# Stochastic Orders and Risk Measures: Consistency and Bounds

Nicole Bäuerle

Institut für Mathematische Stochastik

Universität Hannover, Welfengarten 1

D-30167 Hannover

`baeuerle@stochastik.uni-hannover.de`

Alfred Müller

Institut für Wirtschaftstheorie und Operations Research

Universität Karlsruhe, Geb. 20.21

D-76128 Karlsruhe, Germany

`mueller@wior.uni-karlsruhe.de`

Febr. 2005

**Abstract**

We investigate the problem of consistency of risk measures with respect to usual stochastic order and convex order. It is shown that under weak regularity conditions risk measures are consistent with these stochastic orders. This result is used to derive bounds for risk measures of portfolios. As a by-product, we extend the characterization of Kusuoka (2001) of coherent, law-invariant risk measures with the Fatou property to unbounded random variables.

# 1 Introduction

Risk measures assign a real number to a risk, which is described by a random variable $X : \Omega \to \mathbb{R}$. As usual in actuarial sciences we assume that $X$ describes a potential loss, but we allow $X$ to assume negative values, which means that a gain occurs. Typically risk measures are defined by an axiomatic approach. For pioneering work in actuarial science, where risk measures occur in the context of premium principles, see e.g. Goovaerts et al. (1984). In the financial literature there is a seminal paper by Artzner et al. (1999). Reasonable axioms for risk measures which have been used throughout the literature are monotonicity, i.e. if $X \leq Y$ then $\rho(X) \leq \rho(Y)$ and convexity, i.e. $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda \rho(X) + (1 - \lambda)\rho(Y)$ for all $\lambda \in [0,1]$. Of course monotonicity captures the idea that a larger loss is more risky. The meaning of convexity is a mathematical formulation that diversification pays. Namely the risk of a portfolio which diversifies between two risks $X$ and $Y$ is less, than the sum of the risks in both positions. For more details on convex risk measures we refer the reader to Föllmer and Schied (2002).

On the other hand there is a long tradition of decision theory which faces a similar problem and often uses stochastic orderings to formulate some favorable properties of risks. For example, if $X$ and $Y$ are two risks and $X \leq_{st} Y$, where $\leq_{st}$ is the usual stochastic order, then any rational decision maker would prefer $X$. Another important ordering to compare risks is the convex ordering $\leq_{cx}$ which is often used in the actuarial sciences, and which is related to notions of risk aversion. Here once more it is reasonable to assume that if $X$ and $Y$ are two risks and $X \leq_{cx} Y$, then a rational decision maker would prefer $X$ if she is risk averse, see e.g. Rothschild and Stiglitz (1970).

The problem whether risk measures are monotone with respect to stochastic orderings has been considered in the literature before for special cases. Pflug (1999) has shown that several risk measures are consistent with second order stochastic dominance. Similar results can be found in a series of papers by Ogryczak and Ruszczyński (1999,2001,2002). De Giorgi (2005) uses consistency with respect to second order stochastic dominance as an axiom in his definition of a risk measure. Goovaerts et al. (2004) show preservation of convex ordering for the so called Haezendonck risk measure.

In this paper we study in general the problem of consistency in the following sense: Is it true that a monotone risk measure $\rho$ has the property that $X \leq_{st} Y$ implies $\rho(X) \leq \rho(Y)$ and that a convex risk measure has the property that $X \leq_{cx} Y$ implies $\rho(X) \leq \rho(Y)$? At least for the monotonicity property this seems to be a trivial statement. However, surprisingly we show that these statements can be wrong. The crucial point is the probability space on which the risks are defined. On certain finite probability spaces these implications can be wrong,

whereas they are true on non-atomic probability spaces. Of course for practical purposes this means that the stochastic orderings and the corresponding properties of risk measures are consistent. This holds also for the stop-loss order and for second order stochastic dominance which are combinations of the stochastic and convex order and are often used in the actuarial sciences or the economic literature.

We apply our results to bound the risk of a portfolio, i.e. of the joint financial position $X_1 + \ldots + X_n$. This is of practical importance, since in many situations only the marginal distributions of $X_1, \ldots, X_n$ are known and not the dependence relation between them. This topic has been discussed in Embrechts et al. (2003) for the Value-at Risk. However, the Value-at Risk is not a convex risk measure. For convex risk measures we derive tight upper bounds which can be computed quite efficiently. For example we show that the general Markowitz type portfolio optimization problem with conservative upper bound can be solved by a linear program.

Most of the literature on risk measures deals with bounded risks. However, financial risks are typically unbounded and even heavy-tailed in actuarial applications. This is the reason why we decided to use $L^1$ as a framework for modelling risks. As a consequence we had to extend some results from $L^\infty$ to $L^1$. In particular, we extended the characterization of Kusuoka (2001) of coherent, law-invariant risk measures with the Fatou property to unbounded random variables. This seems to be an interesting result of its own.

Our paper is organized as follows: in section 2 we review the necessary theory of stochastic orderings and in section 3 we deal with risk measures on finite probability spaces. The theory of risk measures on $L^1$ and non-atomic probability spaces is developed in section 4. The remaining sections deal with the applications of obtaining tight upper bounds for the risk of portfolios and with portfolio optimization problems.

# 2  Stochastic orders

Let us recall some basic definitions and results from the theory of stochastic orders. For a comprehensive review we refer to Shaked and Shanthikumar (1994) or Müller and Stoyan (2002). In the following, in inequalities between expectations it is always tacitly assumed that the expectations exist. Throughout the paper we assume that all mentioned random variables have a finite mean and are defined on a common probability space $(\Omega, \mathcal{A}, P)$ unless stated otherwise.

**Definition 2.1.** For given random variables $X, Y$ we define the order relations

$$X \leq_{st} Y, \ \text{if} \ Ef(X) \leq Ef(Y) \ \text{for all increasing} \ f.$$

$$X \leq_{cx} Y, \ \text{if} \ Ef(X) \leq Ef(Y) \ \text{for all convex} \ f.$$

$$X \leq_{icx} Y, \ \text{if} \ Ef(X) \leq Ef(Y) \ \text{for all increasing convex} f.$$

$$X \leq_{icv} Y, \ \text{if} \ Ef(X) \leq Ef(Y) \ \text{for all increasing concave} f.$$

The ordering $\leq_{st}$ is called *usual stochastic ordering* or *first order stochastic dominance (FSD)*. In a financial setting it means that any rational decision maker, having a utility function in the sense of von Neumann-Morgenstern, prefers the return $Y$ to the return $X$. In the following theorem we collect some basic results for the order relation $\leq_{st}$. For more details see Müller and Stoyan (2002), section 1.2.

For any random variable $X$ we denote by $F_X(t) := P(X \leq t), \ t \in \mathbb{R}$ the *distribution function* and by

$$q_X(\alpha) := \inf\{x \in \mathbb{R} \mid P(X \leq x) \geq \alpha\}, \quad 0 < \alpha < 1,$$

the *quantile function*, which is the generalized inverse of the distribution function.

