A Quantitative Comparison of Risk Measures

Alois Pichler[*] and Pia O. Schiller[†]

September 14, 2015

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 evaluate 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, which are developed for norms first, allows tight comparisons of risk measures, and tight lower and upper bounds for risk measures are made available. 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, Stochastic dominance

Classification: 90C15, 60B05, 62P05

1 Introduction

Risk measures can 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.

Risk measures are usually defined on a vector space 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 with the genuine norm of the space. This paper elaborates the converse relations. 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.

In this way the risk measure naturally compares with the genuine norm and with the norm induced. We elaborate corresponding comparisons of norms and continuity relations by establishing lower and upper bounds. We provide tight relations wherever possible, so that stochastic programs can be compared easily for different risk measures. A sensitivity analysis thus can assess a subjective preference of a risk measures.

[*]Norwegian University of Science and Technology, NTNU
Contact: alois.pichler@univie.ac.at
[†]University Weimar

Our results are related to Wozabal (2014), who robustifies risk measures with respect to changing the underlying probability space. López-Díaz et al. (2012) employs a specific norm (called an L_p-mectic there), while Bellini and Caperdoni (2007), in contrast, relate and compare risk measures with stochastic dominance relations. The papers Roman et al. (2007) and Wozabal (2014) contain typical applications in operations research and finance (as portfolio 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 De Lara and Leclère (2015). 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 de 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.

In this paper we follow a two-fold approach by comparing risk functionals with other risk functionals as implicitly contained in Iancu et al. (2013). In addition we consider other convex functions, as norms or functionals derived from coherent risk functionals. These convex functionals perhaps violate some usual axioms of coherent risk measures, but they are useful and eligible for comparisons and efficient in computations.

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.

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

2 Preliminaries

Risk measures are often defined on a linear vector space $L \subseteq L^1(\Omega, \mathcal{F}, P)$ of $\mathbb{R}$-valued random variables, and they are typically specified by a subset of the following axioms. Although well-known we repeat these axioms here in their convex form, as we shall relate them to further axioms on norms on the domain space L.

Definition 1 (Axioms for risk functionals)**.** A *version independent, coherent risk measure*[1] is a mapping $\rho:\, L \to \mathbb{R} \cup \{\infty\}$ satisfying the following axioms:

(RM) Monotonicity: $\rho(Y_1) \le \rho(Y_2)$ whenever $Y_1 \le Y_2$ almost surely;

(RC) Convexity $\rho\big((1-\lambda)\,Y_0 + \lambda Y_1\big) \le (1-\lambda)\rho(Y_0) + \lambda\rho(Y_1)$ for $0 \le \lambda \le 1$;

(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 $\lambda > 0$;

[1] often also risk functional, we use both terms

(RD) Version independence: $\rho(Y_1) = \rho(Y_2)$, provided that Y_1 and Y_2 share the same law, i.e., $P(Y_1 \le y) = P(Y_2 \le y)$ for all $y \in \mathbb{R}$.[2]

We shall call the set

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

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 functionals 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 $\rho(Y) = \operatorname{ess\,sup} Y$. $\rho_Z(Y) := \mathbb{E}\, Z \cdot Y$ (with $\mathbb{E}Z = 1$ and $Z \ge 0$ almost everywhere) is an example of a risk functional, which is not necessarily version independent.

Norms. We relate risk functionals with norms on a Banach 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 space $(L, \|\cdot\|)$.

(NM) Monotonicity: $\|Y_1\| \le \|Y_2\|$ whenever $|Y_1| \le |Y_2|$ 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:[3]

$$\sup_{\|Z\|^* \le 1} \mathbb{E}\,(Y \cdot Z) = \sup_{Z \ge 0,\, \|Z\|^* \le 1} \mathbb{E}\,(Y \cdot Z) \quad \text{whenever } Y \ge 0 \text{ a.s.}, \tag{1}$$

where

$$\|Z\|^* := \sup_{\|Y\| \le 1} \mathbb{E}\,(Y \cdot Z) \tag{2}$$

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

Remark 2. Note that (NP) is not more than the Hahn-Banach theorem for the bi-linear form $(Y, Z) \mapsto \mathbb{E}YZ$, just restricted to the nonnegative orthant L_+.

