The Axiomatic Approach to Risk Measures for Capital Determination

Hans Föllmer
Humboldt-Universität zu Berlin

Stefan Weber
Leibniz Universität Hannover

January 5, 2015

1 Introduction

The quantification of the downside risk of a financial position is a key issue, both for regulators and financial institutions. It is also a challenge from a theoretical point of view. The theory of monetary risk measures, initiated by Artzner, Delbaen, Eber & Heath (1999), provides an axiomatic framework that has helped to clarify the issue, and whose structure has turned out to be surprisingly rich, with strong connections to many other areas in economics, statistics, and mathematics, and in particular to the theory of preferences in the face of risk and uncertainty. In this survey we describe some of the major developments.

The basic idea is to quantify the downside risk as a capital requirement. A financial position is described by its uncertain monetary outcome, that is, as a real-valued function X on some set of possible scenarios. Its risk $\rho(X)$ is defined as the monetary amount that needs to be added to the position to make it acceptable. For a given notion of acceptability, the resulting functional ρ on financial positions then has the properties of a *monetary risk measure* described in Section 2.

The standard example of a monetary risk measure is *Value at Risk*. In this case, a position is considered to be acceptable if the probability of a loss is below a given threshold. In particular it is assumed that this probability exists, and that it is accessible through the use of historical data or Monte Carlo simulation. While it continues to be the industry standard, Value at Risk has serious deficiencies. In particular, it may penalize diversification, and it does not capture the risk of very large losses that may hide behind the threshold. Other drawbacks such as procyclical effects and excessive reliance on specific probabilistic assumptions became apparent during the recent financial crisis; see, for example, in *The Turner Review – A regulatory response to the global banking crisis* (2009*a*). At a more fundamental level, the Turner Review emphasizes the issue of *Knightian uncertainty*, referring to situations where probabilities are not available, or do not even make sense.

The axiomatic approach to risk measures formulates potentially desirable properties of a monetary risk measure, investigates their consequences, and designs and analyzes specific examples. A key requirement is that diversification should not be penalized, in contrast to Value at Risk. This means that a convex combination of acceptable positions should again be acceptable. As a result, the monetary risk measure has the properties of a *convex risk measure* described in Section 2.2. Coherence is a stronger property, but less compelling from a financial point view: In addition to convexity, it requires that any multiple of an acceptable position should still be acceptable.

In this survey, the focus is on monetary risk measures that are convex. A convex risk measure will typically admit a dual representation. This involves a class $\mathcal{Q}$ of probability measures on the underlying space of possible scenarios and a penalty function that assigns different weights to these measures. The probability measures can be viewed as plausible probabilistic models, but they are taken more or less seriously, as specified by the penalty function. For a given position X, the capital requirement $\rho(X)$ is then computed as the worst case of the penalized expected loss over all probabilistic models in the class $\mathcal{Q}$; see Section 3. At this general level, no probabilistic model has to be fixed in advance. Thus, the general theory of monetary risk measures provides a conceptual framework that admits Knightian uncertainty. However, probability measures do enter the stage via convex duality, taking the role of stress tests.

The specific form of the dual representation depends on the space $\mathcal{X}$ of financial positions on which the risk measure ρ is defined and on additional regularity properties of ρ. Section 3 describes various standard settings. Some do not use an underlying probability measure P, others do, but only in order to fix a class of null-sets or a notion of integrability, without using the specific structure of the probabilistic model. The choice of the setting will play an important role in the discussion of m-convexity, elicitability, and robustness in Sections 7, 8, and 9.

In Section 4 we review standard examples of risk measures including Average Value at Risk, also called Expected Shortfall, divergence risk measures where the penalty function quantifies the divergence from some benchmark model, and utility-based shortfall risk where acceptability is defined in terms of expected utility. We describe their properties and provide their dual representation. In Section 5 we use families of convex risk measures to construct a general class of risk functionals that includes other indices of riskiness such as the "economic index of riskiness" proposed by Aumann & Serrano (2008). These functionals are quasi-convex, but they have no longer the translation property of a monetary risk measure.

Section 6 describes the connection between convex risk measures and variational preferences in the face of risk and uncertainty. If preferences satisfy the classical "axioms of rationality" formulated by von Neumann & Morgenstern (1944) and Savage (1954), then they admit a numerical representation in terms of the expected utility $E_P[u(X)]$. If these axioms are relaxed as proposed by Gilboa & Schmeidler (1989) or, more generally, by Maccheroni, Marinacci & Rustichini (2006), then their numerical representation is obtained by replacing the linear expectation $E_P[\cdot]$ by the non-linear functional $-\rho(\cdot)$, where ρ is either a coherent or a convex risk measure. While classical risk aversion corresponds to the concavity of u, the convexity of the risk measure ρ captures a different behavioral feature, described by an axiom of model risk aversion; cf. Maccheroni et al. (2006) and Föllmer, Schied & Weber (2009).

Most of the literature on risk measures assumes that the risk measure is *law-invariant*, or *distribution-based.* This means that positions X are described as random variables on some probability space $(\Omega, \mathcal{F}, P)$, and that the capital requirement $\rho(X)$ only depends on the distribution μ_X of X under P. As discussed in Section 7, the dual representation of a law-invariant convex risk measure can be restated in a more specific form, whose building blocks are provided by Average Value at Risk. As a special case we obtain the class of *spectral risk measures* proposed by Acerbi & Tasche (2002), defined as mixtures of Average Value at Risk. These can be described as Choquet integrals of the loss with respect to some concave distortion of the underlying probability measure P.

Clearly, a distribution-based convex risk measure can be regarded as a statistical functional on the class of distributions μ_X, also called *lotteries.* As such, they respect stochastic dominance both of the first and second kind. At the level of distributions, however, it is plausible to require an additional type of convexity, namely convexity of the level sets of the functional ρ on lotteries. Thus, the same capital is required for a compound lottery obtained by randomizing the choice between two lotteries μ and ν, if the capital requirement is the same for μ and ν. Distribution-based risk measures with this convexity property are characterized in Section 7: As shown by Weber (2006) and Delbaen, Bellini, Bignozzi & Ziegel (2014), they are utility-based, and in the coherent case they reduce to the class of *expectiles.*

From a statistical point of view, it is desirable to require additional properties of a distribution-based risk measure that are useful for backtesting and estimation procedures. In Section 8 we discuss distribution-based risk measures that are *M-estimators* in the sense of Huber (1981); these risk measues are also called *elicitable.* Gneiting (2011) argues that this property is crucial for the purposes of backtesting; see, however, Acerbi & Szekely (2014) and Davis (2013) for alternative points of view. Elicitable functionals also facilitate the application of regression techniques that generalize quantile regression, see e.g. Koenker (2005). As observed by Osband (1985), elicitable monetary functionals have convex level sets on lotteries. Applying the results of Weber (2006) and Delbaen et al. (2014) described in Section 7.5, one obtains a complete characterization of elicitable convex risk measures.

Another desirable property is *robustness*. In the classical formulation of Hampel (1971), robustness requires continuity of the functional with respect to the weak topology on lotteries, hence insensitivity with respect to tail behavior. In the context of robust statistics, it is well known that the median, or any other quantile, is robust in this sense, while the mean is not robust. Accordingly, Value at Risk is Hampel-robust, while convex distribution-based risk measures such as Average Value at Risk are not Hampel-robust; they are, and in fact are meant to be, sensitive with respect to tail behavior.

The perspective changes if the concept of robustness is modified. In Section 9 we describe a new approach to the robustness of risk measures developed by Krätschmer, Schied & Zähle (2014). It involves a restriction of the class of admissible lotteries as in Cont, Deguest & Scandolo (2010), and in addition a refinement of the weak topology that takes into account the tail behavior, for example by replacing the usual Prohorov metric with a suitable Wasserstein metric. As a result, the simple dichotomy "robust or not robust" is replaced by a continuum of degrees of robustness. In particular, Krätschmer, Schied & Zähle (2014) introduce an index of qualitative robustness that takes values in $[0, 1]$ if the distribution-based risk measure ρ is convex. We illustrate its computation for distortion risk measures and for elicitable risk measures. Its maximal value 1 is attained, for example, by Average Value at Risk and by expectiles.

The discussion in Sections 7, 8, and 9 shows that the structure of distribution-invariant risk measures is by now well understood, using different points of view. Obviously, there is no distribution-based risk measure that satisfies all potentially desirable properties. In particular, there is a clear trade-off between convexity and robustness under the assumption of elicitability. If one insists on robustness in the classical sense of Hampel (1971), then the only choice is to give up convexity and to go back to Value at Risk, as proposed by Kou & Peng (2014). If, on the other hand, one insists on convexity, then one can attain the largest possible degree of qualitative robustness in the sense of Krätschmer, Schied & Zähle (2014) by taking an expectile.

In any case, the scope of risk measures is not limited to distribution-invariance, as illustrated by the role of risk measures in the representation of variational preferences in the face of model uncertainty in Section 6. As suggested by the example of "stressed" Value at Risk in Section 2.3, forward-looking methods such as stress testing should be combined with all available information concerning the distributional properties of a financial position. To a large extent, this agenda is still open from a mathematical point of view.

Our presentation in the first sections closely follows Föllmer & Schied (2011), to which we refer for proofs and further details, and the survey Föllmer & Knispel (2012), which contains additional material on the close connection between monetary risk measures and actuarial premium principles.

2 Risk measures as capital requirements

A *financial position* will be described by its monetary outcome. Since the outcome is typically uncertain, it will be modeled as a real-valued measurable function X on some measurable space $(\Omega, \mathcal{F})$ of possible scenarios. The value $X(\omega)$ denotes the discounted net worth of the position at the end of a given trading period if the scenario $\omega \in \Omega$ is realized. The discounted net worth corresponds to the profits and losses of the position and is also called the P&L. Our aim is to quantify the downside risk of the position X as the additional capital $\rho(X)$ that is required to make the position acceptable from the point of view of a supervising agency.

2.1 Monetary risk measures

We fix a linear space $\mathcal{X}$ of financial positions and a subset $\mathcal{A} \subseteq \mathcal{X}$ of positions which are defined to be acceptable. Our focus will be on the downside risk, and thus we require that $Y \in \mathcal{A}$ whenever $Y \geq X$ for some $X \in \mathcal{A}$. We also assume that $\mathcal{X}$ contains the constants, and that a constant position is acceptable if and only if the constant is not smaller than some finite threshold. Then the functional ρ on $\mathcal{X}$ defined by

$$\rho(X) := \inf\{m \in \mathbb{R} | X + m \in \mathcal{A}\} \tag{1}$$

has the following properties of a monetary risk measure.

Definition 2.1. *A functional $\rho : \mathcal{X} \to (-\infty, \infty]$ with $\rho(0) < \infty$ is called a* monetary risk measure *if it is*

(i) monotone*, i. e., $\rho(X) \leq \rho(Y)$ if $X \geq Y$,*

and

(ii) cash-invariant*, i. e., $\rho(X + m) = \rho(X) - m$ for any constant $m \in \mathbb{R}$.*

A monetary risk measure will be called normalized *if $\rho(0) = 0$.*

Any monetary risk measure ρ can be represented in the form (1), using the acceptance set

$$\mathcal{A}_\rho := \{X \in \mathcal{X} | \rho(X) \leq 0\}. \tag{2}$$

It can thus be viewed as a *capital requirement*: $\rho(X)$ is the minimal capital that has to be added to the position X to make it acceptable. If ρ is defined by some acceptance set $\mathcal{A}$ via (1) then we have $\mathcal{A} \subseteq \mathcal{A}_\rho$, and equality holds if and only if $\mathcal{A}$ has the following closure property;

$$X + m \in \mathcal{A} \text{ for all } m > 0 \Longrightarrow X \in \mathcal{A}. \tag{3}$$

Remark 2.2. *We have assumed that all financial positions are already discounted by the risk-free interest rate r. If the risk of the undiscounted position $\tilde{X} := (1 + r)X$ is defined as $\tilde{\rho}(\tilde{X}) := \rho(X)$, then the resulting functional $\tilde{\rho}$ is still monotone. However, cash-invariance now takes the form*

$$\tilde{\rho}(\tilde{X} + (1 + r)m) = \tilde{\rho}(\tilde{X}) - m,$$

that is, adding m units of money to the portfolio at time 0 and investing it in a risk-free asset reduces the capital requirement by m.

Remark 2.3. *We could relax the restriction that the supporting capital can only be invested in a risk-free asset. Suppose that we allow investment in some larger class $\mathcal{V} \subseteq \mathcal{X}$ of "admissible" assets that contains the constant 0. This means that the initial acceptance set $\mathcal{A}$ is enlarged to the set $\mathcal{A}_\mathcal{V}$ of all positions X such that $X + V \geq A$ for some $V \in \mathcal{V}$ and some $A \in \mathcal{A}$; cf. Föllmer & Schied (2002), Föllmer & Schied (2011), Section 4.8, and Artzner, Delbaen & Koch-Medina (2009). Clearly, this enlargement will reduce the capital requirement, that is, the resulting risk measure $\rho_\mathcal{V}$ will satisfy $\rho_\mathcal{V}(X) \leq \rho(X)$ for any position X.*

2.2 Convex and coherent risk measures

In order to capture the idea that diversification should not increase the risk, it is natural to require *quasi-convexity* of the functional ρ, i. e.,

$$\rho(\lambda X + (1 - \lambda)Y) \leq \max\{\rho(X), \rho(Y)\} \tag{4}$$

for $X, Y \in \mathcal{X}$ and $\lambda \in (0, 1)$. In that case, the acceptance set $\mathcal{A}_\rho$ is convex, and this implies that ρ is in fact a convex functional on $\mathcal{X}$; see, e. g., Föllmer & Schied (2011), Proposition 4.6.

Definition 2.4. *A monetary risk measure is called a* convex risk measure *if it satisfies the condition (4) of quasi-convexity and is hence convex, i.e.,*

$$\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$$

for $X, Y \in \mathcal{X}$ and $\lambda \in (0,1)$. A convex risk measure is called coherent *if it is also positively homogeneous, i. e.,*

$$\rho(\lambda X) = \lambda\rho(X)$$

for $X \in \mathcal{X}$ and $\lambda \geq 0$.

Remark 2.5. *In their seminal paper, Artzner et al. (1999) focussed on the coherent case. The subsequent extension to the convex case was proposed independently by Frittelli & Rosazza Gianin (2002), Heath (2000), and Föllmer & Schied (2002).*

Any coherent risk measure ρ is *normalized* and *subadditive*, i. e.,

$$\rho(X+Y) \leq \rho(X) + \rho(Y)$$

for $X, Y \in \mathcal{X}$. More generally: If ρ is a monetary risk measure, then any two of the three properties of convexity, positive homogeneity, and subadditivity imply the remaining third; cf., e. g., Föllmer & Schied (2011), Exercise 4.1.3.

Note, however, that coherence of a monetary risk measure ρ is equivalent to the condition that the acceptance set $\mathcal{A}_\rho$ is not only convex but also a cone. In this case, an acceptable position X remains acceptable if it is multiplied by an arbitrarily large factor $\lambda > 0$. Obviously, this property of a coherent risk measure is questionable from a financial point of view.

In the sequel, our main focus will be on the general convex case. In addition to the preceding argument, this is motivated by the following remark.

Remark 2.6. *Consider the situation of Remark 2.3. Suppose that the initial risk measure ρ is coherent. If the admissible assets are subject to convex trading constraints such that the class $\mathcal{V}$ is convex but not a cone, then the resulting risk measure $\rho_{\mathcal{V}}$ will no longer be coherent, but it will be convex. This was one of the reasons in Föllmer & Schied (2002) to go beyond coherence and to introduce the general notion of a convex risk measure. A similar argument applies to liquidity-adjusted valuation and risk measures, see Remark 3.11 in Weber, Anderson, Hamm, Knispel, Liese & Salfeld (2013).*

Another reason is that we may want to define acceptability in terms of economic preferences, for example in terms of expected utility. Suppose that a position is defined to be acceptable if its expected utility does not fall below some given threshold. If the utility function is strictly concave, then the corresponding risk measure will be convex but not coherent. Such utility-based risk measures will be discussed in Section 4.2 below.

