proposition is straightforward and essentially known: By Föllmer and Schied (2011, Corollary 4.65), a law-invariant convex risk measure with the Fatou property on the space of random variables with finite p-th moment, $p \in [1, \infty]$, is consistent; see also Bäuerle and Müller (2006, Theorem 4.2). From Jouini et al. (2006) and Delbaen (2009, 2012), the Fatou property can be dropped for law-invariant convex risk measures on $\mathcal{X}$. It is then clear from Section 3.1 that the maximum, the minimum or a convex combination of law-invariant convex risk measures is also consistent.

From Theorem 3.1 and Proposition 3.2, we obtain the following characterization of consistent risk measures via convex risk measures.

Theorem 3.3. *A risk measure ρ on $\mathcal{X}$ is a consistent (resp. positively homogeneous consistent) risk measure if and only if it has the following representation*

$$\rho(X) = \min_{\tau \in \mathcal{C}} \tau(X), \quad X \in \mathcal{X}, \tag{3.4}$$

where $\mathcal{C}$ is a set of law-invariant convex (resp. coherent) risk measures on $\mathcal{X}$.

Remark 3.1. It was mentioned in Section 7 of Bäuerle and Müller (2006) that the characterization of risk measures with convexity replaced by consistency is an open question, which is answered by our Theorems 3.1 and 3.3.

The acceptance set of ρ in (3.4) is the union of the acceptance sets of $\tau \in \mathcal{C}$. The reason why a consistent risk measure is not necessarily convex is also revealed by the acceptance sets: the union of convex sets need not be convex; see Example 3.3 below for an explicit construction.

Example 3.3 (A non-convex consistent risk measure)**.** For $X \in \mathcal{X}$, let

$$\rho_1(X) = \frac{1}{2} q_1(X) + \frac{1}{2}\mathbb{E}[X], \quad \rho_2(X) = \psi_{1/2}(X), \quad \text{and} \quad \rho(X) = \min\{\rho_1(X), \rho_2(X)\}.$$

From Proposition 3.2, ρ is a consistent risk measure. To see that ρ is not convex, take $X \sim \mathrm{Bernoulli}(1/2)$, $Y \sim \mathrm{Bernoulli}(1/6)$ such that (X, Y) is comonotonic. We can calculate

$$\rho_1(2X) = \frac{3}{2},\ \rho_1(2Y) = \frac{7}{6},\ \rho_2(2X) = 2,\ \rho_2(2Y) = \frac{2}{3} \text{ and } \rho_1(X+Y) = \rho_2(X+Y) = \frac{4}{3}.$$

Hence, $2\rho(X+Y) = 8/3 > 13/6 = \rho(2X) + \rho(2Y)$, implying that ρ is not convex.

Finally, as both law-invariant convex as well as coherent risk measures are consistent, they can be characterized via their adjustment sets. This yields a new representation for law-invariant convex and coherent risk measures which we could not trace in the literature.

Proposition 3.4. *A law-invariant risk measure ρ on $\mathcal{X}$ is a convex (resp. coherent) risk measure if and only if it has a representation*

$$\rho(X) = \min_{g \in \mathcal{G}} \sup_{\alpha \in [0,1]} \left\{ \psi_\alpha(X) - g(\alpha) \right\}, \quad X \in \mathcal{X}, \tag{3.5}$$

in which $\mathcal{G}$ is a convex set (resp. convex cone) of measurable functions mapping $[0,1]$ to $(-\infty, \infty]$.

Remark 3.2. It is straightforward to see that the statements in Theorems 3.1 and 3.3 and Proposition 3.4 hold true if the minimum in (3.4) or (3.5) is replaced by an infimum, like in (3.1). In that case, the sets $\mathcal{G}$ and $\mathcal{C}$ can always be conveniently chosen as countable sets. A proof of this statement is given in Appendix A.2.

In the next two examples, we illustrate situations where the representation (3.4) appears naturally.

Example 3.4 (Order statistics of risk assessments)**.** Since any order statistic of finitely many objects can be written as a minimax,[9] the infimum in (3.4) can be interpreted as any order statistic of convex risk measures, which does not need to be the smallest one. For a concrete example, suppose that there are 10 experts evaluating a risk, each using a convex risk measure ρ_i, $i \in I = \{1, \dots, 10\}$. Taking the largest of all evaluations, one arrives at a convex risk measure $\rho_{[1]} = \max_{i \in I} \rho_i$. Taking the second largest, one arrives in a consistent risk measure, which may not be convex. Indeed, by noting that the maximum of convex risk measures is still convex, the second-largest value of ρ_i, $i \in I$ can be written in the form of (3.4) as

$$\rho_{[2]}(X) = \min_{j \in I} \max_{i \in I \setminus \{j\}} \rho_i(X), \quad X \in \mathcal{X}.$$

Certainly, $\rho_{[2]}$ may not be as conservative as $\rho_{[1]}$; the disadvantage of $\rho_{[1]}$ is that it may be overly conservative or not robust. In many applications, not necessarily in finance, the largest value of several assessments is discarded (for a non-financial example, the Olympics judges for diving). This shows that the use of consistent risk measures offers a great flexibility from very conservative to moderately conservative risk assessment.

Example 3.5 (Competition)**.** This example shows that (3.4) appears naturally in the presence of competition for risk evaluations. Suppose that several insurance companies offer insurance coverage for a certain type of loss, and each of them uses a law-invariant convex risk measure to price the contract. An individual who seeks insurance coverage naturally compares the prices offered by different insurance companies, and choose the one that is the most favourable, given all other conditions fixed. Therefore, the effective pricing risk measure is the minimum of the risk measures of each company, which belongs to the form of (3.4). As a related observation, convexity of the effective risk measure is often lost in the presence of risk minimization; see Müller et al. (2017) for a similar observation in the context of utility maximization.

Remark 3.3. In the following remarks we discuss the closely related concepts of convexity and consistency.

1. In case convexity is not satisfied (e.g. Examples 3.4-3.5), consistency helps to understand alternative requirements for a reasonable risk measure and its properties. As illustrated by the main results in our paper, many nice properties of a convex risk measure remain if it is replaced by a consistent one. Moreover, we obtain new representation for convex risk measures, which are useful in some optimization problems (see the portfolio problem in Section 5).

[9] Generally, the j-th order statistic of n objects $a_1, \dots, a_n$ can be written as $a_{[j]} = \min_k \max_{i \in I_k} a_i$ where I_k, $k = 1, \dots, \binom{n}{n-j+1}$ are all sets of $n - j + 1$ indices. Also note that, by Proposition 3.2, any order statistic (smallest, largest, second-largest, median, etc.) of finitely many convex risk measures, or a convex combination of them, is a consistent risk measure.

2. Convexity (or quasi-convexity) means that pooling *any losses* together should not increase the total risk. We can compare this requirement with consistency, especially its equivalent formulation (DC) which addresses the issue of diversification. Recall that (DC) corresponds to pooling losses with *any dependence* together should not be worse than pooling comonotonic ones. Hence, consistent risk measures satisfy a weaker form of rewarding diversification, compared to convex risk measures. Note also that (DC) is always more flexible and easier to justify than (CX) because it is a weaker requirement.

3. From the perspective of decision making, convexity (CX) does not directly compare the riskiness of two random losses.[10] On the other hand, all the five equivalent properties described for consistent risk measures in Section 2 refer to direct comparisons of risks satisfying some conditions. Hence, consistent risk measures can be seen as a most natural form of risk measures based on direct risk comparison.

To summarize, if for some reason convexity is dropped, then the classic interpretation that putting any losses together reduces the total risk is lost. The advantages of studying consistent risk measures, on the other hand, are the flexibility to include many other examples, an axiom that is easier to justify, and new representation results of (possibly convex) risk measures (Theorem 3.1 and Proposition 3.4). It should be clear, however, that it is not our intention to suggest dropping convexity in regulation.

3.3 The Fatou property of consistent risk measures

In this section, we show that a consistent risk measure necessarily satisfies the following Fatou property.

(FP) Fatou property: $\liminf_{n\to\infty} \rho(X_n) \geqslant \rho(X)$ if $\{X_n\}_{n\in\mathbb{N}}$ is a bounded sequence in $\mathcal{X}$, and $X_n \to X \in \mathcal{X}$ a.s. as $n \to \infty$.

It is well known (e.g. p.40 of Delbaen (2012)) that, for monotone risk measures, the Fatou property is equivalent to continuity from below, that is,

(CB) Continuity from below: $\lim_{n\to\infty} \rho(X_n) = \rho(X)$ if $\{X_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$, and $X_n \nearrow X \in \mathcal{X}$ a.s. as $n \to \infty$.

The Fatou property is essential to the robust representation of convex risk measures; see e.g. Section 4.2 of Föllmer and Schied (2011). On the set $\mathcal{X}$ of bounded random variables, a law-invariant convex risk measure always satisfies the Fatou property; see Theorem 2.1 of Jouini et al. (2006) for a standard probability

[10]Note that convexity cannot be mathematically described via direct comparison between risks, since consistency with any partial order is preserved by taking a minimum, but convexity is not.

space and Theorem 30 of Delbaen (2012) for a general probability space. In the next result, we show that a consistent risk measure also enjoys this property.

Theorem 3.5. *A consistent risk measure on $\mathcal{X}$ always satisfies the Fatou property.*

Since law-invariant convex risk measures are consistent, Theorem 3.5 implies that law-invariant convex risk measures satisfies the Fatou property, as shown in the literature mentioned above. The proof of Theorem 3.5 relies on the ES-based representation in Theorem 3.1, and in particular, we use the convenient fact that the infimum in (3.1) can be attained. Another ingredient of the proof is the following simple lemma, which we present here as it may be of independent interest. It might be interesting to note that the proof of Theorem 3.5 does not need to use the aforementioned fact that law-invariant convex risk measures always satisfy (FP).

Lemma 3.6. *For $\{X_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ and $X \in \mathcal{X}$, if $X_n \nearrow X \in \mathcal{X}$ a.s. as $n \to \infty$, then $\psi_\alpha(X_n) \to \psi_\alpha(X)$ uniformly in $\alpha \in [0,1]$.*

3.4 Representation results for law-invariant risk measures

In the axiomatic theory of law-invariant risk measures, one of the most elegant results is the Kusuoka representation based on duality, established in Kusuoka (2001) for coherent risk measures and generalized in Frittelli and Rosazza Gianin (2005) for convex risk measures. Consistent risk measures have a similar representation, although not based on duality, and this is already reflected in Theorems 3.1 and 3.3. A key property for the Kusuoka representation is comonotonic additivity.

(CA) Comonotonic additivity: $\rho(X+Y) = \rho(X) + \rho(Y)$ if $(X,Y) \in \mathcal{X}^2$ is comonotonic.

In the following, ρ is a law-invariant risk measure on $\mathcal{X}$. Below, (i)-(iii) are existing representations of risk measures on $\mathcal{X}$, whereas (iv) is new and it is straightforward from Theorem 3.3.

(i) ρ is a comonotonic additive and coherent risk measure if and only if it has a representation (Theorem 7 of Kusuoka (2001)):
$$\rho = \int_0^1 \psi_\alpha \mathrm{d}\mu(\alpha), \tag{3.6}$$
where μ is a probability measure on $[0,1]$. ρ in (3.6) is called a *spectral risk measure*.

(ii) ρ is a coherent risk measure if and only if it has a representation (Theorem 4 of Kusuoka (2001)):
$$\rho = \sup_{\mu\in\mathcal{Q}} \left\{ \int_0^1 \psi_\alpha \mathrm{d}\mu(\alpha) \right\}, \tag{3.7}$$
where $\mathcal{Q}$ is a set of probability measures on $[0,1]$.