**Theorem 2.2.** *For random variables $X$ and $Y$ with distribution functions $F_X$ and $F_Y$ the following statements are equivalent:*

**(i)** $X \leq_{st} Y$;

**(ii)** *there is a probability space $(\Omega', \mathcal{A}', P')$ and random variables $X'$ and $Y'$ on it with the distribution functions $F_X$ and $F_Y$ such that $X'(\omega') \leq Y'(\omega')$ for all $\omega' \in \Omega'$;*

**(iii)** $F_X(t) \geq F_Y(t)$ *for all $t$.*

**(iv)** $q_X(\alpha) \leq q_Y(\alpha)$ *for all $\alpha \in (0,1)$.*

In part (ii) it is always possible to choose $\Omega' = (0,1)$, $\mathcal{A}'$ the Borel-$\sigma$-algebra, $P'$ the Lebesgue-measure, and $X'(\omega) := q_X(\omega)$, $Y'(\omega) := q_Y(\omega)$. If $X$ and $Y$ are random variables on some finite probability space $\Omega$, then it is, however, **not** always possible to choose $\Omega' = \Omega$!

**Example 2.3.** Let $\Omega = \{\omega_1, \omega_2\}$ with $P(\{\omega_1\}) = 1/3, \ P(\{\omega_2\}) = 2/3$, and consider

$$X(\omega_1) = 3, X(\omega_2) = 1, Y(\omega_1) = 2, Y(\omega_2) = 4.$$

Then $X \leq_{st} Y$, but $X(\omega_1) > Y(\omega_1)$. As the probabilities of $\omega_1$ and $\omega_2$ are different, there do not exist two different random variables with the same distribution on $\Omega$. Therefore there are no random variables $X' =_{st} X$ and $Y' =_{st} Y$ on $\Omega$ with $X' \leq Y'$ a.s.! One has to choose some strictly larger sample space $\Omega'$ to find random variables with this property. In this case it is possible to choose $\Omega' = \{\omega_1, \omega_2, \omega_3\}$ with $P'(\{\omega_i\}) = 1/3$, $i = 1, 2, 3$, and $X'(\omega_1) = X'(\omega_2) = 1$, $X'(\omega_3) = 3$, $Y'(\omega_1) = 2$, $Y'(\omega_2) = Y'(\omega_3) = 4$. Then $X' =_{st} X$, $Y' =_{st} Y$, and $X' \leq Y'$ $P'$-a.s.

The ordering $\leq_{icx}$ is also known as *stop-loss order* in actuarial sciences. The reason is that $\leq_{icx}$ holds if and only if the corresponding stop-loss transforms are ordered. The stop-loss transform $\pi_X$ of a random variable $X$ is defined as

$$\pi_X(t) = E(X - t)_+ = E \max\{X - t, 0\} = \int_t^\infty \bar{F}_X(s) \mathrm{d}s, \quad t \in \mathbb{R}.$$

Here the latter equality follows by partial integration. In decision theory $X \leq_{icx} Y$ has the meaning that any risk averse decision maker prefers the risk $X$ to the risk $Y$.

The *increasing concave ordering* $\leq_{icv}$ is the corresponding ordering for returns instead of losses. This is also known as *second order stochastic dominance (SSD)*, especially in the economic literature. If $X \leq_{icv} Y$ holds, where $X$ and $Y$ are risky returns, then any risk averse decision maker prefers $Y$ to $X$, see Rothschild and Stiglitz (1970). Note that $X \leq_{icv} Y$ holds, if and only if $-X \geq_{icx} -Y$. Because of this equivalence, and as we consider here risks, we will state our subsequent results for $\leq_{icx}$. It is easy to translate them into results for $\leq_{icv}$.

The next result is a slight generalization of Theorem 1.5.14 in Müller and Stoyan (2002).

**Theorem 2.4.** *Let $X$ and $Y$ be random variables on a probability space $(\Omega, \mathcal{A}, P)$ with $X \leq_{icx} Y$. Then there is a random variable $Z$ on $(\Omega, \mathcal{A}, P)$ with*

$$X \leq_{st} Z \leq_{cx} Y.$$

*Proof.* The proof of Theorem 1.5.14 in Müller and Stoyan (2002) shows that it is sufficient to find a random variable $Z$ with $\pi_Z(t) = \max\{\pi_X(t), EY - t\}$. If $EX = EY$, then we can choose $Z = X$. Otherwise there is a unique point $a$, where the functions $t \mapsto \pi_X(t)$ and $t \mapsto EY - t$ cross. Then we can define

$$Z(\omega) = \begin{cases} a, & \text{if } X(\omega) \leq a, \\ X(\omega), & \text{else.} \end{cases}$$

$\square$

4

Theorem 2.4 shows that it is sufficient to consider the orderings $\leq_{st}$ and $\leq_{cx}$. If one finds results for these two orders, then one can combine them to find also results for $\leq_{icx}$ or $\leq_{icv}$. For the ordering $\leq_{cx}$ there is a result similar to Theorem 2.2.

**Theorem 2.5.** *For random variables $X$ and $Y$ with distribution functions $F_X$ and $F_Y$ the following statements are equivalent:*

**(i)** $X \leq_{cx} Y$;

**(ii)** *there is a probability space $(\Omega', \mathcal{A}', P')$ and random variables $X'$ and $Y'$ on it with the distribution functions $F_X$ and $F_Y$ such that $E[Y'|X'] = X'$;*

**(iii)** $\pi_X(t) \leq \pi_Y(t)$ *for all $t$, and $EX = EY$.*

**(iv)**

$$\int_\alpha^1 q_X(s)ds \leq \int_\alpha^1 q_Y(s)ds$$

*for all $\alpha \in (0,1)$, and $EX = EY$.*

As in Theorem 2.2 it is in part (ii) always possible to choose $\Omega' = (0,1)$, $\mathcal{A}'$ the Borel-$\sigma$-algebra and $P'$ the Lebesgue-measure. If $X$ and $Y$ are random variables on some finite probability space $\Omega$, then it is, however, **not** always possible to choose $\Omega' = \Omega$! This is not possible even if $P$ is uniform on $\Omega$, as we will show in the next example.

**Example 2.6.** Let $\Omega = \{1, 2, 3, 4\}$ with $P(\{\omega\}) = 1/4$ for all $\omega$, and consider $X(\omega) = \omega$ and

$$Y(1) = Y(2) = 1, Y(3) = Y(4) = 4.$$

It is easy to see that $X \leq_{cx} Y$, and that a random variable $X'$ on $\Omega$ has the same distribution as $X$ if and only if $X'$ is a permutation of the numbers 1,2,3,4. Therefore for any random variable $Y'$ having the same distribution as $Y$ it is $E[Y'|X'] \in \{1, 4\}$ and $E[Y'|X']$ must have the same distribution as $Y$. Therefore $E[Y'|X'] = X'$ is not possible!

In section 4 we will need results on the relation between convex order and the concept of a *fusion*. This topic is considered in Elton and Hill (1992). We will use here a slight variation of the definition of a fusion given there. We use the notation $\mu_P$ for the mean of a probability measure $P$ (which is assumed to exist throughout the paper) and $\delta_c$ for the one-point measure in $c$.

**Definition 2.7.** A probability measure $P$ is said to be a fusion of a probability measure $Q$, if there are probability measures $P_1, P_2$ and a constant $0 \leq \alpha \leq 1$ such that

$$P = \alpha P_1 + (1 - \alpha)\delta_{\mu_{P_2}} \quad \text{and} \quad Q = \alpha P_1 + (1 - \alpha)P_2,$$

In the special case that $P_2 = Q_{|A}$ for some set $A$, our definition coincides with what is called an *elementary fusion* in Elton and Hill (1992). Therefore our definition is more general than the definition of an *elementary fusion* there. On the other hand it is obvious that for our definition it still holds that $P \leq_{cx} Q$, if $P$ is a fusion of $Q$. We can show the following result.