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

$$\|Z\|^* \ge \mathbb{E}\, Z, \tag{3}$$

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

$$\mathbb{E}\, Y \le \|\mathbb{1}\|^* \cdot \|Y\|\,. \tag{4}$$

[2] also *law invariant*, or *distribution based*

[3] (NP) is mnemonic for positive

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

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

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

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

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. *Let $\|\cdot\|$ be a norm satisfying (NM)–(NP). Then, for $c \ge 1$,*[4]

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

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

$$\rho_c^{ho}(Y) = \sup\left\{\mathbb{E}(Y\,Z):\ Z\ge 0,\, \mathbb{E}\,Z = 1 \text{ and } \|Z\|^* \le c\right\}. \tag{7}$$

Remark 4. We mention Krokhmal (2007) for a construction leading to (6). The functional

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

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

Remark 5. The constraint $c \ge 1$ is a necessary condition, as otherwise the risk measure ρ_c^{ho} is degenerate and, e.g., $\rho_c^{ho}(\mathbb{1}) = -\infty$.

Remark 6. The relation $\rho_c^{ho}(Y) \le \rho_{c'}^{ho}(Y)$ – provided that $c \le c'$ – is obvious.

[4] $x_+ := \max\{0, x\}$

Proof of Theorem 3. Observe first that

$$c\cdot\left\|(Y-x)_+\right\| = \sup_{\|Z\|^*\le c} \mathbb{E}\,(Y-x)_+\cdot Z \ge \sup_{\substack{\mathbb{E}Z=1,\\ Z\ge 0,\ \|Z\|^*\le c}} \mathbb{E}\,(Y-x)_+\cdot Z$$

by the Hahn–Banach theorem, and thus

$$x+c\cdot\left\|(Y-x)_+\right\| \ge \sup_{\substack{\mathbb{E}Z=1,\\ Z\ge 0,\ \|Z\|^*\le c}} \mathbb{E}\,\left(x+(Y-x)_+\right)\cdot Z \ge \sup_{\substack{\mathbb{E}Z=1,\\ Z\ge 0,\ \|Z\|^*\le c}} \mathbb{E}\,Y\cdot Z,$$

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

As for the converse inequality observe that

$$\sup_{\substack{\mathbb{E}Z=1,\\ Z\ge 0,\ \|Z\|^*\le c}} \mathbb{E}\,YZ = \sup_{\substack{Z\ge 0,\\ \|Z\|^*\le c}} \inf_{x\in\mathbb{R}} x+\mathbb{E}\,(Y-x)Z,$$

as one may drop the constraint $\mathbb{E}Z=1$: indeed, $\inf_{x\in\mathbb{R}} x+\mathbb{E}\,(Y-x)Z=\mathbb{E}Y+\inf_{x\in\mathbb{R}} x\cdot(1-\mathbb{E}Z)=-\infty$ 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 $\inf_{x\in\mathbb{R}} x+\mathbb{E}\,(Y-x)Z=x_Z+\mathbb{E}\,(Y-x_Z)_+\cdot Z$ for $x_Z:=\operatorname{ess\,inf}\{Y:\ Z>0\}$ (note that this quantity is well-defined by the density assumption (ND)). So it follows that

$$\sup_{\substack{\mathbb{E}Z=1,\\ Z\ge 0,\ \|Z\|^*\le c}} \mathbb{E}\,YZ = \sup_{\substack{Z\ge 0,\\ \|Z\|^*\le c}} x_Z+\mathbb{E}\,(Y-x_Z)_+\cdot Z, \tag{9}$$

and by the Hahn–Banach theorem and (NP) again (as $(Y-x_Z)_+\ge 0$) that

$$(9) = x_Z+c\cdot\left\|(Y-x_Z)_+\right\| \ge \inf_{x\in\mathbb{R}} x+c\cdot\left\|(Y-x)_+\right\|,$$

the remaining inequality. ☐

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 7 (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*

$$\|Y\|\le\rho_c^{ho}\left(|Y|\right)\le c\cdot\|Y\|\,.$$

The lower bound is sharp. Further it holds that $\rho_c^{ho}(Y)\le\|Y\|_\infty$.