(iii) ρ is a convex risk measure if and only if it has a representation (Theorem 7 of Frittelli and Rosazza Gianin (2005)):

$$\rho = \sup_{\mu \in \mathcal{P}} \left\{ \int_0^1 \psi_\alpha \mathrm{d}\mu(\alpha) - v(\mu) \right\}, \tag{3.8}$$

where $\mathcal{P}$ is the set of all probability measures on $[0,1]$, and $v : \mathcal{P} \to \mathbb{R} \cup \{+\infty\}$ is a function.

(iv) ρ is a consistent risk measure if and only if it has a representation:

$$\rho = \inf_{v \in \mathcal{V}} \sup_{\mu \in \mathcal{P}} \left\{ \int_0^1 \psi_\alpha \mathrm{d}\mu(\alpha) - v(\mu) \right\}, \tag{3.9}$$

where $\mathcal{P}$ is the set of all probability measures on $[0,1]$, and $\mathcal{V}$ is a set of functions mapping $\mathcal{P}$ to $\mathbb{R} \cup \{+\infty\}$.

We make the following observations about the new representation (iv).

(a) The representation (3.9) is not based on duality and hence it does not belong to the class of robust representations in the sense of Föllmer and Schied (2011). The representation (3.1) in Theorem 3.1 is a special form of (3.9), implying that (3.9) can be simplified via replacing the set $\mathcal{P}$ of measures on $[0,1]$ by point-masses on $\alpha \in [0,1]$.

(b) $\mathcal{V}$ in (3.9) can be chosen as

$$\mathcal{V} = \left\{ v_X : \mathcal{P} \to \mathbb{R} \cup \{+\infty\},\ \mu \mapsto \int_0^1 \psi_\alpha(X) \mathrm{d}\mu(\alpha) - \rho(X),\ X \in \mathcal{X} \right\}. \tag{3.10}$$

(c) A consistent risk measure ρ is convex if and only if it has a representation (3.9) in which $\mathcal{V}$ is a convex set. The proof is similar to that of Proposition 3.4. Note that for a law-invariant convex risk measure ρ, v in (3.8) can be chosen as (Frittelli and Rosazza Gianin, 2005)

$$v(\mu) = \sup_{X \in \mathcal{X}} \left\{ \int_0^1 \psi_p(X) \mathrm{d}\mu(p) - \rho(X) \right\} = \sup_{\bar{v} \in \mathcal{V}} \bar{v}(\mu) \quad \text{with } \mathcal{V} \text{ as in (3.10).}$$

That is, the infimum and the supremum in (3.9) can be exchanged.

The relationships among these law-invariant risk measures (RM) on $\mathcal{X}$ are summarized below.

$$\text{(TI)+(SC)} = \text{consistent RM} \xrightarrow{\text{+(CX)}} \text{convex RM} \xrightarrow{\text{+(PH)}} \text{coherent RM} \xrightarrow{\text{+(CA)}} \text{spectral RM}.$$

Note that (CA) alone is sufficient for a consistent RM to be a spectral RM, that is,

$$\text{consistent RM} \xrightarrow{\text{+(CA)}} \text{spectral RM},$$

and this was implicitly established in Yaari (1987) in the context of choice under risk. The corresponding representations are summarized below.

$$\begin{aligned}\text{(TI)+(SC)} = \inf_{v\in\mathcal{V}}\sup_{\mu\in\mathcal{P}}\left\{\int_0^1 \psi_\alpha \mathrm{d}\mu(\alpha) - v(\mu)\right\} &\overset{\text{+(CX)}}{\longrightarrow} \sup_{\mu\in\mathcal{P}}\left\{\int_0^1 \psi_\alpha \mathrm{d}\mu(\alpha) - v(\mu)\right\} \\ &\overset{\text{+(PH)}}{\longrightarrow} \sup_{\mu\in\mathcal{Q}}\left\{\int_0^1 \psi_\alpha \mathrm{d}\mu(\alpha)\right\} \\ &\overset{\text{+(CA)}}{\longrightarrow} \int_0^1 \psi_\alpha \mathrm{d}\mu(\alpha),\end{aligned}$$

where μ, $\mathcal{V}$, $\mathcal{Q}$ and $\mathcal{P}$ are some functions and sets described through (3.6)-(3.9) in (i)-(iv).

4 Risk sharing via consistent risk measures

In this section, we investigate risk sharing problems for n agents whose preferences are described by minimizing n different risk measures. Most of the literature on risk sharing focuses on concave utilities or convex risk measures. We are interested in a setting under which risk measures are not necessarily convex, but are consistent.

For $X \in \mathcal{X}$, let $\mathbb{A}_n(X) = \{(X_1, \dots, X_n) \in \mathcal{X}^n : X_1 + \cdots + X_n = X\}$, that is, $\mathbb{A}_n(X)$ is the set of all allocations of a risk X to n agents. As in Barrieu and El Karoui (2005) and Delbaen (2012), the *inf-convolution* of risk measures $\rho_1, \dots, \rho_n$ is defined as

$$\square_{i=1}^n \rho_i(X) = \inf\left\{\sum_{i=1}^n \rho_i(Y_i) : (Y_1, \dots, Y_n) \in \mathbb{A}_n(X)\right\},\ X \in \mathcal{X}. \tag{4.1}$$

It is immediate to check that $\square_{i=1}^n \rho_i$ is a monetary risk measure given that $\rho_1, \dots, \rho_n$ are monetary risk measures and $\square_{i=1}^n \rho_i$ is finite-valued. Let $\rho^*(X) = \square_{i=1}^n \rho_i(X)$, $X \in \mathcal{X}$. For monetary risk measures $\rho_1, \dots, \rho_n$ and $X \in \mathcal{X}$, an allocation $(X_1, \dots, X_n) \in \mathbb{A}_n(X)$ is *optimal*, if $\sum_{i=1}^n \rho_i(X_i) = \rho^*(X)$. This notion is equivalent to *Pareto-optimal* for monetary risk measures; see Jouini et al. (2008).

We summarize the relevant results in Barrieu and El Karoui (2005), Jouini et al. (2008), Filipović and Svindland (2008) and Delbaen (2012) below. If $\rho_1, \dots, \rho_n$ are law-invariant convex risk measures on $\mathcal{X}$, then (a) there exists an optimal allocation $(X_1, \dots, X_n) \in \mathbb{A}_n(X)$ which is comonotonic; (b) ρ^* is a law-invariant convex risk measure, and (c) the *penalty function* (the Fenchel transformation, as in Föllmer and Schied (2002)) of ρ^* is the sum of the penalty functions of $\rho_1, \dots, \rho_n$. The above results (a)-(c) can be generalized to consistent risk measures. Most importantly, the adjustment sets satisfy property (c), namely, the adjustment set of ρ^* is the sum of those of $\rho_1, \dots, \rho_n$.

Theorem 4.1. *Suppose that $\rho_1, \dots, \rho_n$ are consistent risk measures on $\mathcal{X}$.*

(i) For $X \in \mathcal{X}$, let $\mathbb{A}_n^c(X) = \{(X_1, \dots, X_n) \in \mathbb{A}_n(X) : (X_1, \dots, X_n)$ is comonotonic$\}$. Then

$$\rho^*(X) = \inf\left\{\sum_{i=1}^n \rho_i(Y_i) : (Y_1, \dots, Y_n) \in \mathbb{A}_n^c(X)\right\},\ X \in \mathcal{X}.$$

(ii) ρ^* is finite-valued, and for $X \in \mathcal{X}$, there exists a comonotonic optimal allocation in $\mathbb{A}_n(X)$.

(iii) ρ^* is a consistent risk measure with acceptance set $\sum_{i=1}^n \mathcal{A}_i$, where $\mathcal{A}_i$ is the acceptance set of ρ_i, $i = 1, \dots, n$.

(iv) The adjustment set of ρ^* is $\sum_{i=1}^n \mathcal{G}_i$, where $\mathcal{G}_i$ is the adjustment set of ρ_i, $i = 1, \dots, n$. This implies

$$\rho^*(X) = \min_{g \in \mathcal{G}_1 + \dots + \mathcal{G}_n} \sup_{\alpha \in [0,1]} \{\psi_\alpha(X) - g(\alpha)\},\ \ X \in \mathcal{X}. \tag{4.2}$$

Theorem 4.1 shows that the representation in Theorem 3.1 is powerful in the context of risk sharing. Similarly to the case of convex risk measures (Barrieu and El Karoui, 2005 and Jouini et al., 2008), it is generally not easy to find explicitly the optimal allocation, although it is shown to exist in $\mathbb{A}_n^c(X)$. Nonetheless, for an optimal allocation $(X_1, \dots, X_n) \in \mathbb{A}_n(X)$, $\rho_1(X_1) + \dots + \rho_n(X_n)$ is equal to $\rho^*(X)$, which can be calculated using (4.2). This can be used to check the optimality of a specific allocation in $\mathbb{A}_n(X)$ without solving for an optimal allocation.

5 Optimal investment via consistent risk measures

In this section, we consider the problem of minimizing a consistent risk measure in the context of optimal investment. More precisely, we consider the problem

$$\text{to minimize: } \rho(X) \ \text{ over } X \in \mathcal{X} \ \text{ subject to } \ \mathbb{E}^{\mathbb{Q}}[X] \geqslant x_0, \tag{5.1}$$

where x_0 is a constant, $\mathbb{Q} \ll \mathbb{P}$, and ρ is a consistent risk measure.

Remark 5.1. It is well known that Problem (5.1) is equivalent to a problem of optimal portfolio investment in a complete financial market. We briefly explain this connection. Let $(\Omega, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space, $\mathcal{F}_0$ is $\mathbb{P}$-trivial (i.e. either $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$ for all $A \in \mathcal{F}_0$), $\mathbb{Q} \sim \mathbb{P}$ be the unique martingale measure in the financial market, and $\mathcal{X}_T(x)$ be the set of value processes of bounded self-financing portfolios on $[0, T]$ with initial value at most $x \in \mathbb{R}$. Without loss of generality we assume the risk-free interest rate is zero. A classic optimal investment problem is to minimize $\rho(-Y_T)$ over $Y \in \mathcal{X}_T(y_0)$ for some initial budget y_0 of the investor. Via the martingale approach (e.g. Föllmer et al. (2009)), this problem can be translated to (5.1) where $x_0 = -y_0$ and $\mathcal{X}$ is the set of all bounded $\mathcal{F}_T$-measurable random variables.

Problem (5.1) and its variants for convex, coherent, or distortion risk measures has been studied in the literature; see e.g. Schied (2004), He and Zhou (2011), Föllmer and Schied (2011) and Embrechts et al. (2018). Analytical solutions to Problem (5.1) for generic non-convex risk measures are rarely available as it involves non-convex optimization. With the help of Theorem 3.1, we are able to solve Problem (5.1) for consistent risk measures, which include many non-convex risk measures.

We first look at the basic ingredients in the representation of Theorem 3.1. Let $\mathcal{G}^*$ be the set of increasing continuous functions $g : [0,1] \to \mathbb{R}$ such that the function $\hat{g} : [0,1] \to \mathbb{R}$, $\alpha \mapsto (1-\alpha)g(\alpha)$ is concave. Throughout the section, we write L_g as the left-derivative of $\hat{g}$, implying $L_g(\alpha) = (1-\alpha)g'(\alpha) - g(\alpha)$ for $\alpha \in [0,1]$ a.e., noting that g is a.e. differentiable. For $g \in \mathcal{G}^*$, we define a risk measure ρ_g by

$$\rho_g(X) = \sup_{\alpha \in [0,1]} \{\psi_\alpha(X) - g(\alpha)\}, \quad X \in \mathcal{X}.$$

Note that ρ_g is convex, and it is finite-valued on $\mathcal{X}$ because $\mathbb{E}[X] - g(0) \leqslant \rho_g(X) \leqslant \text{ess-sup}(X) - g(0)$.

