

A quantitative comparison of risk measures

Alois Pichler¹

Published online: 13 January 2017
© Springer Science+Business Media New York 2017

Abstract The choice of a risk measure reflects a subjective preference of the decision maker in many managerial or real world economic problem formulations. To assess the impact of personal preferences it is thus of interest to have comparisons with other risk measures at hand. This paper develops a framework for comparing different risk measures. We establish a one-to-one relationship between norms and risk measures, that is, we associate a norm with a risk measure and conversely, we use norms to recover a genuine risk measure. The methods allow tight comparisons of risk measures and tight lower and upper bounds for risk measures are made available whenever possible. In this way we present a general framework for comparing risk measures with applications in numerous directions.

Keywords Risk measures · Dual representation · Fenchel–Young inequality

JEL Classification 90C15 · 60B05 · 62P05

1 Introduction

Risk measures are designed to quantify the risk which is associated with a random, uncertain outcome in a single real number. Risk measures have been considered in insurance first to price insurance contracts, but nowadays they constitute an essential basis for decision making and management in all areas of operations research, management science and (mathematical) finance whenever unobserved, random, future outcomes are involved.

This paper addresses comparisons of risk measures and continuity relations by establishing lower and upper bounds. We give tight relations wherever possible so that stochastic programs, e.g., can be compared easily when employing different risk measures. A sensitivity analysis thus can assess a subjective preference of a particular risk measure.

✉ Alois Pichler
alois.pichler@mathematik.tu-chemnitz.de

¹ Technische Universität Chemnitz, Fakultät für Mathematik, Chemnitz, Germany

Our research is motivated by [Iancu et al. \(2015\)](#), but in addition to the methods outlined in this reference we involve convex functions (as norms or functionals derived from coherent risk measures, e.g.) to obtain explicit bounds. Some of these convex functionals perhaps violate usual axioms of coherent risk measures, but they are useful and eligible for comparisons or efficient in computations.

Risk measures are usually defined on vector spaces of real valued random variables and these vector spaces typically come with a norm. It has been observed that risk measures induce a norm themselves, which is occasionally equivalent to the genuine norm of the domain space. Here, we establish and employ the converse relations, that is, we develop a natural relationship between norms and risk measures so that both concepts can be exchanged against each other or employed alternatively to specify the other. This equivalence is a major tool of our analysis and in this way the risk measure naturally compares with the genuine norm and with the norm induced. To illustrate the methods we provide many examples of precise relations between risk measures and norms, and between different risk measures, thus supporting the analysis of personal preferential choices of risk measures.

Our results are related to [Wozabal \(2014\)](#), who robustifies risk measures with respect to changing the underlying probability space. [López-Díaz et al. \(2012\)](#) employ a specific norm (called an L_p -metric there), while [Bellini and Caperdoni \(2007\)](#), in contrast, relate and compare risk measures with stochastic dominance relations. The papers [Wozabal \(2010, 2014\)](#) contain typical applications in operations research and finance, as asset allocation.

Risk measures impose several difficulties in a multistage stochastic framework. [Cheridito and Kupper \(2011\)](#) investigate a natural composition of risk measures. This concept, or approach, is formalized by [Lara and Leclère \(2016\)](#). [Asamov and Ruszczyński \(2014\)](#), [Ruszczyński \(2010\)](#) and [Miller and Ruszczyński \(2011\)](#) elaborate on risk measures in relation to dynamic, multistage stochastic programming. Applications of these extended concepts of risk measures are finally provided, for example, by [Philpott and Matos \(2012\)](#) and [Philpott et al. \(2013\)](#), who involve compositions of risk measures to formulate and elaborate on multistage hydro-thermal scheduling problems in New Zealand and Brazil.

Outline of the paper. The following section (Sect. 2) repeats the axioms for risk measures and formulates corresponding axioms for norms. The Sects. 3 and 4 present explicit comparisons with norms and with other risk measures. We develop bounds on composite risk measures in Sect. 5, while Sect. 6 concludes with a comprehensive summary.

2 Preliminaries

We consider a linear vector space $L \subseteq L^1(\Omega, \mathcal{F}, P)$ of \mathbb{R} -valued random variables with pointwise ordering. The set L is used in this paper to represent random losses. The vector space $L \subseteq L^1$ is the domain of a risk measure. We state the axioms of risk measure here in convex form, as we shall relate them to further axioms on (convex) norms on the domain space L later.

Definition 2.1 (*Axioms for risk functionals*) A version independent, coherent risk measure is a mapping $\rho : L \rightarrow \mathbb{R} \cup \{\infty\}$ satisfying the following axioms:

- (RM) Monotonicity: $\rho(Y_1) \leq \rho(Y_2)$ whenever $Y_1 \leq Y_2$ almost surely;
- (RC) Convexity $\rho((1 - \lambda)Y_0 + \lambda Y_1) \leq (1 - \lambda)\rho(Y_0) + \lambda\rho(Y_1)$ for $0 \leq \lambda \leq 1$;
- (RT) Translation equivariance: $\rho(Y + c \cdot \mathbb{1}) = \rho(Y) + c$ if $c \in \mathbb{R}$ ($\mathbb{1}(\cdot) = 1$ is the constant random variable);
- (RH) Positive homogeneity: $\rho(\lambda Y) = \lambda \rho(Y)$ whenever $\lambda > 0$;

(RD) Version independence: $\rho(Y_1) = \rho(Y_2)$, provided that Y_1 and Y_2 share the same law, i.e., $P(Y_1 \leq y) = P(Y_2 \leq y)$ for all $y \in \mathbb{R}$.¹

We shall call the set

$$L_+ := \{Y \in L : Y \geq 0 \text{ a.s.}\}$$

the *nonnegative orthant*, or the *nonnegative cone* of L .

If only the properties (RM), (RC) and (RT) are fulfilled, then the functional ρ is often simply called *risk functional* as well, while positively homogeneous risk measures satisfying (RH) are also called *coherent* risk measures. In this paper we shall address the specific properties explicitly in the given context, if this is necessary.

The trivial risk measures satisfying the axioms (RM)–(RD) are the expectation $\rho(Y) = \mathbb{E}[Y]$ and the max-risk functional $\rho(Y) = \text{ess sup } Y$. An example of a risk functional that is not necessarily version independent is $\rho_Z(Y) := \mathbb{E}[Z \cdot Y]$ (with $\mathbb{E}[Z] = 1$ and $Z \geq 0$ almost everywhere).

Norms. We equip the vector space $L \subseteq L^1$ with a norm and relate risk functionals with the normed space $(L, \|\cdot\|)$. In line with the axioms presented above on risk measures we make use of the following assumptions on the norm $\|\cdot\|$ and the (Banach) space $(L, \|\cdot\|)$.

- (NM) Monotonicity: $\|Y_1\| \leq \|Y_2\|$ whenever $|Y_1| \leq |Y_2|$ almost surely;
- (N1) Normalization (scaling): $\|\mathbb{1}\| = 1$;
- (ND) Density: uniformly bounded random variables and L^∞ are dense in L with respect to the norm $\|\cdot\|$;
- (NP) Representation on the nonnegative cone:²

$$\sup_{\|Z\|^* \leq 1} \mathbb{E}[Y \cdot Z] = \sup_{Z \geq 0, \|Z\|^* \leq 1} \mathbb{E}[Y \cdot Z] \text{ whenever } Y \geq 0 \text{ a.s.}, \tag{1}$$

where

$$\|Z\|^* := \sup_{\|Y\| \leq 1} \mathbb{E}[Y \cdot Z] \tag{2}$$

is the norm on the dual, which is denoted $(L^*, \|\cdot\|^*)$.

Remark 2.2 Note that (NP) is not more than the Hahn–Banach theorem for the bi-linear form $(Y, Z) \mapsto \mathbb{E}[Y Z]$, just restricted to the nonnegative orthant L_+ .

From normalization (N1) and (2) it follows that

$$\|Z\|^* \geq \mathbb{E}[Z], \tag{3}$$

and particularly that $\|\mathbb{1}\|^* \geq \mathbb{E}[\mathbb{1}] = 1$. Further it is evident from (1) that

$$\mathbb{E}[Y] \leq \|\mathbb{1}\|^* \cdot \|Y\|. \tag{4}$$

Specifications of the norm. Obvious candidates for a norm satisfying all relations (NM)–(NP) are the norms $\|\cdot\|_p$ on L^p -spaces. Further examples include the Luxembourg norm

$$\|Y\|_\Phi := \inf \left\{ \xi > 0 : \mathbb{E} \left[\Phi \left(\frac{|X|}{\xi} \right) \right] \leq 1 \right\}$$

on Orlicz spaces, where $\Phi : [0, \infty) \rightarrow [0, \infty)$ is a convex function satisfying $\Phi(0) = 0$, $\Phi(1) = 1$ and $\lim_{x \rightarrow \infty} \Phi(x) = \infty$ (cf. Bellini and Rosazza Gianin 2012, for example, for details). Further, we mention Orlicz hearts, the spaces $(L_\sigma, \|\cdot\|_\sigma)$ and their dual $(L_\sigma^*, \|\cdot\|_\sigma^*)$, as well as Lorentz spaces, they satisfy all specified relations as well (cf. Pichler 2013a).

¹ Also law invariant, or distribution based.

² (NP) is mnemonic for positive.

3 Relations between norms and risk measures

Given a risk measure ρ , then one may associate the function

$$\|\cdot\|_\rho := \rho(|\cdot|) \tag{5}$$

with ρ . The risk functional ρ is Lipschitz continuous with respect to the associated norm $\|\cdot\|_\rho$ (cf. Pichler 2013a).

In this chapter we elaborate the converse construction by providing explicitly a risk measure, which is based on a norm. We demonstrate that the risk measure is completely specified by the nonnegative cone L_+ . We further demonstrate in this section that the risk measure obtained is again (Lipschitz-)continuous with respect to the initial norm.

Theorem 3.1 *Let $\|\cdot\|$ be a norm satisfying (NM)–(NP). Then, for $c \geq 1$, the higher order risk measure³*

$$\rho_c^{ho}(Y) := \inf_{x \in \mathbb{R}} \{x + c \cdot \|(Y - x)_+\|\} \tag{6}$$

is a coherent risk measure. $\rho_c^{ho}(\cdot)$ is version independent, iff the norm is version independent. Its dual representation is

$$\rho_c^{ho}(Y) = \sup \{ \mathbb{E}[Y Z] : Z \geq 0, \mathbb{E}[Z] = 1 \text{ and } \|Z\|^* \leq c \}. \tag{7}$$

Remark 3.2 We point out that (7) specifies the risk measure $\rho_c^{ho}(\cdot)$ by considering the ball of radius c in the dual space. This emphasizes the strong relationship between a risk measure and a norm.

Remark 3.3 We mention Krokmal (2007) for a construction leading to (6). The functional

$$\rho_\alpha(Y) := \inf_{x \in \mathbb{R}} \left\{ x + \frac{1}{1 - \alpha} \cdot \|(Y - x)_+\|_\Phi \right\} \tag{8}$$

corresponding to the Luxembourg norm is also called *Haезendonck–Goovaerts* risk measure (or *Haезendonck–Goovaerts* premium in insurance). Bellini and Rosazza Gianin (2008, 2012) involve (8) to define generalized α -quantiles of the random variable Y . Dentecheva et al. (2010) establish the duality relation of (6)–(7) for the L_p -norm $\|\cdot\|_p$ in this journal, which is called *higher order* risk measure (thus the name and superscript, ρ^{ho}).

Remark 3.4 The constraint $c \geq 1$ is a necessary condition, as otherwise the risk measure ρ_c^{ho} is degenerate and, e.g., $\rho_c^{ho}(\mathbb{1}) = -\infty$.

Proof of Theorem 3.1 Observe first that

$$c \cdot \|(Y - x)_+\| = \sup_{\|Z\|^* \leq c} \{ \mathbb{E}[(Y - x)_+ \cdot Z] \} \geq \sup_{\substack{\mathbb{E}[Z] = 1, \\ Z \geq 0, \|Z\|^* \leq c}} \{ \mathbb{E}[(Y - x)_+ \cdot Z] \}$$

³ $x_+ := \max \{0, x\}$.

by the Hahn–Banach theorem, and thus

$$\begin{aligned} x + c \cdot \|(Y - x)_+\| &\geq \sup_{\substack{\mathbb{E}[Z] = 1, \\ Z \geq 0, \|Z\|^* \leq c}} \mathbb{E}[(x + (Y - x)_+) \cdot Z] \\ &\geq \sup_{\substack{\mathbb{E}[Z] = 1, \\ Z \geq 0, \|Z\|^* \leq c}} \mathbb{E}[Y \cdot Z], \end{aligned}$$

as $x + (Y - x)_+ \geq Y$, establishing thus the first inequality relation.