**Theorem 2.8.** *For probability measures $P$ and $Q$ with the same finite mean the following are equivalent:*

**(i)** $P \leq_{cx} Q$*;*

**(ii)** *there is a sequence of probability measures $(Q_n)$ such that $Q_0 = Q$, $Q_{n+1}$ is a fusion of $Q_n$, and $Q_n$ converges weakly to $P$;*

*Moreover, there are random variables $X, X_0, X_1, X_2, ...$ with $X \sim P$ and $X_n \sim Q_n$ such that $E|X - X_n| \to 0$.*

*Proof.* The equivalence of (i) and (ii) can be shown as Theorem 4.1 in Elton and Hill (1992). For a simple constructive proof in the case of distributions on the real line see Müller (1996). To prove the second part of the theorem, let $(X_n)$ be a sequence of random variables with distributions $Q_n$, which form a reversed martingale. According to Strassen's theorem (see equivalence (ii) in Theorem 2.5) such a sequence exists. It follows from the well known martingale convergence theorem that the sequence $(X_n)$ converges to some random variable $X$ almost surely and in $L^1$. Obviously $X$ has distribution $P$. □

# 3 Risk measures on finite probability spaces

In what follows all random variables constitute the outcome of potential losses. Positive values are regarded as losses, negative values as gains. We start with the basic definitions of risk measures on a finite probability space. Let $(\Omega, \mathcal{A}, P)$ be a probability space with $\Omega = \{\omega_1, ..., \omega_n\}$ finite and $\mathcal{A} = \wp(\Omega)$. We assume that all elements of $\Omega$ are relevant, i.e. $P(\{\omega_i\}) > 0$ for $i = 1, ..., n$. We denote by $\mathcal{X}$ the set of all functions $X : \Omega \to \mathbb{R}$. A mapping $\rho : \mathcal{X} \to \mathbb{R}$ is called a *risk measure*, if it fulfills some conditions. The following conditions for $X, Y \in \mathcal{X}$ are of interest:

**Monotonicity:** If $X \leq Y$ then $\rho(X) \leq \rho(Y)$;

**Translation invariance:** if $m \in \mathbb{R}$, then $\rho(X + m) = \rho(X) + m$;

**Subadditivity:** $\rho(X + Y) \leq \rho(X) + \rho(Y)$;

**Positive homogeneity:** if $\lambda > 0$, then $\rho(\lambda X) = \lambda \rho(X)$;

**Convexity:** $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda \rho(X) + (1 - \lambda)\rho(Y)$ for all $\lambda \in [0, 1]$;

**Law invariance:** If $P_X = P_Y$, then $\rho(X) = \rho(Y)$.

According to Artzner et al. (1999) a functional $\rho$ is called a *coherent risk measure*, if it is monotone, translation invariant, subadditive and positively homogeneous. They show that any coherent risk measure has a representation

$$\rho(X) = \sup_{Q \in \mathcal{Q}} E_Q(X), \tag{3.1}$$

where $\mathcal{Q}$ is some set of probability measures. This means that $\rho(X)$ is the worst expected loss under $Q$, where $Q$ varies over some set of probability measures.

Föllmer and Schied (2002) introduced the weaker concept of $\rho$ being a *convex risk measure* if it satisfies the condition of monotonicity, translation invariance and convexity. They show that any convex risk measure is of the form

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \left( E_Q(X) - \alpha(Q) \right),$$

where $\alpha$ is a penalty function, which can be chosen to be convex and lower semi-continuous with $\alpha(Q) \geq -\rho(0)$.

For this section, it is important to note that the notion of law invariance is indeed ambiguous. There are two interpretations: a weak and a strong one. The weak interpretation is that for any random variables $X$ and $Y$ defined on an arbitrary probability space $(\Omega, \mathcal{A}, P)$, then $P_X = P_Y$ implies $\rho(X) = \rho(Y)$, so that $(\Omega, \mathcal{A}, P)$ does not really matter. The strong interpretation is that $(\Omega, \mathcal{A}, P)$ is given and we only consider random variables $X$ and $Y$ on this specific probability space. $P_X = P_Y$ then implies $\rho(X) = \rho(Y)$. In what follows we will always understand law invariance in the strong sense. On certain finite probability spaces, there is a difference between these two notions as we will see in this section. However, on non-atomic probability spaces both notions coincide (see section 4).

In this paper we are interested in consistency of risk measures with stochastic orders. Given a stochastic order $\preceq$ we say that a risk measure is

**consistent with** $\preceq$, if $X \preceq Y$ implies $\rho(X) \leq \rho(Y)$.

As stochastic orders are law invariant, it is clear that law-invariance is a necessary condition for such a property. At first sight, it might seem to be an obvious consequence from Theorem 2.2 that a law invariant monotone risk measure has the property that $X \leq_{st} Y$ implies $\rho(X) \leq \rho(Y)$. This, however, is not true in general. If the probability space has atoms of different size, then law invariance is not a severe restriction, and then it may happen that the implication becomes wrong.

**Example 3.1.** Let us look again at Example 2.3 with $\Omega = \{\omega_1, \omega_2\}$, $P(\{\omega_1\}) = 1/3$, $P(\{\omega_2\}) = 2/3$, and

$$X(\omega_1) = 3, X(\omega_2) = 1, Y(\omega_1) = 2, Y(\omega_2) = 4.$$

Consider the risk measure $\rho(X) = X(\omega_1)$. This obviously is even a coherent risk measure (with $\mathcal{Q}$ in the representation (3.1) containing only $\delta_{\omega_1}$, the one-point measure in $\omega_1$). It is law-invariant in the strong sense. However, $\rho(X) > \rho(Y)$, though $X \leq_{st} Y$.

When $P$ is the uniform distribution on a finite probability space, then there are many different random variables with the same distributions, namely any permutation of $X$ has the same distribution as $X$. In that case, law invariance is a severe restriction that ensures a monotone risk measure to be monotone with respect to usual stochastic order.

**Theorem 3.2.** *Let* $\Omega = \{\omega_1, ..., \omega_n\}$ *with* $P(\omega_i) = 1/n$, $i = 1, ..., n$, *and let* $\rho$ *be a monotone law-invariant risk measure on* $\mathcal{X}$. *Then* $X \leq_{st} Y$ *implies* $\rho(X) \leq \rho(Y)$.

*Proof.* It is sufficient to show that for any $X, Y : \Omega \to \mathbb{R}$ with $X \leq_{st} Y$ we can find $X'$ and $Y'$ on $\Omega$ with $X' =_{st} X$, $Y' =_{st} Y$, and $X'(\omega) \leq Y'(\omega)$ for all $\omega$. Let $x_i := X(\omega_i)$ and $y_i := Y(\omega_i)$, $i = 1, ..., n$, and let

$$x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)} \quad \text{and } y_{(1)} \leq y_{(2)} \leq ... \leq y_{(n)}$$

be the increasing rearrangements of the vectors $(x_1, ..., x_n)$ and $(y_1, ..., y_n)$, respectively. As $x_{(i)} = q_X(i/n)$ and $y_{(i)} = q_Y(i/n)$, it follows from Theorem 2.2 that $x_{(i)} \leq y_{(i)}$ for all $i$. Therefore it is sufficient to define $X'(\omega_i) = x_{(i)}$ and $Y'(\omega_i) = y_{(i)}$ for all $i$. $\square$

Next we consider consistency of risk measures with respect to convex order. The natural conjecture is that $X \leq_{cx} Y$ should imply $\rho(X) \leq \rho(Y)$ for any reasonable risk measure. Again this is not true, if we have a finite probability space with atoms of different size.