2.3 Value at Risk

The most commonly used risk measure in practice is *Value at Risk* at some level $\lambda \in (0,1)$. Its definition requires a probability measure P on the underlying space of scenarios to which we have sufficient access through past observations. A position X is now defined to be acceptable if the probability $P[X < 0]$ of a shortfall does not exceed the level λ. The resulting capital requirement is given by

$$\begin{aligned} \mathrm{VaR}_\lambda(X) &= \inf\{m \in \mathbb{R} \,|\, P[X+m<0] \leq \lambda\} \\ &= -\sup\{c \in \mathbb{R} \,|\, P[X<c] \leq \lambda\} = -q_X^+(\lambda), \end{aligned} \tag{5}$$

where $q_X^+(\lambda)$ is the upper λ-quantile of the random variable X on the probability space $(\Omega, \mathcal{F}, P)$.

Clearly, VaR_λ is a monetary risk measure on the space $\mathcal{X} = L^0(\Omega, \mathcal{F}, P)$ of all random variables on $(\Omega, \mathcal{F}, P)$ that are finite P-almost surely. It is also positively homogeneous. If X is Gaussian with variance $\sigma_P^2(X)$, then we have

$$\mathrm{VaR}_\lambda(X) = E_P[-X] + \Phi^{-1}(1-\lambda)\sigma_P(X),$$

where Φ^{-1} denotes the inverse of the distribution function Φ of the standard normal distribution. Since σ_P^2 is a convex functional on $L^2(\Omega, \mathcal{F}, P)$, VaR_λ can be viewed as a convex risk measure on any Gaussian subspace of $L^2(\Omega, \mathcal{F}, P)$ if $\lambda \leq 0.5$. This remark extends to elliptical distributions, see Embrechts, McNeil & Straumann (2002).

However, VaR_λ is not convex on the full space $L^0(\Omega, \mathcal{F}, P)$. Indeed, take two events A_i $(i = 1, 2)$ such that $P[A_i] \leq \lambda$ but $P[A_1 \cup A_2] > \lambda$. Then the digital positions $X_i = -I_{A_i}$ are both acceptable, but the sum $X = X_1 + X_2$ is not. This shows that VaR_λ is not subadditive, hence not convex. Note also that VaR_λ does not pay attention to extreme losses that occur with small probability. As a result, considerable risks can be "hidden" behind the threshold defined by Value at Risk.

These deficiencies were recognized early on, and they have motivated the axiomatic approach to a general theory of monetary risk measures, which was initiated by Artzner et al. (1999) in the late nineties, and which is the topic of this survey. Other drawbacks became apparent during the recent financial crisis; see the analysis of Value at Risk and its procyclical effects in *The Turner Review – A regulatory response to the global banking crisis* (2009*a*). In particular, the Turner Review points to an excessive reliance on a single probabilistic model P derived from past observations. This point is adressed in the *Revisions to the Basel II market risk framework* (2009*b*) by considering a "Stressed Value at Risk", which involves alternative models $\tilde{P}$ derived from observations and simulations of periods of significant financial stress.

At a more fundamental level, the Turner Review raises the issue of model uncertainty or model ambiguity, often called *Knightian uncertainty.* As illustrated by Sections 3.1 and 3.2, and in particular by the discussion of preferences in the face of risk and uncertainty in Section 6, the theory of risk measures provides a conceptual framework which does not require an underlying probabilistic model, and thus allows one to deal with Knightian uncertainty in mathematical terms.

3 Dual representation of convex risk measures

A convex risk measure ρ typically admits a dual representation of the form

$$\rho(X) = \sup_Q \{E_Q[-X] - \alpha(Q)\}, \tag{6}$$

where the supremum is taken over probability measures Q on $(\Omega, \mathcal{F})$, and where α is a penalty function with values in $[-\ rho(0), \infty]$ defined by

$$\alpha(Q) := \sup_{X \in \mathcal{A}_\rho} E_Q[-X]. \tag{7}$$

The reason is as follows. The space $\mathcal{X}$ will usually be a Banach space, as in the standard settings discussed below, and the risk measure ρ will have additional regularity properties, including lower semi-continuity with respect to the weak topology on $\mathcal{X}$. In such a situation, the Fenchel-Moreau theorem implies that the convex functional ρ admits a dual representation in terms of its Fenchel-Legendre transform, defined on the dual space $\mathcal{X}'$ of continuous linear functionals on $\mathcal{X}$. Moreover, the monetary properties of ρ imply that the relevant linear functionals can be identified as expectations with respect to some probability measure Q, and this yields the specific representation (6). This argument is exemplified below for some standard choices of the space $\mathcal{X}$; for a systematic discussion in a general framework and further references we refer to Frittelli & Rosazza Gianin (2002) and Biagini & Frittelli (2009).

Note that it is enough to take the supremum in (6) over the class $\mathcal{Q}_\rho := \{Q \mid \alpha(Q) < \infty\}$. The probability measures in $\mathcal{Q}_\rho$ can be seen as plausible probabilistic models which are taken more or less seriously, as specified by the penalty function α. The capital requirement $\rho(X)$ is thus computed as the worst case of the penalized expected loss $E_Q[-X]$ over all plausible models $Q \in \mathcal{Q}_\rho$. Note that the general definition of a convex risk measure does not a priori involve the choice of a probabilistic model. But probabilistic models Q come into play via the dual representation (6).

In the *coherent case* the acceptance set $\mathcal{A}_\rho$ is a convex cone, and this implies $\alpha(Q) \in \{0, \infty\}$. A coherent risk measure ρ with dual representation (6) thus takes the form

$$\rho(X) = \sup_{Q \in \mathcal{Q}_\rho} E_Q[-X]$$

with $\mathcal{Q}_\rho = \{Q | \alpha(Q) = 0\}$; cf. Artzner et al. (1999), Delbaen (2000), and Föllmer & Schied (2011), Corollaries 4.19 and 4.37.

We are now going to illustrate our general setting by some standard choices of the space $\mathcal{X}$. In the first two cases, no probability measure is given in advance. In the remaining three cases, we will fix a probability measure P on the space $(\Omega, \mathcal{F})$ of possible scenarios. In the third case only the null-sets of P will matter; in the last two cases P will also determine the notion of integrability required in the definition of $\mathcal{X}$.

3.1 $\mathcal{X} = \mathscr{L}^\infty(\Omega, \mathcal{F})$

Let us denote by $\mathscr{L}^\infty(\Omega, \mathcal{F})$ the Banach space of all bounded measurable functions on $(\Omega, \mathcal{F})$, by $\mathcal{M}_1$ the class of all probability measures on $(\Omega, \mathcal{F})$, and by $\mathcal{M}_{1,f}$ the class of all finitely additive set functions $Q : \mathcal{F} \to [0, 1]$ with $Q[\Omega] = 1$.

A convex risk measure ρ on $\mathcal{X} = \mathscr{L}^\infty(\Omega, \mathcal{F})$ is Lipschitz continuous with respect to the supremum norm, hence weakly lower-semicontinuous on the Banach space $\mathcal{X}$. Applying the Fenchel-Moreau theorem, combined with the monetary properties of ρ and the weak compactness of $\mathcal{M}_{1,f}$, we obtain the representation

$$\rho(X) = \max_{Q \in \mathcal{M}_{1,f}} \{E_Q[-X] - \alpha(Q)\}, \tag{8}$$

where $\alpha(Q)$ is defined as in (7) for any $Q \in \mathcal{M}_{1,f}$, cf. Föllmer & Schied (2011), Theorem 4.16.

The representation (8) reduces to the dual representation (6) in terms of σ-additive probability measures, with max instead of sup, if $\alpha(Q) = \infty$ for any $Q \in \mathcal{M}_{1,f}$ that is not σ-additive. This condition is satisfied if and only if ρ is *continuous from below* in the sense that

$$X_n \nearrow X \quad \text{pointwise on } \Omega \quad \Longrightarrow \quad \rho(X_n) \searrow \rho(X).$$

Moreover, continuity from below is equivalent to the *Lebesgue property*, that is,

$$X_n \to X \quad \text{pointwise on } \Omega \quad \Longrightarrow \quad \rho(X) = \lim_n \rho(X_n) \tag{9}$$

whenever the sequence $(X_n) \subseteq \mathcal{X}$ is uniformly bounded; cf., e. g., Föllmer & Schied (2011), Theorem 4.22 and Exercise 4.2.2.

If the equality on the right hand side of equation (9) is replaced by the inequality $\rho(X) \le \liminf_n \rho(X_n)$, then ρ is said to have the *Fatou property*, and this is equivalent to *continuity from above*. The Fatou property is clearly necessary for a dual representation (6) in terms of probability measures $Q \in \mathcal{M}_1$, but it is in general not sufficient.

3.2 $\mathcal{X} = C_b(\Omega)$

Assume that Ω is a separable metric space and that $\mathcal{F}$ is the corresponding σ-field of Borel sets. Consider a convex risk measure ρ on $\mathscr{L}^\infty(\Omega, \mathcal{F})$, assuming that ρ is *tight* in the sense that there exists an increasing sequence of compact sets $K_n \subseteq \Omega$ such that

$$\lim_{n \uparrow \infty} \rho(\lambda 1_{K_n}) = \rho(\lambda) \quad \text{for any } \lambda \ge 1.$$

The restriction of the convex risk measure ρ to the space $\mathcal{X} = C_b(\Omega)$ of bounded continuous functions on Ω possesses a robust representation of the form

$$\rho(X) = \max_{Q \in \mathcal{M}_1} \{E_Q[-X] - \alpha(Q)\} \quad \text{for any } X \in C_b(\Omega),$$

where

$$\alpha(Q) := \inf\{\alpha(\tilde{Q}) | \tilde{Q} \in \mathcal{M}_{1,f}, E_{\tilde{Q}}[\cdot] = E_Q[\cdot] \text{ on } C_b(\Omega)\}.$$

Moreover, if Ω is a Polish space, then the level sets $\{Q \in \mathcal{M}_1 | \alpha(Q) \leq c\}$ are relatively compact for the weak topology on $\mathcal{M}_1$; cf. Föllmer & Schied (2011), Propositions 4.27 and 4.30.

For the following cases we fix a probability measure P on $(\Omega, \mathcal{F})$ and denote by $\mathcal{M}_1(P)$ the class of all probability measures $Q \in \mathcal{M}_1$ which are absolutely continuous with respect to P.

3.3 $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, P)$

Let ρ be a convex risk measure on $\mathscr{L}^\infty(\Omega, \mathcal{F})$ that respects the null sets of P, i. e., $\rho(X) = \rho(Y)$ whenever the equivalence relation $X = Y$ P-a. s. holds. Then ρ can be regarded as a convex risk measure on the Banach space $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, P)$ of equivalence classes. In this case, the dual representation

$$\rho(X) = \sup_{Q \in \mathcal{M}_1(P)} \{E_Q[-X] - \alpha(Q)\} \tag{10}$$

holds if and only if ρ is *continuous from above*, i. e.,

$$\rho(X_n) \nearrow \rho(X) \quad \text{whenever } X_n \searrow X \ P\text{-a. s.}$$

or, equivalently, iff ρ has the *Fatou property*

$$\rho(X) \leq \liminf_{n \uparrow \infty} \rho(X_n)$$

for any bounded sequence $(X_n)_{n \in \mathbb{N}}$ in $L^\infty(\Omega, \mathcal{F}, P)$ which converges P-a. s. to X. Moreover, the supremum in (10) is attained for each $X \in L^\infty(\Omega, \mathcal{F}, P)$ iff ρ is continuous from below; cf. Delbaen (2002) and Föllmer & Schied (2011), Theorem 4.33 and Corollary 4.35.

3.4 $\mathcal{X} = L^p(\Omega, \mathcal{F}, P)$

In many applications, it is convenient to deal with unbounded instead of bounded random variables. This requires an extension of the theory of risk measures to spaces larger than $L^\infty(\Omega, \mathcal{F}, P)$. Canonical choices are the Banach spaces $L^p(\Omega, \mathcal{F}, P)$ with $1 \leq p < \infty$, cf. Scandolo (2003) and Kaina & Rüschendorf (2009). For a convex risk measure ρ on $\mathcal{X} = L^p(\Omega, \mathcal{F}, P)$ the Fatou property is equivalent to lower-semicontinuity of ρ with respect to the L^p-norm. In this case, the dual representation takes the form

$$\rho(X) = \sup_{Q \in \mathcal{M}^q(P)} \{E_Q[-X] - \alpha(Q)\}, \quad X \in L^p(\Omega, \mathcal{F}, P), \tag{11}$$

with dual exponent $q := p/(p-1) \in (1, \infty]$; here we use the notation

$$\mathcal{M}^q(P) := \{Q \in \mathcal{M}_1(P) | \tfrac{dQ}{dP} \in L^q(\Omega, \mathcal{F}, P)\}.$$

Moreover, if the convex risk measure ρ is *finite* on $L^p(\Omega, \mathcal{F}, P)$, then it is even Lipschitz continuous with respect to the L^p-norm, and the representation (11) holds with max instead of sup. For a systematic discussion of risk measures on $L^p(\Omega, \mathcal{F}, P)$ we refer to Kaina & Rüschendorf (2009).

3.5 Risk measures on Orlicz hearts

The L^p-spaces discussed in the previous section can be embedded into the general framework of Orlicz spaces and Orlicz hearts. The study of convex risk measures on such spaces was initiated by Cheridito & Li (2009). As shown by Krätschmer, Schied & Zähle (2014), this approach is particularly useful if we want to study statistical properties of risk measures such as robustness. We will discuss this issue in Section 9.

Let $\Psi : [0,\infty) \to [0,\infty]$ be a Young function, i. e., a left-continuous, non-decreasing convex function such that $\lim_{x\downarrow 0}\Psi(x) = 0$ and $\lim_{x\uparrow\infty}\Psi(x) = \infty$. The Orlicz space

$$L^\Psi := L^\Psi(\Omega,\mathcal{F},P) = \{X \in L^0 : E[\Psi(c|X|)] < \infty \text{ for some } c > 0\}$$

is a Banach space if endowed with the Luxemburg norm

$$\|X\|_\Psi := \inf\{a > 0 : E[\Psi(|X|/a)] \le 1\}.$$

The Orlicz heart

$$H^\Psi := H^\Psi(\Omega,\mathcal{F},P) = \{X \in L^0 : E[\Psi(c|X|)] < \infty \text{ for all } c > 0\}$$

is a linear subspace of L^Ψ that reduces to the trivial space $\{0\}$ if Ψ attains the value ∞. From now on we assume that Ψ is a real-valued and hence continuous. In this case, we have

$$L^\infty \subseteq H^\Psi \subseteq L^\Psi \subseteq L^1,$$

and the Orlicz heart H^Ψ is a Banach space, namely the closure of L^∞ in L^Ψ with respect to the Luxemburg norm $\|\cdot\|_\Psi$.

The equality $H^\Psi = L^\Psi$ holds if and only if the Young function Ψ satisfies

$$\Psi(2x) \le C\Psi(x) \tag{12}$$

for some constant C and for large enough x. This condition, often called the "Δ_2-condition", will play a crucial role in our discussion of robustness properties of a risk measure in Section 9. It is clearly satisfied by the Young functions $\Psi(x) = x^p/p$ for $p \in [1,\infty)$, and so we have $H^\Psi = L^\Psi = L^p$ in the classical situation of L^p-spaces.

In contrast, the exponential Young function $\Psi(x) = e^x - 1$ does not satify the Δ_2-condition. Here the Orlicz heart

$$H^\Psi = \left\{X \in L^1 : E_P[e^{c|X|}] < \infty \text{ for all } c > 0\right\} \tag{13}$$

is strictly contained in the Orlicz space

$$L^\Psi = \left\{X \in L^1 : E_P[e^{c|X|}] < \infty \text{ for some } c > 0\right\}.$$

The Orlicz heart H^Ψ defined in (13) is the natural domain of the entropic risk measures discussed below in Section 4.4 and Section 9: They are finite on the Orlicz heart but not on the Orlicz space.

For a finite Young function Ψ, its conjugate function $\Psi^*(y) := \sup_{x\ge 0}\{xy - \Psi(x)\}$, $y \ge 0$, is again a Young function, and the conjugate of Ψ^* is given by Ψ.

We can now state the main representation result for convex risk measures which are finite on the Orlicz heart H^Ψ; cf. Cheridito & Li (2009), Theorem 4.3.

Proposition 3.1. *Suppose that ρ is a finite convex risk measure on the Orlicz heart H^Ψ. Then ρ admits the dual representation*

$$\rho(X) = \max_{Q\in\mathcal{M}^{\Psi^*}(P)}\{E_Q[-X] - \alpha(Q)\}, \quad X \in H^\Psi, \tag{14}$$

where

$$\mathcal{M}^{\Psi^*}(P) := \{Q \in \mathcal{M}_1(P) \mid \tfrac{dQ}{dP} \in L^{\Psi^*}(P)\}.$$

Clearly, this result contains the representation (14) of finite convex risk measures on L^p-spaces as a special case.