As for the converse inequality we shall assume first that Y is bounded. It follows that

$$\sup_{\substack{\mathbb{E}[Z] = 1, \\ Z \geq 0, \|Z\|^* \leq c}} \mathbb{E}[YZ] = \sup_{\substack{Z \geq 0, \\ \|Z\|^* \leq c}} \inf_{x \in \mathbb{R}} \{x + \mathbb{E}[(Y - x)Z]\}, \tag{9}$$

as one may drop the constraint $\mathbb{E}[Z] = 1$: indeed,

$$\inf_{x \in \mathbb{R}} \{x + \mathbb{E}[(Y - x)Z]\} = \mathbb{E}[YZ] + \inf_{x \in \mathbb{R}} \{x \cdot (1 - \mathbb{E}[Z])\} = -\infty$$

unless $\mathbb{E}[Z] = 1$, such that the inf does not contribute to the outer sup if $\mathbb{E}[Z] \neq 1$.

Observe next that $\inf_{x \in \mathbb{R}} \{x + \mathbb{E}[(Y - x)Z]\} = x^* + \mathbb{E}[(Y - x^*)_+ \cdot Z]$ for $x^* := \text{ess inf } Y$ (recall that Y is bounded by assumption so that this quantity is well-defined). So it follows that

$$\sup_{\substack{\mathbb{E}[Z] = 1, \\ Z \geq 0, \|Z\|^* \leq c}} \mathbb{E}[YZ] = \sup_{\substack{Z \geq 0, \\ \|Z\|^* \leq c}} x^* + \mathbb{E}[(Y - x^*)_+ \cdot Z], \tag{10}$$

and by the Hahn–Banach theorem and (NP) again (as $(Y - x^*)_+ \geq 0$) that

$$(10) = x^* + c \cdot \|(Y - x^*)_+\| \geq \inf_{x \in \mathbb{R}} \{x + c \cdot \|(Y - x)_+\|\}, \tag{11}$$

the remaining inequality.

If Y is not bounded, then for every $\varepsilon > 0$ there is a bounded random variable Y_ε so that $\|Y - Y_\varepsilon\| < \varepsilon$ due to the assumption (ND). It follows by combining (9) and (11) that

$$\begin{aligned} \sup_{\substack{\mathbb{E}[Z] = 1, \\ Z \geq 0, \|Z\|^* \leq c}} \mathbb{E}[YZ] + c\varepsilon &\geq \sup_{\substack{\mathbb{E}[Z] = 1, \\ Z \geq 0, \|Z\|^* \leq c}} \mathbb{E}[Y_\varepsilon Z] \\ &\geq \inf_{x \in \mathbb{R}} \{x + c \cdot \|(Y_\varepsilon - x)_+\|\} \geq \inf_{x \in \mathbb{R}} \{x + c \cdot \|(Y - x)_+\|\} - c\varepsilon. \end{aligned}$$

The assertion is immediate, as $\varepsilon > 0$ is arbitrary. □

The result (7) of the previous theorem is a dual characterization. The results allows a natural comparison of risk measures by employing the genuine norm. We shall elaborate corresponding bounds (lower and upper bounds) in the following statement. These results provide a tool to prove continuity in the sequel.

Theorem 3.5 (Bounds, and comparison with norms) *Let the risk measure $\rho_c^{ho}(\cdot)$ be induced by a norm as specified in (6). Then it holds that*

$$\|Y\| \leq \rho_c^{ho}(|Y|) \leq c \cdot \|Y\|.$$

The lower bound is sharp.

Proof As for the first inequality recall the elementary inequality $y \leq x + (y - x)_+ \leq x + c(y - x)_+$ for $c \geq 1$, and hence

$$|Y| \leq x + c(|Y| - x)_+.$$

It follows from monotonicity (NM) and the triangle inequality that

$$\|Y\| \leq \|x + c(|Y| - x)_+\| \leq x + c\|(|Y| - x)_+\|.$$

By passing to the infimum thus

$$\|Y\| \leq \rho_c^{ho}(|Y|).$$

Equality moreover is obtained for the constant random variable, $Y = \mathbf{1}$ and $x = 1$ in (6).

We employ the dual representation (7) to verify the second inequality. It follows from the duality relation (NP) that

$$\begin{aligned} \rho_c^{ho}(|Y|) &= \sup \{ \mathbb{E}[|Y|Z] : Z \geq 0, \mathbb{E}[Z] = 1 \text{ and } \|Z\|^* \leq c \} \\ &\leq \sup \{ \|Y\| \cdot \|Z\|^* : \|Z\|^* \leq c \} \\ &= c \cdot \|Y\|, \end{aligned}$$

the remaining assertion. □

Remark 3.6 (The special case $c = 1$: specification on the nonnegative cone L_+) The risk measure ρ_1^{ho} is completely specified by the norm $\|\cdot\|$ on the nonnegative cone, as

$$\rho_1^{ho}(Y) = \|Y\| \text{ for all } Y \in L_+$$

by choosing $c = 1$ in the latter theorem. Together with (5) this equation establishes the one-to-one relationship between risk measures and norms.

It is thus enough to specify a risk measure ρ on the nonnegative cone L_+ , as it extends to the entire space by setting

$$\rho_1^{ho}(Y) := \inf_{x \in \mathbb{R}} \{ x + \rho((Y - x)_+) \}. \tag{12}$$

This setting carries translation equivariance (RT) to the entire space L and it holds that $\rho(Y) = \rho_1^{ho}(Y)$ for $Y \in L_+$ [cf. also the inf convolution in e.g., Pflug and Römisch (2007, Section 2.4.3); a related construction is the homogenization outlined in Shapiro et al. (2009, Section 6.3.2) carrying homogeneity (RH) to the entire domain].

We finally remark that a local comparison on the nonnegative cone L_+ extends to a global comparison on $L \supseteq L_+$. Indeed, if $\rho(Y) \leq \rho'(Y)$ for all $Y \in L_+$, then $\rho(Y) \leq \rho'(Y)$ for $Y \in L$ by (12).

Corollary 3.7 (Continuity) *The risk measure ρ_c^{ho} is Lipschitz continuous with respect to its norm,*

$$\left| \rho_c^{ho}(Y_2) - \rho_c^{ho}(Y_1) \right| \leq c \cdot \|Y_2 - Y_1\|$$

for all $Y_1, Y_2 \in L$.

Proof Just observe that $\rho_c^{ho}(Y_2) = \rho_c^{ho}(Y_2 - Y_1 + Y_1) \leq \rho_c^{ho}(Y_2 - Y_1) + \rho_c^{ho}(Y_1)$ by convexity, and thus

$$\rho_c^{ho}(Y_2) - \rho_c^{ho}(Y_1) \leq \rho_c^{ho}(Y_2 - Y_1) \leq \rho_c^{ho}(\|Y_2 - Y_1\|) \leq c \cdot \|Y_2 - Y_1\|.$$

The assertion follows by interchanging the roles of Y_1 and Y_2 . □

Average Value-at-Risk The (upper) Average Value-at-Risk (AV@R) is the special case of the latter theorem for the norm $\|\cdot\|_1$. The equivalent expressions

$$\begin{aligned} \text{AV@R}_\alpha(Y) &:= \inf_{x \in \mathbb{R}} \left\{ x + \frac{1}{1 - \alpha} \mathbb{E}[(Y - x)_+] \right\} \\ &= \sup \left\{ \mathbb{E}[YZ] : 0 \leq Z \leq \frac{1}{1 - \alpha} \text{ and } \mathbb{E}[Z] = 1 \right\} \end{aligned} \tag{13}$$

are well-known indeed (cf. Rockafellar and Uryasev 2000; Pflug 2000), but here they constitute a special case of Theorem 3.1 for the norm $\|\cdot\|_1$ and its dual $\|\cdot\|_1^* = \|\cdot\|_\infty$.

3.1 Higher order risk measures

Higher order risk measures constitute a special case of the risk measure (6) addressed in Theorem 3.1, as the L^p -norm $\|\cdot\|_p$ takes the role of the general norm $\|\cdot\|$. We denote them by

$$\rho_{c,p}^{ho}(Y) := \inf_{x \in \mathbb{R}} \left\{ x + c \cdot \|(Y - x)_+\|_p \right\}, \tag{14}$$

where $c \geq 1$ and $p \geq 1$.

For this particular choice we have the following extension of Theorem 3.1 with precise bounds.

Theorem 3.8 (Comparison with norms) *Assume that $1 \leq p \leq p'$. It holds that*

(i)
$$\|Y\|_p \leq \rho_{c,p'}^{ho}(\|Y\|), \tag{15}$$

(ii) and
$$\rho_{c,p}^{ho}(Y) \leq c^{p/p'} \cdot \|Y\|_{p'}. \tag{16}$$

Both bounds are moreover sharp.

Proof The first inequality is immediate from Theorem 3.5. For the second inequality recall the dual representation (7) (see also Dentcheva et al. 2010) that

$$\rho_{c,p}^{ho}(\|Y\|) = \sup \left\{ \mathbb{E}[\|Y\|Z] : \|Z\|_q \leq c, Z \geq 0 \text{ and } \mathbb{E}[Z] = 1 \right\},$$

where q is the Hölder conjugate exponent, $\frac{1}{p} + \frac{1}{q} = 1$.

Let q' be the conjugate exponent to p' ($\frac{1}{p'} + \frac{1}{q'} = 1$). It follows that

$$\begin{aligned} \rho_{c,p}^{ho}(\|Y\|) &\leq \sup \left\{ \|Y\|_{p'} \|Z\|_{q'} : \|Z\|_q \leq c, Z \geq 0 \text{ and } \mathbb{E}[Z] = 1 \right\} \\ &= \|Y\|_{p'} \cdot \sup \left\{ \|Z\|_{q'} : \|Z\|_q \leq c, Z \geq 0 \text{ and } \mathbb{E}[Z] = 1 \right\}. \end{aligned}$$

Note now that $\frac{1}{q'} = \frac{1}{\frac{p'}{p-1}} = \frac{1-p'}{p'} + \frac{p'}{p-1} = \frac{1-p'}{p'} + \frac{p'}{q}$, such that by Hölder’s interpolation inequality ($\|Z\|_{q_0} \leq \|Z\|_{q_0}^{1-\theta} \cdot \|Z\|_{q_1}^\theta$ whenever $\frac{1}{q_0} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}$, cf. Wojtaszczyk 1991)

$$\|Z\|_{q'} \leq \|Z\|_1^{1-p/p'} \cdot \|Z\|_q^{p/p'} = c^{p/p'},$$

as $Z \geq 0$ implies that $\|Z\|_1 = \mathbb{E}[Z] = 1$. It follows that

$$\rho_{c,p}^{ho}(|Y|) \leq c^{p/p'} \cdot \|Y\|_{p'}.$$

To see that this bound is sharp consider $Y = c^{p/p'} \cdot \mathbb{1}_A$, where $P(A) = \frac{1}{c^p}$. Then $\|Y\|_{p'} = (c^p \cdot P(A))^{1/p'} = 1$, and

$$\begin{aligned} t + c \cdot \|(Y - t)_+\|_p &= t + c \cdot \left(0 + P(A) \left(c^{p/p'} - t\right)^p\right)^{\frac{1}{p}} \\ &= t + c \cdot \left(\frac{1}{c^p} \left(c^{p/p'} - t\right)^p\right)^{\frac{1}{p}} = t + \left(c^{p/p'} - t\right) = c^{p/p'} \end{aligned}$$

for all $t \in [0, c^{p/p'}]$. As the mapping $t \mapsto t + c \cdot \|(Y - t)_+\|_p$ is convex it follows that the infimum is $\rho(|Y|) = c^{p/p'}$. This proves that the bound is sharp. \square

3.2 Higher order semideviation

The higher order semideviation is a risk measure addressed in Rockafellar et al. (2006) and discussed in Shapiro et al. (2009). In the chosen context of norms the dual representation is natural and straight forward again. We introduce and discuss the general case first before establishing the continuity relations. The results then are further specified to L^p -norms, as it is again possible to give explicit, precise bounds.