**Example 3.3.** Let us consider $\Omega = \{\omega_1, \omega_2\}$, $P(\{\omega_1\}) = 1/3$, $P(\{\omega_2\}) = 2/3$, and define

$$X(\omega_1) = X(\omega_2) = 2, Y(\omega_1) = 4, Y(\omega_2) = 1.$$

Consider the coherent and law-invariant risk measure $\rho(X) = X(\omega_2)$. According to Jensen's inequality it is obviously $X \leq_{cx} Y$, but $\rho(X) > \rho(Y)$.

Here again we get a positive result for the case of $P$ being the uniform distribution on a finite set. In that case the only relevant property of a risk measure is its convexity. In the proof of the next theorem we need some basic facts about majorization and doubly stochastic matrices. We refer the reader to Marshall and Olkin (1979) for details about this topic.

**Theorem 3.4.** *Let $\Omega = \{\omega_1, ..., \omega_n\}$ with $P(\omega_i) = 1/n$, $i = 1, ..., n$, and let $\rho$ be a law-invariant risk measure having the convexity property. Then $X \leq_{cx} Y$ implies $\rho(X) \leq \rho(Y)$.*

*Proof.* Let $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{y} = (y_1, ..., y_n)$ be the vectors with $X(\omega_i) = x_i$ and $Y(\omega_i) = y_i$, $i = 1, ..., n$. It is well known that $X \leq_{cx} Y$ holds, if and only if $\mathbf{x} \leq_M \mathbf{y}$, where $\leq_M$ denotes the majorization order (see e.g. Müller and Stoyan (2002)). According to a theorem of Hardy, Littlewood and Polya (see e.g. Marshall and Olkin (1979), p. 22) this holds if and only if there is a doubly stochastic matrix $\mathbf{A}$ such that

$$\mathbf{x} = \mathbf{A}\mathbf{y}, \tag{3.2}$$

and according to Birkhoff's theorem (see e.g. Marshall and Olkin (1979), p. 19) the doubly stochastic matrix $\mathbf{A}$ is a convex combination of permutation matrices, i.e. there are permutation matrices $\mathbf{B}_j$ and weights $\alpha_j$, $j = 1, ..., m$ with $\alpha_j \geq 0$ and $\sum_{j=1}^m \alpha_j = 1$, such that

$$\mathbf{A} = \sum_{j=1}^m \alpha_j \mathbf{B}_j. \tag{3.3}$$

Define random variables $Z_j : \Omega \to \mathbb{R}$ such that $Z_j(\omega_i)$ is the $i$-th component of $\mathbf{B}_j\mathbf{y}$. Then $Z_j$, $j = 1, ..., m$ all have the same distribution as $Y$, and it follows from (3.2) and (3.3) that

$$X = \sum_{j=1}^m \alpha_j Z_j.$$

Law-invariance and the convexity property of $\rho$ therefore imply

$$\rho(X) = \rho\left(\sum_{j=1}^m \alpha_j Z_j\right) \leq \sum_{j=1}^m \alpha_j \rho(Z_j) = \sum_{j=1}^m \alpha_j \rho(Y) = \rho(Y).$$

$\square$

Notice that the proof does not use any of the properties *monotonicity, translation invariance, subadditivity or positive homogeneity* of a risk measure. Therefore it applies to risk measures as considered in Schied (2004). We also note that our Theorem 3.4 is related to Lemma 2.3 in Schied (2004), which implies that on a finite probability space $\rho(E[Y|X]) \le \rho(Y)$ holds for all law-invariant convex risk measures. In light of Theorem 2.5 one might believe that this implies consistency with respect to $\le_{cx}$. However, it is not possible to prove Theorem 3.4 with this approach, as Example 2.6 reveals.

Also one might be tempted to assume that the converse statement of Theorem 3.4 holds, i.e. suppose $\rho$ is a risk measure which is consistent with $\le_{cx}$. Is then $\rho$ convex in the sense of the postulated axiom? The answer is no! This can be seen from the following counterexample: Let $(\Omega, \mathcal{A}, P)$ be any nontrivial probability space and define

$$\rho(X) = \begin{cases} EX, & \text{if } Var(X) \le 1, \\ EX + 1, & \text{if } Var(X) > 1. \end{cases}$$

So $\rho$ is only a function of the expectation and the variance, thus clearly law-invariant, and increasing in the expectation and the variance. As $X \le_{cx} Y$ implies $EX = EY$ and $Var(X) \le Var(Y)$, this in turn implies $\rho(X) \le \rho(Y)$. However, $\rho$ is not convex. Assume that $X \equiv 0$ and let $Y$ be a random variable with $EY = 0$ and $Var(Y) = 5$. Then $\rho(X) = 0$, $\rho(Y) = 1$, and

$$\rho(\frac{1}{2}X + \frac{1}{2}Y) = \rho(\frac{1}{2}Y) = 1 > \frac{1}{2} = \frac{1}{2}\rho(X) + \frac{1}{2}\rho(Y).$$

For a law-invariant, monotone and convex risk measure we can combine Theorem 3.2 with Theorem 3.4 to get consistency of the risk measure with $\le_{icx}$.

**Theorem 3.5.** *Let $\Omega = \{\omega_1, ..., \omega_n\}$ with $P(\omega_i) = 1/n$, $i = 1, ..., n$, and let $\rho$ be a law-invariant risk measure on $\mathcal{X}$, having the monotonicity and convexity property. Then $X \le_{icx} Y$ implies $\rho(X) \le \rho(Y)$.*

*Proof.* Assume that $X \le_{icx} Y$. According to Theorem 2.4 there is some $Z$ with $X \le_{cx} Z \le_{st} Y$. It follows from Theorem 3.4 that $\rho(X) \le \rho(Z)$ and from Theorem 3.2 that $\rho(Z) \le \rho(Y)$. Hence $\rho(X) \le \rho(Y)$. $\qquad\qquad\square$

# 4  Risk measures on non-atomic probability spaces

We now assume that $(\Omega, \mathcal{A}, P)$ is a standard probability space with $P$ non-atomic. This implies that for any distribution $Q$ on $\mathbb{R}$ there is a random variable $X : \Omega \to \mathbb{R}$ with $Q = P_X$. Indeed, there even exists a sequence of i.i.d. random variables with this distribution. In the

literature on risk measures it is typically assumed that the risk measure is defined on the set $L^\infty(\Omega, \mathcal{A}, P)$ of all bounded random variables, see e.g. Artzner et al. (1999) and Kusuoka (2001). Almost all concrete models in insurance and finance, however, consider unbounded random variables with e.g. normal distributions or even heavy tailed distributions. Therefore we consider it more natural to work with the larger space $\mathcal{X} = L^1(\Omega, \mathcal{A}, P)$ of all random variables having a finite mean under $P$. This makes it necessary to allow a risk measure to assume the value $+\infty$, compare Delbaen (2002). Thus a risk measure is a mapping $\rho : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$. The properties *monotonicity, translation invariance, subadditivity, positive homogeneity, convexity* and *law-invariance* are defined as before. Notice that since the probability space is now rich enough, law-invariance even in the strong sense means that $\rho$ can equivalently be considered as a mapping on the set of all distribution functions of probability measures with a finite mean. In this general setting continuity properties are of additional interest. We first show that we can not expect a risk measure to be continuous with respect to convergence in law.