Lemma 5.1. *For any $g \in \mathcal{G}^*$ and any uniform random variable U on $[0,1]$, the random variable $Y = -L_g(U)$ satisfies $\psi_\alpha(Y) = g(\alpha)$ for $\alpha \in [0,1]$.*

As a consequence of Lemma 5.1, for a function g on $[0,1]$, $g \in \mathcal{G}^*$ if and only if there exists a bounded random variable Y such that $\psi_\alpha(Y) = g(\alpha)$ for $\alpha \in [0,1]$. Hence, the adjustment set in (3.2) of any consistent risk measure is a subset of $\mathcal{G}^*$.

Below we show that Problem (5.1) admits explicit solutions for the risk measure ρ_g. For this, we first fix some notation. Let U be a uniform random variable on $[0,1]$ such that $(U, \mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P})$ is comonotonic. Such a random variable U always exists in an atomless probability space (see Lemma A.28 of Föllmer and Schied (2011)) and U is a.s. unique if $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}$ is continuously distributed. Define two measures ν and μ on $\mathcal{B}([0,1])$ by, for $t \in [0,1]$, $\nu([0,t]) = q_t(\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P})$ and $\frac{\mathrm{d}\mu}{\mathrm{d}\nu}(t) = 1-t$. Via integration by parts, one can check

$$\mu([0,1]) = \int_0^1 (1-\alpha)\mathrm{d}q_\alpha(\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}) + q_0(\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}) = \int_0^1 q_\alpha(\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P})\mathrm{d}\alpha = \mathbb{E}^{\mathbb{Q}}[1] = 1, \tag{5.2}$$

and hence μ is a probability measure on $\mathcal{B}([0,1])$. Note that for any bounded measurable function g on $[0,1]$,

$$\int_0^1 g(\alpha)\mathrm{d}\mu(\alpha) = \int_0^1 g(\alpha)(1-\alpha)\mathrm{d}\nu(\alpha) = g(0)q_0(\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}) + \int_0^1 g(\alpha)(1-\alpha)\mathrm{d}q_\alpha(\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}). \tag{5.3}$$

Proposition 5.2. *For $g \in \mathcal{G}^*$, the problem*

$$\text{to minimize: } \rho_g(X) \ \text{ over } X \in \mathcal{X} \ \text{ subject to } \ \mathbb{E}^{\mathbb{Q}}[X] \geqslant x_0, \tag{5.4}$$

admits a solution

$$X^* = -L_g(U) + x_0 - \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha),$$

and Problem (5.4) has a minimal value $\rho_g(X^*) = x_0 - \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha)$*. Moreover, if* $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}$ *is continuously distributed, then* X^* *is the a.s. unique solution to Problem (5.4).*

With the help of Proposition 5.2, we can establish the solution to Problem (5.1) for consistent risk measures. Recall that, by Theorem 3.1 and Lemma 5.1, a consistent risk measure ρ always has a representation $\rho = \min_{g\in\mathcal{G}} \rho_g$ for some $\mathcal{G} \subset \mathcal{G}^*$.

Theorem 5.3. *Suppose that* ρ *is a consistent risk measure with representation* $\rho = \min_{g\in\mathcal{G}} \rho_g$ *where* $\mathcal{G} \subset \mathcal{G}^*$*.*

(i) Problem (5.1) has an infimum value given by

$$\inf\{\rho(X) : X \in \mathcal{X},\ \mathbb{E}^{\mathbb{Q}}[X] \geqslant x_0\} = x_0 - \sup_{g\in\mathcal{G}} \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha).$$

(ii) Problem (5.1) admits a solution if and only if $\arg\max_{g\in\mathcal{G}} \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha) \neq \varnothing$*.*

(iii) If $g^* \in \arg\max_{g\in\mathcal{G}} \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha)$*, then a solution to Problem (5.1) is given by*

$$X^* = -L_{g^*}(U) + x_0 - \int_0^1 g^*(\alpha)\mathrm{d}\mu(\alpha). \tag{5.5}$$

(iv) If $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}$ *is continuously distributed, then any solution to Problem (5.1) has the form (5.5) a.s. for some* $g^* \in \arg\max_{g\in\mathcal{G}} \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha)$*.*

Below present two examples of Theorem 5.3, one addressing the case of an Expected Shortfall, and the other regarding non-existence of solutions to Problem 5.1.

Example 5.1 (Optimal investment for Expected Shortfall)**.** Suppose that $\rho = \psi_p$ for some $p \in (0,1)$ and for simplicity, we write $q_\alpha = q_\alpha(\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P})$ for $\alpha \in [0,1]$. Problem (5.1) has a constant solution $X^* = x_0$ if $q_1 \leqslant \frac{1}{1-p}$ and it has no solution if $q_1 > \frac{1}{1-p}$. This result is known in Proposition 6 of Embrechts et al. (2018) based on a probabilistic approach. Below we show this statement using Theorem 5.3. Recall that $\psi_p = \min_{g\in\mathcal{G}} \rho_g$ where $\mathcal{G} = \{g \in \mathcal{G}^* : g(p) \leqslant 0\}$ from Example 3.2.

(i) $q_1 \leqslant \frac{1}{1-p}$. For $g \in \mathcal{G}$, $g(p) \leqslant 0$ implies $\hat{g}(p) \leqslant 0$. Together with the fact that $\hat{g}$ is concave and $\hat{g}(0) = 0$, we conclude that $\hat{g}$ is increasing on $[0,p]$. Note that

$$\int_0^1 g(\alpha)\mathrm{d}\mu(\alpha) = \int_0^1 g(\alpha)(1-\alpha)\mathrm{d}\nu(\alpha) = \int_0^1 \hat{g}(\alpha)\mathrm{d}\nu(\alpha). \tag{5.6}$$

Recall that $\nu([0,1]) = q_1$, and (5.2) implies $\int_0^1 \alpha\mathrm{d}\nu(\alpha) = \int_0^1 \mathrm{d}\nu(\alpha) - \int_0^1 (1-\alpha)\mathrm{d}\nu(\alpha) = q_1 - 1$. Since $\hat{g}$ is concave, by Jensen's inequality and, we have

$$\int_0^1 \frac{1}{q_1}\hat{g}(\alpha)\mathrm{d}\nu(\alpha) \leqslant \hat{g}\left(\int_0^1 \frac{1}{q_1}\alpha\mathrm{d}\nu(\alpha)\right) = \hat{g}\left(\frac{1}{q_1}(q_1-1)\right) = \hat{g}\left(1 - \frac{1}{q_1}\right). \tag{5.7}$$

Noting that $1 - \frac{1}{q_1} \leqslant 1 - (1-p) = p$ and $\hat{g}$ is increasing on $[0, p]$, (5.6) and (5.7) give

$$\int_0^1 g(\alpha)\mathrm{d}\mu(\alpha) = \int_0^1 \hat{g}(\alpha)\mathrm{d}\nu(\alpha) \leqslant q_1 \hat{g}\left(1 - \frac{1}{q_1}\right) \leqslant q_1 \hat{g}(p) \leqslant 0.$$

Therefore, the constant function $g^* = 0$ maximizes $\int_0^1 g(\alpha)\mathrm{d}\mu(\alpha)$ over $g \in \mathcal{G}$. By Theorem 5.3 (iii), $X^* = x_0$ is a solution to Problem (5.1).

(ii) $q_1 > \frac{1}{1-p}$. For $n \in \mathbb{N}$, take $g_n(\alpha) = n \min\left\{\frac{1}{1-\alpha}, n\right\} - \frac{n}{1-p}$, $\alpha \in [0,1]$. Clearly, g_n is an increasing function and $\hat{g}_n$ is concave. Hence, $g_n \in \mathcal{G}$. We can calculate

$$\frac{1}{n}\int_0^1 g_n(\alpha)\mathrm{d}\mu(\alpha) = \int_0^1 \min\{1, n(1-\alpha)\}\mathrm{d}\nu(\alpha) - \frac{1}{1-p}.$$

Therefore, by the Bounded Convergence Theorem,

$$\lim_{n\to\infty} \frac{1}{n}\int_0^1 g_n(\alpha)\mathrm{d}\mu(\alpha) = q_1 - \frac{1}{1-p} > 0.$$

As a consequence,

$$\sup_{g\in\mathcal{G}} \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha) \geqslant \lim_{n\to\infty} \int_0^1 g_n(\alpha)\mathrm{d}\mu(\alpha) = \infty.$$

By Theorem 5.3 (ii), Problem (5.1) has no solution.

Example 5.2 (Non-existence of solutions)**.** Problem (5.1) obviously does not have a solution if $\inf\{\rho_g(X) : X \in \mathcal{X},\ \mathbb{E}^{\mathbb{Q}}[X] \geqslant 0\} = -\infty$. In addition, it may not have a solution even if $\inf\{\rho_g(X) : X \in \mathcal{X},\ \mathbb{E}^{\mathbb{Q}}[X] \geqslant 0\} > -\infty$, as we shall see from this example. Assume that $x_0 = 0$, $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P} = 2U$ where U is a uniform random variable on $[0,1]$, and for $n \in \mathbb{N}$, let $g_n(\alpha) = \min\{\frac{1}{1-\alpha}, n\}$, $\alpha \in [0,1]$. Define $\rho = \inf_{n\in\mathbb{N}} \rho_{g_n}$. Using Theorem 5.3, we can calculate $\inf\{\rho_g(X) : X \in \mathcal{X},\ \mathbb{E}^{\mathbb{Q}}[X] \geqslant 0\} > -2$. On the other hand, the above infimum turns out to be not attainable, and hence Problem (5.1) does not have a solution. Details of this example are provided in Appendix A.4.

Remark 5.2. Generally, for a set $\mathcal{G}$ of measurable functions on $[0,1]$, the formula $\rho = \min_{g\in\mathcal{G}} \rho_g$ always defines a consistent risk measure given that ρ is finite-valued (see Theorem 3.1), even if $\mathcal{G}$ is not a subset of $\mathcal{G}^*$. In this case, Proposition 5.2 and Theorem 5.3 may not be directly applied, since the construction of X^* in Proposition 5.2 relies on the fact that $\mathcal{G} \subset \mathcal{G}^*$ so that $\psi_\alpha(-L_g(U)) = g(\alpha)$ for all $\alpha \in [0,1]$. Hence, in order to apply Theorem 5.3, one needs to first write ρ as the minimum of ρ_g for g in a subset of $\mathcal{G}^*$, which is always possible as a consequence of Theorem 3.1 and Lemma 5.1.

6 Discussion

There are extensive recent debates on the desirability of VaR and ES as the standard regulatory risk measure in both banking and insurance; see Embrechts et al. (2014), BCBS (2016) and IAIS (2014). The criterion for a desirable risk measure used in banking and insurance regulation may vary, depending on particular considerations and circumstances. One may need to consider mathematical simplicity, statistical and computational tractability, aggregation and systemic effects, model uncertainty, optimization properties, capital allocation and consistency of risk ranking. To this discussion we add that a suitable risk measure applied in regulatory practice should encourage socially responsible financial decisions, and this can be reflected in consistency with respect to certain risk ordering principles, among which the equivalent notions of risk aversion in this paper serve as a natural candidate. We illustrate with our main results that this requirement is weaker and more flexible than convexity, typically assumed in the literature. For some recent academic discussions on the desirability of VaR, ES and other risk measures in regulation, we refer to Cont et al. (2010), Kou and Peng (2016), Fissler and Ziegel (2016), Embrechts et al. (2018) and the references therein.

Acknowledgements. The authors thank the Editor, an Associate Editor, two referees for helpful comments on earlier versions of the paper. The authors also thank Paul Embrechts, Taizhong Hu, Fabio Maccheroni, Alfred Müller, Cosimo-Andrea Munari, Alexander Schied, and Andreas Tsanakas for various insightful discussions and excellent suggestions that improved the paper. T. Mao was supported by the NNSF of China (grant numbers: 71671176, 71871208, 71921001), the fundamental Research Funds for the Central Universities (WK204160028) and partially by the Department of Statistics and Actuarial Science at the University of Waterloo. R. Wang acknowledges support from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-03823, RGPAS-2018-522590), and the CAE Research Grant from the Society of Actuaries.