Theorem 3.9 For every $0 \leq \lambda \leq 1$,

$$\rho_\lambda^{sd}(Y) := \mathbb{E}[Y] + \lambda \cdot \|(Y - \mathbb{E}[Y])_+\|$$

is a risk functional. If the norm satisfies (NM)–(NP), then the representation

$$\rho_\lambda^{sd}(Y) = \sup_{Z \geq 0, \mathbb{E}[Z]=1} \left(1 - \frac{\lambda}{\|Z\|^*}\right) \mathbb{E}[Y] + \frac{\lambda}{\|Z\|^*} \mathbb{E}[YZ] \tag{17}$$

holds true.

Remark 3.10 Recall from (3) that $\|Z\|^* \geq \mathbb{E}[Z] = 1$, such that (17) is a convex combination with nonnegative weights provided that $0 \leq \lambda \leq 1$. Further, the restriction to $\lambda \leq 1$ ensures monotonicity (RM).

Proof Observe first that

$$\rho_\lambda^{sd}(Y) = \mathbb{E}[Y] + \lambda \cdot \|(Y - \mathbb{E}[Y])_+\| = \sup_{Z \geq 0} \mathbb{E}[Y] + \frac{\lambda}{\|Z\|^*} \mathbb{E}[Z \cdot (Y - \mathbb{E}[Y])], \tag{18}$$

as it is enough to restrict the supremum to $Z \geq 0$ by (NP). Note that Z is in the nominator and the denominator, so by rescaling to $\mathbb{E}[Z] = 1$ it follows that

$$\rho_\lambda^{sd}(Y) = \sup_{\mathbb{E}[Z]=1, Z \geq 0} \left(1 - \frac{\lambda}{\|Z\|^*}\right) \mathbb{E}[Y] + \frac{\lambda}{\|Z\|^*} \mathbb{E}[Z \cdot Y],$$

which is the assertion. \square

Theorem 3.11 (Comparison of the higher order semideviation with the norm) *It holds that*

$$\|Y\|_1 \leq \rho_\lambda^{sd}(|Y|) \leq (\|\mathbb{1}\|^* + \lambda) \cdot \|Y\| \tag{19}$$

and

$$\lambda \cdot \|Y\| \leq \rho_\lambda^{sd}(|Y|) \leq (\|\mathbf{1}\|^* + \lambda) \cdot \|Y\|.$$

Proof The first inequality in (19), $\|Y\|_1 = \mathbb{E}[|Y|] \leq \rho_\lambda^{sd}(|Y|)$, is evident from the definition of ρ_λ^{sd} .

Observe next that $\lambda y \leq t + \lambda(y - t)_+$ whenever $t \geq 0$. Hence $\lambda|Y| \leq t + \lambda(|Y| - t)_+$ and thus

$$\lambda \cdot \|Y\| = \|\lambda \cdot |Y|\| \leq \|t + \lambda(|Y| - t)_+\| \leq t + \lambda \cdot \||Y| - t\|_+$$

by the triangle inequality and the monotonicity assumption (NM) of the norm. By choosing $t := \mathbb{E}[|Y|]$ (note that $t \geq 0$) it follows that

$$\lambda \cdot \|Y\| \leq \mathbb{E}[|Y|] + \lambda \cdot \||Y| - \mathbb{E}[|Y|]\|_+ = \rho_\lambda^{sd}(|Y|).$$

As for the upper bound recall from Theorem 3.9 the representation

$$\begin{aligned} \rho_\lambda^{sd}(Y) &= \sup_{\mathbb{E}[Z]=1, Z \geq 0} \left(1 - \frac{\lambda}{\|Z\|^*}\right) \mathbb{E}[Y] + \frac{\lambda}{\|Z\|^*} \mathbb{E}[YZ] \\ &\leq \sup_{\mathbb{E}[Z]=1, Z \geq 0} \left(1 - \frac{\lambda}{\|Z\|^*}\right) \mathbb{E}[Y] + \frac{\lambda}{\|Z\|^*} \|Y\| \|Z\|^*, \end{aligned}$$

from which we deduce that

$$\rho_\lambda^{sd}(|Y|) \leq \mathbb{E}[|Y|] + \lambda \|Y\| \leq \|\mathbf{1}\|^* \|Y\| + \lambda \|Y\|$$

by (4), completing the proof. □

As in the case for the higher order risk measure it is possible to give better and tight bounds by further specifying the norm. We state the results in what follows.

Definition (*Higher order semideviation*) The higher order semideviation risk measure (for the parameters $0 \leq \lambda < 1$ and $p \geq 1$) is

$$\rho_{\lambda,p}^{sd}(Y) := \mathbb{E}[Y] + \lambda \cdot \|(Y - \mathbb{E}[Y])_+\|_p.$$

The risk measure is also called *mean upper semideviation of order p* in the literature (cf. Shapiro et al. 2009; Pichler and Shapiro 2015). We present the following bounds for this risk measure, which is important in applications.

Theorem 3.12 (Comparison with L^p -norms) *The following holds true:*

(i) if $p = 1$, then

$$\|Y\|_1 \leq \rho_{\lambda,1}^{sd}(|Y|) \leq (1 + \lambda) \cdot \|Y\|_1; \tag{20}$$

(ii) if $1 < p < 2$, then

$$\lambda \cdot \|Y\|_p \leq \rho_{\lambda,p}^{sd}(|Y|) \leq k_{\lambda,p} \cdot \|Y\|_p, \tag{21}$$

where $k_{\lambda,p} > (1 + \lambda^q)^{1/q}$ and q is the conjugate Hölder exponent, $\frac{1}{p} + \frac{1}{q} = 1$;

(iii) if $p \geq 2$, then the inequalities

$$\lambda \cdot \|Y\|_p \leq \rho_{\lambda,p}^{sd}(|Y|) \leq (1 + \lambda^q)^{1/q} \cdot \|Y\|_p, \tag{22}$$

are tight.

All bounds are sharp, the upper bounds are sharp except that there is no closed, explicit expression for $k_{\lambda,p}$ whenever $1 < p < 2$.

Remark 3.13 The inequalities in (i) are useful in practical situation. The lower bounds in (ii) and (iii) are tight, but particularly for small values of λ they are rather loose. We shall address this issue further in the section on compositions of risk measures (Sect. 5) below.

Proof of Theorem 3.12 (i) and the first inequality of (ii) and (iii) are immediate from Theorem 3.11.

To accept that the bounds are sharp consider $Y := \frac{1}{P(A)^{1/p}} \mathbb{1}_A$. Then $\|Y\|_p = 1$, but

$$\begin{aligned} \rho_{\lambda,p}^{sd}(|Y|) &= P(A)^{1-\frac{1}{p}} + \lambda \left(P(A) \left(\frac{1}{P(A)^{1/p}} - P(A)^{1-\frac{1}{p}} \right)^p \right)^{\frac{1}{p}} \\ &= P(A)^{1-\frac{1}{p}} + \lambda P(A)^{\frac{1}{p}} \left(P(A)^{-\frac{1}{p}} - P(A)^{1-\frac{1}{p}} \right) \\ &= P(A)^{1-\frac{1}{p}} + \lambda(1 - P(A)) \xrightarrow{P(A) \rightarrow 0} \lambda, \end{aligned}$$

provided that $p > 1$. It follows that the lower bound in (21) and (22) cannot be improved.

From Eq. (17) in Theorem 3.9 it follows that

$$\begin{aligned} \rho_{\lambda}^{sd}(Y) &= \sup_{\mathbb{E}[Z]=1, Z \geq 0} \mathbb{E} \left[\left(\left(1 - \frac{\lambda}{\|Z\|^*} \right) \mathbb{1} + \frac{\lambda}{\|Z\|^*} Z \right) \cdot Y \right], \\ &\leq \sup_{\mathbb{E}[Z]=1, Z \geq 0} \left\| \left(1 - \frac{\lambda}{\|Z\|^*} \right) \mathbb{1} + \frac{\lambda}{\|Z\|^*} Z \right\|_q \cdot \|Y\|_p \end{aligned} \tag{23}$$

by Hölder’s inequality. Equality in (23), however, is attained for

$$Y = \left(\left(1 - \frac{\lambda}{\|Z\|^*} \right) \mathbb{1} + \frac{\lambda}{\|Z\|^*} Z \right)^{p-1}, \tag{24}$$

such that

$$\sup_{Y \neq 0} \frac{\rho_{\lambda,p}^{sd}(Y)}{\|Y\|_p} = \sup_{\mathbb{E}[Z]=1, Z \geq 0} \left\| \left(1 - \frac{\lambda}{\|Z\|^*} \right) \mathbb{1} + \frac{\lambda}{\|Z\|^*} Z \right\|_q.$$

Consider the function $\sigma_{\beta}(\cdot) := \frac{1}{\beta} \mathbb{1}_{[1-\beta, 1]}(\cdot)$ (for which $\|\sigma_{\beta}\|_q = \frac{1}{\beta^{\frac{q-1}{q}}}$) and the random variable $Z = \sigma_{\beta}(U)$ for some uniformly distributed random variable U .⁴ Then

$$\begin{aligned} \left\| \left(1 - \frac{\lambda}{\|Z\|^*} \right) \mathbb{1} + \frac{\lambda}{\|Z\|^*} Z \right\|_q^q &= \left\| 1 - \frac{\lambda}{\|\sigma_{\beta}\|_q} + \frac{\lambda}{\|\sigma_{\beta}\|_q} \sigma_{\beta} \right\|_q^q \\ &= (1 - \beta) \left(1 - \lambda \beta^{\frac{q-1}{q}} \right)^q + \beta \left(1 - \lambda \beta^{\frac{q-1}{q}} + \lambda \frac{1}{\beta} \beta^{\frac{q-1}{q}} \right)^q \\ &= (1 - \beta) \left(1 - \lambda \beta^{\frac{q-1}{q}} \right)^q + \left(\lambda(1 - \beta) + \beta^{\frac{1}{q}} \right)^q. \end{aligned} \tag{25}$$

By letting $\beta \rightarrow 0$ it follows that $\lim_{\beta \rightarrow 0} (25) = 1 + \lambda^q$, from which we conclude that

$$1 + \lambda \geq \sup_{Y \neq 0} \frac{\rho_{\lambda,p}^{sd}(Y)}{\|Y\|_p} \geq (1 + \lambda^q)^{1/q}.$$

⁴ U is uniformly distributed, iff $P(U \leq u) = u$.

The upper bound can be rewritten as $\sup \left\{ \rho_{\lambda,p}^{sd}(|Y|) : \|Y\|_p \leq 1 \right\}$. As $Y \mapsto \rho_{\lambda,p}^{sd}(|Y|)$ is convex, it is enough to consider Y in the set of extreme points in the unit ball of L^p , which are the random variables $Y = \frac{\mathbb{1}_{A_1} - \mathbb{1}_{A_2}}{P(A_1 \cup A_2)^{1/p}}$ (cf. [Sundaresan 1969](#)), i.e.,

$$|Y| = \frac{\mathbb{1}_A}{P(A)^{1/p}} \tag{26}$$

for $A := A_1 \cup A_2$. Now if $|Y|$ is as in (26), then, because of (24) and (25), the extremum is attained for $\sigma_\beta = \frac{1}{\beta} \mathbb{1}_{[1-\beta, 1]}$, $\beta \in (0, 1)$, which means that it is enough to maximize (25) with respect to $\beta \in [0, 1]$.

But (25) is a strictly decreasing in β for $p \geq 2$ (i.e., $q \leq 2$), such that the maximum as attained for $\beta \rightarrow 0$.