**Theorem 4.1.** *Let $\rho$ be a monotone, convex and law-invariant risk measure, continuous with respect to convergence in law. Then $\rho$ is constant.*

*Proof.* Without loss of generality we assume that $\Omega = (0,1)$, $\mathcal{A}$ is the Borel-$\sigma$-algebra and $P$ is Lebesgue-measure and fix two numbers $a < b$. Define $X_0 \equiv b$ and for $n = 1, 2, 3, ...$ let

$$X_n(\omega) = \begin{cases} 2^n b + (1 - 2^n)a, & \omega \leq 1/2^n, \\ a, & \text{otherwise}, \end{cases}$$

and

$$X_n'(\omega) = \begin{cases} 2^n b + (1 - 2^n)a, & 1/2^n < \omega \leq 2/2^n, \\ a, & \text{otherwise}. \end{cases}$$

Thus $X_n =_{st} X_n'$, $X_n = \frac{1}{2}(X_{n+1} + X_{n+1}')$ and $X_n \to a$ in law. Convexity and law-invariance imply

$$\rho(X_n) = \rho(\frac{1}{2}(X_{n+1} + X_{n+1}')) \leq \frac{1}{2}\rho(X_{n+1}) + \frac{1}{2}\rho(X_{n+1}') = \rho(X_{n+1})$$

Thus $n \mapsto \rho(X_n)$ is an increasing sequence. Therefore monotonicity of the risk measure implies

$$\rho(a) = \lim_{n \to \infty} \rho(X_n) \geq \rho(X_0) = \rho(b) \geq \rho(a)$$

Thus $\rho(a) = \rho(b)$ and by monotonicity we get for any $X$ with $a \leq X \leq b$ that $\rho(X) = \rho(a)$. As $a$ and $b$ were arbitrary, this implies that $\rho$ is constant on the set of all bounded

11

random variables, and approximating an unbounded random variable $Y$ by the sequence $Y_n = (Y \wedge n) \vee -n$ of bounded random variables, this extends also to unbounded random variables. $\qquad\qquad\square$

In what follows we denote by $X_n \to_w X$ weak convergence, $X_n \to_p X$ convergence in probability and by $X_n \to_{L^1} X$ convergence in $L^1$, i.e. $E|X - X_n| \to 0$. In the literature on risk measures, where only random variables in $L^\infty$ are considered, one typically assumes the following lower semi-continuity assumption, called *Fatou property*:

If $(X_n)$ is a uniformly bounded sequence of random variables with $X_n \to_p X$, then

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

As we are working on $L^1$, it is natural to replace this condition by the more natural following condition, that we will also call the *Fatou property*:

If $X, X_1, X_2, ...$ are integrable random variables with $X_n \to_{L^1} X$, then

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

Recall that $X_n \to_{L^1} X$ holds if and only if both $X_n \to_p X$ and $(X_n)$ is uniformly integrable. Therefore a uniformly bounded sequence $(X_n)$ converging to $X$ in probability of course satisfies $X_n \to_{L^1} X$, and thus our condition is more general than the usual one. Moreover, notice that the sequence $(X_n)$ used in the proof of Theorem 4.1 does not satisfy $X_n \to_{L^1} X$.

Now we will show that under reasonable assumptions risk measures on non-atomic probability spaces are consistent with usual stochastic order and convex order.

**Theorem 4.2.** *Assume that $(\Omega, \mathcal{A}, P)$ is a standard probability space with $P$ non-atomic, and let the risk measure $\rho$ be monotone and law-invariant. Then $X \leq_{st} Y$ implies $\rho(X) \leq \rho(Y)$.*

*Proof.* As all non-atomic standard probability spaces are Borel isomorph, we can assume without loss of generality that $\Omega = (0,1)$ and that $P$ is the Lebesgue-measure. If $X$ and $Y$ are random variables with $X \leq_{st} Y$, then there are random variables $X'$ and $Y'$ on $(\Omega, \mathcal{A}, P)$ with the same distribution satisfying $X' \leq Y'$, see Theorem 2.2 and the subsequent remark. Therefore monotonicity and law-invariance of the risk measure imply

$$\rho(X) = \rho(X') \leq \rho(Y') = \rho(Y).$$

$\qquad\qquad\square$

**Theorem 4.3.** *Assume that $(\Omega, \mathcal{A}, P)$ is a standard probability space with $P$ non-atomic, and assume that the risk measure $\rho$ is convex, law-invariant and has the Fatou property. Then $X \leq_{cx} Y$ implies $\rho(X) \leq \rho(Y)$.*

*Proof.* We first define $X_1 := X$, $X_2 := Y$ and assume that $P_{X_1}$ is a fusion of $P_{X_2}$. This means that there are probability measures $P_1$ and $P_2$ such that

$$P_{X_1} = \alpha P_1 + (1 - \alpha)\delta_{\mu_{P_2}} \quad \text{and} \quad P_{X_2} = \alpha P_1 + (1 - \alpha)P_2,$$

Let $W, I, Y_1, Y_2, \dots$ be random variables with $P(I = 1) = \alpha = 1 - P(I = 0)$, $W$ having distribution $P_1$ and $Y_1, Y_2, \dots$ being a sequence of i.i.d. random variables having distribution $P_2$. Define $V_n := IW + (1 - I)Y_n$, $n = 1, 2, \dots$. Then $V_n$ has the same distribution as $X_2$ and $V_0 := IW + (1 - I)EY_1$ has the same distribution as $X_1$. Since

$$\frac{1}{n}\sum_{i=1}^{n} V_i = IW + (1 - I)\frac{1}{n}\sum_{i=1}^{n} Y_i,$$

it follows from the law of large numbers that

$$E\left|V_0 - \frac{1}{n}\sum_{i=1}^{n} V_i\right| = E\left|(1 - I)\frac{1}{n}\sum_{i=1}^{n}(Y_i - EY_1)\right|$$

$$\leq E\left|\frac{1}{n}\sum_{i=1}^{n}(Y_i - EY_1)\right| \to 0.$$

Thus law-invariance, convexity and the Fatou-property imply

$$\rho(X_1) = \rho(V_0) \leq \liminf_{n\to\infty} \rho(\frac{1}{n}\sum_{i=1}^{n} V_i) \leq \liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \rho(V_i) = \rho(V_1) = \rho(X_2).$$

Now consider arbitrary $X_1$ and $X_2$ with $X_1 \leq_{cx} X_2$. According to Theorem 2.8 there is a sequence of probability measures $(P_n)$ such that $P_0 = P_{X_2}$, $P_{n+1}$ is a fusion of $P_n$, and $P_n \to_w P_{X_1}$, and there exists a sequence $\{Z_n\}$ of random variables with distributions $\{P_n\}$ such that $Z_n \to_{L^1} Z$ for some $Z$ with the same distribution as $X_1$. It follows from the first part of the proof that $\rho(Z_{n+1}) \leq \rho(Z_n)$, and applying the Fatou-property once more yields

$$\rho(X_1) = \rho(Z) \leq \liminf_{n\to\infty} \rho(Z_n) \leq \rho(Z_0) = \rho(X_2).$$

$\square$

As in Theorem 3.5 we can combine Theorem 4.2 and Theorem 4.3 to get the following result.