A Proofs of theorems, lemmas and propositions

A.1 Proofs in Section 2

Proof of Theorem 2.1. (i) We first show that each of (SC), (DM) and (DC) implies (LD).

1. (SC)⇒(LD): if $X, Y \in \mathcal{X}$ satisfy $X \overset{\mathrm{d}}{=} Y$, then $X \prec_{\mathrm{sd}} Y$ and $Y \prec_{\mathrm{sd}} X$. Hence $\rho(X) = \rho(Y)$.
2. (DM)+(M)+(TI)⇒(LD): by Cherny and Grigoriev (2007, Theorem 1.1), a dilatation monotone and $\|\cdot\|_\infty$-continuous (that is, $\lim_{n\to\infty} \rho(X_n) = \rho(X)$ for any sequence $X_n \in \mathcal{X}$ satisfying $\|X_n - X\|_\infty := \text{ess-sup}|X_n - X| \to 0$ as $n \to \infty$) function on $\mathcal{X}$ is law-invariant, and we note that a monetary risk measure is always $\|\cdot\|_\infty$-continuous.
3. (DC)⇒(LD): take $X, Y \in \mathcal{X}$ and $X \overset{\mathrm{d}}{=} Y$. Since $(Y, 0)$ is comonotonic, (DC) implies that $\rho(X + 0) \leqslant \rho(Y + 0)$, and thus $\rho(X) \leqslant \rho(Y)$. By symmetry, $\rho(Y) \leqslant \rho(X)$. This shows that ρ satisfies

(LD).

(ii) (MC)⇔(SC): (M) and (MC) are equivalent to (SC) by the Separation Theorem in Shaked and Shanthikumar (2007, Theorem 4.A.6).

(iii) (SC)⇔(EI): (SC) and (EI) are equivalent by Shaked and Shanthikumar (2007, Theorem 4.A.3).

(iv) (MC)⇔(DM): It suffices to show that (MC) is equivalent to (DM)+(LD), which follows from the martingale representation of convex order in Shaked and Shanthikumar (2007, Theorem 3.A.4).

(v) (SC)⇔(DC): It is easy to verify (SC)⇒(DC) since $X+Y \prec_{\text{mps}} X^c+Y^c$ for all $(X,Y,X^c,Y^c)\in\mathcal{X}^4$ such that $X \overset{\text{d}}{=} X^c$, $Y \overset{\text{d}}{=} Y^c$, and (X^c,Y^c) is comonotonic; see Dhaene et al. (2002). To show (DC)⇒(SC), let X, $Y \in \mathcal{X}$ be two random variables such that $X \prec_{\text{mps}} Y$. We shall show $\rho(X) \leqslant \rho(Y)$ for the following three cases.

Case 1. Assume that X and Y take values in finite sets and satisfy

$$Y = X - \delta_1 1_{A_1} + \delta_2 1_{A_2}, \tag{A.1}$$

where $\delta_1, \delta_2 > 0$ are constants, X takes respective constant values x_1 and x_2 in A_1 and A_2 such that $x_1 \leqslant x_2$, and $A_1,\ A_2 \in \mathcal{F}$ are two disjoint sets such that $X \prec_{\text{mps}} Y$, i.e., $\delta_1\mathbb{P}(A_1) = \delta_2\mathbb{P}(A_2)$. In this case, Y is called a simple mean-preserving spread of X (see Müller and Scarsini, 2001).

(1a) If $\delta_1 = \delta_2 =: \delta$, i.e., $\mathbb{P}(A_1) = \mathbb{P}(A_2)$, denote $A_0 = \{X \leqslant x_2\} \cap A_1^c \cap A_2^c$, $A_3 = \{x_2 < X \leqslant x_2+\delta\}$, $A_4 = \{X > x_2+\delta\}$, and define random variables W_1, W_2, W_1^c and W_2^c as

$$W_1 = W_1^c = (X+\delta)1_{A_0} + x_1 1_{A_1} + (x_2+\delta)1_{A_2\cup A_3} + X1_{A_4},$$

$$W_2 = -\delta 1_{A_0\cup A_2} + (X-x_2-\delta)1_{A_3},\ \ W_2^c = -\delta 1_{A_0\cup A_1} + (X-x_2-\delta)1_{A_3}.$$

It is easy to see $X = W_1+W_2$, $Y = W_1^c+W_2^c$ a.s., $W_i \overset{d}{=} W_i^c$, $i=1,2$ from $\mathbb{P}(A_1)=\mathbb{P}(A_2)$. Also, notice that $\cup_{i=0}^4 A_i = \Omega$, and

$$[W_1^c|A_0\cup A_1] \leqslant x_2+\delta = [W_1^c|A_2\cup A_3] \leqslant [W_1^c|A_4],\ \text{a.s.},$$

$$[W_2^c|A_0\cup A_1] = -\delta \leqslant [W_2^c|A_3] \leqslant 0 = [W_2^c|A_2\cup A_4],\ \text{a.s.}$$

It follows that W_1^c and W_2^c are comonotonic. Hence, we have

$$\rho(X) = \rho(W_1+W_2) \leqslant \rho(W_1^c+W_2^c) = \rho(Y).$$

(1b) If $\delta_1 \neq \delta_2$, let $A_{11} \subset A_1$ and $A_{21} \subset A_2$ such that $\mathbb{P}(A_{11}) = \mathbb{P}(A_{21}) = \min\{\mathbb{P}(A_1), \mathbb{P}(A_2)\}$, and define X_1 as

$$X_1 = X - \min\{\delta_1, \delta_2\} 1_{A_{11}} + \min\{\delta_1, \delta_2\} 1_{A_{21}}.$$

That is, if $\delta_1 < \delta_2$, then $A_{21} = A_2$, $X_1(\omega) = x_1 - \delta_1 = y_1 = Y(\omega),\ \omega \in A_{11}$, and

$$Y = X_1 - \delta_1 1_{A_1 \setminus A_{11}} + (\delta_2 - \delta_1) 1_{A_{21}}, \tag{A.2}$$

while if $\delta_1 > \delta_2$, then $A_{11} = A_1$ and $X_1(\omega) = x_2 + \delta_2 = y_2 = Y(\omega),\ \omega \in A_{21}$, and

$$Y = X_1 - (\delta_1 - \delta_2) 1_{A_{11}} + \delta_2 1_{A_2 \setminus A_{21}}. \tag{A.3}$$

It is obvious that in both cases Y is still a mean-preserving spread of X_1 (either (A.2) or (A.3) holds, both having the same form as (A.1)) such that the difference between X_1 and Y is strictly smaller than that between X and Y. Specifically, we have that one value difference ($|\delta_1 - \delta_2|$) becomes smaller while the other one set difference ($A_2 \setminus A_{21}$ or $A_1 \setminus A_{11}$) becomes smaller. Repeating the above procedure, we recursively define the random variable X_n, $n \in \mathbb{N}$ such that the sequence $\{X_n,\ n \in \mathbb{N}\}$ satisfies

$$X_{n+1} = X_n - \beta_n 1_{A_{1n}} + \beta_n 1_{A_{2n}}, \quad n \in \mathbb{N},$$

where $A_{1n} \subset A_1 \setminus (\cup_{i=1}^{n-1} A_{1i})$ and $A_{2n} \subset A_2 \setminus (\cup_{i=1}^{n-1} A_{2i})$ are two measurable sets such that $\mathbb{P}(A_{1n}) = \mathbb{P}(A_{2n}) = \min\{\mathbb{P}(A_1 \setminus (\cup_{i=1}^{n-1} A_{1i})), \mathbb{P}(A_2 \setminus (\cup_{i=1}^{n-1} A_{2i}))\}$, and $\beta_n \geqslant 0$ is the biggest constant such that Y is still a mean-preserving spread of X_{n+1}. From the construction of X_n, $n \in \mathbb{N}$, we know

$$\lim_{n \to \infty} ||X_n - Y||_\infty \leqslant \lim_{m_n \to \infty} \frac{1}{m_n} (\delta_1 \vee \delta_2) = 0,$$

where $\{m_n,\ n \in \mathbb{N}\}$ is some sequence which converges to infinity as $n \to \infty$. Hence, from (1a), we have

$$\rho(X) \leqslant \rho(X_1) \leqslant \ldots \leqslant \rho(X_n) \leqslant \rho(Y) + \rho(||X_n - Y||_\infty).$$

Since ρ is a monetary risk measure, which is continuous with respect to $\|\cdot\|_\infty$, this implies $\rho(X) \leqslant \rho(Y)$ by letting $n \to \infty$ in the above inequality.

Case 2. Assume that X and Y take values in finite sets. By Theorem 6.2 of Müller and Scarsini (2001) (see also Rothschild and Stiglitz (1970)), there exist Z_k, $k = 0, \ldots, m$, such that $X = Z_0$, $Y = Z_m$, and Z_{k+1} is a simple mean-preserving spread of Z_k, $k = 0, \ldots, m-1$. From Case 1, we have

$$\rho(X) \leqslant \rho(Z_1) \leqslant \ldots \leqslant \rho(Z_m) = \rho(Y).$$

Case 3. For general random variables $X, Y \in \mathcal{X}$, there exist two sequences of random variables X_n^*, $n \in \mathbb{N}$ and Y_n^*, $n \in \mathbb{N}$ with finite outcomes such that $X_n^* \prec_{\text{mps}} Y_n^*$ for $n \in \mathbb{N}$ and $\|X_n^* - X\|_\infty \to 0$ and $\|Y_n^* - Y\|_\infty \to 0$ as $n \to \infty$ (for this assertion, see the proof of Theorem 3.1 of Mao and Hu (2015)). Since ρ is $\|\cdot\|_\infty$-continuous, and by the result in Case 2, we have $\rho(X) \leqslant \rho(Y)$.

Combining the above three cases, we conclude that ρ preserves MPS. □

A.2 Proofs in Section 3

Proof of Theorem 3.1. First it is easy to verify that a risk measure ρ defined by (3.1) satisfies (TI) and (SC) by noting that the risk measures $\psi_\alpha - g(\alpha)$, $\alpha \in [0,1]$, $g \in \mathcal{G}$ satisfies (TI) and (SC). For the other direction, suppose that ρ is a consistent risk measure and denote its acceptance set by $\mathcal{A}_\rho$, that is, $\mathcal{A}_\rho = \{X \in \mathcal{X} : \rho(X) \leqslant 0\}$. Since ρ satisfies (SC), we have that if $X \in \mathcal{A}_\rho$, $Y \in \mathcal{X}$ and $Y \prec_{\text{sd}} X$, then $Y \in \mathcal{A}_\rho$. As a consequence,

$$\mathcal{A}_\rho = \bigcup_{X \in \mathcal{A}_\rho} \{X\} = \bigcup_{X \in \mathcal{A}_\rho} \{Y \in \mathcal{X} : Y \prec_{\text{sd}} X\}. \tag{A.4}$$

By the definition of the acceptance set, (A.4) is equivalent to

$$\begin{aligned}\rho(Z) &= \inf \left\{ x \in \mathbb{R} : Z - x \in \bigcup_{X \in \mathcal{A}_\rho} \{Y \in \mathcal{X} : Y \prec_{\text{sd}} X\} \right\} \\ &= \inf_{X \in \mathcal{A}_\rho} \inf\{x \in \mathbb{R} : Z - x \prec_{\text{sd}} X\}, \quad Z \in \mathcal{X}. \end{aligned} \tag{A.5}$$

Note that second order stochastic dominance can be characterized in terms of ψ_α, $\alpha \in [0,1]$ as in (2.2). Then it follows from (A.5) that

$$\rho(X) = \inf_{Y \in \mathcal{A}_\rho} \inf \left\{ x \in \mathbb{R} : \psi_\alpha(X) - x \leqslant \psi_\alpha(Y) \text{ for all } \alpha \in [0,1] \right\} = \inf_{Y \in \mathcal{A}_\rho} \sup_{\alpha \in [0,1]} \{\psi_\alpha(X) - \psi_\alpha(Y)\}.$$

Finally, for any $X \in \mathcal{X}$, we take $Z = X - \rho(X)$. Clearly, $Z \in \mathcal{A}_\rho$, and

$$\rho(X) = \sup_{\alpha \in [0,1]} \{\psi_\alpha(X) - \psi_\alpha(Z)\} = \min_{Y \in \mathcal{A}_\rho} \sup_{\alpha \in [0,1]} \{\psi_\alpha(X) - \psi_\alpha(Y)\}.$$

Thus, we obtain (3.1) and (3.2). □

Proof of Proposition 3.4. We show the conclusion for convex risk measures. The case for coherent risk measures is similar.