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

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

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

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

By passing to the infimum thus

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

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

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

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

As for the second assertion choose $x = \|Y\|_\infty$ in (6). ☐

Remark 8 (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

$$\rho_1^{ho}(Y) := \inf_{x \in \mathbb{R}} x + \rho\big((Y - x)_+\big), \tag{10}$$

and it holds that $\rho(Y) = \rho_1^{ho}(Y)$ for $Y \in L_+$.

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 (10).

Corollary 9 (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

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

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

$$\mathsf{AV@R}_\alpha(Y) := \inf_{x\in\mathbb{R}} x + \frac{1}{1-\alpha}\mathbb{E}(Y-x)_+ = \sup\left\{\mathbb{E}YZ:\, 0\le Z\le \frac{1}{1-\alpha} \text{ and } \mathbb{E}Z=1\right\}$$

are well-known indeed (cf. Rockafellar and Uryasev (2000) and Pflug (2000)), but here they are just a special case of Theorem 3 for the norm $\|\cdot\|_1$ and its dual $\|\cdot\|_1^* = \|\cdot\|_\infty$. Anticipating the terminology of the following section (Section 3.1 below) the Average Value-at-Risk at level α is a higher order measure, $\mathsf{AV@R}_\alpha(\cdot) = \rho^{ho}_{\frac{1}{1-\alpha}}(\cdot)$ with $c=\frac{1}{1-\alpha}$ and $p=1$.

Note as well the global bound

$$\mathsf{AV@R}_\alpha(Y) \le \rho^{ho}_{\frac{\|\mathbb{1}\|^*}{1-\alpha}}(Y), \tag{11}$$

a simple consequence of (4).

Further it follows from Theorem 7 readily that

$$\mathbb{E}Y \le \mathsf{AV@R}_\alpha(|Y|) \le \frac{1}{1-\alpha}\mathbb{E}\,|Y|\,. \tag{12}$$

Lipschitz continuity

$$|\mathsf{AV@R}_\alpha(Y) - \mathsf{AV@R}_\alpha(Y')| \le \frac{1}{1-\alpha}\,\|Y-Y'\|_1$$

is finally a consequence of the previous theorem (Corollary 9).

3.1 Higher order risk measures

Higher order risk measures constitute a special case of the risk measure (6) addressed in Theorem 3, as the L^p-norm $\|\cdot\|_p$ takes the role of the general norm $\|\cdot\|$. We denote them by

$$\rho^{ho}_{c,p}(Y) := \inf_{x\in\mathbb{R}} x + c\cdot\left\|(Y-x)_+\right\|_p, \tag{13}$$

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

For this particular choice we have the following extension of Theorem 3 with precise bounds.

Theorem 10 (Comparison with norms)**.** *Assume that $1\le p\le p'$. It holds that*

(i)

$$\|Y\|_p \le \rho^{ho}_{c,p'}(|Y|)\,, \tag{14}$$

(ii) and

$$\rho^{ho}_{c,p}(Y) \le c^{p/p'}\cdot\|Y\|_{p'}\,. \tag{15}$$

Both bounds are moreover sharp.

Proof. The first inequality is immediate from Theorem 3. For the second inequality recall the dual representation (7) (see also Dentcheva et al. (2010)) that

$$\rho^{ho}_{c,p}(|Y|) = \sup\left\{\mathbb{E}YZ:\, \|Z\|_q\le c,\, Z\ge 0 \text{ and } \mathbb{E}Z=1\right\},$$

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

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

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

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

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

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

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

3.2 Higher order semideviation

The higher order semideviation is a risk measure discussed in Shapiro et al. (2009). Again, the dual representation is straight forward and natural in the chosen context of norms. 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 precise, explicit bounds.

Theorem 11. *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

$$\rho_\lambda^{sd}(Y)=\sup_{Z\geq 0,\,\mathbb{E}Z=1}\left(1-\frac{\lambda}{\|Z\|^*}\right)\mathbb{E}Y+\frac{\lambda}{\|Z\|^*}\mathbb{E}YZ \tag{16}$$

holds true.