For $p > 2$, the maximum in (25) is attained in the interior, at some $\beta \in (0, 1)$. However, a closed analytic expression is not available. □

Example 3.14 (Dutch risk measure (cf. also Example 4.8 below)) The Dutch risk measure (or Dutch premium principle, cf. [Heerwaarden and Kaas 1992](#)) is the higher order semideviation risk measure

$$\rho_{1,\lambda}^{sd}(Y) := \mathbb{E}[Y] + \lambda \cdot \mathbb{E}[(Y - \mathbb{E}[Y])_+], \tag{27}$$

where $0 \leq \lambda < 1$. [Shapiro \(2013\)](#) provides the Kusuoka representation

$$\rho_{1,\lambda}^{sd}(Y) := \sup_{\kappa \in [0, 1]} (1 - \kappa\lambda) \cdot \mathbb{E}[Y] + \kappa\lambda \cdot \text{AV@R}_{1-\kappa}(Y). \tag{28}$$

It follows from (13) that $\text{AV@R}_{1-\kappa}(Y) \leq \frac{1}{\kappa} \mathbb{E}[|Y|]$, such that

$$\mathbb{E}[Y] \leq \rho_{1,\lambda}^{sd}(Y) \leq \sup_{\kappa \in [0, 1]} (1 - \kappa\lambda) \cdot \mathbb{E}[Y] + \lambda \mathbb{E}[|Y|] \leq (1 + \lambda) \mathbb{E}[|Y|].$$

These estimates are sharp, and in line with the bounds (20) presented in Theorem 3.12.

4 Comparisons of risk measures, and nonlinear bounds

The previous section presents relations between a norm and a risk measure. This section compares risk measures by establishing direct relations. We start with the following remark to distinguish local and global comparisons. Finally we describe bounds which involve the random variable in a nonlinear and in a non-homogeneous way.

Global comparisons Suppose a comparison of the risk measures ρ_1 and ρ_2 is of the global form

$$\rho_1(Y) \leq K \cdot \rho_2(Y) \quad \text{for all } Y. \tag{29}$$

By considering $Y + c \cdot \mathbb{1}$ instead of Y it follows from translation equivariance (RT) that $\rho_1(Y) \leq \rho_2(Y) + (K - 1)c$ for all $c \in \mathbb{R}$. This inequality is impossible unless $K = 1$. A comparison of risk functionals in the general, global form (29) thus can only read

$$\rho_1(Y) \leq \rho_2(Y) \quad \text{for all } Y \tag{30}$$

(i.e., $K = 1$ in (29)).

Only a few risk measures allow a general, global comparison as in (30). The inequality for the Average Value-at-Risk,

$$\text{AV@R}_\alpha(Y) \leq \text{AV@R}_{\alpha'}(Y)$$

for $\alpha \leq \alpha'$ is an example of a global comparison—a special case of $\rho_c^{ho}(Y) \leq \rho_{c'}^{ho}(Y)$ for $c \leq c'$. In what follows we state the global relations, if eligible.

Comparisons on the nonnegative cone To allow for more than global comparisons of risk measures we recall from the previous section that it is enough to know the risk measure—or the associated norm—on the nonnegative cone L_+ . Instead of the global form (30) we thus consider

$$\rho(Y) \leq K \cdot \rho'(Y) \quad \text{for all } Y \geq 0 \text{ a.s.} \tag{31}$$

as well, that is to involve only *nonnegative* random variables or the norms associated with the risk measure in order to obtain useful and non-trivial comparisons.

Many applications consider only nonnegative outcomes, and considering only nonnegative random variables is not as restrictive as it seems. Every payoff function of an insurance policy, for example, is always nonnegative (i.e., $Y \geq 0$). Some applications interested in financial risk occasionally ignore the profit by considering the loss $\max\{0, Y\}$ instead of Y . Finally, instead of bounded random variables it is often enough to consider $c + Y$ (for some appropriate $c > 0$) instead of Y in order to have the comparison on the nonnegative cone (31) available. This approach is in line with the result reported in Theorem 3.5 and particularly with Remark 3.6.

4.1 Comparison of distortion risk functionals

Distortion risk functionals constitute a basic and elementary ingredient for risk functionals, as every general version independent risk functional is the supremum over a class of distortion risk functionals (cf. Kusuoka 2001; Shapiro 2013; Noyan and Rudolf 2014; they are related to extreme points in the dual set). The distortion risk functional (also spectral risk measure, cf. Acerbi 2002; Denneberg 1990) is defined by

$$\rho_\sigma(Y) := \int_0^1 \sigma(u) F_Y^{-1}(u) du,$$

where $F_Y^{-1}(u) := \inf\{y : P(Y \leq y) \geq u\}$ is the generalized inverse of the cumulative distribution function (cdf) $F_Y(y) := P(Y \leq y)$. $\sigma : [0, 1) \rightarrow [0, \infty)$ is a nonnegative, nondecreasing function satisfying $\int_0^1 \sigma(u) du = 1$. The function $\sigma(\cdot)$ is called *distortion functional*. For convenience we shall associate the function $\Sigma(p) := \int_p^1 \sigma(u) du$ (i.e., its negative antiderivative) with $\sigma(\cdot)$.

On the nonnegative cone we have the following additional formula for distortion risk functionals, which does not involve the quantile function F_Y^{-1} , but Y 's cumulative distribution function (cdf) F_Y directly.

Proposition 4.1 *For a random variable $Y \in L_+$ in the nonnegative cone the distortion risk functional ρ_σ can be evaluated by the alternative expression*

$$\rho_\sigma(Y) = \int_0^\infty \Sigma(F_Y(y)) dy.$$

Proof Cf. Denneberg (1990) or Pflug and Pichler (2014, Corollary 3.15). □

Theorem 4.2 (Comparison of spectral risk measures) *Suppose that*

$$K := \sup_{0 \leq \alpha < 1} \frac{\int_\alpha^1 \sigma_1(u) du}{\int_\alpha^1 \sigma_2(u) du} \tag{32}$$

is finite ($K < \infty$), then

$$\rho_{\sigma_1}(Y) \leq K \cdot \rho_{\sigma_2}(Y) \quad (Y \geq 0 \text{ a.s.}), \tag{33}$$

the bound is sharp. It holds moreover that $K \geq \limsup_{\alpha \uparrow 1} \frac{\sigma_1(\alpha)}{\sigma_2(\alpha)}$.

Proof Pichler (2013a, Theorem 14) provides a proof. □

We provide a few examples to illustrate the strength of the previous result.

Example 4.3 (Average Value-at-Risk) The distortion function of the Average Value-at-Risk is

$$\sigma(u) = \begin{cases} 0 & \text{if } u \leq \alpha, \\ \frac{1}{1-\alpha} & \text{if } u > \alpha \end{cases} \text{ and } \Sigma_\alpha(u) := \min \left\{ 1, \frac{1-u}{1-\alpha} \right\}.$$

For $\alpha_1 \leq \alpha_2$ the non-trivial constant is $K = \frac{1-\alpha_1}{1-\alpha_2}$ and as a particular consequence of (33) it follows that

$$\text{AV@R}_{\alpha_1}(|Y|) \leq \text{AV@R}_{\alpha_2}(|Y|) \leq \frac{1-\alpha_1}{1-\alpha_2} \text{AV@R}_{\alpha_1}(|Y|). \tag{34}$$

Particularly, $\text{AV@R}_\alpha(|Y|) \leq \frac{1}{1-\alpha} \mathbb{E}[|Y|]$.

The risk measure $\rho(Y) := \beta \cdot \mathbb{E}[Y] + (1-\beta) \cdot \text{AV@R}_\alpha(Y)$ is popular in applications (referred to as *risk measure for integrated risk management*, cf. Pflug and Ruszczyński 2005), as it is a natural combination of elementary risk measures (compare also the Dutch risk measure (28)). Its Σ -function is $\Sigma_{\alpha,\beta}(u) = \min \left\{ 1 - \beta u, \frac{1-\alpha\beta}{1-\alpha}(1-u) \right\}$. The global upper bound

$$\beta \cdot \mathbb{E}[Y] + (1-\beta) \cdot \text{AV@R}_\alpha(Y) \leq \text{AV@R}_{\frac{\alpha-\alpha\beta}{1-\alpha\beta}}(Y) \tag{35}$$

follows by elementary computations and Remark 3.6, where the adapted risk level $\alpha' := \frac{\alpha-\alpha\beta}{1-\alpha\beta}$ is smaller than the initial risk level α , $\alpha' \leq \alpha$. The global bound (35) is sharp again, and the risk level α' cannot be improved.

A lower, sharp bound expressed for a general risk level $\tilde{\alpha}$ is

$$\text{AV@R}_{\tilde{\alpha}}(Y) \leq \frac{1}{\Sigma_{\alpha,\beta}(\tilde{\alpha})} (\beta \cdot \mathbb{E}[Y] + (1-\beta) \cdot \text{AV@R}_\alpha(Y)) \quad (Y \geq 0). \tag{36}$$

Recall that the comparisons (34) and (36) are not valid in case that $Y \notin L_+$. Finally, as a special case we note that $\frac{1}{\Sigma_{\alpha,\beta}(\alpha')} = \frac{1-\alpha\beta}{1-\alpha\beta(2-\beta)}$ for the critical risk level α' specified in (35).

Example 4.4 (Proportional hazards risk measure) The proportional hazards risk measure (adapted from insurance, cf. Young 2006), is

$$\rho_c^H(Y) = \int_0^\infty (1 - F_Y(u))^c du$$

on the nonnegative cone, where $0 < c \leq 1$ is a parameter accounting for risk aversion. The risk functional compares with the Average Value-at-Risk by

$$\text{AV@R}_\alpha(Y) \leq \frac{1}{(1-\alpha)^c} \cdot \rho_c^H(Y) \quad (Y \in L_+)$$

for every $c > 0$, again an immediate consequence of Proposition 4.1 and Theorem 4.2 with $\Sigma(\alpha) = (1-\alpha)^c$.

Example 4.5 (Wang transform) Wang’s risk measure (cf. Wang 1995) employs the parametric family $\Sigma_\lambda^W(u) = \Phi(\lambda + \Phi^{-1}(1 - u))$ for $\lambda \geq 0$, it is defined as

$$\rho_\lambda^W(Y) = \int_0^\infty \Sigma_\lambda^W(F_Y(y)) dy$$

on the nonnegative cone ($\Phi(\cdot)$ is the cdf of the normal distribution, cf. Proposition 4.1). It is a convex risk measure for $\lambda \geq 0$. It is comparably easy to see that $\sup_{u < 1} \frac{\Sigma_\alpha(u)}{\Sigma_\lambda^W(u)} = \frac{1}{\Sigma_\lambda^W(\alpha)}$ and it follows that

$$AV@R_\alpha(Y) \leq \frac{1}{\Sigma_\lambda^W(\alpha)} \rho_\lambda^W(Y) \quad (Y \geq 0).$$

The proportional hazards risk measure dominates the Wang transform for every combination of λ and c ,

$$\rho_\lambda^W(Y) \leq K_{\lambda,c} \cdot \rho_c^H(Y), \tag{37}$$

although an explicit expression for the constant $K_{\lambda,c}$ is not available. A converse inequality to (37) is not possible.

4.2 General version independent risk measures

This section provides bounds for risk functionals of different type. We start with comparing general version independent risk functionals and then give a general bound to compare higher order risk measures with higher order semideviations.

For a class S of distortion functions it is immediate that $\rho_S(Y) := \sup_{\sigma \in S} \rho_\sigma(Y)$ is a risk functional satisfying all Axioms (RM)–(RD). To obtain a comparison of the form

$$\rho_{S_1}(Y) \leq K \cdot \rho_{S_2}(Y)$$

on the nonnegative cone L_+ one may choose

$$K := \sup_{\sigma_1 \in S_1} \inf_{\sigma_2 \in S_2} \sup_{\alpha < 1} \frac{\int_\alpha^1 \sigma_1(u) du}{\int_\alpha^1 \sigma_2(u) du}, \tag{38}$$

which is an immediate consequence of (33) in Theorem 4.2, although K is possibly not the best constant. It is evident from the max–min inequality that

$$K \leq \inf_{\sigma_2 \in S_2} \sup_{\sigma_1 \in S_1} \sup_{\alpha < 1} \frac{\int_\alpha^1 \sigma_1(u) du}{\int_\alpha^1 \sigma_2(u) du} \leq \sup_{\sigma_1 \in S_1, \sigma_2 \in S_2, \alpha < 1} \frac{\int_\alpha^1 \sigma_1(u) du}{\int_\alpha^1 \sigma_2(u) du},$$

and the latter are sometimes easier to evaluate in situations of practical relevance.