⇒: Suppose that ρ is a law-invariant convex risk measure. From Theorem 3.1, ρ has representation (3.5) and we can choose $\mathcal{G}$ as in (3.2). It suffices to show that $\mathcal{G}$ is a convex set. Write $g_Y : [0,1] \to \mathbb{R}$, $\alpha \mapsto \psi_\alpha(Y)$

for $Y \in \mathcal{X}$. For $Y_1, Y_2 \in \mathcal{A}_\rho$, let Y_2^* be such that $Y_2^* \overset{\mathrm{d}}{=} Y_2$ and (Y_1, Y_2^*) is comonotonic. For $\alpha, \lambda \in [0, 1]$, we have

$$\begin{aligned}\lambda g_{Y_1}(\alpha) + (1-\lambda) g_{Y_2}(\alpha) &= \lambda \psi_\alpha(Y_1) + (1-\lambda)\psi_\alpha(Y_2)\\ &= \lambda \psi_\alpha(Y_1) + (1-\lambda)\psi_\alpha(Y_2^*)\\ &= \psi_\alpha(\lambda Y_1 + (1-\lambda) Y_2^*),\end{aligned}$$

where the last equality is due to the comonotonic additivity of ψ_α. Note that $Y_2^* \in \mathcal{A}_\rho$ as ρ is law-invariant. Since ρ is convex, $\mathcal{A}_\rho$ is a convex set, and therefore $\lambda Y_1 + (1-\lambda) Y_2^* \in \mathcal{A}_\rho$. As a consequence, $\lambda g_{Y_1} + (1-\lambda) g_{Y_2} = g_{\lambda Y_1 + (1-\lambda) Y_2^*} \in \mathcal{G}$ and $\mathcal{G}$ is a convex set.

$\Leftarrow$: Suppose that a risk measure ρ has representation (3.5) in which $\mathcal{G}$ is convex. It suffices to show that $\mathcal{A}_\rho$ is a convex set. Let

$$\mathcal{A}^* = \bigcup_{g \in \mathcal{G}} \left\{ X \in \mathcal{X} : \sup_{\alpha \in [0,1]} \{\psi_\alpha(X) - g(\alpha)\} \leqslant 0 \right\}.$$

From (3.5), we have that $\mathcal{A}_\rho = \overline{\mathcal{A}^*}$ where $\overline{A}$ denotes the $\mathcal{X}$-closure of a set A. To show that $\mathcal{A}_\rho$ is a convex set, it suffices to show that $\mathcal{A}^*$ is a convex set. For any $X_1, X_2 \in \mathcal{A}^*$, there exist $g_1, g_2 \in \mathcal{G}$ such that

$$\psi_\alpha(X_i) \leqslant g_i(\alpha), \text{ for all } \alpha \in [0, 1], \quad i = 1, 2.$$

It follows that for all $\alpha, \lambda \in [0, 1]$,

$$\psi_\alpha(\lambda X_1 + (1-\lambda) X_2) \leqslant \lambda \psi_\alpha(X_1) + (1-\lambda)\psi_\alpha(X_2) \leqslant \lambda g_1(\alpha) + (1-\lambda) g_2(\alpha) =: g^*(\alpha).$$

Note that $g^* \in \mathcal{G}$ since $\mathcal{G}$ is convex. This shows $\lambda X_1 + (1-\lambda) X_2 \in \mathcal{A}^*$, that is, $\mathcal{A}^*$ is convex. □

Proof of the statement in Remark 3.2. We only show that $\mathcal{G}$ in (3.1) can be taken as a countable set. The cases of (3.4) and (3.5) are similar. Let $\mathcal{H}_{\mathbb{Q}}$ denote a set of simple non-decreasing and bounded functions mapping $[0, 1]$ to the set of rational numbers $\mathbb{Q}$, specifically,

$$\mathcal{H}_{\mathbb{Q}} = \bigcup_{n \in \mathbb{N}} \mathcal{H}_{n,\mathbb{Q}}$$

with

$$\mathcal{H}_{n,\mathbb{Q}} = \left\{ h : [0,1] \to \mathbb{Q} : h(x) = \sum_{i=1}^{2^n} \frac{a_i}{2^n} 1_{\{\frac{i-1}{2^n} < x \leqslant \frac{i}{2^n}\}},\ a_i \in \mathbb{N}, |a_i| \leqslant n 2^n,\ a_{i-1} < a_i,\ i = 2, \dots, 2^n \right\}.$$

It can be verified that $\mathcal{H}_{\mathbb{Q}}$ is a countable set since $\mathcal{H}_{n,\mathbb{Q}}$ is a finite set for each $n \in \mathbb{N}$. Denote

$$\mathcal{G}_{\mathbb{Q}} = \{ h \in \mathcal{H}_{\mathbb{Q}} : h \leqslant g \text{ for some } g \in \mathcal{G} \}$$

where $\mathcal{G}$ is defined by (3.2). By Theorem 3.1 (i), it is easy to see

$$\rho(X) \leqslant \inf_{h \in \mathcal{G}_{\mathbb{Q}}} \sup_{\alpha \in [0,1]} \{\psi_\alpha(X) - h(\alpha)\} =: \rho_{\mathbb{Q}}(X), \quad X \in \mathcal{X}.$$

For the other direction, for $X \in \mathcal{X}$ and any $\varepsilon > 0$, there exists $g \in \mathcal{G}$ and $\alpha \in [0,1]$ such that

$$\rho(X) + \varepsilon > \psi_\alpha(X) - g(\alpha).$$

For such g, which is a bounded function, there exists $h \in \mathcal{G}_{\mathbb{Q}}$ such that $g < h + \varepsilon$. Hence, $\rho(X) + \varepsilon > \psi_\alpha(X) - g(\alpha) > \psi_\alpha(X) - h(\alpha) - \varepsilon \geqslant \rho_{\mathbb{Q}}(X) - \varepsilon$. Letting $\varepsilon \downarrow 0$ yields $\rho(X) \geqslant \rho_{\mathbb{Q}}(X)$. Therefore, we conclude $\rho(X) = \rho_{\mathbb{Q}}(X)$; thus, $\mathcal{G}$ in (3.1) can be chosen as $\mathcal{G}_{\mathbb{Q}}$. □

Proof of Theorem 3.5. Suppose that $\{X_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ and $X \in \mathcal{X}$ satisfy $X_n \nearrow X \in \mathcal{X}$ a.s. as $n \to \infty$. Using Lemma 3.6 and Theorem 3.1, we have, writing $\mathcal{G}$ as the adjustment set of ρ,

$$\begin{aligned}
|\rho(X_n) - \rho(X)| &= \left| \min_{g\in\mathcal{G}} \sup_{\alpha\in[0,1]} \{\psi_\alpha(X_n) - g(\alpha)\} - \min_{g\in\mathcal{G}} \sup_{\alpha\in[0,1]} \{\psi_\alpha(X) - g(\alpha)\} \right| \\
&\leqslant \sup_{g\in\mathcal{G}} \sup_{\alpha\in[0,1]} |\psi_\alpha(X_n) - \psi_\alpha(X)| \\
&= \sup_{\alpha\in[0,1]} |\psi_\alpha(X_n) - \psi_\alpha(X)| \to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

This shows that ρ satisfies (CB), and hence (FP). □

Proof of Lemma 3.6. It is obvious that $\psi_\alpha(X_n)$, $n \in \mathbb{N}$ and $\psi_\alpha(X)$ are continuous and increasing function in $\alpha \in [0,1]$. It is well known that if a sequence of continuous and increasing functions converges to a continuous function point-wise on a compact set, then the convergence is uniform. It suffices to show that $\psi_\alpha(X_n) \to \psi_\alpha(X)$ point-wise. One can easily check that ψ_α for each $\alpha \in [0,1]$ satisfies (CB) (or, directly use Theorem 30 of Delbaen (2012)). Therefore, $\psi_\alpha(X_n) \to \psi_\alpha(X)$ point-wise, which implies uniform convergence as mentioned above. □

A.3 Proofs in Section 4

Proof of Theorem 4.1. (i) It is obvious that $\rho^*(X) \geqslant \inf\{\sum_{i=1}^n \rho_i(Y_i) : (Y_1,\dots,Y_n) \in \mathbb{A}_n^c(X)\}$. We only need to show the opposite direction of the inequality. For any $(X_1,\dots,X_n) \in \mathbb{A}_n(X)$, by the comonotone improvement in the form of Filipović and Svindland (2008, Proposition 5.1) and Ludkovski and Rüschendorf (2008, Theorem 2), there exists $(Y_1,\dots,Y_n) \in \mathbb{A}_n^c(X)$ such that $Y_i \prec_{\text{mps}} X_i$, $i = 1,\dots,n$. This implies $\rho_i(Y_i) \leqslant \rho_i(X_i)$ since ρ_i is consistent, $i = 1,\dots,n$. Taking infimums on both sides yields

$$\inf\left\{\sum_{i=1}^{n}\rho_i(Y_i):(Y_1,\dots,Y_n)\in\mathbb{A}_n^c(X)\right\}\leqslant\inf\left\{\sum_{i=1}^{n}\rho_i(X_i):(X_1,\dots,X_n)\in\mathbb{A}_n(X)\right\}$$
$$=\rho^*(X).$$

(ii) To show that ρ^* is finite-valued, note that, using (SC) and (TI), for $(X_1,\dots,X_n)\in\mathbb{A}_n(X)$,

$$\sum_{i=1}^{n}\rho_i(X_i)\geqslant\sum_{i=1}^{n}\rho_i(\mathbb{E}[X_i])=\sum_{i=1}^{n}\rho_i(0)+\mathbb{E}[X].$$

Therefore, $\rho^*(X)\geqslant\sum_{i=1}^{n}\rho_i(0)+\mathbb{E}[X]>-\infty$. To show the existence of an optimal allocation, based on part (i) and the fact that $\rho^*(X)$ is finite, it suffices to show that every sequence in the set $\mathbb{A}_n^c(X)$ has a subsequence which converges to an element of $\mathbb{A}_n^c(X)$ in L^∞-norm. This result is given in the proof of Theorem 3.2 of Jouini et al. (2008).

(iii) It is easy to verify that ρ^* is a monetary risk measure. To show that ρ^* is consistent, from Theorem 2.1, it suffices to show that ρ^* satisfies (DM). Let $X,Y\in\mathcal{X}$ be two random variables such that $X=\mathbb{E}[Y|X]$. For any $(Y_1,\dots,Y_n)\in\mathbb{A}_n(Y)$, let $X_i=\mathbb{E}[Y_i|X]$, $i=1,\dots,n$. It is obvious that $(X_1,\dots,X_n)\in\mathbb{A}_n(X)$. For each $i=1,\dots,n$, since ρ_i satisfies (DM) and (X_i,Y_i) is a martingale, we have $\rho_i(X_i)\leqslant\rho_i(Y_i)$. This shows

$$\rho^*(X)=\inf\left\{\sum_{i=1}^{n}\rho_i(X_i):(X_1,\dots,X_n)\in\mathbb{A}_n(X)\right\}$$
$$\leqslant\inf\left\{\sum_{i=1}^{n}\rho_i(Y_i):(Y_1,\dots,Y_n)\in\mathbb{A}_n(Y)\right\}=\rho^*(Y).$$

That is, ρ^* satisfies (DM).