Remark 12. Recall from (3) that $\|Z\|^*\geq\mathbb{E}Z=1$, such that (16) 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

$$\rho_\lambda^{sd}(Y) = \mathbb{E}Y + \lambda \cdot \|(Y - \mathbb{E}Y)_+\| = \sup_{Z \geq 0} \mathbb{E}Y + \frac{\lambda}{\|Z\|^*} \mathbb{E}Z \cdot (Y - \mathbb{E}Y), \tag{17}$$

as it is enough to restrict the supremum to $Z \geq 0$ by (NP). Note that Z is in the nominator and the deonomiator, 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. □

Theorem 13 (Comparison of the higher order semideviation with the norm)**.** *It holds that*

$$\|Y\|_1 \leq \rho_\lambda^{sd}(|Y|) \leq (\|\mathbb{1}\|^* + \lambda) \cdot \|Y\| \tag{18}$$

and

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

Proof. The first inequality in (18), $\|Y\|_1 = \mathbb{E}\,|Y| \leq \rho_p^{sd}(|Y|)$, 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

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

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

$$\lambda \cdot \|Y\| \leq \mathbb{E}\,|Y| + \lambda \cdot \left\|(|Y| - \mathbb{E}\,|Y|)_+\right\| = \rho_p^{sd}(|Y|).$$

As for the upper bound recall from Theorem 11 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

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

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

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

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 14 (Comparison with L^p-norms)**.** *The following holds true:*

(i) if $p = 1$, then
$$\|Y\|_1 \leq \rho^{sd}_{\lambda,1}(|Y|) \leq (1+\lambda)\cdot\|Y\|_1\,; \tag{19}$$

(ii) if $1 < p < 2$, then
$$\lambda\cdot\|Y\|_p \leq \rho^{sd}_{\lambda,p}(|Y|) \leq k_{\lambda,p}\cdot\|Y\|_p\,, \tag{20}$$

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
$$\lambda\cdot\|Y\|_p \leq \rho^{sd}_{p}(|Y|) \leq (1+\lambda^q)^{1/q}\cdot\|Y\|_p\,, \tag{21}$$

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 15. 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 (Section 5) below.

Proof of Theorem 14. (i) and the first inequality of (ii) and (iii) are immediate from Theorem 13.

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

$$\begin{aligned}\rho^{sd}_p(|Y|) &= P(A)^{1-\frac{1}{p}} + \lambda\left(P(A)\left(\frac{1}{P(A)^{1/p}} - P(A)^{1-\frac{1}{p}}\right)^p\right)^{\frac{1}{p}}\\ &= P(A)^{1-\frac{1}{p}} + \lambda P(A)^{\frac{1}{p}}\left(P(A)^{-\frac{1}{p}} - P(A)^{1-\frac{1}{p}}\right)\\ &= P(A)^{1-\frac{1}{p}} + \lambda\big(1-P(A)\big) \xrightarrow[P(A)\to 0]{} \lambda,\end{aligned}$$

provided that $p > 1$. It follows that the lower bound in (20) and (21) cannot be improved.

From Equation (16) in Theorem 11 it follows that

$$\begin{aligned}\rho^{sd}_\lambda(Y) &= \sup_{\mathbb{E}Z=1,\,Z\geq 0}\mathbb{E}\left(\Big(1-\frac{\lambda}{\|Z\|^*}\Big)\mathbb{1}+\frac{\lambda}{\|Z\|^*}Z\right)\cdot Y,\\ &\leq \sup_{\mathbb{E}Z=1,\,Z\geq 0}\left\|\Big(1-\frac{\lambda}{\|Z\|^*}\Big)\mathbb{1}+\frac{\lambda}{\|Z\|^*}Z\right\|_q\cdot\|Y\|_p\end{aligned} \tag{22}$$

by Hölder's inequality. Equality in (22), however, is attained for

$$Y = \left(\Big(1-\frac{\lambda}{\|Z\|^*}\Big)\mathbb{1}+\frac{\lambda}{\|Z\|^*}Z\right)^{p-1}, \tag{23}$$

such that

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

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.[5] Then

$$\begin{aligned}\left\|\Big(1-\frac{\lambda}{\|Z\|^*}\Big)\mathbb{1}+\frac{\lambda}{\|Z\|^*}Z\right\|_q^q &= \left\|1-\frac{\lambda}{\|\sigma_\beta\|_q}+\frac{\lambda}{\|\sigma_\beta\|_q}\sigma_\beta\right\|_q^q\\ &= (1-\beta)\left(1-\lambda\beta^{\frac{q-1}{q}}\right)^q+\beta\left(1-\lambda\beta^{\frac{q-1}{q}}+\lambda\frac{1}{\beta}\beta^{\frac{q-1}{q}}\right)^q\\ &= (1-\beta)\left(1-\lambda\beta^{\frac{q-1}{q}}\right)^q+\left(\lambda(1-\beta)+\beta^{\frac{1}{q}}\right)^q. \end{aligned}\tag{24}$$