4 Examples

In this section we fix a probability measure P on $(\Omega, \mathcal{F})$, and we assume that the probability space $(\Omega, \mathcal{F}, P)$ is atomless. The following risk measures are all *distribution-based* in the sense that the capital requirement $\rho(X)$ only depends on the distribution of X, viewed as a random variable on $(\Omega, \mathcal{F}, P)$. They will first be considered on the space $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, P)$, but they will all admit a canonical extension to the larger space $L^1(\Omega, \mathcal{F}, P)$; see Theorem 7.3 below.

4.1 Average Value at Risk

We have seen that Value at Risk at some fixed level α is positively homogeneous but not convex, and thus it is not a coherent risk measure. However, if we take the average

$$\mathrm{AVaR}_\lambda(X) := \tfrac{1}{\lambda} \int_0^\lambda \mathrm{VaR}_\alpha(X)\, d\alpha \tag{15}$$

up to some level $\lambda \in (0,1]$, then we obtain the basic example of a coherent risk measure, known as *Average Value at Risk*, *Conditional Value at Risk*, *Tail Value at Risk*, or *Expected Shortfall*. Since VaR_α is decreasing in α, we have

$$\mathrm{AVaR}_\lambda(X) \geq \mathrm{VaR}_\lambda(X), \tag{16}$$

that is, AVaR_λ prescribes higher capital requirements than VaR_λ. In fact, AVaR_λ can be characterized as the smallest distribution-based convex risk measure that dominates VaR_λ.

For any $\lambda \in (0,1]$, AVaR_λ is a coherent risk measure whose dual representation takes the form

$$\mathrm{AVaR}_\lambda(X) = \max_{Q \in \mathcal{Q}_\lambda} E_Q[-X] \tag{17}$$

with

$$\mathcal{Q}_\lambda := \{Q \ll P \mid \tfrac{dQ}{dP} \leq \tfrac{1}{\lambda}\};$$

cf., e. g., Föllmer & Schied (2011), Theorem 4.52. Average Value at Risk can also be written as

$$\mathrm{AVaR}_\lambda(X) = \tfrac{1}{\lambda} E_P[(q_X(\lambda) - X)^+] - q_X(\lambda) = \tfrac{1}{\lambda} \inf_{z \in \mathbb{R}} \{E_P[(z - X)^+] - \lambda z\} \tag{18}$$

for any λ-quantile $q_X(\lambda)$ of X, cf. Föllmer & Schied (2011), Lemma 4.51.

All these representations yield a natural extension of Average Value at Risk from $L^\infty(\Omega, \mathcal{F}, P)$ to the space $L^1(\Omega, \mathcal{F}, P)$. If X has a continuous distribution, then we have $q_X(\lambda) = \mathrm{VaR}_\lambda(X)$ and $\lambda = P[X \leq q_X(\lambda)]$. In this case, the first equation in (18) reduces to

$$\mathrm{AVaR}_\lambda(X) = E_P[-X \mid -X > \mathrm{VaR}_\lambda(X)], \tag{19}$$

that is, $\mathrm{AVaR}_\lambda(X)$ can be described as the conditional expectation of the loss $-X$, given that the loss exceeds the level $\mathrm{VaR}_\lambda(X)$.The second representation in (18) is useful in the context of the computation of efficient risk-return-frontiers, see Uryasev & Rockafellar (2001).

Average Value at Risk plays a prominent role in the *Swiss Solvency Test*, see e.g. Filipović & Vogelpoth (2008). From a theoretical point of view, it provides the main building block for general distribution-based convex risk measures, as will be explained in Section 7.

4.2 Utility-based risk measures

For a given convex and increasing *loss function* $\ell : \mathbb{R} \to \mathbb{R}$ we define the *shortfall risk* of a position X as

$$R(X) = E_P[\ell(-X)].$$

The functional R is convex and monotone on $\mathcal{X}$, and it is thus a *risk functional* in the sense of Section 5 below. However, R is not cash-invariant, and hence it is not a convex risk measure.

Remark 4.1. *In the special case $\ell(x) = x^+$ we have $\ell(-X) = X^- = -\min(X, 0)$, and so we recover the classical actuarial definition*

$$R(X) = E_P[X^-] \tag{20}$$

of mean risk ("mittleres Risiko")*, as it was introduced in 1868 by K. F. W. Hattendorf; cf. Hattendorff (1868). Interpreted as the price of a put option, the same functional is proposed in Jarrow (2002), where it is called the* Put Option Premium.

Let us fix a threshold l_0 in the interior of the range of ℓ, and let us define the acceptance set

$$\mathcal{A} := \{X \in \mathcal{X} | R(X) = E_P[\ell(-X)] \leq l_0\}.$$

Then $\mathcal{A}$ is convex, and the resulting capital requirement according to (1) is a convex risk measure that will be denoted by ρ^ℓ; it was introduced in Föllmer & Schied (2002). For a given position X, the amount $m = \rho^\ell(X)$ is a solution to the equation

$$R(X + m) = E_P[\ell(-X - m)] = l_0, \tag{21}$$

and the solution is unique if ℓ is strictly increasing. Hence, one can use stochastic root finding techniques for its numerical computation; see Dunkel & Weber (2010).

Note that the acceptance set can be rewritten as

$$\mathcal{A} = \{X \in \mathcal{X} | E_P[u(X)] \geq u_0\}$$

in terms of the utility function $u(x) := -\ell(-x)$, with $u_0 = -l_0$. For this reason, the convex risk measure ρ^ℓ is called a *utility-based risk measure*, or *utility-based shortfall risk*.

Proposition 4.2. *A utility-based risk measure ρ^ℓ admits a dual representation (10) with minimal penalty function*

$$\alpha(Q) = \inf_{\lambda > 0} \tfrac{1}{\lambda}(l_0 + E_P[\ell^*(\lambda \tfrac{dQ}{dP})]), \tag{22}$$

where ℓ^* denotes the Fenchel-Legendre transform of ℓ.

For the proof see Föllmer & Schied (2011), Theorem 4.115. In the special case of a power loss function $\ell(x) = \frac{1}{p} x^p 1_{(0,\infty)}(x)$ with $p \geq 1$, the penalty function is given by

$$\alpha(Q) = (p\, l_0)^{\frac{1}{p}} \left\| \tfrac{dQ}{dP} \right\|_q$$

in terms of the dual exponent $q = p/(p-1)$; cf. Föllmer & Schied (2011), Example 4.118.

As we shall see in Section 8 below, utility-based risk measures play a key role as soon as we take a statistical point of view and require a risk measure to be "elicitable".

4.3 Divergence risk measures and optimized certainty equivalents

For a lower semicontinuous convex function $g : \mathbb{R}_+ \to \mathbb{R} \cup \{\infty\}$ with $g(1) < \infty$ and superlinear growth $g(x)/x \to \infty$ as $x \uparrow \infty$, the corresponding *g-divergence* of a probability measure $Q \in \mathcal{M}_1$ with respect to P is defined by

$$D_g(Q|P) := E_P\left[g\left(\tfrac{dQ}{dP}\right)\right]$$

if $Q \ll P$ and by $D_g(Q|P) = \infty$ otherwise. By Jensen's inequality we have $D_g(Q|P) \geq D_g(P|P) = g(1)$.

If P is viewed as a benchmark model, then one can choose as penalty function a g-divergence with respect to P. The resulting convex risk measure ρ_g defined by

$$\rho_g(X) = \sup_{Q \in \mathcal{M}_1(P)} \{E_Q[-X] - D_g(Q|P)\}$$

is called a *divergence risk measure*. Denoting by $g^*(y) := \sup_{x>0}\{xy - g(x)\}$ the convex conjugate function of g, the risk measure ρ_g can also be represented by the variational identity

$$\rho_g(X) = \inf_{y\in\mathbb{R}}\{E_P[g^*(y-X)] - y\}; \tag{23}$$

cf., e. g., Föllmer & Schied (2011), Theorem 4.122.

Divergence risk measures arise naturally in a number of cases:

- For $\lambda \in (0,1]$ and the function g defined by $g(x) = 0$ for $x \le \lambda^{-1}$ and $g(x) = \infty$ otherwise, the divergence risk measure ρ_g coincides with Average Value at Risk at level λ. Here we have $g^*(y) = \frac{1}{\lambda}y\,1_{(0,\infty)}(y)$, and hence the variational identity (23) coincides with formula (18).
- For $g(x) = x\log x$, the g-divergence reduces to the relative entropy $H(Q|P)$ of Q with respect to P, and then ρ_g is also called an *entropic risk measure*; see Section 4.4.
- Inserting the penalty function (22) into the dual representation (10), the utility-based risk measure ρ^ℓ with loss function ℓ and threshold l_0 can be rewritten as

$$\begin{aligned}\rho^\ell(X) &= \sup_{Q\in\mathcal{M}_1(P)}\{E_Q[-X] - \inf_{\lambda>0}\{l_0 + E_P[\ell^*(\lambda\tfrac{dQ}{dP})]\}\}\\ &= \sup_{\lambda>0}\sup_{Q\in\mathcal{M}_1(P)}\{E_Q[-X] - E_P[g_\lambda(\tfrac{dQ}{dP})]\}\end{aligned}$$

  in terms of the convex functions $g_\lambda(y) := \frac{1}{\lambda}(\ell^*(\lambda y) + l_0)$. It can thus be described as the supremum of certain divergence risk measures, namely

$$\rho^\ell(X) = \sup_{\lambda>0}\rho_{g_\lambda}(X).$$

- For a utility function $u : \mathbb{R} \to \mathbb{R}\cup\{-\infty\}$ with $u(0) = 0$, Ben-Tal & Teboulle (1987) and Ben-Tal & Teboulle (2007) introduced the *optimized certainty equivalent* of a financial position $X \in \mathcal{X}$, defined as

$$S_u(X) := \sup_{\eta\in\mathbb{R}}\{\eta + E_P[u(X-\eta)]\}.$$

  This can be interpreted as the present value of an optimal split of the uncertain future income X into a certain amount η that is made available right now and an uncertain future amount $X-\eta$. Denote by $g(z) := \sup_{x\in\mathbb{R}}\{xz - \ell(x)\}$ the convex conjugate function of the loss function ℓ associated to u via $\ell(x) = -u(-x)$. Then the variational identity (23) yields

$$S_u(X) = -\rho_g(X).$$

  Thus, the optimized certainty equivalent coincides, up to a change of sign, with the divergence risk measure ρ_g.

Numerical procedures for the computation of divergence risk measures are discussed in Hamm, Salfeld & Weber (2013) and Drapeau, Kupper & Papapantoleon (2014).

4.4 Entropic risk measures

For any probability measure Q on $(\Omega,\mathcal{F})$, the *relative entropy* of Q with respect to P is defined as

$$H(Q|P) := \begin{cases} E_Q\left[\log\frac{dQ}{dP}\right] & \text{if } Q \ll P,\\ +\infty & \text{otherwise.}\end{cases}$$

Note that the relative entropy can be interpreted as a g-divergence for the function $g(x) = x\log x$.

Definition 4.3. *For any constant $\gamma > 0$, the convex risk measure e_γ defined by*

$$e_\gamma(X) := \sup_{Q \in \mathcal{M}_1} \{E_Q[-X] - \tfrac{1}{\gamma} H(Q|P)\}$$

is called the entropic risk measure *with parameter γ.*

Using the well-known variational principle

$$H(Q|P) = \sup_{X \in L^\infty(\Omega,\mathcal{F},P)} \{E_Q[-X] - \log E_P[e^{-X}]\}$$

for the relative entropy, it follows that e_γ takes the explicit form

$$e_\gamma(X) = \tfrac{1}{\gamma} \log E_P[e^{-\gamma X}]; \tag{24}$$

cf., e. g., Föllmer & Schied (2011), Example 4.34. Clearly, e_γ is well-defined as a finite convex risk measure on the Orlicz heart (13) with respect to the function $\Psi(x) = e^x - 1$. It is easy to see that $e_\gamma(X)$ is increasing in γ, with limits

$$e_0(X) := \lim_{\gamma \downarrow 0} e_\gamma(X) = E_P[-X] \quad \text{and} \quad e_\infty(X) := \lim_{\gamma \uparrow \infty} e_\gamma(X) = \operatorname{ess\,sup}(-X). \tag{25}$$

Thus, the limit e_0 is the *risk-neutral capital requirement*, and the risk measure e_∞ is the *worst case risk measure*, defined by zero tolerance for losses, that is, by the acceptance set $\mathcal{A} = \{X \in \mathcal{X} \mid P[X < 0] = 0\}$.

In view of (24), the acceptance set (2) of e_γ for $\gamma \in (0, 1)$ can be written as

$$\mathcal{A} = \{X | E_P[\ell_\gamma(-X)] \leq 1\} = \{X | E_P[u_\gamma(X)] \geq 0\}$$

in terms of the loss function $\ell_\gamma(x) = e^{\gamma x}$ or the utility function $u_\gamma(x) = -e^{-\gamma x}$. This shows that e_γ is a utility-based risk measure with respect to exponential utility.

Formula (24) also shows that the functional $-e_\gamma$ is a *certainty equivalent* with respect to the exponential utility function u_γ, that is,

$$u_\gamma(-e_\gamma(X)) = E_P[u_\gamma(X)].$$

The entropic risk measures can actually be characterized by this property: If ρ is a monetary risk measure and $-\rho$ is a certainty equivalent with respect to some strictly increasing continuous function u, then ρ must be a entropic risk measure e_γ for some parameter $\gamma \geq 0$, including the linear case $\gamma = 0$. This follows from a classical result of Bruno de Finetti; cf. De Finetti (1931) or, e. g., Example 4.13 in Föllmer & Schied (2011).

For any $\gamma \in [0, \infty]$, the risk measure e_γ is *additive on independent positions*, i. e.,

$$e_\gamma(X + Y) = e_\gamma(X) + e_\gamma(Y) \tag{26}$$

if X and Y are independent under P. Clearly, this additivity is preserved for any risk measure of the form

$$\rho(X) = \int_{[0,\infty]} e_\gamma(X)\, \nu(d\gamma)$$

with some probability measure ν on $[0, \infty]$. For an axiomatic characterization of such mixtures we refer to Goovaerts, Kaas, Laeven & Tang (2004) and Goovaerts & Laeven (2008); cf. also Gerber & Goovaerts (1981), where an analogous problem is discussed in the context of actuarial premium principles.

5 Indices of riskiness

In this survey we focus on convex risk measures and on their interpretation as capital requirements. Let us now show how they are connected to more general notions of the downside risk which are obtained by dropping the monetary condition of cash-invariance in Definition 2.1.

Definition 5.1. *A functional $R : \mathcal{X} \to (-\infty, \infty]$ such that $R \not\equiv \infty$ and $\lim_{m\downarrow-\infty} R(m) = \infty$ will be called a* quasi-convex risk functional *if it is monotone as in Definition 2.1 and quasi-convex as in (4).*

Without the condition of cash-invariance we can no longer conclude that the functional R is convex. However, R can be identified with a *family of convex risk measures* in the following manner. For each $r > \underline{r} := \inf R$ the level set

$$\mathcal{A}_r := \{X \in \mathcal{X} \mid R(X) \leq r\}$$

is a convex acceptance set, and thus it defines via (1) a convex risk measure ρ_r given by

$$\rho_r(X) = \inf\{m \mid X + m \in \mathcal{A}_r\}. \tag{27}$$

The acceptance sets $(\mathcal{A}_r)$ are clearly increasing in r, the risk measures (ρ_r) are decreasing, and R can be reconstructed from $(\mathcal{A}_r)$ or from (ρ_r) via

$$R(X) = \inf\{r > \underline{r} \mid X \in \mathcal{A}_r\} = \inf\{r > \underline{r} \mid \rho_r(X) \leq 0\}. \tag{28}$$

Thus any quasi-convex risk functional R corresponds to a family (ρ_r) of convex risk measures.

Typically, each ρ_r will admit a dual representation (6) with minimal penalty function $\alpha(\cdot, r)$. In this case we have $\rho_r(X) \leq 0$ iff $E_Q[-X] \leq \alpha(Q, r)$ for any $Q \in \mathcal{M}$. Thus (28) yields the following representation of the risk functional R:

$$R(X) = \sup_{Q\in\mathcal{M}_1} r(Q, E_Q[-X]), \tag{29}$$

where

$$s \mapsto r(Q, s) := \inf\{r \in \mathbb{R} \mid s \leq \alpha(Q, r)\}$$

denotes the left inverse of the increasing function $\alpha(Q, \cdot)$. We refer to Drapeau (2010), Brown, De Giorgi & Sim (2010), and Drapeau & Kupper (2013) for a systematic discussion and a wide variety of case studies.