Denote by $\mathcal{A}^*$ the acceptance set of ρ^* and it remains to show $\mathcal{A}^*=\sum_{i=1}^{n}\mathcal{A}_i$. For any $X_i\in\mathcal{A}_i$, that is, $\rho_i(X_i)\leqslant 0$, $i=1,\dots,n$, we have that for $X=X_1+\dots+X_n$

$$\square_{i=1}^{n}\rho_i(X)=\inf\left\{\sum_{i=1}^{n}\rho_i(Y_i):\ (Y_1,\dots,Y_n)\in\mathbb{A}_n(X)\right\}\leqslant\sum_{i=1}^{n}\rho_i(X_i)\leqslant 0.$$

We have $X\in\mathcal{A}^*$, and thus $\sum_{i=1}^{n}\mathcal{A}_i\subset\mathcal{A}^*$. For the other direction, let $X\in\mathcal{A}^*$, i.e., $\square_{i=1}^{n}\rho_i(X)\leqslant 0$. By (ii), there exists $(X_1,\dots,X_n)\in\mathbb{A}_n^c(X)$ such that

$$\rho_1(X_1)+\dots+\rho_n(X_n)=\square_{i=1}^{n}\rho_i(X)\leqslant 0.$$

By (TI) of ρ_i, $i=1,\dots,n$, we can safely take $\rho_i(X_i)\leqslant 0$, $i=1,\dots,n$. Hence, we have that $X=\sum_{i=1}^{n}X_i\in\sum_{i=1}^{n}\mathcal{A}_i$. It follows that $\mathcal{A}^*\subset\sum_{i=1}^{n}\mathcal{A}_i$, and in summary, $\mathcal{A}^*=\sum_{i=1}^{n}\mathcal{A}_i$.

(iv) We only show the case $n = 2$; for $n \geqslant 3$ the proof is similar. First, we introduce the functions g_Y, g_{Y_1,Y_2} and $\hat{g}_{Y_1,Y_2}$. For $Y \in \mathcal{X}$, let $g_Y : [0,1] \to \mathbb{R}$, $\alpha \mapsto \psi_\alpha(Y)$ and for $(Y_1, Y_2) \in (\mathcal{X})^2$, let $g_{Y_1,Y_2} : [0,1] \to \mathbb{R}$, $\alpha \mapsto \psi_\alpha(Y_1 + Y_2)$ and $\hat{g}_{Y_1,Y_2} : [0,1] \to \mathbb{R}$, $\alpha \mapsto \psi_\alpha(Y_1) + \psi_\alpha(Y_2)$. Note that by the comonotonic additivity of ψ_α, $g_{Y_1,Y_2} = \hat{g}_{Y_1,Y_2}$ if (Y_1, Y_2) is comonotonic.

Let $\mathcal{A}_i$ be the acceptance set of ρ_i, $i = 1, 2$. By (iii), we know that the acceptance set of ρ^* is $\mathcal{A}_1 + \mathcal{A}_2$. From (iii) and Theorem 3.1, ρ^* has an adjustment set

$$\mathcal{G} = \{g_Y : Y \in \mathcal{A}_1 + \mathcal{A}_2\} = \{g_{Y_1,Y_2} : (Y_1, Y_2) \in \mathcal{A}_1 \times \mathcal{A}_2\}.$$

On the other hand, by definition of the adjustment sets $\mathcal{G}_1$ and $\mathcal{G}_2$,

$$\mathcal{G}_1 + \mathcal{G}_2 = \{g_{Y_1} : Y_1 \in \mathcal{A}_1\} + \{g_{Y_2} : Y \in \mathcal{A}_2\} = \{\hat{g}_{Y_1,Y_2} : (Y_1, Y_2) \in \mathcal{A}_1 \times \mathcal{A}_2\}.$$

Since ρ_1 and ρ_2 are law-invariant, for any $(Y_1, Y_2) \in \mathcal{A}_1 \times \mathcal{A}_2$ we can find $(Y_1^c, Y_2^c) \in \mathcal{A}_1 \times \mathcal{A}_2$ such that $Y_i^c \overset{\mathrm{d}}{=} Y_i$, $i = 1, 2$ and (Y_1^c, Y_2^c) is comonotonic. Then $\hat{g}_{Y_1,Y_2} = \hat{g}_{Y_1^c,Y_2^c} = g_{Y_1^c,Y_2^c} \in \mathcal{G}$. This shows $\mathcal{G}_1 + \mathcal{G}_2 \subset \mathcal{G}$.

For the other direction of the inclusion, by using the comonotone improvement again, for any $(Y_1, Y_2) \in \mathcal{A}_1 \times \mathcal{A}_2$, there exists $(Z_1, Z_2) \in \mathbb{A}_2^c(Y_1 + Y_2)$ such that $Z_i \prec_{\mathrm{mps}} Y_i$, $i = 1, 2$. Since ρ_1 and ρ_2 are consistent, $(Z_1, Z_2) \in \mathcal{A}_1 \times \mathcal{A}_2$. It follows that $g_{Y_1,Y_2} = g_{Z_1,Z_2} = \hat{g}_{Z_1,Z_2} \in \mathcal{G}_1 + \mathcal{G}_2$. This shows $\mathcal{G} \subset \mathcal{G}_1 + \mathcal{G}_2$. Now we can conclude that $\mathcal{G} = \mathcal{G}_1 + \mathcal{G}_2$. □

A.4 Proofs in Section 5

Proof of Lemma 5.1. Note that $-L_g$ is an increasing function. It is easy to check, for $\alpha \in (0, 1)$,

$$\psi_\alpha(Y) = \psi_\alpha(-L_g(U)) = \frac{1}{1-\alpha} \int_\alpha^1 (-L_g(u))\mathrm{d}u = \frac{1}{1-\alpha}(\hat{g}(\alpha) - \hat{g}(1)) = g(\alpha).$$

Further, by continuity of $\psi_\alpha(Y)$ and $\hat{g}(\alpha)$ for $\alpha \in [0, 1]$, we know $\psi_\alpha(Y) = g(\alpha)$ for all $\alpha \in [0, 1]$. □

Proof of Proposition 5.2. By Lemma 5.1 and translation-invariance of ψ_α, we have, for $\alpha \in [0, 1]$,

$$\psi_\alpha(X^*) = \psi_\alpha(-L_g(U)) + x_0 - \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha) = g(\alpha) + x_0 - \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha). \tag{A.6}$$

Therefore,

$$\rho_g(X^*) = \sup_{\alpha \in [0,1]} \{\psi_\alpha(X^*) - g(\alpha)\} = x_0 - \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha).$$

Take $X \in \mathcal{X}$ be such that $\mathbb{E}^{\mathbb{Q}}[X] \geqslant x_0$. By the Hardy-Littlewood inequality (e.g. Remark 3.25 of Rüschendorf (2013)), we have

$$\mathbb{E}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} X\right] \leqslant \int_0^1 q_\alpha\left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right) q_\alpha(X)\mathrm{d}\alpha. \tag{A.7}$$

Integration by parts leads to

$$\begin{aligned}\int_0^1 q_\alpha\left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right) q_\alpha(X)\mathrm{d}\alpha &= \int_0^1 q_\alpha\left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right) \mathrm{d}\left(-(1-\alpha)\psi_\alpha(X)\right)\\ &= \psi_0(X)q_0\left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right) + \int_0^1 (1-\alpha)\psi_\alpha(X)\mathrm{d}q_\alpha\left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right)\\ &= \int_0^1 \psi_\alpha(X)\mathrm{d}\mu(\alpha). \end{aligned} \tag{A.8}$$

Putting (A.7) and (A.8) together, we obtain

$$x_0 \leqslant \mathbb{E}^{\mathbb{Q}}[X] = \mathbb{E}\left[\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}X\right] \leqslant \int_0^1 q_\alpha\left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right) q_\alpha(X)\mathrm{d}\alpha = \int_0^1 \psi_\alpha(X)\mathrm{d}\mu(\alpha).$$

On the other hand,

$$\begin{aligned}\rho_g(X) &= \sup_{\alpha\in[0,1]} \{\psi_\alpha(X) - g(\alpha)\}\\ &\geqslant \int_0^1 (\psi_\alpha(X) - g(\alpha))\mathrm{d}\mu(\alpha) && \text{(A.9)}\\ &= \int_0^1 \psi_\alpha(X)\mathrm{d}\mu(\alpha) - \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha) \geqslant x_0 - \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha) = \rho(X^*). && \text{(A.10)}\end{aligned}$$

This shows the optimality of X^* as well as the corresponding minimal value of Problem (5.4).

Next we show uniqueness of the solution. If X is a solution to Problem (5.4), then all inequalities above are equalities. First, the inequality in (A.7) is an equality if and only if $(X, \mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P})$ is comonotonic. Second, the inequality in (A.9) is an equality if and only if $\psi_\alpha(X)-g(\alpha)$ is a constant c for μ-a.s. $\alpha \in [0,1]$. Note that $q_t(\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P})$ is strictly increasing for $t \in [0,1]$ since $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}$ is continuously distributed. As a consequence, $\mu([0,t])$ is strictly increasing for $t \in [0,1]$. Therefore, $\psi_\alpha(X) - g(\alpha) = c$ holds μ-a.s. means it holds for a.e. $\alpha \in [0,1]$, hence for all $\alpha \in [0,1]$ due to continuity. Finally, the inequality in (A.10) is an equality if and only if $c = x_0 - \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha)$. Summarizing the above three statements, any solution X to Problem (5.4) is an increasing function of $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}$, and it satisfies

$$\psi_\alpha(X) = g(\alpha) + x_0 - \int_0^1 g(\alpha)\mathrm{d}\mu(\alpha),\ \alpha \in [0,1].$$

Together with (A.6), we know that X is identically distributed as X^*. Since both of them are increasing functions of the continuously distributed random variable $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}$, we conclude that $X = X^*$ a.s. □

Proof of Theorem 5.3. Write $\mathcal{X}_0 = \{X \in \mathcal{X} : \mathbb{E}^Q[X] \geqslant x_0\}$.

(i) By Proposition 5.2 and exchanging the order of two infima, we have

$$\begin{aligned}\inf_{X\in\mathcal{X}_0}\rho(X) &= \inf_{X\in\mathcal{X}_0}\min_{g\in\mathcal{G}}\rho_g(X)\\ &= \inf_{g\in\mathcal{G}}\inf_{X\in\mathcal{X}_0}\rho_g(X)\\ &= \inf_{g\in\mathcal{G}}\left\{x_0-\int_0^1 g(\alpha)\mathrm{d}\mu(\alpha)\right\} = x_0-\sup_{g\in\mathcal{G}}\int_0^1 g(\alpha)\mathrm{d}\mu(\alpha).\end{aligned}$$

(ii) Suppose that Y is a solution to Problem (5.1). Because $\rho(Y)=\min_{g\in\mathcal{G}}\rho_g(Y)$, there exists $g^*\in\mathcal{G}$ such that $\rho(Y)=\rho_{g^*}(Y)$. By part (i), we have

$$\rho_{g^*}(Y)=\rho(Y)=x_0-\sup_{g\in\mathcal{G}}\int_0^1 g(\alpha)\mathrm{d}\mu(\alpha).$$

By Proposition 5.2,

$$\rho_{g^*}(Y)\geqslant x_0-\int_0^1 g^*(\alpha)\mathrm{d}\mu(\alpha)\geqslant x_0-\sup_{g\in\mathcal{G}}\int_0^1 g(\alpha)\mathrm{d}\mu(\alpha).$$

Therefore, $\int_0^1 g^*(\alpha)\mathrm{d}\mu(\alpha) = \sup_{g\in\mathcal{G}}\int_0^1 g(\alpha)\mathrm{d}\mu(\alpha)$, and hence g^* maximizes $\int_0^1 g(\alpha)\mathrm{d}\mu(\alpha)$ over $g\in\mathcal{G}$. The other direction of the statement is implied by (iii).