By letting $\beta\to 0$ it follows that $\lim_{\beta\to 0}(24)=1+\lambda^q$, from which we conclude that

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

The upper bound can be rewritten as $\sup\left\{\rho^{sd}(|Y|):\,\|Y\|_p\leq 1\right\}$. As $Y\mapsto\rho^{sd}(|Y|)$ 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.,

$$|Y|=\frac{\mathbb{1}_A}{P(A)^{1/p}}\tag{25}$$

(put $A:=A_1\cup A_2$). Now if $|Y|$ is as in (25), then, because of (23) and (24), 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 (24) with respect to $\beta\in[0,1]$.

But (24) is a strictly decreasing in β for $p\geq 2$ (i.e., $q\leq 2$), such that the maximum as attained for $\beta\to 0$.

For $p>2$, the maximum in (24) is attained in the interior, at some $\beta\in(0,1)$. However, a closed analytic expression is not available. □

Example 16 (Dutch risk measure (cf. also Example 25 below))**.** The Dutch risk measure (or Dutch premium principle, cf. van Heerwaarden and Kaas (1992)) is the higher order semideviation risk measure

$$\rho_{1,\lambda}^{sd}(Y):=\mathbb{E}Y+\lambda\cdot\mathbb{E}\left(Y-\mathbb{E}Y\right)_+,\tag{26}$$

where $0\leq\lambda<1$. Shapiro (2013) provides the additional representation

$$\rho_{1,\lambda}^{sd}(Y):=\sup_{\kappa\in[0,1]}(1-\kappa\lambda)\cdot\mathbb{E}Y+\kappa\lambda\cdot\mathsf{AV@R}_{1-\kappa}(Y).\tag{27}$$

It follows from (12) that $\mathsf{AV@R}_{1-\kappa}(Y)\leq\frac{1}{\kappa}\mathbb{E}\,|Y|$, 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 (19) presented in Theorem 14.

[5] U is uniformly distributed, iff $P(U\leq u)=u$.

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 a non-homogeneous way.

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

$$\rho_1(Y) \leq K \cdot \rho_2(Y) \qquad \text{for all } Y. \tag{28}$$

Then it follows from translation equivariance (RT) that $\rho_1(Y) + c = \rho_1(Y + c) \leq K \cdot \rho_2(Y + c) = K \cdot c + K \cdot \rho_2(Y)$, or $\rho_1(Y) \leq \rho_2(Y) + (K - 1)c$ for all $c \in \mathbb{R}$. This inequality is impossible unless $K = 1$. A comparison of risk functionals in the general, global form (28) thus can only read

$$\rho_1(Y) \leq \rho_2(Y) \qquad \text{for all } Y, \tag{29}$$

(i.e., $K = 1$ in (28)).

Only a few risk measures allow a general, global comparison as in (29). Examples of a global comparison include the relation (11) in the previous section, and a standard inequality for the Average Value-at-Risk,

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

where $\alpha \leq \alpha'$ — a special case of $\rho_c^{ho}(Y) \leq \rho_{c'}^{ho}(Y)$ for $c \leq c'$ (cf. Remark 6). In what follows we state the global relations, if available.

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 (29) we thus consider

$$\rho(Y) \leq K \cdot \rho'(Y) \qquad \text{for all } Y \geq 0 \text{ a.s.} \tag{30}$$

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.

Remark 17. 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 (30) available. This approach is in line with the result reported in Theorem 7, and particularly with Remark 8.