Any convex risk measure is clearly a quasi-convex risk functional in the sense of Definition 5.1. We have already seen one example of a risk functional that is not a monetary risk measure, namely the shortfall risk $R(X) = E_P[\ell(-X)]$ discussed in Section 4.2; in this case the functional is actually convex.

Remark 5.2. *More generally, we could replace the expectation $E_P[\cdot]$ in the definition of shortfall risk by $\rho(-\cdot)$, where ρ is a normalized convex risk measure. In terms of the utility function $u(x) = -\ell(-x)$, the resulting risk functionals take the form*

$$R(X) = \rho(u(X));$$

the corresponding utility functionals $U = -R$ will be characterized in Section 6. Let us now apply this extension to the loss function $\ell(x) = x^+$ that appears in Hattendorf's classical definition (20) *of* mean risk*; see Remark 4.1. The resulting risk functionals*

$$R(X) = \rho(\min(X, 0))$$

are convex and monotone, they satisfy the condition $\rho(-m) = m$ for $m \geq 0$, and they only depend on the loss $\min(X, 0)$*. This class of* loss-based risk measures *was proposed and analyzed by Cont, Deguest & He (2013).*

Let us now introduce a different class of examples that includes the "economic index of riskiness" R^{AS} proposed by Aumann & Serrano (2008) and the "operational measure of riskiness" R^{FH} proposed by Foster & Hart (2009).

We begin by fixing a convex acceptance set $\mathcal{A} \subseteq \mathcal{X}$ in the general setting of Section 2. We assume that $\mathcal{A}$ has the closure property (3) and contains 0. Instead of asking how much capital is needed to make a given position X acceptable, we now focus on the *maximal acceptable exposure*, defined as

$$\lambda(X) := \sup\{\lambda \geq 0 | \lambda X \in \mathcal{A}\} \in [0, \infty].$$

Definition 5.3. *The functional $R : \mathcal{X} \to [0, \infty]$ defined by*

$$R(X) = \lambda(X)^{-1}$$

is called the index of riskiness *corresponding to the acceptance set $\mathcal{A}$.*

Proposition 5.4. *The index of riskiness R is a quasi-convex risk functional in the sense of Definition 5.1, and it is positively homogeneous.*

Indeed, the functional R is clearly monotone and positively homogeneous, and it is easy to check that its level sets $\{R(\cdot) \leq r\}$ are convex.

Let us now focus on the case where the acceptance set is defined in terms of *shortfall risk*, that is,

$$\mathcal{A} := \{X \in \mathcal{X} | E_P[\ell(-X)] \leq l_0\}.$$

for some convex loss function ℓ and some threshold l_0 as in Section 4.2. In this case, the index of riskiness $R(X)$ is the inverse of the unique solution $\lambda(X)$ of

$$E_P[\ell(-\lambda X)] = l_0$$

or, alternatively, of

$$E_P[u(\lambda X)] = u_0,$$

where $u(x) = -\ell(-x)$ and $u_0 = -l_0$. Clearly, the index R is now *distribution-based*, that is, $R(X)$ only depends on the distribution of X, viewed as a random variable on the probability space $(\Omega, \mathcal{F}, P)$. Moreover, it is easy to see that R is monotone with respect to stochastic dominance both of the first and the second kind, cf. Section 7.3.

Example 5.5. *For exponential utility $u(x) = 1 - e^x$ with threshold $u_0 = 0$, we obtain the* economic index of riskiness *$R^{AS}(X)$ proposed by Aumann & Serrano (2008). It is defined by the equation*

$$E_P[\exp(-(R^{AS}(X))^{-1}X)] = 1,$$

and its inverse $(R^{AS})^{-1}$ can be interpreted as the critical risk aversion level *for the position X with respect to exponential utility.*

Example 5.6. *For logarithmic utility $u(x) = \log(1 + x)$ with threshold $u_0 = 0$, we obtain the* operational index of riskiness *R^{FH} introduced by Foster & Hart (2009). It is defined by the equation*

$$E_P[\log(X + R^{FH})] = \log R^{FH}.$$

Thus it can be viewed as the critical wealth level *for the position X with respect to logarithmic utility.*

For both examples, Hart (2011) has shown that the index of riskiness R can be characterized as the numerical representation of a suitably defined preference order $\succeq_U$ of "uniform dominance", that is,

$$R(X) \leq R(Y) \Longleftrightarrow X \succeq_U Y. \tag{30}$$

Take the class $\mathcal{U}$ of smooth utility functions u with decreasing absolute risk aversion, increasing relative risk aversion, and such that a given position X is not accepted at every level of wealth, that is, $E[u(X+w)] \leq u(w)$ for some $w > 0$. In order to characterize the index R^{AS} of Aumann and Serrano, Hart (2011) defines *wealth-uniform dominance* $X \succeq_{WU} Y$ by the condition that, for any $u \in \mathcal{U}$,

$$E_P[u(X+w)] \leq u(w) \text{ for all } w > 0 \Longrightarrow E_P[u(Y+w)] \leq u(w) \text{ for all } w > 0$$

that is, rejection of X at every level of wealth implies rejection of Y at any level of wealth. In order to characterize the index R^{FH} of Foster and Hart, *utility-uniform dominance* $X \succeq_{UU} Y$ is defined by the condition that, at any level of wealth $w > 0$,

$$E_P[u(X+w)] \leq u(w) \text{ for all } u \in \mathcal{U} \Longrightarrow E_P[u(Y+w)] \leq u(w) \text{ for all } u \in \mathcal{U}$$

that is, rejection of $X+w$ by all $u \in \mathcal{U}$ implies rejection of $Y+w$ by all $u \in \mathcal{U}$.

The indices R^{AS} and R^{FH} both satisfy the equivalence (30) with the corresponding choice of the preference order. Moreover, any positively homogeneous risk functional that represents the preference order must be a positive multiple of the index; cf. Hart (2011) and Aumann & Serrano (2008).

6 Convex risk measures and variational preferences

In this section we describe the connection between convex risk measures and the numerical representation of preferences in the face of model uncertainty.

Consider a preference order $\succeq$ on the space $\mathcal{X} = \mathscr{L}^\infty(\Omega, \mathcal{F})$ of financial positions, and suppose that $\succeq$ admits a numerical representation, that is,

$$X \succeq Y \quad \Longleftrightarrow \quad U(X) \geq U(Y)$$

where U is some functional on $\mathcal{X}$ with values in $[-\infty, \infty)$. In the paradigm of *expected utility*, the functional U takes the form

$$U(X) = E_P[u(X)] = \int u(x)\, \mu_X(dx)$$

where P is a probability measure on $(\Omega, \mathcal{F})$, u is continuous and strictly increasing, and μ_X denotes the distribution of X under P, also called a "lottery". The classical *axioms of rationality* as formulated by von Neumann & Morgenstern (1944) characterize such preferences at the level of lotteries μ_X; a corresponding characterization at the level of positions X was given by Savage (1954). Moreover, risk aversion of the preference order in the sense that $E_P[X] \succeq X$ is characterized by concavity of the function u.

Lotteries and positions can be seen as special cases of stochastic kernels $\tilde{X}$ from $(\Omega, \mathcal{F})$ to $\mathbb{R}$, often called *acts* or *horse race lotteries*; cf., e. g., Kreps (1988). Indeed, a lottery μ corresponds to the kernel $\tilde{X}(\omega, \cdot) \equiv \mu$, a position X to the kernel $\tilde{X}(\omega, \cdot) = \delta_{X(\omega)}$. Let us now fix the class $\tilde{\mathcal{X}}$ of all stochastic kernels $\tilde{X}(\omega, dx)$ for which there exists some constant c such that $\tilde{X}(\omega, [-c, c]) = 1$ for all $\omega \in \Omega$. For preferences $\succeq$ on $\tilde{\mathcal{X}}$, Anscombe & Aumann (1963) have formulated a version of the rationality axioms which is equivalent to a numerical representation by an expected utility functional of the form

$$\tilde{U}(\tilde{X}) = E_P\left[\int u(x)\, \tilde{X}(\cdot, dx)\right]. \tag{31}$$

Let now ρ be a convex risk measure on $\mathcal{L}^\infty(\Omega, \mathcal{F})$ with dual representation (6), and take some increasing continuous function $u : \mathbb{R} \to \mathbb{R}$. Consider the utility functional $\tilde{U} : \tilde{\mathcal{X}} \to \mathbb{R}$ defined by

$$\tilde{U}(\tilde{X}) := -\rho\left(\int u(x)\, \tilde{X}(\cdot, dx)\right) = \inf_{Q \in \mathcal{M}_1} \left\{ E_Q\left[\int u(x)\, \tilde{X}(\cdot, dx)\right] + \alpha(Q) \right\}, \tag{32}$$

where the linear expectation functional E_P in (31) is replaced by the concave functional $-\rho$. The "variational" preference order $\succeq$ on $\tilde{\mathcal{X}}$ defined by

$$\tilde{X} \succeq \tilde{Y} \quad :\Longleftrightarrow \quad \tilde{U}(\tilde{X}) \geq \tilde{U}(\tilde{Y})$$

satisfies the following properties:

- *Monotonicity*: The preference order $\succeq$ on $\tilde{\mathcal{X}}$ is monotone with respect to the embedding of the space of standard lotteries $\mathcal{M}_{1,c}(\mathbb{R})$ in $\tilde{\mathcal{X}}$, i. e.,

  $$\tilde{X}(\omega,\cdot) \succeq_{(1)} \tilde{Y}(\omega,\cdot) \text{ for all } \omega \in \Omega \quad \Longrightarrow \quad \tilde{X} \succeq \tilde{Y},$$

  where $\succeq_{(1)}$ denotes first order stochastic dominance; cf. Section 7.3.

- *Archimedian axiom*: For $\tilde{X}, \tilde{Y}, \tilde{Z} \in \tilde{\mathcal{X}}$ with $\tilde{Z} \succeq \tilde{Y} \succeq \tilde{X}$ there exist $\alpha, \beta \in (0,1)$ such that

  $$\alpha\tilde{Z} + (1-\alpha)\tilde{X} \succeq \tilde{Y} \succeq \beta\tilde{Z} + (1-\beta)\tilde{X}.$$

- *Weak certainty independence*: If for $\tilde{X}, \tilde{Y} \in \tilde{\mathcal{X}}$ and for some $\nu \in \mathcal{M}_{1,c}(\mathbb{R})$ and $\alpha \in (0,1]$ we have $\alpha\tilde{X} + (1-\alpha)\nu \succeq \alpha\tilde{Y} + (1-\alpha)\nu$, then

  $$\alpha\tilde{X} + (1-\alpha)\mu \succeq \alpha\tilde{Y} + (1-\alpha)\mu \quad \text{for all } \mu \in \mathcal{M}_{1,c}(\mathbb{R}).$$

- *Uncertainty aversion*: If $\tilde{X}, \tilde{Y} \in \tilde{\mathcal{X}}$ are equivalent under $\succeq$, then

  $$\alpha\tilde{X} + (1-\alpha)\tilde{Y} \succeq \tilde{X} \quad \text{for all } \alpha \in [0,1].$$

Conversely, Maccheroni et al. (2006) have shown that the preceding four axioms imply that preferences can be represented by a utility functional $\tilde{U}$ of the form (32), where ρ is a convex risk measure on $\mathcal{L}^\infty(\Omega, \mathcal{F})$; cf. also Föllmer, Schied & Weber (2009) and Föllmer & Schied (2011), Theorem 2.88. While risk aversion corresponds to concavity of the function u, the convexity of the risk measure ρ captures a different behavioral feature, namely uncertainty aversion.

To illustrate the axiom of uncertainty aversion, consider two acts $\tilde{X}$ and $\tilde{Y}$ on $\Omega := \{0,1\}$ such that $\tilde{X}(\omega) = \tilde{Y}(1-\omega)$. Under Knightian uncertainty, with no information about probabilities for the two possible scenarios, it is natural to assume that the two acts $\tilde{X}$ and $\tilde{Y}$ are equivalent with respect to the given preference order $\succeq$ on $\tilde{\mathcal{X}}$. In the case of uncertainty aversion, a mixture $\tilde{Z} = \alpha\tilde{X} + (1-\alpha)\tilde{Y}$ is preferred over both $\tilde{X}$ and $\tilde{Y}$. To explain why, consider the simple special case $\tilde{X}(\omega) = \delta_\omega$. Then we have $\tilde{Z}(1) = \alpha\delta_1 + (1-\alpha)\delta_0$ and $\tilde{Z}(0) = (1-\alpha)\delta_1 + \alpha\delta_0$. Thus model uncertainty is reduced in favor of risk, since the unknown probability of success is now known to be bounded by α and $1-\alpha$. For $\alpha = \frac{1}{2}$, the resulting lottery $\tilde{Z}(\omega) = \frac{1}{2}(\delta_1 + \delta_0)$ is completely independent of the scenario ω, i. e., Knightian uncertainty is completely replaced by the classical risk of a simple coin toss.

Weak certainty independence can be strengthened to

- *Full certainty independence*: For all $\tilde{X}, \tilde{Y} \in \tilde{\mathcal{X}}$, $\mu \in \mathcal{M}_{1,c}(\mathbb{R})$, and $\alpha \in (0,1]$ we have

  $$\tilde{X} \succeq \tilde{Y} \quad \Longrightarrow \quad \alpha\tilde{X} + (1-\alpha)\mu \succeq \alpha\tilde{Y} + (1-\alpha)\mu.$$

In this case the risk measure ρ in (32) is actually coherent, and (32) reduces to the utility functional

$$\tilde{U}(\tilde{X}) = \inf_{Q \in \mathcal{Q}_\rho} E_Q[\int u(x)\, \tilde{X}(\cdot, dx)];$$

cf. Gilboa & Schmeidler (1989) and, e. g., Föllmer & Schied (2011), Theorem 2.86. Under the additional assumptions of law-invariance and comonotonicity, the right-hand side can be described as a Choquet integral with respect to a concave distortion of an underlying probability measure P; cf. Section 7.2 and also Yaari's "dual theory of choice", c.f. Yaari (1987).

Thus convex and coherent risk measures play a crucial role in recent advances in the theory of preferences in the face of risk and uncertainty. For further extensions, where the cash-invariance of ρ is replaced by a weaker condition of cash-subadditivity, we refer to Cerreia-Vioglio, Maccheroni, Marinacci & Montrucchio (2011) and Drapeau & Kupper (2013).

7 Distribution-based Risk Measures

From now on we fix a probability measure P on $(\Omega, \mathcal{F})$. As in Section 4 we assume that the probability space $(\Omega, \mathcal{F}, P)$ is atomless. Hence it supports a random variable U with uniform distribution on $(0,1)$. Thus, any probability distribution μ on the real line can be represented as the distribution of a random variable X on $(\Omega, \mathcal{F}, P)$; for example, we can take $X = q(U)$, where q denotes a quantile function of μ.

Definition 7.1. *A monetary risk measure ρ on $\mathcal{X} \subseteq L^0(\Omega, \mathcal{F}, P)$ is called* law-invariant *or* distribution-based *if $\rho(X)$ only depends on the distribution of X under P, i. e., if $\rho(X) = \rho(Y)$ whenever X and Y have the same distribution under P.*

Value at Risk as defined in (5) is clearly distribution-based, and the same is true for the risk measures *Average Value at Risk*, *utility-based shortfall risk*, and the *divergence risk measures* discussed in Section 4.

7.1 The Kusuoka representation

As shown by Jouini, Schachermayer & Touzi (2006), law-invariance of a *convex* risk measure ρ on $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, P)$ implies continuity from above, and therefore ρ admits a dual representation of the form (10). Moreover, the minimal penalty function $\alpha(Q)$ depends only on the law of the density $\frac{dQ}{dP}$ under P. It follows that the general dual representation can be restated in the following more explicit form, often called the *Kusuoka representation*; cf. Kusuoka (2001) in the coherent case and Kunze (2003), Dana (2005) and Frittelli & Rosazza Gianin (2005) in the general convex case.