(iii) Using Proposition 5.2 and part (i), we have

$$\rho_{g^*}(X^*)=x_0-\int_0^1 g^*(\alpha)\mathrm{d}\mu(\alpha)=x_0-\sup_{g\in\mathcal{G}}\int_0^1 g(\alpha)\mathrm{d}\mu(\alpha)=\inf_{X\in\mathcal{X}_0}\rho(X).$$

Hence, X^* is a solution to Problem (5.1).

(iv) Let X^* be a solution to Problem (5.1). Using the same argument as in part (ii), we know that there exists $g^*\in\mathcal{G}$ such that

$$\rho_{g^*}(X^*)=x_0-\int_0^1 g^*(\alpha)\mathrm{d}\mu(\alpha).$$

Therefore, by Proportion 5.2 again, we know that X has the form (5.5) a.s. □

Details in Example 5.2. Using Proposition 5.2, we have

$$\begin{aligned}\min\{\rho_{g_n}(X):X\in\mathcal{X},\ \mathbb{E}^{\mathbb{Q}}[X]\geqslant 0\} &= -\int_0^1 g_n(\alpha)\mathrm{d}\mu(\alpha)\\ &= -\int_0^1 \min\left\{\frac{1}{1-\alpha},n\right\}2(1-\alpha)\mathrm{d}\alpha=\frac{1}{n}-2.\end{aligned}$$

Therefore,

$$\inf\{\rho_g(X):X\in\mathcal{X},\ \mathbb{E}^{\mathbb{Q}}[X]\geqslant 0\}=\inf_{n\in\mathbb{N}}\min\{\rho_{g_n}(X):X\in\mathcal{X},\ \mathbb{E}^{\mathbb{Q}}[X]\geqslant 0\}=-2.$$

We shall see that this infimum value is not attainable. Suppose, for the purpose of contradiction, that there exists $X \in \mathcal{X}$ such that $\mathbb{E}^{\mathbb{Q}}[X] \geqslant 0$ and $\rho_g(X) = -2$. Note that

$$\rho_g(X) = \inf_{n\in\mathbb{N}} \sup_{\alpha\in[0,1]} \left\{ \psi_\alpha(X) - \min\left\{ \frac{1}{1-\alpha}, n \right\} \right\} \geqslant \sup_{\alpha\in[0,1]} \left\{ \psi_\alpha(X) - \frac{1}{1-\alpha} \right\}.$$

This gives $\psi_\alpha(X) \leqslant \frac{1}{1-\alpha} - 2$ for all $\alpha \in [0,1]$. On the other hand, by (A.8),

$$0 \leqslant \mathbb{E}^{\mathbb{Q}}[X] \leqslant \int_0^1 \psi_\alpha(X)\mathrm{d}\mu(\alpha) \leqslant \int_0^1 \left(\frac{1}{1-\alpha} - 2 \right) 2(1-\alpha)\mathrm{d}\alpha = 0.$$

As a consequence, it must be $\psi_\alpha(X) = \frac{1}{1-\alpha} - 2$ because $\psi_\alpha(X)$ is continuous in α. This leads to a contradiction, since $\lim_{\alpha\to 1}(1-\alpha)\psi_\alpha(Y) = 0$ for any integrable random variable Y. □

B Generalization of the main results to $\mathcal{X} = L^p$, $p \in [1,\infty)$

In this part of the Appendix, we discuss our main results in the more general framework of unbounded random variables. Let L^p denote the space of random variables in $(\Omega, \mathcal{A}, \mathbb{P})$ with finite p-th moment, $p \in [0,\infty)$, and L^∞ the space of essentially bounded random variables. $\mathcal{X} = L^\infty$ corresponds to the setting used throughout the main text. For regulatory purposes, considering the sets L^p of risks is indeed necessary since for finance and insurance one often encounters models with unbounded risks.

As discussed by Filipović and Svindland (2012), the canonical space of convex risk measures is L^1, and as such we consider $\mathcal{X} = L^p$, $p \in [1,\infty)$ in the following. A risk measure on L^p is a functional mapping L^p to $(-\infty,\infty]$ with $\rho(c) < \infty$ for $c \in \mathbb{R}$. Recall that a consistent risk measure satisfies (TI) and (SC).

B.1 Results in Section 3 for $\mathcal{X} = L^p$

We first generalize the representation results in Section 3 to L^p. When $\mathcal{X} = L^p$, $p \in [1,\infty)$, we need the L^p-Fatou property for a convex risk measure to be consistent:

(FP) L^p-Fatou property: $\liminf_{n\to\infty} \rho(X_n) \geqslant \rho(X)$ if $X, X_1, X_2, \dots \in \mathcal{X} = L^p$ and $X_n \xrightarrow{L^p} X$ as $n \to \infty$.

Similarly to the case of $\mathcal{X} = L^\infty$, the Fatou property is necessary and sufficient for the Kusuoka representation of convex risk measures in Section 3.4 to hold in L^p; see Kaina and Rüschendorf (2009), Delbaen (2009) and Filipović and Svindland (2012). A real-valued convex risk measure always satisfies the Fatou property (in fact, it is always L^p-continuous); see e.g. Rüschendorf (2013).

Proposition B.1. *A law-invariant convex risk measure on L^p, $p \in [1,\infty)$ with the Fatou property (FP) is a consistent risk measure, and so is the maximum, the minimum or a convex combination of law-invariant convex risk measures with (FP).*

Proposition B.1 is the L^p-version of Proposition 3.2; it is a direct result of the Kusuoka representation in L^p; see Filipović and Svindland (2012).

Theorem B.2. *For a risk measure ρ on $\mathcal{X} = L^p$, $p \in [1,\infty)$, the following are equivalent:*

(i) *ρ is a consistent risk measure;*

(ii) *ρ has the following representation*

$$\rho(X) = \min_{g\in\mathcal{G}} \sup_{\alpha\in[0,1)} \{\psi_\alpha(X) - g(\alpha)\}\,, \quad X \in \mathcal{X}, \tag{B.1}$$

where $\mathcal{G}$ is a set of measurable functions mapping $[0,1)$ to $(-\infty,\infty]$;

(iii) *ρ has the following representation*

$$\rho(X) = \min_{\tau\in\mathcal{C}} \tau(X), \quad X \in \mathcal{X}, \tag{B.2}$$

where $\mathcal{C}$ is a set of law-invariant convex risk measures on $\mathcal{X}$ with the Fatou property.

Theorem B.2 is the L^p-version of Theorem 3.1 and Theorem 3.3. The implication (i)⇒(ii) shares a similar proof to that of Theorem 3.1; the only difference is that $\alpha \in [0,1]$ is replaced by $\alpha \in [0,1)$ to avoid $\infty - \infty$. The implication (ii)⇒(iii) holds by noting that $\psi_\alpha - g(\alpha)$ for each $\alpha \in [0,1]$ is a convex risk measure with the Fatou property. The implication (iii)⇒(i) follows from Proposition B.1.

Although a consistent risk measure ρ can be represented by the infimum of convex risk measures on L^p with the Fatou property, ρ itself does not necessarily have the Fatou property, which can be illustrated by the following example. This is in contrast to the case of $\mathcal{X} = L^\infty$ in Theorem 3.5.

Example B.1. Suppose that $p \in [1,\infty)$ and $(\Omega, \mathcal{F}, \mathbb{P}) = ((0,1), \mathcal{B}(0,1), \lambda)$ where $\mathcal{B}(0,1)$ is the set of all Borel sets on $(0,1)$ and λ is the Lebesgue measure. Define the random variable X as $X(\omega) = \omega^{-1/(2p)}$, $\omega \in (0,1)$ and a sequence of random variables $\{X_n, n \in \mathbb{N}\}$ with $X_n(\omega) = n^{1/(2p)} 1_{\{\omega\leqslant 1/n\}} + X(\omega) 1_{\{\omega\geqslant 1/n\}}$, $\omega \in (0,1)$. It is obvious that $X_n \to X$ a.s. as $n \to \infty$ and $|X_n| \leqslant X \in L^p$. It follows that $X_n \overset{L^p}{\to} X$ as $n \to \infty$. Define the risk measure ρ as

$$\rho(Y) = \inf_{n\geqslant 1} \rho_n(Y), \quad Y \in L^p,$$

where $\rho_n(Y) = \inf\{m \in \mathbb{R} : Y - m \prec_{\rm sd} X_n\}$, $n \in \mathbb{N}$. It is easy to verify that ρ is a consistent risk measure. Note that for any $n \in \mathbb{N}$ and $x \in \mathbb{R}$, $X - x \prec_{\rm sd} X_n$ does not hold. It follows that $\rho_{X_n}(X) = \infty$ for all $n \in \mathbb{N}$, which further implies $\rho(X) = \infty$. However, $\rho(X_n) \leqslant 0$ for all $n \in \mathbb{N}$, implying ρ does not satisfy the Fatou property.

B.2 Results in Section 2 for $\mathcal{X} = L^p$

Below we generalize Theorem 2.1 to the space L^p. To show this result, we require the Fatou property, which is typical in the literature of risk measures on L^p. Note that the most technical part of the following Theorem, (DC) implying (MC), has a proof that is completely different from the case of bounded random variables. These two proofs cannot be combined, since the proof for $\mathcal{X} = L^\infty$ relies on the norm continuity of ρ and a construction leading to L^∞ convergence, which is not possible for L^p, and the proof for $\mathcal{X} = L^p$ relies on the Fatou property of ρ and a different construction leading to L^p convergence, which is not useful for L^∞.

Theorem B.3. *For a monetary risk measure ρ on L^p, $p \in [1, \infty)$ with the Fatou property, the properties (SC), (MC), (DM), (DC) and (EI) are all equivalent. In turn, they imply that ρ satisfies (LD).*

Proof. By Theorem 2.1 and Remark 2.2 of Svindland (2014), (DM) implies (LD) on L^p, $p \in [1, \infty)$. Following the same proof as in Theorem 2.1, (MC), (SC) and (EI) are equivalent, (MC) implies both (DC) and (DM), and (DM)+(LD) implies (MC). It remains to show that (DC) implies (MC).