Example 4.6 (Entropic Value-at-Risk) The Entropic Value-at-Risk introduced in Ahmadi-Javid (2012) is

$$EV@R_\alpha(Y) := \inf_{z > 0} \left\{ \frac{1}{z} \log \left(\frac{1}{1 - \alpha} \mathbb{E} \left[e^{zY} \right] \right) \right\}. \tag{39}$$

The $EV@R$ is the tightest upper bound that can be obtained from the Chernoff inequality for the Value-at-Risk. Delbaen (2015) elaborates the Kusuoka representation

$$EV@R_\alpha(Y) = \sup \left\{ \rho_\sigma(Y) : \int_0^1 \sigma(u) \log \sigma(u) du \leq \log \frac{1}{1 - \alpha} \right\}$$

using Kullback–Leiber divergence. We refer to the original reference [Ahmadi-Javid \(2012, Proposition 3.2\)](#) for the global comparison

$$\text{EVaR}_\alpha(Y) \leq \text{AV@R}_\alpha(Y),$$

whenever $\alpha \in [0, 1)$. Equality is attained, for example, for the random variable $Y = \mathbb{1}_{[\alpha, 1]}(U)$, but strict inequality holds for $Y = \mathbb{1}_{[\beta, 1]}(U)$, $\beta \neq \alpha$.

Relations between higher order risk measures and the higher order semideviation We provide the following tight bounds to compare the risk measures discussed and introduced in the previous Sect. 3.

Theorem 4.7 (*Comparison of different risk measures*) *The relation*

$$\rho_\lambda^{sd}(Y) \leq \rho_{1+\lambda}^{ho}(Y) \tag{40}$$

is tight and holds globally ($0 \leq \lambda \leq 1$).

For arbitrary chosen $0 \leq \lambda \leq 1$ and $c \geq 1$ it holds that

$$\rho_c^{ho}(Y) \leq \frac{c}{\lambda} \cdot \rho_\lambda^{sd}(|Y|) \tag{41}$$

and conversely

$$\rho_\lambda^{sd}(Y) \leq (\|\mathbb{1}\|^* + \lambda) \|Y\| \leq (\|\mathbb{1}\|^* + \lambda) \cdot \rho_c^{ho}(|Y|).$$

Proof Assume first that $Y \geq 0$. Recall from (18) the representation

$$\rho_\lambda^{sd}(Y) = \sup_{Z \geq 0} \mathbb{E}[Y] + \frac{\lambda}{\|Z\|^*} [\mathbb{E}[Z \cdot (Y - \mathbb{E}[Y])]]$$

of the higher order risk measure. Note that $\mathbb{E}[Y] \geq 0$ by assumption, so

$$\rho_\lambda^{sd}(Y) \leq \sup_{Z \geq 0, \mathbb{E}[Z]=1} \mathbb{E} \left[Y \frac{Z}{\|Z\|^*} \right] + \frac{\lambda}{\|Z\|^*} \mathbb{E}[Z \cdot Y].$$

Hence

$$\begin{aligned} \rho_\lambda^{sd}(Y) &\leq \sup_{Z \geq 0, \mathbb{E}[Z]=1} \mathbb{E} \left[Y \frac{(1 + \lambda) \cdot Z}{\|Z\|^*} \right] \\ &= \sup_{Z \geq 0, \mathbb{E}[Z]=1, \|Z\|^* \leq 1 + \lambda} \mathbb{E}[YZ] = \rho_{1+\lambda}^{ho}(Y) \end{aligned}$$

by (7), which is the desired estimate on L_+ . The general assertion for $L \supseteq L_+$ follows from Remark 3.6.

Recall finally from Theorems 3.5 and 3.11 that

$$\rho_c^{ho}(Y) \leq c \cdot \|Y\| \leq c \cdot \frac{1}{\lambda} \rho_\lambda^{sd}(|Y|)$$

and

$$\rho_\lambda^{sd}(Y) \leq (\|\mathbb{1}\|^* + \lambda) \|Y\| \leq (\|\mathbb{1}\|^* + \lambda) \rho_c^{ho}(|Y|)$$

to establish the remaining relations. □

Example 4.8 (Dutch risk measure, cf. Example 3.14) As an application for the general constant (38) consider the Dutch risk measure (27). By applying the relations (27) and (40) from the previous theorem to the norm $\|\cdot\|_1$ it follows that

$$\mathbb{E}[Y] + \lambda \cdot \mathbb{E}[(Y - \mathbb{E}[Y])_+] \leq \text{AV@R}_{\frac{\lambda}{1+\lambda}}(Y). \tag{42}$$

The *critical risk level* $\frac{\lambda}{1+\lambda}$ is the smallest possible risk level, for which (42) holds true. In contrast to the estimates already presented in Example 3.14 this is a global and tight upper bound for the Dutch risk measure.

To obtain a tight lower bound one may apply (38) and (36) to compute the constant K (cf. (28)), resulting in

$$\text{AV@R}_\alpha(Y) \leq \frac{1}{(1-\alpha)(1+\alpha\lambda)} \cdot \rho_{\lambda,1}^{sd}(|Y|) \quad (Y \in L_+).$$

This bound is tight, as can be seen by considering $Y = \mathbb{1}_A$ with $P(A^c) = \alpha$.

For the crucial risk level $\alpha = \frac{\lambda}{1+\lambda}$ the lower, tight bound is

$$\text{AV@R}_{\frac{\lambda}{1+\lambda}}(Y) \leq \frac{1 + 2\lambda + \lambda^2}{1 + \lambda + \lambda^2} \cdot (\mathbb{E}[|Y|] + \lambda \cdot \mathbb{E}[(|Y| - \mathbb{E}[|Y|])_+]).$$

This results further in the sandwich inequality

$$\frac{3}{4} \text{AV@R}_\alpha(Y) \leq \mathbb{E}[|Y|] + \frac{\alpha}{1-\alpha} \cdot \mathbb{E}[(|Y| - \mathbb{E}[|Y|])_+] \leq \text{AV@R}_\alpha(|Y|)$$

for $\alpha \leq \frac{1}{2}$, which demonstrates that the Dutch risk measure is basically—up to a multiplicative error of 25%—an Average Value-at-Risk with critical risk level α . This may give rise in some applications for replacing the more complicated AV@R in optimization problems by the simple semideviation risk measure, or vice versa.

4.3 The associated functional

Associated with the distortion risk measure and the distortion function $\sigma(\cdot)$ is the functional

$$\rho_\sigma^*(Y) := \sup_{\alpha < 1} \frac{\text{AV@R}_\alpha(Y)}{\frac{1}{1-\alpha} \int_\alpha^1 \sigma(u) du}. \tag{43}$$

The associated functional naturally provides the upper bound

$$\text{AV@R}_\alpha(Y) \leq \frac{1}{1-\alpha} \int_\alpha^1 \sigma(u) du \cdot \rho_\sigma^*(|Y|),$$

which holds for all $\alpha < 1$ (note that $\rho_\sigma^*(Y)$ is not necessarily nonnegative).

The functional (43) satisfies the axioms of a risk measures (Definition 2.1), except translation equivariance (RT) and thus is not a risk measure.

As a corollary of Theorem 4.2 we have the following relations for the constant ρ_σ^* .

Corollary 4.9 For $Y \geq 0$ it holds that

$$\rho_{\sigma_2}^*(Y) \leq K \cdot \rho_{\sigma_1}^*(Y),$$

where K is the constant (32) (Theorem 4.2).

Proof $\|Y\|_\sigma := \rho_\sigma(|Y|)$ and $\|Z\|_\sigma^* := \rho_\sigma^*(|Z|)$ are norms, which are dual to each other (cf. Pichler 2013a) with respect to the bilinear form $(Y, Z) \mapsto \mathbb{E}[YZ]$. From the Hahn–Banach theorem it follows that

$$\begin{aligned} \rho_{\sigma_2}^*(|Z|) &= \sup \{ \mathbb{E}[YZ] : \|Y\|_{\sigma_2} \leq 1 \} = \sup \{ \mathbb{E}[YZ] : \rho_{\sigma_2}(|Y|) \leq 1 \} \\ &\leq \sup \{ \mathbb{E}[YZ] : \rho_{\sigma_1}(|Y|) \leq K \} = K \cdot \rho_{\sigma_1}^*(|Z|), \end{aligned}$$

which is the assertion. □

4.4 Nonlinear upper bounds for the distortion risk functional

Approximations of the Average Value-at-Risk, which have been proposed in the literature, include

$$\text{AV@R}_\alpha^\tau(Y) := \inf_{x \in \mathbb{R}} \left\{ x + \frac{1}{1 - \alpha} \mathbb{E}[(Y - x)_{\tau,+}] \right\}$$

where $y_{\tau,+} = \frac{1}{2} \left(y + \sqrt{y^2 + 4\tau^2} \right)$ (we refer to Luna et al. 2016 discussing also other variants of approximating y_+ by some $y_{\tau,+}$). In this setting it holds that $x_+ \leq x_{\tau,+} \leq \tau + x_+$. This approximation apparently generalizes for the higher order risk measure

$$\rho_{p,\tau}^{ho}(Y) := \inf_{x \in \mathbb{R}} \left\{ x + c \cdot \|(Y - x)_{\tau,+}\|_p \right\},$$

and consequently $\rho_p^{ho}(Y) \leq \rho_{p,\tau}^{ho}(Y) \leq \rho_p^{ho}(Y) + \tau c$. The functional $\rho_{p,\tau}^{ho}(Y)$ is translation equivariant, but not positively homogeneous. The particular advantage of the functional $\rho_{p,\tau}^{ho}(\cdot)$ is given by the fact that the approximation $y \mapsto y_{\tau,+}$ is strictly convex and differentiable, while $y \mapsto y_+$ is only subdifferentiable.

Distortions—upper bounds derived from Fenchel–Young inequality The upper bound

$$\rho_\sigma(Y) \leq \mathbb{E}[h(Y)] + \int_0^1 h^*(\sigma(u))du \tag{44}$$

is valid for every (measurable) function $h : \mathbb{R} \rightarrow \mathbb{R}$. (44) follows from the Fenchel–Young inequality $\sigma(U) \cdot Y \leq h(Y) + h^*(\sigma(U))$, where $h^*(\sigma) := \sup_{y \in \mathbb{R}} \sigma \cdot y - h(y)$ is the usual convex conjugate function of the function h . Pichler (2013b) demonstrates that the relation (44) is sharp, that is, for every random variable $Y \in L$ there exists a function h such that equality holds in (44).

5 Composite and conditional risk measures

The risk measures considered in the previous sections are defined on a probability space with a sigma algebra. Extensions to filtered probability spaces are considered in several places, to the best of our knowledge the earliest occurrence is Artzner et al. (2007) in this journal. These risk measures are designed to capture the evolution of risk in a multistage environment, such that general results as Bellman’s principle can be used and adapted to multistage situations. Examples are given in Ruszczyński and Shapiro (2006), Ruszczyński and Yao (2015) and Densing (2014). Shapiro (2016) elaborates properties of these risk measures in the context of convex analysis, while Dentcheva et al. (2016) investigate statistical properties of composite risk measures.

Continuing the intention of the previous sections we develop bounds for composite risk measures by using the relations with norms developed in the previous sections. Having practical implementations in mind we restrict ourselves to higher order measures and the higher order semideviation. We employ conditional expectations to handle these conditional risk measures efficiently and to compare them with the respective norms.

Throughout this section we assume that \mathcal{F}_1 is a sub-sigma algebra, $\mathcal{F}_1 \subseteq \mathcal{F}$. With $\mathcal{F}_0 := \{\emptyset, \Omega\}$ we denote the trivial sigma algebra, so that $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}$.

5.1 Higher order measures

For the definition of higher order risk measures we refer to [Ruszczynski \(2010, Example 3\)](#) and the references given therein.

Definition 5.1 The conditional higher order measure is

$$\rho_{c,p}^{ho}(Y, \mathcal{F}_1) := \text{ess inf}_{x \triangleleft \mathcal{F}_1} \left\{ x + c \cdot \mathbb{E}[(Y - x)_+^p | \mathcal{F}_1]^{1/p} \right\}, \tag{45}$$

where $c \triangleleft \mathcal{F}_1$. We write $x \triangleleft \mathcal{F}_1$ ($c \triangleleft \mathcal{F}_1$, resp.) for x (c , resp.) being measurable with respect to the sigma algebra \mathcal{F}_1 . For convenience it is accepted in the literature to write also $\rho_{c,p}^{ho}(Y | \mathcal{F}_1) := \rho_{c,p}^{ho}(Y, \mathcal{F}_1)$, which is in line with the notation of conditional expectation.