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); they are ralated to extreme points in the dual set). The distortion risk functional (also spectral risk measure, cf. Acerbi (2002)) is defined by

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

where $F_Y^{-1}(u) := \inf\{y:\ P(Y \le y) \ge u\}$ is the generalized inverse of the cumulative distribution function (cdf) $F_Y(y) := P(Y \le y)$. $\sigma : [0,1) \to [0,\infty)$ is a nonnegative, nondecreasing function satisfying $\int_0^1 \sigma(u)\mathrm{d}u = 1$. The function $\sigma(\cdot)$ is called *distortion functional*. For convenience we shall associate the function $\Sigma(p) := \int_p^1 \sigma(u)\mathrm{d}u$ (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 18. *For a random variable $Y \in L_+$ in the nonnegative cone the distortion risk functional ρ_σ can be evaluated by the alternative expression*

$$\rho_\sigma(Y) = \int_0^\infty \Sigma\big(F_Y(y)\big)\mathrm{d}y.$$

Proof. It follows from Riemann–Stieltjes integration by parts that

$$\begin{aligned} \rho_\sigma(Y) &= \int_0^1 \sigma(u) F_Y^{-1}(u)\mathrm{d}u = -\int_0^1 F_Y^{-1}(u)\mathrm{d}\Sigma(u), \\ &= -F_Y^{-1}(u)\cdot \Sigma(u)\big|_{u=0}^1 + \int_0^1 \Sigma(u)\mathrm{d}F_Y^{-1}(u). \end{aligned}$$

As $Y \ge 0$ by assumption, we have that $F_Y^{-1}(0)\cdot\Sigma(0) = 0\cdot 1 = 0$, and thus

$$\rho_\sigma(Y) = \int_0^1 \Sigma(u)\mathrm{d}F_Y^{-1}(u) = \int_0^\infty \Sigma\big(F_Y(y)\big)\mathrm{d}y$$

after changing the variables. This completes the proof. □

The following result is essential to compare distortion risk functionals.

Theorem 19 (Compararison of spectral risk measures)**.** *Suppose that*

$$K := \sup_{0\le\alpha<1} \frac{\int_\alpha^1 \sigma_1(u)\,\mathrm{d}u}{\int_\alpha^1 \sigma_2(u)\,\mathrm{d}u} \tag{31}$$

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

$$\rho_{\sigma_1}(Y) \le K \cdot \rho_{\sigma_2}(Y), \tag{32}$$

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

Proof. Notice first that $\int_\alpha^1 \sigma_1(u)\,\mathrm{d}u \le K\cdot\int_\alpha^1 \sigma_2(u)\,\mathrm{d}u$ for all $\alpha\ge 0$. It follows from Riemann–Stieltjes integration by parts and as the quantile function $u\mapsto F_Y^{-1}(u)$ is nondecreasing, that

$$\begin{aligned}
\rho_{\sigma_1}(Y) &= \int_0^1 F_Y^{-1}(u)\,\sigma_1(u)\,\mathrm{d}u = -\int_0^1 F_Y^{-1}(u)\,\mathrm{d}\Sigma_1(u)\\
&= -\left.F_Y^{-1}(u)\,\Sigma_1(u)\right|_0^1 + \int_0^1 \Sigma_1(u)\,\mathrm{d}F_Y^{-1}(u) = F_Y^{-1}(0) + \int_0^1 \Sigma_1(u)\,\mathrm{d}F_Y^{-1}(u)\\
&\le F_Y^{-1}(0) + K\cdot \int_0^1 \Sigma_2(u)\,\mathrm{d}F_Y^{-1}(u)\\
&= F_Y^{-1}(0) + K\cdot \left.F_Y^{-1}(u)\,\Sigma_2(u)\right|_0^1 - K\cdot\int_0^1 F_Y^{-1}(u)\,\mathrm{d}\Sigma_2(u)\\
&= -F_Y^{-1}(0)\,(K-1) + K\cdot\int_0^1 F_Y^{-1}(u)\,\sigma_2(u)\,\mathrm{d}u.
\end{aligned}$$

Note next that $K\ge 1$ (choose $\alpha=0$ in (31)) and $F_Y^{-1}(0)\ge 0$, and it readily follows that

$$\rho_{\sigma_1}(Y)\le K\cdot\int_0^1 F_Y^{-1}(u)\,\sigma_2(u)\,\mathrm{d}u = K\cdot\rho_{\sigma_2}(Y).$$

To accept that K is the smallest constant satisfying (32) just consider the random variable $Y=\mathbb{1}_{A^c}$, for which $\rho_\sigma(Y)=\rho_\sigma(\mathbb{1}_{A^c})=\int_{P(A)}^1\sigma(u)\,\mathrm{d}u$. The assertion follows, as the measurable set A may be chosen arbitrarily.