**Theorem 4.4.** *Assume that $(\Omega, \mathcal{A}, P)$ is a standard probability space with $P$ non-atomic, and let $\rho$ be a monotone, convex and law-invariant risk measure with the Fatou property. Then $X \leq_{icx} Y$ implies $\rho(X) \leq \rho(Y)$.*

Of course the question arises, whether there exist risk measures on $L^1$ with these properties, in particular the Fatou property and how they look like. We will show now that it is possible to extend the characterization of Kusuoka (2001) of coherent, law invariant risk measures with the Fatou property from $L^\infty$ to $L^1$. To this end, we recall the definition of the risk measure *Conditional Value at Risk* (CVaR), which is defined for some fixed value $\alpha \in (0, 1]$ as

$$\rho_\alpha(X) := \frac{1}{\alpha} \int_{1-\alpha}^1 q_X(u) du.$$

We also set

$$\rho_0(X) := ess.sup(X).$$

Note that $\rho_\alpha : L^1 \to \mathbb{R} \cup \{\infty\}$, $\alpha \in [0, 1]$ is well-defined, $\rho_1(X) = EX$ and $\rho_\alpha(X)$ is decreasing in $\alpha$ (see e.g. Acerbi and Tasche (2002)). We obtain the following theorem:

**Theorem 4.5.** *Assume that $(\Omega, \mathcal{A}, P)$ is a standard probability space with $P$ non-atomic. A risk measure $\rho : L^1(\Omega, \mathcal{A}, P) \to \mathbb{R}$ is a law invariant coherent risk measure with the Fatou property if and only if the following representation holds:*

$$\rho(X) = \sup\{\int_0^1 \rho_\alpha(X)\mu(d\alpha); \mu \in P_0\}, \tag{4.1}$$

*where $P_0$ is a compact, convex set of probability measures on $[0, 1]$.*

*Proof.* First assume that $\rho$ is of the given form. Obviously $\rho$ is law invariant. Acerbi and Tasche (2002) have shown that $\rho_\alpha$ is coherent. As coherence is preserved under taking convex combinations and suprema, it follows that $\rho$ is coherent, too. Thus, it remains to show that $\rho$ satisfies the Fatou property, i.e. for $X, X_1, X_2, \ldots \in L^1$ with $X_n \to_p X$ and $\{X_n\}$ uniformly integrable, it follows that

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

First note that $X_n \to_p X$ implies $q_{X_n}(\alpha) \to q_X(\alpha)$ for almost all $\alpha \in (0, 1)$ and that $q_X(U)$ has distribution function $F_X$, where $U$ is a random variable with uniform distribution on $(0, 1)$. Thus, we obtain by assumption that $\{q_{X_n}(U)\}$ is uniformly integrable. Applying a standard result from probability theory we obtain for $\alpha \in (0, 1]$

$$\rho_\alpha(X) = \frac{1}{\alpha} \int_{1-\alpha}^1 q_X(u) du = \frac{1}{\alpha} \int_{1-\alpha}^1 \lim_{n \to \infty} q_{X_n}(u) du$$
$$= \lim_{n \to \infty} \rho_\alpha(X_n).$$

Also note that $X_n \to_p X$ implies $\liminf_{n \to \infty} \rho_0(X_n) \geq \rho_0(X)$ (but not necessarily equality). Next let $\mu \in P_0$ be arbitrary. Recall that $\alpha \mapsto \rho_\alpha(X)$ is decreasing with $\rho_1(X) = EX$.

14

Suppose that $N(\varepsilon)$ is such that $|EX_n - EX| \leq \varepsilon$ for all $n \geq N(\varepsilon)$. Thus, for all $n \geq N(\varepsilon)$ and all $\alpha \in [0,1]$ we have $\rho_\alpha(X_n) \geq EX - \varepsilon$ and the Lemma of Fatou implies

$$\liminf_{n \to \infty} \int_0^1 \rho_\alpha(X_n)\mu(d\alpha) \geq \int_0^1 \rho_\alpha(X)\mu(d\alpha).$$

Obviously we have for all $\mu' \in P_0$:

$$\sup\{\int_0^1 \rho_\alpha(X_n)\mu(d\alpha); \mu \in P_0\} \geq \int_0^1 \rho_\alpha(X_n)\mu'(d\alpha).$$

Taking the limes inferior on both sides yields

$$\liminf_{n \to \infty} \rho(X_n) \geq \liminf_{n \to \infty} \int_0^1 \rho_\alpha(X_n)\mu'(d\alpha) \geq \int_0^1 \rho_\alpha(X_n)\mu'(d\alpha).$$

Since this inequality holds for all $\mu' \in P_0$, taking the supremum over all $\mu' \in P_0$ implies the Fatou property of the risk measure $\rho$.

Now we have to show that any law invariant coherent risk measure $\rho'$ with the Fatou property is of the given form. Consider the restriction of $\rho'$ to $L^\infty$. By Theorem 4 in Kusuoka (2001) there is a representation as in (4.1), which we will denote by $\rho$. We have to show that $\rho = \rho'$ also holds for unbounded random variables in $L^1$. Suppose first $X \in L^1$ and $X$ is bounded from below. Consider the sequence $X_n := X \wedge n$. Thus, $X_n$ are bounded, $X_n \uparrow X$ and $\{X_n\}$ is uniformly integrable. From the monotonicity and the Fatou property of $\rho$ and $\rho'$ and the fact that $\rho = \rho'$ on $L^\infty$, we obtain

$$\rho(X) \leq \liminf_{n \to \infty} \rho(X_n) = \liminf_{n \to \infty} \rho'(X_n) \leq \rho'(X) \leq \liminf_{n \to \infty} \rho'(X_n) = \liminf_{n \to \infty} \rho(X_n) \leq \rho(X).$$

Thus, $\rho(X) = \rho'(X)$ for all $X \in L^1$ which are bounded from below. Next let $X \in L^1$ be arbitrary and consider $X_n := X \vee (-n)$. Obviously we have $X_n \downarrow X$, $X_n$ are bounded from below and $\{X_n\}$ is uniformly integrable. First notice that for this special sequence $\{X_n\}$ we have $\lim_{n \to \infty} \rho(X_n) = \rho(X)$. This can be seen as follows: fix $n$ large. Then we have $q_{X_n}(u) = q_X(u)$ for all $u \geq \alpha^*$ and $\rho_0(X_n) = \rho_0(X)$. Thus, $\rho_\alpha(X_n) = \rho_\alpha(X)$ for all $\alpha \leq 1 - \alpha^*$. Now for $\alpha > 1 - \alpha^*$ we obtain

$$|\rho_\alpha(X_n) - \rho_\alpha(X)| = \frac{1}{\alpha}\left|\int_{1-\alpha}^1 q_{X_n}(u) - q_X(u)du\right| \leq \frac{1}{1-\alpha^*}|EX - EX_n|.$$

The last expression can be made small enough by choosing $n$ large enough. Hence, $\rho_\alpha(X_n) \to \rho_\alpha(X)$ uniformly for $\alpha \in [0,1]$, which implies for all $\varepsilon > 0$ that there is an $N(\varepsilon)$ such that for all $n \geq N(\varepsilon)$ and $\mu \in P_0$:

$$\int_0^1 \rho_\alpha(X_n)\mu(d\alpha) \leq \int_0^1 \rho_\alpha(X)\mu(d\alpha) + \varepsilon.$$

15

Taking the supremum over all $\mu \in P_0$, letting $\varepsilon \downarrow 0$ and combining this with the Fatou property yields

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

and thus $\lim_{n \to \infty} \rho(X_n) = \rho(X)$.