Remark 5.2 For the trivial sigma algebra $\mathcal{F}_0 = \{\emptyset, \Omega\}$ it holds that $\rho_{c,p}^{ho}(Y) = \rho_{c,p}^{ho}(Y, \mathcal{F}_0)$, this is the higher order measure introduced in (14). The composite risk measure $Y \mapsto \rho_{c,p}^{ho}(\rho_{c,p}^{ho}(Y, \mathcal{F}_1))$ is a functional satisfying (RM)–(RH), but it is not version independent (i.e., (RD) is not satisfied).

We have the following comparison with the norm.

Theorem 5.3 For the composite risk measure it holds that

$$\|Y\|_p \leq \rho_{c_0,p}^{ho}(\rho_{c_1,p}^{ho}(|Y|, \mathcal{F}_1)) \leq c_0 \cdot \|c_1\|_q \cdot \|Y\|_p,$$

where $c_0 \triangleleft \mathcal{F}_0$ is deterministic, $c_1 \triangleleft \mathcal{F}_1$ and q is the exponent conjugate to p , i.e., $\frac{1}{p} + \frac{1}{q} = 1$.

Proof One may repeat the computations from the proof of [Theorem 3.8](#) conditioned on \mathcal{F}_1 to see that

$$\mathbb{E}(|Y|^p | \mathcal{F}_1)^{1/p} \leq \rho_{c_1,p}^{ho}(|Y|, \mathcal{F}_1) \leq c_1 \cdot \mathbb{E}(|Y|^p | \mathcal{F}_1)^{1/p}. \tag{46}$$

It follows that

$$\begin{aligned} \|Y\|_p^p &= \mathbb{E}[\mathbb{E}(|Y|^p | \mathcal{F}_1)] = \left\| \mathbb{E}(|Y|^p | \mathcal{F}_1)^{1/p} \right\|_p^p \\ &\leq \left\| \rho_{c_1,p}^{ho}(|Y|, \mathcal{F}_1) \right\|_p^p \leq \rho_{c_0,p}^{ho}(\rho_{c_1,p}^{ho}(|Y|, \mathcal{F}_1))^p \end{aligned}$$

by (46) and (15), and this is the first claim. As for the second observe that

$$\begin{aligned} \rho_{c_0,p}^{ho}(\rho_{c_1,p}^{ho}(|Y|, \mathcal{F}_1)) &\leq c_0 \left\| \rho_{c_1,p}^{ho}(|Y|, \mathcal{F}_1) \right\|_p \leq c_0 \left\| c_1 \mathbb{E}(|Y|^p | \mathcal{F}_1)^{1/p} \right\|_p \\ &\leq c_0 \cdot \|c_1\|_q \cdot (\mathbb{E}[\mathbb{E}(|Y|^p | \mathcal{F}_1)])^{1/p} = c_0 \cdot \|c_1\|_q \cdot \|Y\|_p \end{aligned}$$

by (16) and (46) again and Hölder’s inequality. This proves the second assertion. □

Example 5.4 (Composition of AV@R) An example of a conditional higher order risk measure, which is been frequently addressed in the literature (cf., for example, [Shapiro 2010](#) or [Philpott and Matos 2012](#)), is the conditional Average Value-at-Risk,

$$AV@R_\alpha(Y, \mathcal{F}_1) := \rho_{c_1,1}^{ho}(Y, \mathcal{F}_1), \tag{47}$$

where c_1 is the constant function $c_1(\cdot) = \frac{1}{1-\alpha}$.

The Average Value-at-Risk is monotone, provided that the risk level is adapted to the stage. We have the following comparison (cf. [Xin and Shapiro 2012](#)).

Proposition 5.5 For $\alpha, \beta \in [0, 1)$ it holds that

$$AV@R_\alpha(AV@R_\beta(Y|\mathcal{F}_1)) \leq AV@R_{\alpha+\beta-\alpha\beta}(Y).$$

Proof The statement is immediate from

$$AV@R_\alpha(AV@R_\beta(Y|\mathcal{F}_1)) = \sup_{\mathbb{E}Z=1, 0 \leq Z \leq \frac{1}{1-\alpha}} \mathbb{E} \left(Z \cdot \text{ess sup}_{\mathbb{E}(Z'|\mathcal{F}_1)=1, 0 \leq Z' \leq \frac{1}{1-\beta}} \mathbb{E}(YZ'|\mathcal{F}_1) \right) \tag{48}$$

(cf. (13)). Note that Z satisfying the constraints can be chosen \mathcal{F}_1 adapted (as the essential supremum is \mathcal{F}_1 -adapted), and it follows thus that $\mathbb{E}[ZZ'] = \mathbb{E}\mathbb{E}(ZZ'|\mathcal{F}_1) = \mathbb{E}[Z\mathbb{E}(Z'|\mathcal{F}_1)] = 1$. It obviously holds that $0 \leq ZZ' \leq \frac{1}{1-\alpha} \frac{1}{1-\beta}$ and thus the assertion. \square

Comparisons with the Average Value-at-Risk The nested Average Value-at-Risk allows a comparison with the Average Value-at-Risk on the positive cone. Indeed, for α, β and $\gamma \in [0, 1)$ it holds that

$$(1 - \gamma) \cdot AV@R_\gamma(Y) \leq AV@R_\alpha(AV@R_\beta(Y|\mathcal{F}_1)) \leq \frac{1}{1 - \beta} AV@R_\beta(Y) \tag{49}$$

provided that $Y \geq 0$ (the bounds in (49) are known for *not* being sharp).

To accept the inequality recall first that $\mathbb{E}Y \leq AV@R_\alpha(Y) \leq \frac{1}{1-\alpha}\mathbb{E}Y$. Then the first inequality follows by monotonicity

$$\begin{aligned} (1 - \gamma) \cdot AV@R_\gamma(Y) &\leq \mathbb{E}Y = \mathbb{E}(\mathbb{E}(Y|\mathcal{F}_1)) \leq \mathbb{E}AV@R_\beta(Y|\mathcal{F}_1) \\ &\leq AV@R_\alpha(AV@R_\beta(Y|\mathcal{F}_1)), \end{aligned}$$

while for the second inequality we find that

$$AV@R_\alpha(AV@R_\beta(Y|\mathcal{F}_1)) \leq AV@R_\alpha \left(\frac{1}{1 - \beta} \mathbb{E}(Y|\mathcal{F}_1) \right) \leq \frac{1}{1 - \beta} AV@R_\beta(Y),$$

the latter inequality being due to [Dentcheva and Ruszczyński \(2003, Corollary 6.30\)](#). A simpler version of the right hand inequality in (49) can be found in [Xin and Shapiro \(2012\)](#).

5.2 Higher order semideviation

Higher order risk measures are defined by [Ruszczyński \(2010, Example 2\)](#) and [Collado et al. \(2012, \(3\)\)](#).

Definition 5.6 The conditional higher order semideviation measure is

$$\rho_{\lambda,p}^{sd}(Y, \mathcal{F}_1) := \mathbb{E}(Y|\mathcal{F}_1) + \lambda \cdot \mathbb{E}((Y - \mathbb{E}(Y|\mathcal{F}_1))_+^p | \mathcal{F}_1)^{1/p}.$$

Remark 5.7 The higher order semideviation risk measure introduced in Theorem 3.9 is $\rho_{\lambda,p}^{sd}(Y) = \rho_{\lambda,p}^{sd}(Y, \mathcal{F}_0)$, where again $\mathcal{F}_0 := \{\emptyset, \Omega\}$ is the trivial sigma algebra.

Remark 5.8 $\rho(Y) := \rho_{\lambda,p}^{sd}(\rho_{\lambda,p}^{sd}(Y, \mathcal{F}_1))$ is a risk functional satisfying (RM)–(RH), but not version independent (i.e., (RD) is not valid).

We have the following multistage analog of Theorem 3.12.

Theorem 5.9 For $\lambda_0 \triangleleft \mathcal{F}_0$ deterministic and $\lambda_1 \triangleleft \mathcal{F}_1$ it holds that

$$\|Y\|_1 \leq \rho_{\lambda_1}^{sd}(\rho_{\lambda_1}^{sd}(Y, \mathcal{F}_1)) \leq (1 + \lambda_0) \cdot \|1 + \lambda_1\|_\infty \cdot \|Y\|_1$$

and

$$\lambda_0 \cdot (\text{ess inf } \lambda_1) \cdot \|Y\|_p \leq \rho_{\lambda_0,p}^{sd}(\rho_{\lambda_1,p}^{sd}(Y, \mathcal{F}_1)) \leq k_{\lambda_0,p} \cdot \|k_{\lambda_1,p}\|_q \cdot \|Y\|_p,$$

where the constants $k_{\lambda,p}$ are as in Theorem 3.12.

Proof The first claim is analogously to the second, so we prove only the second claim.

One may repeat the computations from the proof of Theorem 3.12 conditioned on \mathcal{F}_1 to see that

$$\mathbb{E}(|Y| | \mathcal{F}_1) \leq \rho_{\lambda_1,1}^{sd}(|Y|, \mathcal{F}_1) \leq k_{\lambda_0,p} \cdot \mathbb{E}(|Y| | \mathcal{F}_1)$$

and

$$\lambda_1 \cdot \mathbb{E}(|Y|^p | \mathcal{F}_1)^{1/p} \leq \rho_{\lambda_1,p}^{sd}(|Y|, \mathcal{F}_1) \leq (1 + \lambda_1^q)^{1/q} \cdot \mathbb{E}(|Y|^p | \mathcal{F}_1)^{1/p}. \tag{50}$$

It follows that

$$\begin{aligned} \lambda_0^p \cdot \text{ess inf } \lambda_1^p \cdot \|Y\|_p^p &= \lambda_0^p \cdot \text{ess inf } \lambda_1^p \cdot \mathbb{E}[\mathbb{E}(|Y|^p | \mathcal{F}_1)] \\ &\leq \lambda_0^p \cdot \left\| \lambda_1 \cdot \mathbb{E}(|Y|^p, \mathcal{F}_1)^{1/p} \right\|_p^p \\ &\leq \lambda_0^p \cdot \left\| \rho_{\lambda_1,p}^{sd}(|Y|, \mathcal{F}_1) \right\|_p^p \leq \rho_{\lambda_0,p}^{sd}(\rho_{\lambda_1,p}^{sd}(|Y|, \mathcal{F}_1))^p \end{aligned}$$

by (50) and (21), and this is the first inequality. As for remaining inequality observe that

$$\begin{aligned} \rho_{\lambda_0,p}^{sd}(\rho_{\lambda_1,p}^{sd}(|Y|, \mathcal{F}_1)) &\leq k_{\lambda_0,p} \cdot \left\| \rho_{\lambda_1,p}^{ho}(|Y|, \mathcal{F}_1) \right\|_p \\ &\leq k_{\lambda_0,p} \cdot \left\| k_{\lambda_1,p} \mathbb{E}(|Y|^p | \mathcal{F}_1)^{1/p} \right\|_p \\ &\leq k_{\lambda_0,p} \cdot \|k_{\lambda_1,p}\|_q \cdot \|Y\|_p \end{aligned}$$

by (22) and (50), the second assertion. □

Aggregation of risk— independent sigma algebras The inequalities in Theorems 5.3 and 5.9 are given for a general random variable Y and a sigma algebra \mathcal{F}_1 . The constants in these inequalities improve significantly, if the random variable Y is independent from the sigma algebra \mathcal{F}_1 (i.e., $P(A \cap B) = P(A) \cdot P(B)$ for every $A \in \mathcal{F}_1$ and $B \in \sigma(Y)$, the sigma algebra generated by Y) and $\lambda \triangleleft \mathcal{F}_0$ constant. In this situation it holds that

$$\begin{aligned} \rho_{\lambda,p}^{sd}(Y, \mathcal{F}_1) &= \mathbb{E}(Y | \mathcal{F}_1) + \lambda \cdot \mathbb{E}((Y - \mathbb{E}(Y | \mathcal{F}_1))_+^p | \mathcal{F}_1)^{1/p} \\ &\equiv \mathbb{E}[Y] + \lambda \cdot \mathbb{E}((Y - \mathbb{E}[Y])_+^p)^{1/p} = \rho_{\lambda,p}^{sd}(Y) \end{aligned}$$

and

$$\rho_{c,p}^{ho}(Y, \mathcal{F}_1) \equiv \rho_{c,p}^{ho}(Y)$$

(for $c \triangleleft \mathcal{F}_0$ constant), i.e., the conditional risk measures are *constant* variables.