Theorem 7.2. *A convex risk measure ρ on $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, P)$ is distribution-based if and only if*

$$\rho(X) = \sup_{\mu \in \mathcal{M}_1((0,1])} \left(\int_{(0,1]} \mathrm{AVaR}_\lambda(X)\, \mu(d\lambda) - \beta(\mu) \right), \tag{33}$$

where β denotes the minimal penalty function given by

$$\beta(\mu) = \sup_{X \in \mathcal{A}_\rho} \int_{(0,1]} \mathrm{AVaR}_\lambda(X)\, \mu(d\lambda).$$

The Kusuoka representation shows that Average Value at Risk provides the basic building blocks for a distribution-based convex risk measure. We refer to Föllmer & Schied (2011), Theorem 4.62, for the proof, and also to Drapeau, Kupper & Reda (2011) for an extension of the representation theorem to distribution-based risk functionals in the sense of Definition 5.1.

The Kusuoka representation also implies that a distribution-based convex risk measure ρ admits a canonical extension to a distribution-based convex risk measure ρ on L^1 with values in $\mathbb{R} \cup \{+\infty\}$. It is given by (33), or by the representation (11) with $p = 1$ and $q = \infty$. Moreover, the continuity properties of the extended risk measure ρ can be described precisely in terms of the Orlicz hearts $H^\Psi \subseteq L^1$ on which ρ takes finite values. The following theorem summarizes results of Filipovic & Svindland (2012) and Cheridito & Li (2009).

Theorem 7.3. *Let ρ be a distribution-based convex risk measure on $L^\infty(\Omega, \mathcal{F}, P)$. Then there exists a unique extension of ρ to a distribution-based convex risk measure*

$$\rho : L^1(\Omega, \mathcal{F}, P) \to \mathbb{R} \cup \{+\infty\}$$

that is lower-semicontinuous with respect to the L^1-norm. Moreover, if the extension is finite on some Orlicz heart H^Ψ with finite Young function Ψ, then it is continuous on H^Ψ with respect to the Luxemburg norm $\|\cdot\|_\Psi$.

From now on, any distribution-based convex risk measure ρ will be considered as a functional on $\mathcal{X} = L^1(\Omega, \mathcal{F}, P)$ with values in $(-\infty, \infty]$, if not stated otherwise.

In the *coherent* case the acceptance set $\mathcal{A}_\rho$ is a cone, and hence the penalty function β can only take the values in 0 or ∞. Thus a coherent risk measure ρ is distribution-based if and only if

$$\rho(X) = \sup_{\mu \in \mathcal{M}} \int_{(0,1]} \mathrm{AVaR}_\lambda(X)\, \mu(d\lambda) \tag{34}$$

for some class $\mathcal{M} \subseteq \mathcal{M}_1((0,1])$, and the maximal representing class is given by $\mathcal{M} = \{\mu | \beta(\mu) = 0\}$. Note that ρ reduces to the linear case $\rho(X) = E_P[-X]$ if and only if the representing set in (34) reduces to $\mathcal{M} = \{\delta_1\}$. In any other case ρ will charge a *risk premium* on top of the expected loss:

$$\rho(X) > E_P[-X] \tag{35}$$

for any non-constant position $X \in L^1(\Omega, \mathcal{F}, P)$. In particular, we have $E_P[X] > 0$ whenever $0 \neq X \in \mathcal{A}_\rho$.

7.2 Choquet integrals and concave distortions

Let us now focus on the special class of distribution-based risk measures which can be represented as mixtures of Average Value at Risk, i. e.,

$$\rho_\mu(X) := \int_{[0,1]} \mathrm{AVaR}_\lambda(X)\, \mu(d\lambda) \tag{36}$$

for some probability measure μ on the unit interval $[0,1]$. Such a risk measure is coherent, and it is also called a *spectral risk measure*; cf. Acerbi & Tasche (2002).

Let g denote the increasing and concave function on the unit interval defined by $g(0) = 0$ and the right-hand derivative

$$g'_+(t) = \int_{(t,1]} s^{-1}\, \mu(ds), \quad 0 < t < 1, \tag{37}$$

and note that $g(1) = 1$. Using Fubini's theorem, we can rewrite the mixture (36) in terms of the function g as

$$\rho_\mu(X) = \rho_g(X) := g(0+)\, \mathrm{ess\ sup}(-X) + \int_0^1 \mathrm{VaR}_\lambda(X)\, g'_+(\lambda) d\lambda. \tag{38}$$

Alternatively, $\rho_\mu = \rho_g$ can be written as a *Choquet integral*; cf. Föllmer & Schied (2011), Theorem 4.70.

Theorem 7.4. *Any spectral risk measure $\rho_\mu = \rho_g$ can be written as the Choquet integral*

$$\int (-X)\, dc := \int_0^\infty c[-X > x]\, dx + \int_{-\infty}^0 (c[-X > x] - 1)\, dx \tag{39}$$

of the loss $-X$ with respect to the capacity c defined as the concave distortion $c = g \circ P$ *of the underlying probability measure P. For this reason, ρ_g is also called a* distortion risk measure *with distortion function g.*

Conversely, take any concave distortion function g on the unit interval, that is, an increasing concave function g on $[0,1]$ with $g(0) = 0$ and $g(1) = 1$. Then there is a unique probability measure μ on $[0,1]$ such that g is given by (37); cf., e. g., Föllmer & Schied (2011), Lemma 4.69. Thus any Choquet integral $\int (-X)\, dc$ with respect to a concave distortion $c = g \circ P$ can be represented in the form (36) as a mixture of Average Value at Risk.

Example 7.5. *Average Value at Risk at level $\lambda \in (0,1)$ can be viewed as a spectral risk measure ρ_μ with $\mu = \delta_\lambda$. Thus it is a distortion risk measure with distortion function $g(x) = (x/\lambda) \wedge 1$.*

Example 7.6. *For a positive integer n consider the concave distortion function $g(x) = 1-(1-x)^n$. The corresponding risk measure ρ_g was proposed in Cherny & Madan (2009) and is sometimes called MINVAR. For independent copies $X_1, \ldots, X_n$ of X we get*

$$\rho(X) = E_P[\max\{-X_1, \ldots, -X_n\}] = -E_P[\min\{X_1, \ldots, X_n\}],$$

i. e., $\rho(X)$ can be described as the expectation under P of the maximal loss occurring in the portfolio $X_1, \ldots, X_n$. The risk measures MAXVAR, MAXMINVAR and MINMAXVAR are defined in the same manner by the distortion functions $x^{1/n}$, $(1-(1-x)^n)^{1/n}$ and $1-(1-x^{1/n})^n$. We can also replace the positive integer n by any real parameter $\beta \geq 1$; cf. Cherny & Madan (2009) and Madan & Cherny (2010).

Any risk measure ρ_μ of the form (36) is *comonotonic*, i. e., ρ_μ satisfies

$$\rho_\mu(X+Y) = \rho_\mu(X) + \rho_\mu(Y) \tag{40}$$

whenever two positions X and Y are *comonotone* in the sense that

$$(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geq 0 \quad \text{for all } (\omega, \omega') \in \Omega \times \Omega;$$

cf., e. g., Föllmer & Schied (2011), Theorem 4.88 combined with Corollary 4.77. Conversely, any distribution-based and comonotonic convex risk measure can be written as a spectral risk measure (36) for some probability measure μ on $[0,1]$. Equivalently, it can thus be viewed as the Choquet integral (39) of the loss with respect to some concave distortion $c = \psi \circ P$; cf., e. g., Föllmer & Schied (2011), Theorem 4.93, or Delbaen (2000) and Delbaen (2002).

Remark 7.7. *Consider a general distortion function g on the unit interval, that is, g is increasing with $g(0) = 0$ and $g(1) = 1$. Then the distortion risk measure ρ_g defined by the Choquet integral $\rho_g(X) = \int(-X)dc$ in (39) with respect to the capacity $c = g \circ P$ is still a positively homogeneous and comonotonic monetary risk measure. However, if g is not concave then ρ_g is no longer convex, and hence not coherent. For example,* VaR_α *can be viewed as the distortion risk measure with non-concave distortion function $g(x) = I_{(\alpha,1]}(x)$. For a characterization of this general class of distortion risk measures see Schmeidler (1986) or Kou & Peng (2014).*

7.3 Law-invariance and stochastic dominance

For random variables X and Y in $\mathcal{X} = L^1(\Omega, \mathcal{F}, P)$ with distributions μ_X and μ_Y, distribution functions F_X and F_Y, and quantile functions q_X and q_Y, consider the two partial orders

$$\begin{aligned} X \succeq_{(1)} Y \quad &:\Longleftrightarrow \quad F_X(x) \leq F_Y(x) \quad \text{for all } x \in \mathbb{R} \\ &\Longleftrightarrow \quad q_X(\alpha) \geq q_Y(\alpha) \quad \text{for all } \alpha \in (0,1) \end{aligned}$$

and

$$\begin{aligned} X \succeq_{(2)} Y \quad &:\Longleftrightarrow \quad \int_{-\infty}^x F_X(z)\,dz \leq \int_{-\infty}^x F_Y(z)\,dz \quad \text{for all } x \in \mathbb{R} \\ &\Longleftrightarrow \quad \int_0^\lambda q_X(\alpha)\,d\alpha \geq \int_0^\lambda q_Y(\alpha)\,d\alpha \quad \text{for all } \lambda \in (0,1), \end{aligned} \tag{41}$$

often called *stochastic dominance* of the first and the second kind. Equivalently we can write

$$X \succeq_{(i)} Y \quad :\Longleftrightarrow \quad \int u\,d\mu_X \geq \int u\,d\mu_Y \quad \text{for all } u \in \mathcal{U}^{(i)},$$

where $\mathcal{U}^{(1)}$ (resp. $\mathcal{U}^{(2)}$) denotes the class of all increasing (resp. all increasing and concave) functions $u : \mathbb{R} \to \mathbb{R}$.

Now let ρ be a monetary risk measure ρ on $\mathcal{X}$. Then ρ is distribution-based if and only if it respects stochastic dominance of the first kind, i. e.,

$$X \succeq_{(1)} Y \quad \Longrightarrow \quad \rho(X) \leq \rho(Y).$$

Indeed, since $(\Omega, \mathcal{F}, P)$ is assumed to be atomless, there exists a random variable U with uniform distribution on $(0,1)$, and law-invariance together with monotonicity of the monetary risk measure ρ implies

$$\rho(X) = \rho(q_X(U)) \leq \rho(q_Y(U)) = \rho(Y), \tag{42}$$

since $q_X \geq q_Y$ if $X \succeq_{(1)} Y$. Conversely, (42) implies $\rho(X) = \rho(Y)$ whenever X and Y have the same distribution.

Proposition 7.8. *If ρ is distribution-based and convex, then it respects stochastic dominance both of the first and the second kind, that is,*

$$X \succeq_{(i)} Y \quad \Longrightarrow \quad \rho(X) \leq \rho(Y). \tag{43}$$

for $i = 1, 2$.

As to stochastic dominance of the second kind, note first that the second part of (41) translates into the equivalence

$$X \succeq_{(2)} Y \quad \Longleftrightarrow \quad \mathrm{AVaR}_\lambda(X) \leq \mathrm{AVaR}_\lambda(Y) \quad \text{for all } \lambda \in (0,1), \tag{44}$$

since $\mathrm{VaR}_\alpha(X) = -q_X(\alpha)$ for almost all $\alpha \in (0,1)$. The representation (33) of distribution-based convex risk measures in terms of Average Value at Risk shows that the inequality is preserved for any convex and distribution-based risk measure ρ. Moreover, we obtain the equivalence

$$X \succeq_{(2)} Y \quad \Longleftrightarrow \quad \rho(X) \leq \rho(Y) \quad \text{for any } \rho \in \mathcal{R}, \tag{45}$$

where $\mathcal{R}$ can be an arbitrary class of distribution-based convex risk measures that contains AVaR_λ for any $\lambda \in (0,1)$. For example we could take the class of comonotonic risk measures ρ_μ in the preceding subsection; cf., e. g., Dhaene, Kukush & Pupashenko (2006) and Dhaene, Vanduffel, Goovaerts, Kaas, Tang & Vyncke (2006).

Conversely, convexity of ρ follows from (43) combined with the following property of *comonotonic convexity*: For $X, Y \in \mathcal{X}$ and $\lambda \in (0,1)$,

$$\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y) \text{ if } X \text{ and } Y \text{ are comonotonic;} \tag{46}$$

cf. Song & Yan (2009), Theorem 3.6 and also Proposition 7.10 below. This yields the following criterion for convexity of a distribution-based monetary risk measure.

Proposition 7.9. *A distribution-based monetary risk measure ρ on $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, P)$ is convex if and only if it satisfies both (43) and (46), that is, ρ respects stochastic dominance of the second kind and has the property of comonotonic convexity.*

In the same way, distribution-based *coherent* risk measures on $\mathcal{X}$ can be characterized as positively homogeneous monetary risk measures on $\mathcal{X}$ which respect stochastic dominance of the second kind and satisfy the following property of *comonotonic subadditivity*:

$$\rho(X+Y) \leq \rho(X) + \rho(Y) \quad \text{if } X \text{ and } Y \text{ are comonotone;} \tag{47}$$

cf. Song & Yan (2009), Theorem 3.2.

7.4 Risk measures on lotteries

Let ρ be a distribution-based monetary risk measure on $\mathcal{X}$. We denote by μ_X the distribution of $X \in \mathcal{X}$ under P, and by

$$\mathcal{M}(\mathcal{X}) = \{\mu_X | X \in \mathcal{X}\}$$

the resulting class of probability measures on $\mathbb{R}$. For all our choices of $\mathcal{X}$, the set $\mathcal{M}(\mathcal{X})$ will be convex, due to our assumption that the underlying probability space is atomless. For $\mathcal{X} = L^p$ with $p \in (0, \infty)$, $\mathcal{M}(\mathcal{X})$ is the class of all probability measures μ such that $\int |x|^p \mu(dx) < \infty$; for $p = 0$ we get the class $\mathcal{M}_1(\mathbb{R})$ of all probability measures on $\mathbb{R}$, for $p = \infty$ the class $\mathcal{M}_{1,c}$ of probability measures on $\mathbb{R}$ with compact support. The probability measures in $\mathcal{M}(\mathcal{X})$ will also be called *lotteries*.

Since $\rho(X)$ only depends on μ_X, the risk measure ρ can be identified with the functional $\mathcal{R}$ on $\mathcal{M}(\mathcal{X})$ defined by

$$\mathcal{R}(\mu_X) = \rho(X).$$

The monetary properties of ρ translate into the following properties of the functional $\mathcal{R}$:

- $\mathcal{R}$ is *monotone* with respect to stochastic dominance of the first kind:

$$\mu \preceq_{(1)} \nu \quad \Longrightarrow \quad \mathcal{R}(\mu) \geq \mathcal{R}(\nu). \tag{48}$$

- $\mathcal{R}$ has the *translation property*

$$\mathcal{R}(T_m\mu) = \mathcal{R}(\mu) - m, \tag{49}$$

  where $T_m\mu$ denotes the shifted measure $T_m\mu(A) := \mu(A - m)$.

In order to characterize convexity of the monetary risk measure ρ on $\mathcal{X}$ in terms of the functional $\mathcal{R}$ on $\mathcal{M}(\mathcal{X})$, note first that the lotteries μ_X can be identified with their upper quantile functions q_X, defined as right-continuous inverses of their distribution functions. The following criterion translates the criterion of Song and Yan in Proposition 7.9 into properties of the functional $\mathcal{R}$:

Proposition 7.10. *The distribution-based monetary risk measure ρ is convex if and only if the functional $\mathcal{R}$ has the following two properties:*
i) $\mathcal{R}$ respects second order stochastic dominance, i.e.,

$$\mu \preceq_{(2)} \nu \quad \Longrightarrow \quad \mathcal{R}(\mu) \geq \mathcal{R}(\nu).$$

ii) For two measures μ_0 and μ_1 in $\mathcal{M}(\mathcal{X})$ with quantile functions q_0 and q_1 and for any $\lambda \in (0, 1)$, we have

$$\mathcal{R}(\mu_\lambda) \leq \lambda\mathcal{R}(\mu_1) + (1 - \lambda)\mathcal{R}(\mu_0)$$

where μ_λ denotes the measure corresponding to the convex combination $q_\lambda = \lambda q_1 + (1 - \lambda)q_0$ of the quantile functions, that is, μ_λ is the distribution of $q_\lambda(U) \in \mathcal{X}$, where U is uniformly distributed on $(0, 1)$.