Finally notice that $K\ge\limsup_{\alpha\to 1}\frac{\int_\alpha^1\sigma_1(u)\mathrm{d}u}{\int_\alpha^1\sigma_2(u)\mathrm{d}u}=\limsup_{\alpha\to 1}\frac{\sigma_1(\alpha)}{\sigma_2(\alpha)}$ by L'Hôpital's rule. ☐

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

Example 20 (Average Value-at-Risk)**.** The distortion function of the Average Value-at-Risk is

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

For $\alpha_1\le\alpha_2$ the non-trivial constant is $K=\frac{1-\alpha_1}{1-\alpha_2}$ and as a particular consequence of (32) it follows that

$$\mathsf{AV@R}_{\alpha_1}(|Y|)\le\mathsf{AV@R}_{\alpha_2}(|Y|)\le\frac{1-\alpha_1}{1-\alpha_2}\mathsf{AV@R}_{\alpha_1}(|Y|)\,. \tag{33}$$

Particularly, $\mathsf{AV@R}_\alpha(|Y|)\le\frac{1}{1-\alpha}\mathbb{E}\,|Y|$.

The risk measure $\rho(Y):=\beta\cdot\mathbb{E}Y+(1-\beta)\cdot\mathsf{AV@R}_\alpha(Y)$ 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 (27)). 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

$$\beta\cdot\mathbb{E}Y+(1-\beta)\cdot\mathsf{AV@R}_\alpha(Y)\le\mathsf{AV@R}_{\frac{\alpha-\alpha\beta}{1-\alpha\beta}}(Y) \tag{34}$$

follows by elementary computations and Remark 8, where the adapted risk level is smaller than the initial risk level α, $\frac{\alpha-\alpha\beta}{1-\alpha\beta}\leq\alpha$. The global bound (34) is sharp again. A lower, sharp bound expressed for a general risk level α' is

$$\mathsf{AV@R}_{\alpha'}(Y)\leq\frac{1}{\Sigma_{\alpha,\beta}(\alpha')}\big(\beta\cdot\mathbb{E}Y+(1-\beta)\cdot\mathsf{AV@R}_{\alpha}(Y)\big)\qquad(Y\geq 0).\tag{35}$$

(Recall that the comparisons (33) and (35) are not valid in case that $Y\notin L_+$. Finally, as a special case we note that $\frac{1}{\Sigma_{\alpha,\beta}\left(\frac{\alpha-\alpha\beta}{1-\alpha\beta}\right)}=\frac{1-\alpha\beta}{1-\alpha\beta(2-\beta)}$ for the critical risk level in (34).)

Example 21 (Proportional hazards risk measure)**.** The proportional hazards risk measure (adapted from insurance, cf. Young (2006)), is

$$\rho_c^H(Y)=\int_0^\infty\big(1-F_Y(u)\big)^c\mathrm{d}u$$

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

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

for every $c>0$, again an immediate consequence of Theorem 19 with $\Sigma(\alpha)=(1-\alpha)^c$.

Example 22 (Wang transform)**.** Wang's risk measure (cf. Wang (1995)) employs the parametric family $\Sigma_\lambda^W(u)=\Phi\left(\lambda+\Phi^{-1}(1-u)\right)$ for $\lambda\geq 0$, it is defined as

$$\rho_\lambda^W(Y)=\int_0^\infty\Sigma_\lambda^W\big(F_Y(y)\big)\mathrm{d}y$$

on the nonnegative cone ($\Phi(\cdot)$ is the cdf of the normal distribution). 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

$$\mathsf{AV@R}_{\alpha}(Y)\leq\frac{1}{\Sigma_\lambda^W(\alpha)}\rho_\lambda^W(Y)\qquad(Y\geq 0).$$

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

$$\rho_\lambda^W(Y)\leq K_{\lambda,c}\cdot\rho_c^H(Y),\tag{36}$$

although an explicit expression for the constant $K_{\lambda,c}$ is not available. A converse inequality to (36) 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

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

on the nonnegative cone L_+ one may choose

$$K := \sup_{\sigma_1\in S_1}\inf_{\sigma_2\in S_2}\sup_{\alpha<1}\frac{\int_\alpha^1 \sigma_1(u)\mathrm{d}u}{\int_\alpha^1 \sigma_2(u)\mathrm{d}u}, \tag{37}$$

which is an immediate consequence of (32) in Theorem 19, althought K is possibly not the best constant. It is evident from the max-min inequality that