For Y_2 independent from \mathcal{F}_1 and c_2 constant it thus holds that

$$\rho_{c_1,p}^{ho}\left(Y_1 + \rho_{c_2,p}^{ho}(Y_2 | \mathcal{F}_1)\right) = \rho_{c_1,p}^{ho}(Y_1) + \rho_{c_2,p}^{ho}(Y_2) \tag{51}$$

by (RT), and this identity gives rise to apply conditional risk functionals in a dynamic setting. If, in addition, Y_1 is measurable with respect to \mathcal{F}_1 ($Y_1 \triangleleft \mathcal{F}_1$) and $c_1, c_2 \triangleleft \mathcal{F}_0$, then the equalities extend to

$$\begin{aligned} \rho_{c_1,p}^{ho}\left(\rho_{c_2,p}^{ho}(Y_1 + Y_2, \mathcal{F}_1)\right) &= \rho_{c_1,p}^{ho}\left(Y_1 + \rho_{c_2,p}^{ho}(Y_2, \mathcal{F}_1)\right) \\ &= \rho_{c_1,p}^{ho}(Y_1 + \rho_{c_2,p}^{ho}(Y_2)) = \rho_{c_1,p}^{ho}(Y_1) + \rho_{c_2,p}^{ho}(Y_2), \end{aligned}$$

the risk can be aggregated in an additive way. Apparently, the same is true for the semi-deviation risk measure $\rho_{\lambda,p}^d$.

5.3 The conditional Entropic Value-at-Risk

The Entropic Value-at-Risk is a further risk measure, which allows a conditional version based on conditional expectation by setting

$$\text{EV@R}_\alpha(Y | \mathcal{F}_1) = \text{ess inf}_{t>0, t \triangleleft \mathcal{F}_1} \frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E}\left(e^{tY} | \mathcal{F}_1\right)$$

(cf. (39)).

By altering the risk level α we have the following inequality on the composition of Entropic Value-at-Risks, which is in line with Proposition 5.5.

Proposition 5.10 (Composition of entropic risk measures) *For deterministic risk levels $\alpha, \beta < 1$ we have that*

$$\text{EV@R}_\alpha(\text{EV@R}_\beta(Y | \mathcal{F}_t)) \leq \text{EV@R}_{\alpha+\beta-\alpha\beta}(Y).$$

Proof Let t be optimal in (39) to compute the Entropic Value-at-Risk for the trivial sigma algebra $\mathcal{F} = \{\emptyset, \Omega\}$. Then it follows from the tower property of the conditional expected value that

$$\begin{aligned} \text{EV@R}_{\alpha+\beta-\alpha\beta}(Y) &= \frac{1}{\gamma} \log \frac{1}{1-\alpha-\beta+\alpha\beta} \mathbb{E}e^{\gamma Y} \\ &= \frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E} \exp\left(t \cdot \frac{1}{t} \log \frac{1-\alpha}{(1-\alpha)(1-\beta)} \mathbb{E}(e^{tY} | \mathcal{F}_1)\right) \\ &= \frac{1}{\gamma} \log \frac{1}{1-\alpha} \mathbb{E} \exp\left(\gamma \cdot \frac{1}{\gamma} \log \frac{1}{1-\beta} \mathbb{E}(e^{\gamma Y} | \mathcal{F}_1)\right) \\ &\geq \frac{1}{\gamma} \log \frac{1}{1-\alpha} \mathbb{E} \exp\left(\gamma \cdot \text{ess inf}_{0 \leq \gamma' \triangleleft \mathcal{F}} \frac{1}{\gamma'} \log \frac{1}{1-\beta} \mathbb{E}(e^{\gamma' Y} | \mathcal{F}_1)\right) \\ &= \frac{1}{\gamma} \log \frac{1}{1-\alpha} \mathbb{E} \exp(\gamma \cdot \text{EV@R}_\beta(Y | \mathcal{F}_1)) \\ &\geq \text{EV@R}_\alpha(\text{EV@R}_\beta(Y | \mathcal{F}_1)), \end{aligned}$$

which concludes the proof. □

5.4 The general conditional risk measure

The composite risk measures in the previous subsections are based on conditional expectation. This is enough to introduce a conditional version of the Average Value-at-Risk (cf. (47)). The methods addressed, however, do not constitute a general rule of defining a conditional risk measure based on a given risk measure. But to extend the theory of the previous sections it is necessary to have conditional risk functionals available. Cheridito and Kupper (2011) describe conditional risk functionals and Asamov and Ruszczyński (2014) give characterizations of coherent (and time-consistent) risk measures, but the definitions in Pflug and Pichler (2016), for example, differ, although the conditional Average Value-at-Risk is the same in all references given. For this reason we add a constructive definition of a general, conditional risk functional here, which is again useful in providing estimates.

The Average Value-at-Risk, as introduced in Example 5.4, is well-defined for α a \mathcal{F}_1 -measurable random variable. By noting that the norm corresponding to the AV@R is $\|\cdot\|_1$ with dual $\|\cdot\|_1^* = \|\cdot\|_\infty$ one may repeat the arguments in Theorem 3.1 and obtain the alternative formulation

$$AV@R_\alpha(Y | \mathcal{F}_1) = \text{ess sup} \left\{ \mathbb{E}(YZ | \mathcal{F}_1) : Z \geq 0, \mathbb{E}(Z | \mathcal{F}_1) = \mathbb{1} \text{ and } Z \leq \frac{1}{1 - \alpha} \right\}. \tag{52}$$

The dual representation of a general coherent risk measure,

$$\rho_S(Y) = \sup_{\sigma \in S} \left\{ \mathbb{E}[YZ] : Z \geq 0, \mathbb{E}[Z] = 1 \text{ and } AV@R_\alpha(Z) \leq \frac{1}{1 - \alpha} \int_\alpha^1 \sigma(u)du, \right. \\ \left. 0 \leq \alpha < 1 \right\},$$

can be employed to define the following variant of a conditional risk functional.

Definition 5.11 The conditional risk measure $\rho_S(\cdot | \mathcal{F}_1)$ corresponding to the risk measure $\rho_S(Y) := \sup_{\sigma \in S} \int_0^1 \sigma(u)F_Y^{-1}(u)du$ and the sigma algebra \mathcal{F}_1 is

$$\rho_S(Y | \mathcal{F}_1) := \text{ess sup}_{\sigma \in S} \left\{ \mathbb{E}(YZ | \mathcal{F}_1) \left| \begin{array}{l} Z \geq 0, \mathbb{E}(Z | \mathcal{F}_1) = 1 \text{ and} \\ AV@R_\alpha(Z | \mathcal{F}_1) \leq \frac{1}{1 - \alpha} \int_\alpha^1 \sigma(u)du \text{ for all } 0 \leq \alpha < 1 \end{array} \right. \right\}. \tag{53}$$

We shall write $\rho_\sigma(Y | \mathcal{F}_1)$ if $S = \{\sigma\}$.

The idea in the preceding definition is to repeat the same risk profile on every atom of \mathcal{F}_1 . This corresponds to the definition of the conditional Average Value-at-Risk, for example, where the risk level α is repeated at every node. With this definition the results of Sect. 4 extend analogously and naturally, in line with the results already mentioned for the higher order measure (Sect. 5.1) and the higher order semideviation (Sect. 5.2).

We finally have the following estimates for compositions of risk measures. Note, that all these estimates are actually global comparisons of risk functionals.

Theorem 5.12 (Global comparison of conditional risk functionals) *It holds that*

- (i) $\mathbb{E}[\rho_\sigma(Y | \mathcal{F}_1)] \leq \rho_\sigma(Y)$,
- (ii) $\rho_S(\mathbb{E}(Y | \mathcal{F}_1)) \leq \rho_S(Y)$, and
- (iii) $\text{ess sup } AV@R_\alpha(Y | \mathcal{F}_1) \leq \|Y\|_\infty$.

Remark 5.13 The relation (i) holds for distortion functions ρ_σ and the Average Value-at-Risk, but it does not extend to general risk functionals ρ_S .

Proof We prove the statement for the (conditional) Average Value-at-Risk first. For $\alpha < \mathcal{F}_0$ a fixed constant it follows with (45) that

$$\begin{aligned} \text{AV@R}_\alpha(Y | \mathcal{F}_1) &= \text{ess inf}_{x < \mathcal{F}_1} \left\{ x + \frac{1}{1 - \alpha} \cdot \mathbb{E}((Y - x)_+ | \mathcal{F}_1) \right\} \\ &\leq \text{ess inf}_{x \in \mathbb{R}} \left\{ x + \frac{1}{1 - \alpha} \cdot \mathbb{E}((Y - x)_+ | \mathcal{F}_1) \right\}, \end{aligned}$$

and by taking expectations that

$$\begin{aligned} \mathbb{E}[\text{AV@R}_\alpha(Y | \mathcal{F}_1)] &\leq \mathbb{E} \left[\text{ess inf}_{x \in \mathbb{R}} \left\{ x + \frac{1}{1 - \alpha} \cdot \mathbb{E}((Y - x)_+ | \mathcal{F}_1) \right\} \right] \\ &\leq \inf_{x \in \mathbb{R}} \left\{ \mathbb{E} \left(x + \frac{1}{1 - \alpha} \cdot \mathbb{E}((Y - x)_+ | \mathcal{F}_1) \right) \right\} = \text{AV@R}_\alpha(Y). \end{aligned} \tag{54}$$

Now recall Kusuoka’s theorem and that $\rho_\sigma(Y) = \int_0^1 \text{AV@R}_\alpha(Y) d\mu(\alpha)$ for the measure $\mu(A) := \sigma(0)\delta_0(A) + \int_A 1 - u d\sigma(u)$ [cf. Pflug and Pichler (2014, Section 3.2)]. By integrating (54) with respect to the measure μ and by interchanging the order of integration we obtain

$$\begin{aligned} \mathbb{E}[\rho_\sigma(Y | \mathcal{F}_1)] &= \mathbb{E} \left[\int_0^1 \text{AV@R}_\alpha(Y | \mathcal{F}_1) \mu(d\alpha) \right] \\ &= \int_0^1 \mathbb{E}[\text{AV@R}_\alpha(Y | \mathcal{F}_1)] \mu(d\alpha) \leq \int_0^1 \text{AV@R}_\alpha(Y) \mu(d\alpha) = \rho_\sigma(Y), \end{aligned}$$

the assertion.

The other assertions are obvious. □

Remark 5.14 It is worth noticing that $\text{AV@R}_\alpha(Y) \not\leq \|\text{AV@R}_\alpha(Y | \mathcal{F}_1)\|_\infty$. As a counterexample consider $P(Y = 0) = 75\%$, $P(Y = 20) = 5\%$ and $P(Y = 10) = 20\%$, such that $\text{AV@R}_{75\%}(Y) = 12$. But with $\mathcal{F}_1 = \sigma(\{Y \neq 10\}, \{Y = 10\})$ it holds that $\text{AV@R}_{75\%}(Y | Y \neq 10) = 5$ and $\text{AV@R}_{75\%}(Y | Y = 10) = 10$.

6 Conclusion and summary

This paper elaborates a natural relation between risk measures and norms of a corresponding Banach space. The relation established is one-to-one in the sense that every risk measure defines a norm, and conversely, every appropriate norm specifies a risk measure. We use these observations to establish continuity relations between risk functionals and norms.

A comprehensive collection of examples of important risk measures is included, for which precise upper and lower bounds are given, if available. In this way we obtain (tight) inequalities, relating all commonly used risk functionals. Applications involving risk functionals thus can be simplified by replacing the risk functional by another one, which is more convenient or simpler in implementations.

The second part of the paper addresses compositions of conditional risk functionals. Based on bounds of its components we elaborate multiplicative bounds for composite risk measures.

Acknowledgements We would like to thank the editor of the journal and the referees for their commitment to assess and improve the paper.