To check that properties i) and ii) of the functional $\mathcal{R}$ imply convexity of the risk measure ρ, take $X_i \in \mathcal{X}$ $(i = 0, 1)$ and $\lambda \in (0, 1)$. Define $\tilde{X}_i := q_i(U)$, where q_i is a quantile function for X_i and U is uniformly distributed on $(0, 1)$, and denote by μ_λ and $\tilde{\mu}_\lambda$ the distributions of the convex combinations $X_\lambda := \lambda X_1 + (1 - \lambda)X_0$ and $\tilde{X}_\lambda := \lambda\tilde{X}_1 + (1 - \lambda)\tilde{X}_0$. Then we have

$$\begin{aligned} AVaR_\alpha(X_\lambda) \leq& AVaR_\alpha(\lambda X_1) + AVaR_\alpha((1 - \lambda)X_0) \\ =& AVaR_\alpha(\lambda\tilde{X}_1) + AVaR_\alpha((1 - \lambda)\tilde{X}_0) \\ =& AVaR_\alpha(\tilde{X}_\lambda) \end{aligned}$$

for any $\alpha \in (0.1)$; here we have used subadditivity of AVaR in the first step, law-invariance in the second, and comonotonicity in the third. This implies

$$\tilde{\mu}_\lambda \preceq_{(2)} \mu_\lambda,$$

due to our characterization (44) of second order stochastic dominance in terms of Average Value at Risk. Using first property i) and then property ii) of $\mathcal{R}$, we obtain

$$\begin{aligned} \rho(X_\lambda) = \mathcal{R}(\mu_\lambda) &\leq \mathcal{R}(\tilde{\mu}_\lambda) \\ &\leq \lambda \mathcal{R}(\tilde{\mu}_1) + (1-\lambda)\mathcal{R}(\tilde{\mu}_0) \\ &= \lambda \rho(X_1) + (1-\lambda)\rho(X_0), \end{aligned}$$

and thus ρ is indeed a convex risk measure.

7.5 Mixture-Convexity

Let ρ be a distribution-based monetary risk measure on $\mathcal{X}$, and let $\mathcal{R}$ denote the corresponding functional on $\mathcal{M}(\mathcal{X})$ defined by 7.4. In Proposition 7.10 we have characterized convexity of the risk measure ρ on $\mathcal{X}$ in terms of the functional $\mathcal{R}$ on $\mathcal{M}(\mathcal{X})$. At the level of distributions, however, we may also want to consider the following convexity properties which are defined in terms of mixtures of probability measures.

Definition 7.11. *Let us say that the risk measure ρ is* mixture-convex*, or simply* m-convex*, if the functional $\mathcal{R}$ is convex on $\mathcal{M}(\mathcal{X})$, that is,*

$$\mathcal{R}(\alpha\mu + (1-\alpha)\nu) \leq \alpha\mathcal{R}(\mu) + (1-\alpha)\mathcal{R}(\nu)$$

for all $\mu, \nu \in \mathcal{M}(\mathcal{X})$ and any $\alpha \in (0,1)$. More generally, we say that ρ is m-quasi-convex *if $\mathcal{R}$ is quasi-convex on $\mathcal{M}(\mathcal{X})$, that is, $\mathcal{R}$ has convex lower level sets $\{\mathcal{R}(\cdot) \leq r\}$. In the same way, ρ will be called* m-quasi-concave *if $\mathcal{R}$ has convex upper level sets $\{\mathcal{R}(\cdot) \geq r\}$.*

Due to the translation property (49), m-quasi-convexity of ρ is equivalent to the condition that the acceptance set

$$\mathcal{A}_{\mathcal{R}} := \{\mu \in \mathcal{M}(\mathcal{X}) \,|\, \mathcal{R}(\mu) \leq 0\},$$

now defined at the level of distributions, is a convex set of probability measures. In other words, if two lotteries μ and ν are acceptable, then any compound lottery $\alpha\mu + (1-\alpha)\nu$ obtained by randomizing the choice between μ and ν with some probability $\alpha \in (0,1)$ should also be acceptable. Similarly, ρ is *m-quasi-concave* if and only if the rejection set $\mathcal{A}^c_{\mathcal{R}}$ is convex; cf. Bellini & Bignozzi (2014) Lemma 2.2.

Example 7.12. *The risk measure $AVaR_\lambda$ is coherent, and in particular convex, at the level of random variables. However, its acceptance set at the level of distributions is* not *convex. Thus $AVaR_\lambda$ is not m-quasi-convex; see Weber (2006).*

Now consider any utility-based risk measure ρ^ℓ defined as in Subsection 4.2 for some increasing loss function ℓ, that is,

$$\rho^\ell(X) = \inf\{m \,|\, E_P[\ell(-X-m)] \leq l_0\}. \tag{50}$$

At the level of distributions, the corresponding acceptance set

$$\{\mu \in \mathcal{M}(\mathcal{X}) \,|\, \int \ell(-x)\,\mu(dx) \leq l_0\} \tag{51}$$

is clearly convex, and so is its complement. Thus utility-based shortfall risk ρ^ℓ is both m-convex and m-concave. Moreover, convexity of the loss function ℓ is equivalent to the first kind of convexity of the risk measure, namely to the convexity of the functional ρ^ℓ on $\mathcal{X}$.

Conversely, as shown by Weber (2006) under some mild regularity conditions, a distribution-based monetary risk measure ρ must be of the form (50) for some increasing loss function ℓ as soon as both the acceptance set $\mathcal{A}_{\mathcal{R}}$ at the level of distributions and its complement $\mathcal{A}^c_{\mathcal{R}}$ are convex. In particular, a distribution-based monetary risk measure ρ is utility-based as soon as it is convex and also both m-quasi-convex and m-quasi-concave.

In order to state this characterization theorem more precisely, we describe the appropriate topological setting; this will also be needed for our discussion of robustness properties in Section 9. Recall first that the *weak topology* on $\mathcal{M}_1(\mathbb{R})$ is generated by the maps $\mu \to \int f d\mu$ for all $f \in C_b(\mathbb{R})$, and that it is metrizable by the Prohorov metric

$$d_{\text{Proh}}(\mu, \nu) = \inf\{\epsilon > 0 : \mu(A) \leq \nu(A^\epsilon) + \epsilon \text{ for all Borel sets } A \subseteq \mathbb{R}\}, \tag{52}$$

where $A^\epsilon := \{x \in \mathbb{R} : \inf_{a \in A} |x - a| \leq \epsilon\}$ denotes the ϵ-hull of A.

To increase sensitivity to the tails, we refine the weak topology as follows. Choose a continuous "gauge function" $\psi : \mathbb{R} \to [1, \infty)$. We denote by C^ψ the space of all continuous functions on the real line such that $|f| \leq c|\psi|$ for some $c > 1$, and we define the class of probability measures

$$\mathcal{M}_1^\psi := \{\mu \in \mathcal{M}_1(\mathbb{R}) \mid \int \psi d\mu < \infty\}.$$

Definition 7.13. *The* ψ-weak topology *is defined as the topology on $\mathcal{M}_1^\psi$ that is generated by the maps $\mu \to \int f d\mu$ for $f \in C^\psi(\mathbb{R})$ or, equivalently, by the* Prohorov ψ-metric

$$d_\psi(\mu, \nu) := d_{Proh}(\mu, \nu) + \left|\int \psi d\mu - \int \psi d\nu\right|. \tag{53}$$

Example 7.14. *For the polynomial gauge function $\psi_p(x) = 1 + |x|^p$ with $p \in [1, \infty)$, the ψ_p-weak topology is also generated by the* Wasserstein metric of order p, *that is,*

$$d_{W,p}(\mu, \nu) := \left(\int_0^1 |F_\mu(u) - F_\nu(u)|^p du\right)^{\frac{1}{p}},$$

where F_μ and F_ν denote the distribution functions of μ and ν; see Krätschmer, Schied & Zähle (2014) and the references therein.

The following theorem and its corollary are due to Weber (2006).

Theorem 7.15. *Let ρ be a distribution-based monetary risk measure on $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, P)$, and let $\mathcal{A}_{\mathcal{R}}$ denote the corresponding acceptance set at the level of distributions. Assume there exists $x \in \mathbb{R}$ with $\delta_x \in \mathcal{A}_{\mathcal{R}}$ such that for any $y \in \mathbb{R}$ with $\delta_y \in \mathcal{A}^c_{\mathcal{R}}$ we have*

$$(1 - \alpha)\delta_x + \alpha\delta_y \in \mathcal{A}_{\mathcal{R}}$$

for sufficiently small $\alpha > 0$. Then the following statements are equivalent:

(i) $\mathcal{A}_{\mathcal{R}}$ is ψ-weakly closed for some gauge function ψ, and the sets $\mathcal{A}_{\mathcal{R}}$ and $\mathcal{A}^c_{\mathcal{R}}$ are both convex.

(ii) There exists a left-continuous loss function $\ell : \mathbb{R} \to \mathbb{R}$ and some scalar l_0 in the interior of the convex hull of the range of ℓ such that

$$\mathcal{A}_{\mathcal{R}} = \{\mu \in \mathcal{M}_{1,c} \mid \int \ell(-x)\mu(dx) \leq l_0\}.$$

The theorem shows that utility-based risk measures play a central role as soon as we insist on law-invariance and on convexity of both the acceptance set $\mathcal{A}_{\mathcal{R}}$ and the rejection set $\mathcal{A}^c_{\mathcal{R}}$. More precisely:

Corollary 7.16. *Assume that the monetary distribution-based risk measure ρ on $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, P)$ satisfies the equivalent conditions of the preceding theorem.*

(i) The risk measure ρ is convex if and only if the loss function ℓ is convex, and in this case ρ is a utility-based risk measure*.*

(ii) The risk measure ρ is coherent if and only if $\ell(x) = l_0 + \alpha x^+ - \beta x^-$ with $\alpha \geq \beta > 0$.

Definition 7.17. *The coherent risk measures in Corollary 7.16 (ii) are also called* expectiles*; see Newey & Powell (1987) and Section 8 below.*

Remark 7.18. *Theorem 7.15 does not require that the risk measure is convex. In particular, it includes* VaR_α *as a special case, with non-convex loss function $\ell(x) = I_{(0,\infty)}(x)$ and threshold $l_0 = \alpha$. For a* convex *risk measure ρ, Delbaen et al. (2014) have shown that only the assumption of convex level sets $\{\mathcal{R}(\cdot) = r\}$ is needed in order to conclude that ρ is utility-based. However, in this general situation the loss function may become infinite.*

8 Elicitability

Capital regulation requires banks and insurance companies to project their balance sheets into the future and to compute capital requirements from the distribution of balance sheet items. As functionals of unknown future distributions, capital requirements prescribed by distribution-based risk measures are predicted quantities, which are estimated within probabilistic models on the basis of available data. Backtesting refers to a comparison of these predictions with past experience, with the aim to validate the models and the forecast procedures that are used.

An approach that is commonly used in practice consists in computing the average score

$$\hat{S} = \frac{1}{n}\sum_{i=1}^{n} S(x_i, y_i), \tag{54}$$

given a history of observations y_i with unknown distribution μ from some class of distributions $\mathcal{M}$, and a history of predictions x_i for the relevant statistical functional $T(\mu)$. Performance is considered to be good if the average score $\hat{S}$ is small.

As argued by Gneiting (2011), such a performance criterion will be meaningful only if the scoring function S is consistent with the functional T in the following sense. For large n and for a constant estimate x, the average score will typically be close to the expectation $\int S(x, y)\mu(dy)$ if the data are stationary and driven by the distribution μ. Therefore, for any $\mu \in \mathcal{M}$, the corresponding value $T(\mu)$ of the functional should minimize the expected score $\int S(\cdot, y)d\mu$. In this case, the scoring function S is called *consistent* for the functional T on $\mathcal{M}$. The functional T is called *elicitable on* $\mathcal{M}$ if it admits a consistent scoring function; cf. Gneiting (2011). For example, the mean $T(\mu) = \int y d\mu(y)$ is elicitable with scoring function $S(x, y) = (y - x)^2$.

Since elicitability might be desirable from a statistical point of view, it is important to know which distribution-based risk measures ρ are elicitable in the sense that the associated functional $\mathcal{R}$ admits a consistent scoring function. Value at Risk at some level α is elicitable. More precisely, the scoring function $S(x, y) = \alpha(y - x)^+ + (1 - \alpha)(y - x)^-$ is consistent for the α-quantile, viewed as a multi-valued functional on distributions; cf. Gneiting (2011). In contrast, Average Value at Risk is *not* elicitable; see Remark 8.5 below.

For a general convex distribution-based risk measure ρ, the question of elicitability was studied, e.g., in Gneiting (2011), Ziegel (2014), Embrechts & Hofert (2014), Emmer, Kratz & Tasche (2013), Kou & Peng (2014), Bellini & Bignozzi (2014) and Bellini, Klar, Müller & Rosazza Gianin (2014). In this section we describe the main results. The crucial observation is that elicitability of ρ implies that ρ is both m-convex and m-concave, as defined in Section 7. But in view of Theorem 7.15 and Corollary 7.16, this means that ρ must be utility-based, and that it will be an expectile if we also require that ρ is coherent.

8.1 Elicitable functionals

The following definition of a scoring function and the additional regularity requirements are due to Bellini & Bignozzi (2014).

Definition 8.1. *A function $S : \mathbb{R}^2 \to [0, \infty)$ is called a* scoring function *if it has the following properties:*

(i) $S(x, y) = 0$ if and only if $x = y$;

(ii) for any y, $x \mapsto S(x, y)$ is increasing for $x > y$ and decreasing for $x < y$;

(iii) for any y, $S(x, y)$ is continuous in x.

Let us say that a scoring function S is regular *if, in addition, it is continuous in y and satisfies, for all x in some neighborhood of 0, the inequality $S(x, y) \leq \psi(y)$ with some gauge function ψ.*

Now consider a real-valued functional T defined on some convex set of probability measures $\mathcal{M} \subseteq \mathcal{M}_1$.

Definition 8.2. *The functional T is called* elicitable *on $\mathcal{M}$, if there exists a scoring function S such that*

$$T(\mu) = \operatorname{argmin}_{x\in\mathbb{R}} \int S(x, y)\mu(dy) \tag{55}$$

for any $\mu \in \mathcal{M}$, that is, $T(\mu)$ is the unique minimizer of the function $x \mapsto \int S(x, y)\mu(dy)$. In this case, we say that T is elicited by the scoring function S, and that S is strictly consistent for T with respect to $\mathcal{M}$.

In the context of robust statistics, an elicitable functional is also called an *M-estimator*; see Huber (1964) or Huber (1981).

The following observation is due to Osband (1985).

Lemma 8.3. *If T is elicitable on $\mathcal{M}$ then T has convex level sets, that is,*

$$T(\mu) = T(\nu) \quad \Longrightarrow \quad T(\alpha\mu + (1 - \alpha)\nu) = T(\mu) = T(\nu)$$

for all $\mu, \nu \in \mathcal{M}$ and any $\alpha \in [0, 1]$.

Together with Theorem 7.15 and Corollary 7.16, this observation provides the key to the characterization of elicitable convex risk measures described below.

8.2 Elicitable Risk Measures

Let ρ be a distribution-based monetary risk measure on $\mathcal{L}^\infty$, and let $\mathcal{R}$ denote the corresponding functional on $\mathcal{M}(L^\infty)$. Recall from Section 7.4 that the monetary properties of ρ translate into monotonicity of $\mathcal{R}$ with respect to stochastic dominance of the first kind and the translation property (49).

Now assume that the functional $\mathcal{R}$ is elicitable on $\mathcal{M}(X)$; in this case, also the risk measure ρ will be called *elicitable*. Combining the monetary properties of $\mathcal{R}$ with Osband's Lemma 8.3, we easily obtain the following convexity properties of $\mathcal{R}$ at the level of distributions; cf. Bellini & Bignozzi (2014), Lemma 2.2.

Corollary 8.4. *If the functional $\mathcal{R}$ is elicitable, then it is both m-convex and m-concave, that is, the acceptance set $\mathcal{A}_{\mathcal{R}} := \{\mu : \mathcal{R}_\rho(\mu) \leq 0\}$ and the rejection set $\mathcal{A}^c_{\mathcal{R}} = \{\mu : \mathcal{R}_\rho(\mu) > 0\}$ at the level of distributions are both convex.*

Remark 8.5. *The Corollary shows that Average Value at Risk is* not *elicitable since it is not m-convex; see Example 7.12. See, however, Acerbi & Szekely (2014) for a discussion of back-testing in this case.*

Combining Corollory 8.4 with Theorem 7.15 and Corollary 7.16, we see that elicitability of ρ typically implies that ρ must be utility-based. More precisely, Bellini & Bignozzi (2014) show that the regularity conditions of Theorem 7.15 are satisfied if ρ is elicited by a scoring function S that is regular in the sense of Definition 8.1. In this way, they obtain the following classification of distribution-based convex risk measures that are elicitable, cf. Bellini & Bignozzi (2014), Theorem 4.9.