Now we show that indeed $\rho(X) = \rho'(X)$. We have $X_n = X + Y_n$ with $Y_n := (-X - n) \vee 0$. By subadditivity of $\rho'$ we get

$$\rho(X_n) = \rho'(X_n) \leq \rho'(X) + \rho'(Y_n) = \rho'(X) + \rho(Y_n).$$

Taking $n \to \infty$ and noting that $Y_n \to_{L^1} 0$ and $\rho(0) = 0$ we obtain $\rho(X) \leq \rho'(X)$. The Fatou property of $\rho'$ now implies

$$\rho'(X) \leq \liminf_{n \to \infty} \rho'(X_n) = \liminf_{n \to \infty} \rho(X_n) = \rho(X)$$

and $\rho = \rho'$ follows. $\qquad\square$

Now in case of a coherent risk measure an alternative proof of Theorem 4.4 is possible, based on the characterization in Theorem 4.5.

**Corollary 4.6.** *Assume that $(\Omega, \mathcal{A}, P)$ is a standard probability space with $P$ non-atomic. Let $\rho : L^1(\Omega, \mathcal{A}, P) \to \mathbb{R}$ be a law invariant coherent risk measure with the Fatou property. Then $X \leq_{icx} Y$ implies $\rho(X) \leq \rho(Y)$.*

*Proof.* It follows from part (iv) in Theorem 2.5 and Theorem 2.2 resp., that $X \leq_{st} Y$ $[X \leq_{cx} Y]$ implies $\rho_\alpha(X) \leq \rho_\alpha(Y)$ for all $\alpha$, and therefore Theorem 4.5 together with Theorem 2.4 immediately implies the result. $\qquad\square$

A result similar to Corollary 4.6 can be found in Leitner (2004). There a so called *balayage ordering* is considered, which is some sort of a distribution free version of convex ordering.

# 5 Bounds for Coherent Risk Measures

The results of the previous sections can be used to bound the risk of portfolios. Suppose $X_1, \ldots, X_n \in L^1$ constitute financial risks. Recall that $X_1, \ldots, X_n$ are called *comonotone*, if there exist increasing functions $f_1, \ldots, f_n$ and a random variable $U$, uniformly distributed over $(0, 1)$ such that $(X_1, \ldots, X_n) =_{st} (f_1(U), \ldots, f_n(U))$. In what follows we denote for arbitrary random variables $X_1, \ldots, X_n$ by $X_1^c, \ldots, X_n^c$ some comonotonic random variables having the same marginal distributions.

**Lemma 5.1.** *For all monotone, convex and law invariant risk measures $\rho$ with the Fatou property and all functions $\psi : \mathbb{R}^n \to \mathbb{R}$ increasing and supermodular we have*

$$\rho(\psi(X_1, \ldots, X_n)) \leq \rho(\psi(X_1^c, \ldots, X_n^c)).$$

*Proof.* It is well-known that $(X_1, \ldots, X_n) \leq_{sm} (X_1^c, \ldots, X_n^c)$ which implies that

$$\psi(X_1, \ldots, X_n) \leq_{icx} \psi(X_1^c, \ldots, X_n^c)$$

for $\psi$ increasing and supermodular (see e.g. Müller and Stoyan (2002) section 3.9). The result thus follows from Theorem 4.4. □

Of course, as far as portfolios are concerned, the function $\psi(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i$ is of particular interest. In this case we can dispense with the assumption of monotonicity of the risk measure. With a proof similar to the one of Lemma 5.1 the get the following result.

**Lemma 5.2.** *For all convex, law invariant risk measures $\rho$ with the Fatou property we have*

$$\rho\left(\sum_{i=1}^{n} X_i\right) \leq \rho\left(\sum_{i=1}^{n} X_i^c\right).$$

Thus, we obtain a tight upper bounds for convex risk measures in the following sense: Suppose $X = (X_1, \ldots, X_n)$ has distribution function $F$ and $F \in \mathcal{F}(F_1, \ldots, F_n)$, i.e. the marginal distributions $F_1, \ldots, F_n$ are known, but not the copula of $X$. This is a situation encountered very often in practice. Then $\rho(\sum_{i=1}^{n} X_i) \leq \rho(\sum_{i=1}^{n} X_i^c)$ and the upper bounds are attained for an $F \in \mathcal{F}(F_1, \ldots, F_n)$, namely if the copula of $F$ is the Fréchet-Hoeffding upper bound $M(x_1, \ldots, x_n) = \min\{x_1, \ldots, x_n\}$ (for a treatment of copulas, see e.g. Nelsen (1999)). As shown by Embrechts et al. (2003), the situation is much more complicated for the very popular risk measure Value-at-Risk, which is not convex.

Of course in the case of a coherent risk measure, the subadditivity axiom gives us a trivial upper bound for the sum:

$$\rho\left(\sum_{i=1}^{n} X_i\right) \leq \sum_{i=1}^{n} \rho(X_i).$$

One is tempted to be satisfied with this bound. But how good is it really? The following example shows that it is not tight and can even be arbitrary bad.

**Example 5.3.** We consider as a risk measure the so-called standard semideviation:

$$\rho(X) = EX + \alpha \left( E\left\{ (X - EX)^+ \right\}^2 \right)^{1/2}, \quad 0 < \alpha \leq 1.$$

$\rho$ is law-invariant, coherent and obviously satisfies the Fatou property (see e.g. Ogryczak and Ruszczyński (2002)). However, $\rho$ is not comonotone additive: take e.g. $X \sim \mathcal{N}(0,1)$ and $Y := X^3$. Then $X$ and $Y$ are comonotone and

$$\rho(X) + \rho(Y) - \rho(X + Y) =: \varepsilon > 0.$$

Thus, if we take $\lambda > 0$, then $\lambda X$ and $\lambda Y$ are comonotone and

$$\rho(\lambda X) + \rho(\lambda Y) - \rho(\lambda X + \lambda Y) = \lambda \varepsilon$$

which tends to $\infty$ if $\lambda$ tends to $\infty$.

Nevertheless, there are some important cases where both bounds coincide. Trivially this holds true if $\rho$ is comonotone additive. This property is for example satisfied by the CVaR $\rho_\alpha$ for all $\alpha \in (0,1]$. On the other hand, if we have some further information about the class of distributions $X$ is in, the bounds can also coincide, irrespective of the chosen risk measure $\rho$.

**Example 5.4.** Suppose $X$ has distribution function $F$ and $F \in \mathcal{F}(X_1, \ldots, X_n) \cap E_n$, i.e. $X$ has an elliptical distribution and the marginal distributions are known (for details about elliptical distributions see Fang et al. (1990)). How does the comonotone version of $X$ look like? If $X$ is elliptically distributed then the marginals are necessary mixtures of normal distributions, i.e. we have for $i = 1, \ldots, n$:

$$X_i = SY_i + \mu_i$$

where $Y_i \sim \mathcal{N}(0, \sigma_i^2), \mu_i \in \mathbb{R}$ and $S$ is an arbitrary random variable with support $[0, \infty)$. Thus, the comonotone version is of the form

$$\begin{aligned} X_1^c &= SY_1 + \mu_1 \\ X_i^c &= c_i(SY_1 + \frac{\mu_i}{c_i}) = c_i X_1 + d_i \end{aligned}$$

where $c_i = \frac{\sigma_i}{\sigma_1} > 0$ and $d_i = \mu_i - \mu_1 c_i$, for $i = 1, \ldots, n$. This means that the $X_i^c$ are linear transformations of $X_1^c$ and thus we obtain for any coherent law invariant risk measure

$$\rho\left(\sum_{i=1}^n X_i^c\right) = \rho(cX_1 + d) = c\rho(X_1) + d = \sum_{i=1}^n \rho(X_i),$$

where $c = \sum c_i$ and $d = \sum d_i$.