References

- Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26, 1505–1518. doi:[10.1016/S0378-4266\(02\)00281-9](https://doi.org/10.1016/S0378-4266(02)00281-9).
- Ahmadi-Javid, A. (2012). Entropic Value-at-Risk: A new coherent risk measure. *Journal of Optimization Theory and Applications*, 155(3), 1105–1123. doi:[10.1007/s10957-011-9968-2](https://doi.org/10.1007/s10957-011-9968-2).
- Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., & Ku, H. (2007). Coherent multiperiod risk adjusted values and Bellman's principle. *Annals of Operations Research*, 152, 5–22. doi:[10.1007/s10479-006-0132-6](https://doi.org/10.1007/s10479-006-0132-6).
- Asamov, T., & Ruszczyński, A. (2014). Time-consistent approximations of risk-averse multistage stochastic optimization problems. *Mathematical Programming*, 153(2), 1–35. doi:[10.1007/s10107-014-0813-x](https://doi.org/10.1007/s10107-014-0813-x).
- Bellini, F., & Caperdoni, C. (2007). Coherent distortion risk measures and higher-order stochastic dominances. *North American Actuarial Journal*, 11(2), 35–42. doi:[10.1080/10920277.2007.10597446](https://doi.org/10.1080/10920277.2007.10597446).
- Bellini, F., & Rosazza Gianin, E. (2008). On Haezendonck risk measures. *Journal of Banking & Finance*, 32(6), 986–994. doi:[10.1016/j.jbankfin.2007.07.007](https://doi.org/10.1016/j.jbankfin.2007.07.007).
- Bellini, F., & Rosazza Gianin, E. (2012). Haezendonck–Goovaerts risk measures and Orlicz quantiles. *Insurance: Mathematics and Economics*, 51(1), 107–114. doi:[10.1016/j.insmatheco.2012.03.005](https://doi.org/10.1016/j.insmatheco.2012.03.005).
- Cheridito, P., & Kupper, M. (2011). Composition of time-consistent dynamic monetary risk measures in discrete time. *International Journal of Theoretical and Applied Finance*, 14(1), 137–162. doi:[10.1142/S0219024911006292](https://doi.org/10.1142/S0219024911006292).
- Collado, R. A., Papp, D., & Ruszczyński, A. (2012). Scenario decomposition of risk-averse multistage stochastic programming problems. *Annals of Operations Research*, 200(1), 147–170. doi:[10.1007/s10479-011-0935-y](https://doi.org/10.1007/s10479-011-0935-y).
- De Lara, M., & Leclère, V. (2016). Building up time-consistency for risk measures and dynamic optimization. *European Journal of Operational Research*, 249, 177–187. doi:[10.1016/j.ejor.2015.03.046](https://doi.org/10.1016/j.ejor.2015.03.046).
- Delbaen, F. (2015). Remark on the paper "Entropic Value-at-Risk: A new coherent risk measure" by Amir Ahmadi-Javid. In P. Barrieu (Ed.), *Risk and stochastic*. World Scientific, ISBN 978-1-78634-194-5.
- Denneberg, D. (1990). Distorted probabilities and insurance premiums. *Methods of Operations Research*, 63, 21–42.
- Densing, M. (2014). Stochastic programming of time-consistent extensions of AVaR. *SIAM Journal on Optimization*, 24(3), 993–1010. doi:[10.1137/130905046](https://doi.org/10.1137/130905046).
- Dentcheva, D., & Ruszczyński, A. (2003). Optimization with stochastic dominance constraints. *SIAM Journal on Optimization*, 14(2), 548–566. doi:[10.1137/S1052623402420528](https://doi.org/10.1137/S1052623402420528).
- Dentcheva, D., Penev, S., & Ruszczyński, A. (2010). Kusuoka representation of higher order dual risk measures. *Annals of Operations Research*, 181, 325–335. doi:[10.1007/s10479-010-0747-5](https://doi.org/10.1007/s10479-010-0747-5).
- Dentcheva, D., Penev, S., & Ruszczyński, A. (2016). Statistical estimation of composite risk functionals and risk optimization problems. *Annals of the Institute of Statistical Mathematics*. doi:[10.1007/s10463-016-0559-8](https://doi.org/10.1007/s10463-016-0559-8).
- Iancu, D. A., Petrik, M., & Subramanian, D. (2015). Tight approximations of dynamic risk measures. *Mathematics of Operations Research*, 40(3), 655–682. doi:[10.1287/moor.2014.0689](https://doi.org/10.1287/moor.2014.0689).
- Krokhmal, P. A. (2007). Higher moment coherent risk measures. *Quantitative Finance*, 7(4), 373–387. doi:[10.1080/14697680701458307](https://doi.org/10.1080/14697680701458307).
- Kusuoka, S. (2001). On law invariant coherent risk measures. In *Advances in mathematical economics*, Chapter 4 (Vol. 3, pp. 83–95). Springer. doi:[10.1007/978-4-431-67891-5](https://doi.org/10.1007/978-4-431-67891-5).
- López-Díaz, M., Sordo, M. A., & Suárez-Llorens, A. (2012). On the L_p -metric between a probability distribution and its distortion. *Insurance: Mathematics and Economics*, 51, 257–264. doi:[10.1016/j.insmatheco.2012.04.004](https://doi.org/10.1016/j.insmatheco.2012.04.004).
- Luna, J. P., Sagastizábal, C., & Solodov, M. (2016). An approximation scheme for a class of risk-averse stochastic equilibrium problems. *Mathematical Programming*, 157(2), 451–481. doi:[10.1007/s10107-016-0988-4](https://doi.org/10.1007/s10107-016-0988-4).
- Miller, N., & Ruszczyński, A. (2011). Risk-averse two-stage stochastic linear programming: Modeling and decomposition. *Operations Research*, 59, 125–132. doi:[10.1287/opre.1100.0847](https://doi.org/10.1287/opre.1100.0847).
- Noyan, N., & Rudolf, G. (2014). Kusuoka representations of coherent risk measures in general probability spaces. *Annals of Operations Research*, 229, 591–605. doi:[10.1007/s10479-014-1748-6](https://doi.org/10.1007/s10479-014-1748-6). (ISSN 0254-5330).
- Pflug, G. C. (2000). Some remarks on the Value-at-Risk and the Conditional Value-at-Risk, Chapter 15. In S. Uryasev (Ed.), *Probabilistic constrained optimization* (Vol. 49, pp. 272–281). New York: Springer.
- Pflug, G. C., & Pichler, A. (2014). *Multistage stochastic optimization*. Springer Series in Operations Research and Financial Engineering: Springer. ISBN 978-3-319-08842-6. doi:[10.1007/978-3-319-08843-3](https://doi.org/10.1007/978-3-319-08843-3).
- Pflug, G. C., & Pichler, A. (2016). Time-consistent decisions and temporal decomposition of coherent risk functionals. *Mathematics of Operations Research*, 41(2), 682–699. doi:[10.1287/moor.2015.0747](https://doi.org/10.1287/moor.2015.0747).

- Pflug, G. C., & Römisch, W. (2007). *Modeling, measuring and managing risk*. River Edge, NJ: World Scientific. doi:[10.1142/9789812708724](https://doi.org/10.1142/9789812708724).
- Pflug, G. C., & Ruszczyński, A. (2005). Measuring risk for income streams. *Computational Optimization and Applications*, 32(1–2), 161–178. ISSN 0926-6003. doi:[10.1007/s10589-005-2058-3](https://doi.org/10.1007/s10589-005-2058-3).
- Philpott, A. B., & de Matos, V. L. (2012). Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. *European Journal of Operational Research*, 218(2), 470–483. doi:[10.1016/j.ejor.2011.10.056](https://doi.org/10.1016/j.ejor.2011.10.056).
- Philpott, A. B., de Matos, V. L., & Finardi, E. (2013). On solving multistage stochastic programs with coherent risk measures. *Operations Research*, 61(4), 957–970. doi:[10.1287/opre.2013.1175](https://doi.org/10.1287/opre.2013.1175).
- Pichler, A. (2013). The natural Banach space for version independent risk measures. *Insurance: Mathematics and Economics*, 53(2), 405–415. doi:[10.1016/j.insmatheco.2013.07.005](https://doi.org/10.1016/j.insmatheco.2013.07.005).
- Pichler, A. (2013). Premiums and reserves, adjusted by distortions. *Scandinavian Actuarial Journal*, 2015(4), 332–351. doi:[10.1080/03461238.2013.830228](https://doi.org/10.1080/03461238.2013.830228).
- Pichler, A., & Shapiro, A. (2015). Minimal representations of insurance prices. *Insurance: Mathematics and Economics*, 62, 184–193. doi:[10.1016/j.insmatheco.2015.03.011](https://doi.org/10.1016/j.insmatheco.2015.03.011).
- Rockafellar, R. T., & Uryasev, S. (2000). Optimization of Conditional Value-at-Risk. *Journal of Risk*, 2(3), 21–41. doi:[10.21314/JOR.2000.038](https://doi.org/10.21314/JOR.2000.038).
- Rockafellar, R.T., Uryasev, S., & Zabarankin, M. (2006). Generalized deviations in risk analysis. *Finance and Stochastics*, 10, 51–74. ISSN 0949-2984. doi:[10.1007/s00780-005-0165-8](https://doi.org/10.1007/s00780-005-0165-8).
- Ruszczyński, A. (2010). Risk-averse dynamic programming for Markov decision processes. *Mathematical Programming Series B*, 125, 235–261.
- Ruszczyński, A., & Shapiro, A. (2006). Conditional risk mappings. *Mathematics of Operations Research*, 31(3), 544–561. doi:[10.1287/moor.1060.0204](https://doi.org/10.1287/moor.1060.0204).
- Ruszczyński, A., & Yao, J. (2015). A risk-averse analog of the Hamilton–Jacobi–Bellman equation. In *Proceedings of the Conference on Control and its Applications*, Chapter 62 (pp. 462–468). Society for Industrial & Applied Mathematics (SIAM). doi:[10.1137/1.9781611974072.63](https://doi.org/10.1137/1.9781611974072.63).
- Shapiro, A. (2010). Analysis of stochastic dual dynamic programming method. *European Journal of Operational Research*, 209, 63–72.
- Shapiro, A. (2013). On Kusuoka representation of law invariant risk measures. *Mathematics of Operations Research*, 38(1), 142–152. doi:[10.1287/moor.1120.0563](https://doi.org/10.1287/moor.1120.0563).
- Shapiro, A. (2016). Rectangular sets of probability measures. *Operations Research*, 64(2), 528–541. doi:[10.1287/opre.2015.1466](https://doi.org/10.1287/opre.2015.1466).
- Shapiro, A., Dentcheva, D., & Ruszczyński, A. (2009). Lectures on stochastic programming. In *MOS-SIAM series on optimization*. SIAM. doi:[10.1137/1.9780898718751](https://doi.org/10.1137/1.9780898718751).
- Sundaresan, K. (1969). Extreme points of the unit cell in Lebesgue–Bochner function spaces. *Proceedings of the American Mathematical Society*, 23(1), 179–184. doi:[10.2307/2037513](https://doi.org/10.2307/2037513).
- van Heerwaarden, A. E., & Kaas, R. (1992). The Dutch premium principle. *Insurance: Mathematics and Economics*, 11, 223–230. doi:[10.1016/0167-6687\(92\)90049-H](https://doi.org/10.1016/0167-6687(92)90049-H).
- Wang, S. S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics*, 17, 43–54. doi:[10.1016/0167-6687\(95\)00010-P](https://doi.org/10.1016/0167-6687(95)00010-P).
- Wojtaszczyk, P. (1991). *Banach spaces for analysts*. Cambridge: Cambridge University Press.
- Wozabal, D. (2010). A framework for optimization under ambiguity. *Annals of Operations Research*, 193(1), 21–47. doi:[10.1007/s10479-010-0812-0](https://doi.org/10.1007/s10479-010-0812-0).
- Wozabal, D. (2014). Robustifying convex risk measures for linear portfolios: A nonparametric approach. *Operations Research*, 62(6), 1302–1315. doi:[10.1287/opre.2014.1323](https://doi.org/10.1287/opre.2014.1323).
- Xin, L., & Shapiro, A. (2012). Bounds for nested law invariant coherent risk measures. *Operations Research Letters*, 40, 431–435. doi:[10.1016/j.orl.2012.09.002](https://doi.org/10.1016/j.orl.2012.09.002).
- Young, V. R. (2006). *Premium Principles*. Encyclopedia of Actuarial Science. Wiley Pennsylvania State University. ISBN 9780470012505. doi:[10.1002/9780470012505.tap027](https://doi.org/10.1002/9780470012505.tap027).