Theorem 8.6. *Let ρ be a distribution-based monetary risk measure on L^∞, and assume that ρ is elicitable with a regular scoring function S.*

(i) If ρ is convex, then ρ is utility-based shortfall risk with some convex loss function ℓ.

(ii) If ρ is coherent, then ρ is an expectile with $\alpha \geq \frac{1}{2}$, as described in Corollary 7.16 (ii).

(iii) If ρ is coherent and comonotonic, then ρ reduces to the linear risk measure $\rho(X) = E_P[-X]$.

Conversely, a utility-based risk measure ρ^ℓ with convex loss function ℓ and threshold l_0 in the interior of the range of ℓ is always elicitable; cf. Bellini & Bignozzi (2014), Theorem 4.6 & Remark 4. Assuming without loss of generality $\ell(0) = 0 = l_0$, ρ^ℓ can be elicited by the regular scoring function

$$S(x, y) = \int_0^{y-x} \ell(u)du. \tag{56}$$

Remark 8.7. *Instead of insisting on convexity, Kou & Peng (2014) focus on the class of distortion risk measures ρ_g defined with respect to a general distortion function g, that is, $\rho_g(X)$ is the Choquet integral $\int(-X)dc$ with respect to the capacity $c = g \circ P$; see Remark 7.7. They show that, within this class, the only elicitible risk measures are the non-convex risk measures VaR_α with some $\alpha \in (0, 1]$ or the risk-neutral risk measure $E_P[-X]$.*

Remark 8.8. *We refer to Acerbi & Szekely (2014) and Davis (2013) for a discussion of the question to which extent the concept of elicitability is relevant for backtesting in the financial context.*

9 Robustness

For a financial position X, the capital requirement prescribed by a law-invariant convex risk measure is obtained by applying the functional $\mathcal{R}$ to the distribution μ_X of X. Typically, this distribution is estimated from historical data or from Monte Carlo simulation, that is, the empirical distribution generated by the data is used as a proxy for μ_X. The capital requirement is then estimated by applying the functional $\mathcal{R}$ to the empirical distribution.

The question is whether this "plug-in" method produces a good approximation of the capital requirement. To begin with, the risk measurement procedure should be *consistent* in the usual statistical sense, that is, the risk estimates should converge to $\mathcal{R}(\mu_X)$. Moreover, the procedure should be *robust*, that is, the distribution of the risk estimates should not be perturbed too much by small changes of the underlying probability law that produces the data.

Robustness is usually defined in terms of the weak topology and the corresponding Prohorov metric, following Hampel (1971). By Hampel's famous theorem, this classical notion of robustness can be characterized by continuity properties of the functional $\mathcal{R}$ with respect to the weak topology. However, as first observed by Cont, Deguest & Scandolo (2010) and Kou, Peng & Heyde (2013), distribution-based convex risk measures such as Average Value of Risk typically are *not* continuous in this sense, and so they are not Hampel-robust. This has raised serious questions concerning the practical relevance of risk measures other than Value at Risk, which clearly satisfies the robustness requirement.

In their recent extensions of Hampel's approach, Cont, Deguest & Scandolo (2010) and Krätschmer, Schied & Zähle (2014) have raised the level at which these issues can be discussed. While Cont, Deguest & Scandolo (2010) discuss the robustness of risk measures with respect to restricted classes of lotteries, Krätschmer et al. (2014) also use a refinement of the weak topology. As a result, robustness is no longer a question of yes or no. Instead, there are degrees of robustness, depending on the choice of admissible lotteries and on the choice of the topology.

In the current section we review the results of Krätschmer, Schied & Zähle (2014) on consistency and robustness, and we illustrate them for different classes of distribution-based risk measures. For additional results on sensitivity in terms of Hampel's influence function we refer to Cont, Deguest & Scandolo (2010).

9.1 Consistency and Continuity

Let ρ be a distribution-based convex risk measure on L^1, and let $\mathcal{R}$ denote the corresponding functional on $\mathcal{M}(L^1)$. If we want to estimate the capital requirement $\rho(X)$ for a position $X \in L^1$ from historical data or from Monte Carlo simulation, the question of consistency arises.

Let $X_1, X_2, \dots$ denote a stationary and ergodic sequence of random variables having the same law as X. Given the observations $X_1(\omega), \dots X_n(\omega)$, it is natural to estimate the value $\rho(X) = \mathcal{R}(\mu_X)$ by

$$\hat{\rho}_n(\omega) := \mathcal{R}(\hat{\mu}_n(\omega)), \tag{57}$$

where

$$\hat{\mu}_n(\omega) = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i(\omega)}$$

denotes the empirical distribution generated by the observations. The ergodic theorem guarantees that, P-almost surely, the empirical distributions μ_n converge weakly to the distribution μ_X of X as n tends to ∞. Under additional integrability assumptions, the convergence may even hold in a stronger sense.

The question is whether we have *consistency of the risk estimates* $\hat{\rho}_n$, that is, whether $\hat{\rho}_n = \mathcal{R}(\hat{\mu}_n)$ converges to the true value $\rho(X) = \mathcal{R}(\mu_X)$. Clearly, this will require continuity properties of the functional $\mathcal{R}$ that correspond to the convergence behavior of the empirical distributions. For example, Value at Risk behaves well with respect to weak convergence, since it is simply a quantile, up to a change of sign, and so we get consistency in this case. However, general distribution-based convex risk measures, and in particular their building blocks AVaR_λ, are not continuous with respect to the weak topology. In order to obtain consistency, we will therefore need a refined notion of weak convergence, both for the convergence of empirical distributions and for the continuity properties of the risk measure.

To this end, let $\Psi : [0, \infty) \to [0, \infty)$ denote a finite Young function. Recall from Section 3.5 that the corresponding Orlicz space L^Ψ and the Orlicz heart H^Ψ satisfy $L^\infty \subseteq H^\Psi \subseteq L^\Psi \subseteq L^1$. We denote by

$$\mathcal{M}(H^\Psi) = \{\mu_X | X \in H^\Psi\}$$

the family of distributions generated by the Orlicz heart H^Ψ. If Ψ satisfies the Δ_2-condition (12) then we have $H^\Psi = L^\Psi$, and $\mathcal{M}(H^\Psi)$ coincides with the class $\mathcal{M}_1^\psi$ defined by (51) for the gauge function $\psi(x) = 1 + \Psi(|x|)$. In particular, this holds in the classical case $\Psi(x) = x^p/p$ for $p \in [1, \infty)$, where $H^\Psi = L^\Psi = L^p$.

Now assume that the risk measure ρ is finite on the Orlicz heart H^Ψ. Then we know from Theorem 7.3, that ρ is continuous on H^Ψ with respect to the Luxemburg norm $\| \cdot \|_\Psi$. This is the key to the following consistency result; cf. Krätschmer, Schied & Zähle (2014), Theorem 2.6.

Theorem 9.1. *Suppose that ρ takes finite values on the Orlicz heart H^Ψ. Take $X \in H^\Psi$, and let $X_1, X_2, \dots$ be a stationary and ergodic sequence of random variables with the same law as X. Then (57) defines a strongly consistent estimator for $\rho(X)$, that is,*

$$\lim_{n\to\infty} \hat{\rho}_n = \rho(X) \quad P-\text{a.s.}.$$

If Ψ satisfies the Δ_2-condition then the continuity properties of ρ and $\mathcal{R}$ can be described more precisely as follows; cf. Krätschmer, Schied & Zähle (2014), Theorem 2.10.

Theorem 9.2. *Suppose that Ψ satisfies the Δ_2-condition (12). Then the following conditions are equivalent:*

(i) ρ is finite on the Orlicz heart H^Ψ.

(ii) If $(X_n) \subseteq L^\infty$ with $\|X_n\|_\Psi \to 0$ as $n \to \infty$, then $\rho(X_n) \to \rho(0)$.

(iii) The functional $\mathcal{R}$ is finite and continuous with respect to the ψ-weak topology on $\mathcal{M}(L^\infty)$.

(iv) The functional $\mathcal{R}$ is finite and continuous with respect to the ψ-weak topology on $\mathcal{M}^1_\psi = \mathcal{M}(H^\Psi)$.

Conversely, if every distribution-based convex risk measure ρ that is finite on the Orlicz heart H^Ψ satisfies the continuity condition (iii), then Ψ must satisfy the Δ_2-condition; cf. Krätschmer, Schied & Zähle (2014), Theorem 2.8.

Example 9.3. *The entropic risk measure e_γ discussed in Section 4.4 is finite on the Orlicz heart H^Ψ of the Young function $\Psi(x) = e^x - 1$. Clearly, Ψ does not satisfy the Δ_2-condition, and in fact e_γ does not satisfy the continuity conditions (iii) and (iv).*

9.2 Qualitative Robustness

We now turn to the issue of qualitative robustness, a notion that was introduced by Hampel (1971) in terms of the weak topology. In this section we describe the refined version of qualitative robustness that was developed by Krätschmer, Schied & Zähle (2014), extending the analysis of Cont, Deguest & Scandolo (2010). This refinement turns out to be crucial for a deeper understanding of the robustness properties of risk measures.

Let ρ be a distribution-based convex risk measure on L^1, and let $\mathcal{R}$ denote the corresponding functional on $\mathcal{M}(L^1)$. For a statistical functional such as $\mathcal{R}$, qualitative robustness means that the laws of the estimates $\hat{\rho}_n = \mathcal{R}(\hat{\mu}_n)$ in (57) do not change too much if there is a sufficiently small change in the underlying probability law that drives the data. To formulate this idea precisely, we need two metrics, a metric d_A for the underlying probabilistic models and a second metric d_B for the laws of the estimates.

Suppose that the observations in (57) are generated by an i.i.d. sequence of random variables with a common distribution μ that belongs to some class $\mathcal{M} \subseteq \mathcal{M}_1$. In this case, we can choose the classical model of an infinite product space. That is, we take $\Omega = \mathbb{R}^\mathbb{N}$, $X_i(\omega) = \omega(i)$, $\mathcal{F} = \sigma(X_1, X_2, \dots)$, and the product measure

$$P_\mu := \mu^{\otimes \mathbb{N}}$$

on $(\Omega, \mathcal{F})$. Thus the underlying probabilistic models P_μ are parametrized by the probability measures $\mu \in \mathcal{M}$, and hence they can be compared in terms of a metric d_A on the set $\mathcal{M} \subseteq \mathcal{M}_1$. For a given μ, the law of the estimator in (57) is given by the image $P_\mu \circ \hat{\rho}_n^{-1}$ of P_μ under the map $\hat{\rho}_n : \Omega \to \mathcal{M}_1$, and these laws will be compared in terms of a metric d_B on $\mathcal{M}_1$.

Definition 9.4. *Let $\mathcal{M} \subseteq \mathcal{M}_1$ be endowed with a metric d_A, and let d_B be a metric on $\mathcal{M}_1$. The risk functional $\mathcal{R}$ is called* robust on $\mathcal{M}$ with respect to d_A and d_B *if for any $\mu \in \mathcal{M}$ and any $\epsilon > 0$ there exists $\delta > 0$ and $n_0 \in \mathbb{N}$ such that*

$$\mu \in \mathcal{M},\ d_A(\mu, \nu) \leq \delta \implies d_B(P_\mu \circ \hat{\rho}_n^{-1}, P_\nu \circ \hat{\rho}_n^{-1}) \leq \epsilon \quad \text{for all } n \geq n_0. \tag{58}$$

In Hampel's classical formulation of qualitative robustness, the focus is on the weak topology on $\mathcal{M} = \mathcal{M}_1$, and he uses the Prohorov metric (52) at both levels, that is, both for the underlying probability measures μ and for the laws of the estimates; see Hampel (1971). In other words, he chooses

$$d_A = d_B = d_{Proh}.$$

In their extension of Hampel's formulation, Krätschmer, Schied & Zähle (2014) use a more flexible approach: They continue to use the Prohorov metric for the laws of the estimates, but at the underlying level they take a Prohorov ψ-metric of the form (53). In other words, they choose

$$d_A = d_\psi, \qquad d_B = d_{Proh}. \tag{59}$$

As a result, the simple dichotomy "robust or not" is replaced by a graded picture with varying degrees of robustness.

From now on we fix a finite Young function Ψ and use the corresponding gauge function ψ defined by $\psi(x) := 1 + \Psi(|x|)$. Since ψ is unbounded, the uniform continuity required in (58) cannot be expected without further restrictions on the set $\mathcal{M} \subseteq \mathcal{M}_\infty$. This is why the following definition involves subsets $\mathcal{N} \subseteq \mathcal{M}$ that are *uniformly ψ-integrating*, that is,

$$\lim_{c \to \infty} \sup_{\nu \in \mathcal{V}} \int_{\{\psi \geq c\}} \psi d\nu = 0.$$

Definition 9.5. *The functional $\mathcal{R}$ is called* Ψ-robust *on a set $\mathcal{M} \subseteq \mathcal{M}_1$, if it is robust with respect to d_ψ and d_{Proh} on every uniformly ψ-integrating subset $\mathcal{N} \subseteq \mathcal{M}$.*

Remark 9.6. *Since it is defined in terms of the Prohorov metric, classical Hampel-robustness requires that only the bulk of the distributions μ and ν must be close in order to get close risk estimates. In contrast to the Prohorov metric, the ψ-Prohorov metric also controls the distance in the tails. For Ψ-robustness, the risk estimates are thus required to be close only if both the bulk* and *the tails of μ and ν are close.*

As shown by Hampel (1971), classical robustness can be characterized by continuity properties of the functional ρ with respect to the weak topology. The following result generalizes Hampel's theorem to Ψ-robustness; cf. Krätschmer, Schied & Zähle (2014), Theorem 2.16.

Theorem 9.7. *Suppose that Ψ satisfies the Δ_2-condition (12). Then the following conditions are equivalent:*

(i) $\mathcal{R}$ is Ψ-robust on $\mathcal{M}^1_\psi = \mathcal{M}(H^\Psi)$.

(ii) $\mathcal{R}$ is Ψ-robust on $\mathcal{M}(L^\infty)$.

(iii) ρ is finite on the Orlicz heart H^Ψ.

Combined with Theorem 9.2, this yields a characterization of Ψ-robustness in terms of continuity properties of the functional $\mathcal{R}$ with respect to the ψ-weak topology or, equivalently, of the risk measure ρ with respect to the Luxemburg norm $\|\cdot\|_\Psi$.

9.3 Degrees of Robustness

Let us now focus on the classical case $\Psi_p(x) = x^p/p$ for $1 \leq p < \infty$. The Δ_2-condition is clearly satisfied, and so we can apply the preceding theorem. Thus the functional $\mathcal{R}$ is Ψ_p-robust on $\mathcal{M}(L^p)$ if and only if the risk measure ρ is finite on L^p. Moreover, to verify Ψ_p-robustness of $\mathcal{R}$ on $\mathcal{M}(L^p)$ it is enough to check it on $\mathcal{M}(L^\infty)$. The following index of qualitative robustness was introduced by Krätschmer, Schied & Zähle (2014).

Definition 9.8. *For a distribution-based convex risk measure ρ on L^∞ with associated functional $\mathcal{R}$, the* index of qualitative robustness *is defined as*

$$\mathrm{IQR}(\rho) \;=\; \Big(\inf\{p \in [1,\infty) : \mathcal{R} \text{ is } \Psi_p\text{−robust on } \mathcal{M}(L^\infty)\}\Big)^{-1}. \tag{60}$$

For a distribution-based convex risk measure ρ the index $IQR(\rho)$ takes values in $[0,1]$. In view of Theorem 9.7, the index can also be written as

$$\mathrm{IQR}(\rho) \;=\; \Big(\inf\{p \in [1,\infty) : \rho \text{ is finite on } L^p\}\Big)^{-1}, \tag{61}$$

and thus it attains its maximal value 1 if the risk measure ρ is finite on L^1.

Remark 9.9. *For $p \in (0,1)$, Ψ_p is no longer a Young function, but the notion of Ψ_p-robustness is still well defined in terms of the Prohorov ψ_p-metric with $\psi_p(x) = 1 + \frac{1}{p}|x|^p$. Thus we could admit the values $p \in (0,1)$ in our definition (61) of the index* $\mathrm{IQR}(\rho)$. *For a distribution-based convex risk measure ρ there is no difference, because ρ cannot be Ψ_p-robust for $p < 1$, as shown in Krätschmer, Schied & Zähle (2014), Proposition 2.14. For the non-convex monetary risk measure Value at Risk, however, the infimum in (61) will become* 0, *and so we get* $\mathrm{IQR}(\mathrm{VaR}_\alpha) = \infty$ *for any $\alpha \in (0,1)$.*

Let us now illustrate the computation of the index for two special classes of distribution-based convex risk measures, the distortion risk measures in Section 7.2 and the elicitable risk measures in Section 8.