# 6 Portfolio Optimization

Let us look at the classical portfolio optimization problem of Markowitz type where we replace the variance by a coherent risk measure $\rho$. The random variables $X_1, \ldots, X_N$ represent the outcome of $N$ financial positions which can be chosen for investment. Thus we have "risks" $-X_1, \ldots, -X_N$. In what follows we suppose that only the marginal distributions of $X_1, \ldots, X_N$ are known and that the set $P_0$ of $\rho$ is the convex hull of finitely many probability measures $P_0 = conv\{\mu_1, \ldots, \mu_M\}, M \in \mathbb{N}$, where $\mu_i, i = 1, \ldots, M$ are discrete probability measures on $0, \frac{1}{K}, \ldots, \frac{K-1}{K}, 1$. Note that every compact, convex set $P_0$ of probability measures on $[0,1]$ can be approximated arbitrarily close by such a set. Since the joint distribution of $\mathbf{X} = (X_1, \ldots, X_N)$ is not known, we look at the conservative optimization problem which minimizes the upper bound, i.e.

$$\begin{cases} -E(\boldsymbol{\lambda}'\mathbf{X}) + \rho(-\boldsymbol{\lambda}'\mathbf{X}^c) & \to \min \\ \sum_{i=1}^{N} \lambda_i = 1 \\ \lambda_i \geq 0 & i = 1, \ldots, N, \end{cases}$$

with $\rho(X) = \sup\{\int_0^1 \rho_\alpha(X)\mu(d\alpha); \mu \in P_0\}$ and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)$. In what follows define $\rho_{ij} := \rho_{\frac{j}{K}}(-X_i)$ and $\mu_{jl} := \mu_l(\frac{j}{K})$ for $i = 1, \ldots, N, \; j = 0, \ldots, K, \; l = 1, \ldots, M$. The computation of $\rho_{ij}$ if $X_i$ is discrete can be done quite efficiently, see e.g. Rockafellar and Uryasev (2002), Proposition 8. Thus, we obtain the optimization problem

$$\begin{cases} -\sum_{i=1}^{N} \lambda_i EX_i + \sup\{\sum_{i=1}^{N}\sum_{j=0}^{K} \lambda_i \rho_{ij}\mu(\frac{j}{K}); \mu \in P_0\} & \to \min \\ \sum_{i=1}^{N} \lambda_i = 1 \\ \lambda_i \geq 0, \; i = 1, \ldots, N, \end{cases}$$

which is similar to a matrix game and can be solved by the following linear program:

$$\begin{cases} -\sum_{i=1}^{N} \lambda_i EX_i + w & \to \min \\ \sum_{i=1}^{N}\sum_{j=0}^{K} \rho_{ij}\lambda_i\mu_{jl} \leq w, & l = 1, \ldots, M \\ \sum_{i=1}^{N} \lambda_i = 1 \\ \lambda_i \geq 0 & i = 1, \ldots, N, \end{cases}$$

with variables $\lambda_1, \ldots, \lambda_N$ and $w$. Thus, in the common situation that only the distribution of $X_1, \ldots, X_n$ is known but not the copula, the general Markowitz type portfolio optimization problem with conservative upper bound can be solved quite efficiently by a linear program.

# 7 Conclusion

As a main result in this paper we have shown that under weak assumptions on the underlying probability space convexity of a risk measure implies consistency with respect to convex order. As convex ordering has the intuitively appealing meaning that any risk averse decision maker prefers the risk which is smaller in this ordering, we consider this consistency an important property of a risk measure, which is more intuitive than the convexity of a risk measure as introduced by Föllmer and Schied (2002). Therefore it could be an interesting topic of future research to replace the axiom of convexity by the weaker axiom of consistency with respect to convex order, and to investigate properties and characterizations of the resulting class of risk measures.

# References

ACERBI, C. and TASCHE, D. (2002) On the coherence of Expected Shortfall. *J. Banking and Finance* **26**, 1487–1503.

ARTZNER, P., DELBAEN, F., EBER, J.-M., and HEATH, D. (1999) Coherent measures of risk. *Math. Finance* **9**, 203–228.

DE GIORGI, E. (2005) Reward-Risk Portfolio Selection and Stochastic Dominance. *J. Banking and Finance* to appear.

DELBAEN, F. (2002) Coherent risk measures on general probability spaces. In *Advances in finance and stochastics*, 1–37. Springer, Berlin.

ELTON, J. and HILL, T. P. (1992) Fusions of a probability distribution. *Ann. Probab.* **20**, 421–454.

EMBRECHTS, P., HÖING, A., and JURI, A. (2003) Using copulae to bound the value-at-risk for functions of dependent risks. *Finance Stoch.* **7**, 145–167.

FANG, K. T., KOTZ, S., and NG, K. W. (1990) *Symmetric multivariate and related distributions*, volume 36 of *Monographs on Statistics and Applied Probability*. Chapman and Hall Ltd., London.

FÖLLMER, H. and SCHIED, A. (2002) Convex measures of risk and trading constraints. *Finance Stoch.* **6**, 429–447.

GOOVAERTS, M. J., DE VIJLDER, F., and HAEZENDONCK, J. (1984) *Insurance Premiums.* North Holland, Amsterdam.

GOOVAERTS, M. J., KAAS, R., DHAENE, J., and TANG, Q. (2004) Some new classes of consistent risk measures. *Insurance Math. Econom.* **34**, 505–516.

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

LEITNER, J. (2004) Balayage monotonous risk measures. *Int. J. Theor. Appl. Fin.* **7**, 887–900.

MARSHALL, A. W. and OLKIN, I. (1979) *Inequalities: theory of majorization and its applications.* Academic Press, New York.

MÜLLER, A. (1996) Orderings of risks: a comparative study via stop-loss transforms. *Insurance Math. Econom.* **17**, 215–222.

MÜLLER, A. and STOYAN, D. (2002) *Comparison Methods for Stochastic Models and Risks.* Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester.

NELSEN, R. B. (1999) *An Introduction to Copulas*, volume 139 of *Lecture Notes in Statistics.* Springer-Verlag, New York.

OGRYCZAK, W. and RUSZCZYŃSKI, A. (1999) From stochastic dominance to mean-risk models: Semideviations as risk measures. *European J. Oper. Res.* **116**, 33–50.

OGRYCZAK, W. and RUSZCZYŃSKI, A. (2001) On consistency of stochastic dominance and mean-semideviation models. *Math. Program.* **89**, 217–232.

OGRYCZAK, W. and RUSZCZYŃSKI, A. (2002) Dual stochastic dominance and related mean-risk models. *SIAM J. Optim.* **13**, 60–78.

PFLUG, G. C. (1999) How to measure risk? In LEOPOLD-WILDBURGER, U., FEICHTINGER, G., and KISTNER, K.-P. (eds.), *Modelling and decisions in economics*, 39–59. Physica-Verlag, Heidelberg.

ROCKAFELLAR, R. and URYASEV, S. (2002) Conditional value-at-risk for general loss distributions. *J. Banking and Finance* **26**, 1443–1471.

ROTHSCHILD, M. and STIGLITZ, J. E. (1970) Increasing risk. I. A definition. *J. Econom. Theory* **2**, 225–243.

SCHIED, A. (2004) On the Neyman-Pearson problem for law-invariant risk measures and robust utility functionals. *Ann. Appl. Probab.* **14**, 1398–1423.

SHAKED, M. and SHANTHIKUMAR, J. G. (1994) *Stochastic Orders and their Applications.* Probability and Mathematical Statistics. Academic Press Inc., Boston, MA.