Note that (DC) implies (LD) based on the same proof as in Theorem 2.1. Let X and Y be two random variables in L^p such that $X \prec_{\rm mps} Y$. We aim to show $\rho(X) \leqslant \rho(Y)$. Via the martingale representation of convex order in Shaked and Shanthikumar (2007, Theorem 3.A.4), we can assume $X = \mathbb{E}[Y|X]$. For each $n \in \mathbb{N}$, define

$$X_n = x^{(n)}_{-n^2-1} 1_{\{X<-n\}} + \sum_{i=-n^2}^{n^2-1} x^{(n)}_i 1_{\{\frac{i}{n} \leqslant X < \frac{i+1}{n}\}} + x^{(n)}_{n^2} 1_{\{X \geqslant n\}} =: \sum_{i=-n^2-1}^{n^2} x^{(n)}_i 1_{A^{(n)}_i}, \quad n \in \mathbb{N},$$

where $x^{(n)}_{-n^2-1} = \mathbb{E}[X|X < -n]$, $x^{(n)}_i = \mathbb{E}[X|\frac{i}{n} \leqslant X < \frac{i+1}{n}]$, $i = -n^2, \dots, n^2 - 1$, and $x^{(n)}_{n^2} = \mathbb{E}[X|X \geqslant n]$. It is easy to check that $X_n \to X$ a.s. as $n \to \infty$ and $\{|X_n|^p\}_{n\in\mathbb{N}}$ is uniformly integrable because Jensen's inequality gives $\mathbb{E}[|X_n|^p] \leqslant \mathbb{E}[|X|^p]$ for each $n \in \mathbb{N}$. Hence, X_n converges to X in L^p as $n \to \infty$. By the Fatou property of ρ, we have

$$\rho(X) \leqslant \liminf_{n\to\infty} \rho(X_n). \tag{B.3}$$

For $A^{(n)}_{-n^2-1}, \dots, A^{(n)}_{n^2}$ defined above, we have

$$Y = \sum_{i=-n^2-1}^{n^2} Y 1_{A^{(n)}_i} \overset{\rm d}{=} \sum_{i=-n^2-1}^{n^2} Y^{(n)}_i 1_{A^{(n)}_i},$$

where $Y^{(n)}_{-n^2-1}, \dots, Y^{(n)}_{n^2}$ are $2n^2+1$ independent random variables, independent of the sets $A^{(n)}_{-n^2-1}, \dots, A^{(n)}_{n^2}$ and satisfying $Y^{(n)}_i \overset{\rm d}{=} [Y|A^{(n)}_i]$ for $i = -n^2 - 1, \dots, n^2$. For $k \in \mathbb{N}$, let $(Y^{(n)}_{-n^2-1,j}, \dots, Y^{(n)}_{n^2,j})$, $j =$

$1, \ldots, 2^k$, be 2^k iid random vectors identically distributed as $(Y^{(n)}_{-n^2-1}, \ldots, Y^{(n)}_{n^2})$. For $k \in \mathbb{N}$, we define

$$Y^{(n,j)} = \sum_{i=-n^2-1}^{n^2} Y^{(n)}_{i,j} 1_{A_i^{(n)}}, \quad j = 1, \ldots, 2^k.$$

It is easy to see that $Y^{(n,1)}, \ldots, Y^{(n,2^k)}$ are identically distributed such that $Y^{(n,j)} \overset{d}{=} Y$, $j = 1, \ldots, 2^k$, and $Y^{(n,1)} + \cdots + Y^{(n,2^{k-1})} \overset{\mathrm{d}}{=} Y^{(n,2^{k-1}+1)} + \cdots + Y^{(n,2^k)}$ by symmetry. By (DC), we know for any two identically distributed random variables X_1 and X_2, $\rho((X_1 + X_2)/2) \leqslant \rho(X_1)$. Using this relation repeatedly, we have

$$\rho\left(\frac{1}{2^k}\sum_{j=1}^{2^k} Y^{(n,j)}\right) \leqslant \rho\left(\frac{1}{2^{k-1}}\sum_{j=1}^{2^{k-1}} Y^{(n,j)}\right) \leqslant \ldots \leqslant \rho(Y^{(n,1)}) = \rho(Y), \tag{B.4}$$

where the last equality is due to (LD). On the other hand, note that

$$Z_{n,k} := \frac{1}{2^k}\sum_{j=1}^{2^k} Y^{(n,j)} = \sum_{i=-n^2-1}^{n^2} 1_{A_i^{(n)}} \left(\frac{1}{2^k}\sum_{j=1}^{2^k} Y^{(n)}_{i,j}\right).$$

By the Law of Large Numbers, we have $\frac{1}{2^k}\sum_{j=1}^{2^k} Y^{(n)}_{i,j}$ converges to $x_i^{(n)} = \mathbb{E}[Y|A_i^{(n)}]$ a.s. as $k \to \infty$ for $i = -n^2 - 1, \ldots, n^2$. This implies that $Z_{n,k}$ converges to X_n a.s. as $k \to \infty$. Also, by the C_r-inequality,

$$\mathbb{E}[|Z_{n,k}|^p] = \sum_{i=-n^2-1}^{n^2} \mathbb{P}(A_i^{(n)})\mathbb{E}\left[\left|\frac{1}{2^k}\sum_{j=1}^{2^k} Y^{(n)}_{i,j}\right|^p\right] \leqslant \sum_{i=-n^2-1}^{n^2} \mathbb{P}(A_i^{(n)})\mathbb{E}\left[\left|Y_i^{(n)}\right|^p\right] = \mathbb{E}[|Y|^p],$$

showing that $|Z_{n,k}|^p$ is uniformly integrable. It then follows that $Z_{n,k}$ converges to X_n in L^p as $k \to \infty$. Then by the Fatou property and (B.4), we have

$$\rho(X_n) \leqslant \liminf_{k\to\infty} \rho(Z_{n,k}) \leqslant \rho(Y).$$

Combined with (B.3), we have $\rho(X) \leqslant \rho(Y)$. This completes the proof. □

B.3 Results in Sections 4 and 5 for $\mathcal{X} = L^p$

We claim that theorem 4.1 (i)-(iv) hold when $\rho_1, \ldots, \rho_n$ are consistent risk measures on $\mathcal{X} = L^p$. The proofs of (i) and (iv) are essentially the same. To generalize (ii) to the case of L^p, one can use the same arguments in the proof of Theorem 2.5 of Filipović and Svindland (2008). To generalize (iii), it suffices to show that ρ^* is law-invariant so that (DM) is equivalent to (SC). For any $X, Y \in L^p$ such that $X \overset{\mathrm{d}}{=} Y$ and any $(X_1, \ldots, X_n) \in \mathbb{A}_n^c(X)$, by Denneberg's Lemma (Denneberg, 1994) we can write $(X_1, \ldots, X_n) = \mathbf{f}(X)$ for a component-wise non-decreasing function $\mathbf{f} : \mathbb{R} \to \mathbb{R}^n$. Then, $\mathbf{f}(Y) \in \mathbb{A}_n^c(Y)$ and $(X_1, \ldots, X_n) \overset{\mathrm{d}}{=} \mathbf{f}(Y)$. From (i), we have $\rho^*(X) = \rho^*(Y)$, that is, ρ^* is law-invariant.

To extend the results in Section 5 to L^p, it suffices to assume $\mathbb{E}^{\mathbb{Q}}[X] < \infty$ for all $X \in L^p$, or equivalently, $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P} \in L^q$ where $1/p + 1/q = 1$, and to modify the definition of $\mathcal{G}^*$ by allowing $g(1) = \infty$ and requiring $\hat{g}(1) = 0$ for $g \in \mathcal{G}^*$. All results follow from the same proof.

References

Acharya, V.V. (2009). A theory of systemic risk and design of prudential bank regulation. *Journal of Financial Stability*, **5**(3), 224–255.

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**, 269–298.

BCBS (2016). *Standards. Minimum capital requirements for Market Risk. January 2016.* Basel Committee on Banking Supervision. Basel: Bank for International Settlements.

Bäuerle, N. and Müller, A. (2006). Stochastic orders and risk measures: Consistency and bounds. *Insurance: Mathematics and Economics*, **38**(1), 132–148.

Bénéplanc, G. and Rochet, J. C. (2011). *Risk Management in Turbulent Times*. OUP USA.

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., and Montrucchio, L. (2011). Risk measures: rationality and diversification. *Mathematical Finance*, **21**(4), 743–774.

Cherny, A. S. and Grigoriev, P. G. (2007). Dilatation monotone risk measures are law invariant. *Finance and Stochastics*, **11**(2), 291–298.

Cont, R., Deguest, R. and Scandolo, G. (2010). Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance*, **10**(6), 593–606.

Delbaen, F. (2009). Risk measures for non-integrable random variables. *Mathematical Finance*, **19**(2), 329–333.

Delbaen, F. (2012). *Monetary Utility Functions*. Osaka University Press.

Denneberg, D. (1994). *Non-additive Measure and Integral.* Kluwer, London.

Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. and Vyncke, D. (2002). The concept of comonotonicity in actuarial science and finance: Theory. *Insurance: Mathematics and Economics*, **31**(1), 3-33.

Embrechts, P., Liu, H. and Wang, R. (2018). Quantile-based risk sharing. *Operations Research*, **66**(4), 936–949.

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., Schied, A. and Wang, R. (2018). Robustness in the optimization of risk measures. *SSRN*: 3254587.

El Karoui, N. and Ravanelli, C. (2009). Cash subadditive risk measures and interest rate ambiguity. *Mathematical Finance*, 19(4), 561–590.

Filipović, D. and Svindland, G. (2008). Optimal capital and risk allocations for law- and cash-invariant convex functions. *Finance and Stochastics*, **12**, 423–439.

Filipović, D. and Svindland, G. (2012). The canonical model space for law-invariant convex risk measures is L^1. *Mathematical Finance*, **22**(3), 585–589.

Fissler, T. and Ziegel, J. F. (2016). Higher order elicitability and Osband's principle. *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. (2011). *Stochastic Finance: An Introduction in Discrete Time.* Walter de Gruyter, Third Edition.

Föllmer, H., Schied, A. and Weber, S. (2009). Robust preferences and robust portfolio choice. In: *Mathematical Modelling and Numerical Methods in Finance*, (eds.: P. Ciarlet, A. Bensoussan, Q. Zhang), Handbook of Numerical Analysis 15, 29–88.

Frittelli M. and Rosazza Gianin E. (2002). Putting order in risk measures. *Journal of Banking and Finance*, **26**,1473–1486.

Frittelli M. and Rosazza Gianin E. (2005). Law invariant convex risk measures. *Advances in Mathematical Economics*, **7**, 33–46.

He, X. D. and Zhou, X. Y. (2011). Portfolio choice via quantiles. *Mathematical Finance*, **21**(2), 203–231.

IAIS (2014). *Consultation Document December 2014. Risk-based Global Insurance Capital Standard*. International Association of Insurance Supervisors.

Jia, J. and Dyer, J. S. (1996). A standard measure of risk and risk-value models. *Management Science*, **42**(12), 1691–1705.

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.

Kaina, M. and Rüschendorf, L. (2009). On convex risk measures on L^p-spaces. *Mathematical Methods in Operations Research*, **69**(3), 475–495.

Kou, S. and Peng, X. (2016). On the measurement of economic tail risk. *Operations Research*, **64**(5), 1056–1072.

Kusuoka, S. (2001). On law invariant coherent risk measures. *Advances in Mathematical Economics*, **3**, 83–95.

Leitner, J. (2004). Balayage monotonous risk measures. *International Journal of Theoretical and Applied Finance*, **7**(7), 887–900.

Ludkovski, M. and Rüschendorf, L. (2008). On comonotonicity of Pareto optimal risk sharing. *Statistics and Probability Letters*, **78**(10), 1181–1188.

Mao, T. and Hu, T. (2015). Relations between the spectral measures and dependence. *Extremes*, **18**(1), 65–84.

McNeil, A. J., Frey, R. and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*. Revised Edition. Princeton, NJ: Princeton University Press.

Müller, A. and Scarsini, M. (2001). Stochastic comparison of random vectors with a common copula. *Mathematics of Operations Research*, **26**(4), 723–740.

Müller, A., Scarsini, M., Tsetlin, I. and Winkler. R. L. (2017). Between first and second-order stochastic dominance. *Management Science*, **63**(9), 2933–2947.

Müller, A. and Stoyan, D. (2002). *Comparison Methods for Stochastic Models and Risks*. Wiley, England.

Puccetti, G. and Wang R. (2015). Extremal dependence concepts. *Statistical Science*, **30**(4), 485–517.

Rothschild, M. and Stiglitz, J. (1970). Increasing risk I: a definition. *Journal of Economic Theory*, **2**, 225–243.

Rüschendorf, L. (2013). *Mathematical Risk Analysis. Dependence, Risk Bounds, Optimal Allocations and Portfolios*. Springer, Heidelberg.

Schied, A. (2004). On the Neyman–Pearson problem for law-invariant risk measures and robust utility functionals. *Annals of Applied Probability*, **14**(3), 1398–1423.

Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders*. Springer Series in Statistics.

Svindland, G. (2014). Dilatation monotonicity and convex order. *Mathematics and Financial Economics*, **8**(3), 241–247.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, **55**(1), 95–115.

Zank, H. (2001). Cumulative prospect theory for parametric and multiattribute utilities. *Mathematics of Operations Research*, **26**(1), 67–81.