9.3.1 Distortion risk measures

Recall from Section 7.2 the distortion risk measure

$$\rho_g(X) = g(0+)\,\mathrm{ess.sup}(-X) + \int_0^1 \mathrm{VaR}_\lambda(X)\, g'_+(\lambda)d\lambda \tag{62}$$

defined in terms of some concave distortion function g on the unit interval.

Let Ψ be a finite Young function with conjugate Young function Ψ^*; see Section 3.5. Then the risk measure ρ_g is finite on the Orlicz heart H^Ψ if and only if the function g'_+ belongs to the Orlicz space L^{Ψ^*} defined with respect to Lebesgue measure on the unit interval. In the context of L^p-spaces, it follows that the index of qualitative robustness of ρ_g is given by

$$\mathrm{IQR}(\rho_g) = \frac{q^* - 1}{q^*}, \tag{63}$$

where

$$q^* = \sup\{q \geq 1 \big| \int_0^1 (g'_+(t))^q dt < \infty\};$$

cf. Krätschmer, Schied & Zähle (2014), Proposition 2.22.

Example 9.10. *Average Value at Risk at level $\lambda \in (0,1)$ has index* $\mathrm{IQR}(AVaR_\lambda) = 1$, *since it is finite on L^1. Note that $AVaR_\lambda$ can be viewed as the distortion risk measure with distortion function $g(x) = (x/\lambda) \wedge 1$; see Example 7.5. Since the right-hand derivative of g is bounded, we have $q^* = \infty$, and thus* $\mathrm{IQR}(AVaR_\lambda) = 1$ *also follows from (63). More generally, a distortion risk measure ρ_g with distortion function $g(x) = (x/\lambda)^\beta \wedge 1$ for some $\beta \in (0,1]$ has index* $\mathrm{IQR}(\rho_g) = \beta$; *cf. Krätschmer, Schied & Zähle (2014), Example 2.23.*

Example 9.11. *Recall from Example 7.6 the special distortion risk measures proposed by Cherny & Madan (2009). Here the index of qualitative robustness is easily computed via equation* (63)*: For $\beta > 1$, the risk measure MAXVAR with distortion function $1 - (1-x)^\beta$ has index* 1, *while MAXVAR, MAXMINVAR and MINMAXVAR with distortion functions $x^{1/\beta}$, $(1-(1-x)^\beta)^{1/\beta}$ and $1 - (1 - x^{1/\beta})^\beta$ all have index $1/\beta$; see Krätschmer, Schied & Zähle (2014), Example 2.24.*

9.3.2 Elicitable risk measures

As we have seen in Section 8, a distribution-based convex risk measure ρ is utility-based as soon as it is elicitable; cf. Theorem 8.6 and also Delbaen et al. (2014). In other words, there is some convex, increasing loss function ℓ and some threshold l_0 in the interior of the range of ℓ such that

$$\rho(X) = \inf\{m \in \mathbb{R} : E[\ell(-X-m)] \leq l_0\}.$$

Consider the Young function $\Psi(x) = \ell(x) - \ell(0)$ and the gauge function $\psi(x) = 1 + \Psi(|x|)$. Suppose that Ψ is finite and satisfies the Δ_2-condition. Then the functional $\mathcal{R}$ corresponding to ρ is continuous with respect to the ψ-weak topology, and it is Ψ-robust; see Example 2.5 in Krätschmer, Schied & Zähle (2014).

Example 9.12. *In the case of a power loss function $\ell(x) = x^p \mathbf{1}_{\{x\geq 0\}}$ with threshold $l_0 > 0$, the associated risk measure ρ is finite on L^p, and $\mathcal{R}$ is continuous with respect to the ψ_p-weak topology and Ψ_p-robust. Moreover, the index of qualitative robustness is equal to $1/p$.*

Example 9.13. *In the coherent case, elicitability implies that the risk measure ρ is an* expectile. *In other words, the loss function is of the form $\ell(x) = l_0 + \alpha x^+ - \beta x^-$ with parameters $\alpha \geq \beta > 0$; see Section 7.5. An expectile is clearly finite on L^1, and hence its index of qualitative robustness is equal to 1, as in the case of Average Value at Risk.*

Example 9.14. *An entropic risk measure e_γ with $\gamma > 0$ is utility-based and hence elicitable. But it does not stay finite on any L^p-space; see Section 4.4. Thus, its index of qualitative robustness is equal to* 0.

References

(2009*a*), The Turner review - A regulatory response to the global banking crisis, published by the British Financial Services Authority, `www.fsa.gov.uk/pubs/other/turner_review.pdf`.

(2009*b*), Revisions to the Basel II market risk framework, published by the Basel Committee on Banking Supervision, `www.bis.org/publ/bcbs158.pdf`.

Acerbi, C. & B. Szekely (2014), 'Back-testing expected shortfall', *Risk* .

Acerbi, C. & D. Tasche (2002), 'On the coherence of expected shortfall', *Journal of Banking & Finance* **26**, 1487–1503.

Anscombe, F. J. & R. J. Aumann (1963), 'A definition of subjective probability', *Ann. Math. Statist.* **34**, 199–205.

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

Artzner, P., F. Delbaen & P. Koch-Medina (2009), 'Risk measures and efficient use of capital', *ASTIN Bulletin* **39**(1), 101–116.

Aumann, R. J. & R. Serrano (2008), 'An economic index of riskiness', *Journal of Political Economy* **116**, 810 – 836.

Bellini, F., B. Klar, A. Müller & E. Rosazza Gianin (2014), 'Generalized quantiles as risk measures', *Insurance Math. Econom.* **54**, 41–48.

Bellini, F. & V. Bignozzi (2014), Elicitable risk measures. Quantitative Finance. To appear.

Ben-Tal, A. & M. Teboulle (1987), 'Penalty functions and duality in stochastic programming via ϕ-divergence functionals', *Math. Oper. Res.* **12**(2), 224–240.

Ben-Tal, A. & M. Teboulle (2007), 'An old-new concept of convex risk measures: the optimized certainty equivalent', *Math. Finance* **17**(3), 449–476.

Biagini, S. & M. Frittelli (2009), On the extension of the Namioka-Klee theorem and on the Fatou property for risk measures, *in* 'Optimality and risk—modern trends in mathematical finance', Springer, Berlin, pp. 1–28.

Brown, D., E. De Giorgi & M. Sim (2010), 'Dual representation of choice and aspirational preferences', Preprint, available under: `http://papers.ssrn.com/sol3/papers.cfm?abstractid=1406399`.

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

Cheridito, P. & T. Li (2009), 'Risk measures on orlicz hearts', *Mathematical Finance* **19**(2), 189–214.

Cherny, A. & D. Madan (2009), 'New measures for performance evaluation', *The Review of Financial Studies* **22**(7), 2571–2606.

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

Cont, R., R. Deguest & X. Dong He (2013), 'Loss-based risk measures', *Stat. Risk. Model.* **30**(2), 133–167.

Dana, R.-A. (2005), 'A representation result for concave Schur concave functions', *Math. Finance* **15**(4), 613–634.

Davis, M. (2013), Consistency of risk measure estimates. Preprint.

De Finetti, B. (1931), 'Sul concetto di media', *Giorn. Ist. Ital. Attuari* **2**, 369–396.

Delbaen, F. (2000), *Coherent risk measures*, Cattedra Galileiana., Scuola Normale Superiore, Classe di Scienze, Pisa.

Delbaen, F. (2002), Coherent risk measures on general probability spaces, *in* 'Advances in Finance and Stochastics', Springer, Berlin, pp. 1–37.

Delbaen, F., F. Bellini, V. Bignozzi & J. F. Ziegel (2014), Risk measures with the CxLS property. Unpublished.

Dhaene, J., A. Kukush & M. Pupashenko (2006), 'On the characterization of premium principle with respect to pointwise comonotonicity', *Theory Stoch. Process.* **12**(3-4), 26–42.

Dhaene, J., S. Vanduffel, M. J. Goovaerts, R. Kaas, Q. Tang & D. Vyncke (2006), 'Risk measures and comonotonicity: a review', *Stoch. Models* **22**(4), 573–606.

Drapeau, S. (2010), 'Risk preferences and their robust representation', *PhD thesis, Humboldt-Universität zu Berlin*, available under: `http://edoc.hu-berlin.de/dissertationen/drapeau-samuel-2010-04-30/PDF/drapeau.pdf`.

Drapeau, S. & M. Kupper (2013), 'Risk preferences and their robust representation', To appear in: *Math. Oper. Res.*

Drapeau, S., M. Kupper & A. Papapantoleon (2014), 'A fourier approach to the computation of cv@r and optimized certainty equivalents', *Journal of Risk* **16**(6), 3?29.

Drapeau, S., M. Kupper & R. Reda (2011), 'A note on robust representations of law-invariant quasi-convex functions', *Adv. Math. Econ.* **15**, 27–39.

Dunkel, J. & S. Weber (2010), 'Stochastic root finding and efficient estimation of convex risk measures', *Oper. Res.* **58**(5), 1505–1521.

Embrechts, P., A. McNeil & D. Straumann (2002), Correlation and dependence in risk management: properties and pitfalls, *in* M.Dempster, ed., 'Risk Management: Value at Risk and Beyond', Cambridge University Press.

Embrechts, P. & M. Hofert (2014), 'Statistics and quantitative risk management for banking and insurance', *Annual Review of Statistics and Its Application* **1**(1), 493–514.

Emmer, S., M. Kratz & D. Tasche (2013), What is the best risk measure in practice? a comparison of standard measures. Preprint.

Filipovic, D. & G. Svindland (2012), 'The canonical model space for law-invariant convex risk measures is l1', *Mathematical Finance* **22**(3), 585–589.

Filipović, D. & N. Vogelpoth (2008), 'A note on the Swiss solvency test risk measure', *Insurance Math. Econom.* **42**(3), 897–902.

Föllmer, H. & A. Schied (2002), 'Convex measures of risk and trading constraints', *Finance and Stochastics* **6**(4), 429–448.

Föllmer, H. & A. Schied (2011), *Stochastic finance–An introduction in discrete time*, Graduate Textbook Series, 3rd edn, Walter de Gruyter & Co., Berlin.

Föllmer, H., A. Schied & S. Weber (2009), Robust preferences and robust portfolio choice, *in* 'Bensoussan, A., Zhang, (eds.) Handbook of Numerical Analysis, Mathematical Modeling and Numerical Methods in Finance', pp. 29–89.

Föllmer, H. & T. Knispel (2012), Convex capital requirements for large portfolios, *in* 'Stochastic Analysis and its Applications to Mathematical Finance, Essays in Honour of Jia-an Yan, Eds. T. Zhang and X. Y. Zhou', World Scientific, pp. 169–195.

Foster, D.P. & S. Hart (2009), 'An operational measure of riskiness', *Journal of Political Economy* **117**, 758 – 814.

Frittelli, M. & E. Rosazza Gianin (2002), 'Putting order in risk measures', *J. Banking & Finance* **26**, 1473–1486.

Frittelli, M. & E. Rosazza Gianin (2005), 'Law invariant convex risk measures', *Adv. Math. Econ.* **7**, 33–46.

Gerber, H. U. & M. J. Goovaerts (1981), 'On the representation of additive principles of premium calculation', *Scand. Actuar. J.* **4**, 221–227.

Gilboa, I. & D. Schmeidler (1989), 'Maxmin expected utility with nonunique prior', *J. Math. Econom.* **18**(2), 141–153.

Gneiting, T. (2011), 'Making and evaluating point forecasts', *Journal of the American Statistical Association* **106**(494), 746–762.

Goovaerts, M. J., R. Kaas, R. J. A. Laeven & Q. Tang (2004), 'A comonotonic image of independence for additive risk measures', *Insurance Math. Econom.* **35**(3), 581–594.

Goovaerts, M. J. & R. J. A. Laeven (2008), 'Actuarial risk measures for financial derivative pricing', *Insurance Math. Econom.* **42**(2), 540–547.

Hamm, A.-M., T. Salfeld & S. Weber (2013), ‘Stochastic root finding for optimized certainty equivalents’, *Proceedings of the 2013 Winter Simulation Conference* pp. 922–932.

Hampel, F. R. (1971), ‘A general qualitative definition of robustness’, *The Annals of Mathematical Statistics* **42**(6), 1887–1896.

Hart, S. (2011), ‘Comparing risks by acceptance and rejection’, *Journal of Political Economy* **119**, 617–638.

Hattendorff, K. F. W. (1868), ‘Über die Berechnung der Reserven und des Risico bei der Lebensversicherung’, *Masius’ Rundschau der Versicherungen* **18**, 169–183.

Heath, D. (2000), ‘Back to the future’, Plenary Lecture at the First World Congress of the Bachelier Society, Paris.

Huber, P. J. (1964), ‘Robust estimation of a location parameter’, *The Annals of Mathematical Statistics* **35**(1), 73–101.

Huber, P. J. (1981), *Robust statistics*, John Wiley & Sons Inc., New York. Wiley Series in Probability and Mathematical Statistics.

Jarrow, R. A. (2002), ‘Put option premiums and coherent risk measures’, *Mathematical Finance* **12**, 135–142.

Jouini, E., W. Schachermayer & N. Touzi (2006), ‘Law invariant risk measures have the Fatou property’, *Adv. Math. Econ.* **9**, 49–71.

Kaina, M. & L. Rüschendorf (2009), ‘On convex risk measures on L^p-spaces’, *Math. Methods Oper. Res.* **69**(3), 475–495.

Koenker, R. (2005), *Quantile Regression*, Econometric Society Monographs, Cambridge University Press.

Kou, S. & X. Peng (2014), On the measurement of economic tail risk. Working Paper.

Kou, S., X. Peng & C. C. Heyde (2013), ‘External risk measures and Basel accords’, *Mathematics of Operations Research* **38**(3), 393–417.

Krätschmer, V., A. Schied & H. Zähle (2014), ‘Comparative and qualitative robustness for law-invariant risk measures’, *Finance and Stochastics* **18**(2), 271–295.

Kreps, D. M. (1988), *Notes on the theory of choice*, Westview Press, Boulder.

Kunze, M. (2003), ‘Verteilungsinvariante konvexe Risikomaße’, *Diploma thesis, Humboldt-Universität zu Berlin*.

Kusuoka, S. (2001), ‘On law invariant coherent risk measures’, *Adv. Math. Econ.* **3**, 83–95.

Maccheroni, F., M. Marinacci & A. Rustichini (2006), ‘Ambiguity aversion, robustness, and the variational representation of preferences’, *Econometrica* **74**(6), 1447–1498.

Madan, D. & A. Cherny (2010), ‘Markets as a counterparty: An introduction to conic finance’, *International Journal of Theoretical and Applied Finance (IJTAF)* **13**(08), 1149–1177.

Newey, Whitney K. & James L. Powell (1987), ‘Asymmetric least squares estimation and testing’, *Econometrica* **55**(4), 819–847.

Osband, K. H. (1985), Providing Incentives for Better Cost Forecasting, PhD thesis, University of California, Berkely.

Savage, L. J. (1954), *The foundations of statistics*, John Wiley & Sons Inc., New York.

Scandolo, G. (2003), Risk measures in a dynamic setting. PhD thesis, University of Milan.

Schmeidler, David (1986), 'Integral representation without additivity', *Proc. Amer. Math. Soc.* **97**(2), 255–261.

Song, Y. & J.-A. Yan (2009), 'Risk measures with comonotonic subadditivity or convexity and respecting stochastic orders', *Insurance Math. Econom.* **45**(3), 459–465.

Uryasev, S. & R. T. Rockafellar (2001), Conditional value-at-risk: optimization approach, *in* 'Stochastic optimization: algorithms and applications (Gainesville, FL, 2000)', Vol. 54 of *Appl. Optim.*, Kluwer Acad. Publ., Dordrecht, pp. 411–435.

von Neumann, J. & O. Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, New Jersey.

Weber, S. (2006), 'Distribution-invariant risk measures, information, and dynamic consistency', *Mathematical Finance* **16**(2), 419–442.

Weber, S., W. Anderson, A.-M. Hamm, T. Knispel, M. Liese & T. Salfeld (2013), 'Liquidity-adjusted risk measures', *Mathematics and Financial Economics* **7**(1), 69–91.

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

Ziegel, J. F. (2014), 'Coherence and elicitability', *Mathematical Finance* .