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

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

Example 23 (Entropic Value-at-Risk)**.** The entropic Value-at-Risk introduced in Ahmadi-Javid (2011) is

$$\mathsf{EV@R}_\alpha(Y) := \inf_{z>0}\frac{1}{z}\log\left(\frac{1}{1-\alpha}\mathbb{E}e^{zY}\right).$$

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

$$\mathsf{EV@R}_\alpha(Y) = \sup\left\{\rho_\sigma(Y):\int_0^1\sigma(u)\log\sigma(u)\mathrm{d}u \le \log\frac{1}{1-\alpha}\right\}$$

using Kullback-Leiber divergence. We refer to the original reference Ahmadi-Javid (2011, Proposition 3.2) for the global comparison

$$\mathsf{EVaR}_\alpha(Y) \le \mathsf{AV@R}_\alpha(Y),$$

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 $Y=\mathbb{1}_{[\beta,1]}(U)$, $\beta\ne\alpha$.

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

Theorem 24 (Comparison of *different* risk measures)**.** *The relation*

$$\rho_\lambda^{sd}(Y) \le \rho_{1+\lambda}^{ho}(Y) \tag{38}$$

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

For arbitrary chosen $0\le\lambda\le 1$ *and* $c\ge 1$ *it holds that*

$$\rho_c^{ho}(Y) \le \frac{c}{\lambda}\cdot\rho_\lambda^{sd}(|Y|) \tag{39}$$

and conversely

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

Proof. Assume first that $Y \geq 0$. Recall from (17) 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

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

Hence

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

by (7), which is the desired estimate on L_+. The general assertion for $L \supseteq L_+$ follows from Remark 8.

Recall finally from Theorem 7 and Theorem 13 that

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

and

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

to establish the remaining relations. □

Example 25 (Dutch risk measure, cf. Example 16)**.** As an application for the general constant (37) consider the Dutch risk measure (26). By applying the relations (26) and (38) from the previous theorem to the norm $\|\cdot\|_1$ it follows that

$$\mathbb{E}\, Y + \lambda \cdot \mathbb{E}(Y - \mathbb{E}\, Y)_+ \leq \mathsf{AV@R}_{\frac{\lambda}{1+\lambda}}(Y). \tag{40}$$

The *cricial risk level* $\frac{\lambda}{1+\lambda}$ is the smallest possible risk level, for which (40) holds true. In contrast to the estimates already presented in Example 16 this is a global and tight upper bound for the Dutch risk measure.

To obtain a tight lower bound one may apply (37) and (35) to compute the constant K (cf. (27)), resulting in

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

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

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

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

This results further in the sandwich inequality

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

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 $\mathsf{AV@R}$ 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

$$\rho_\sigma^*(Y) := \sup_{\alpha<1} \frac{\mathsf{AV@R}_\alpha(Y)}{\frac{1}{1-\alpha}\int_\alpha^1 \sigma(u)\mathrm{d}u}. \tag{41}$$

The associated functional naturally provides the upper bound

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

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

The functional (41) satisfies the axioms of a risk measures (Definition 1), except translation equivariance (RT) and thus is not a risk measure.

As a corollary of Theorem 19 we have the following relations for the constant ρ_σ^*.

Corollary 26. *For $Y \geq 0$ it holds that*

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

where K is the constant (31) (Theorem 19).

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

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

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

$$\mathsf{AV@R}_\alpha^\tau(Y) := \inf_{t\in\mathbb{R}}\ x + \frac{1}{1-\alpha}\mathbb{E}(Y-x)_{\tau,+}$$

where $y_{\tau,+} = \frac{1}{2}\left(y + \sqrt{y^2 + 4\tau^2}\right)$ (we refer to Luna et al. (2015) 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

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

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 not translation equivariant, and 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 differentiable, while $y \mapsto y_+$ is only subdifferentiable.

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

$$\rho_\sigma(Y) \leq \mathbb{E}\, h(Y) + \int_0^1 h^*(\sigma(u))\mathrm{d}u \tag{42}$$

is valid for every (measurable) function $h : \mathbb{R} \to \mathbb{R}$. (42) 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 (2015) demonstrates that the relation (42) is sharp, that is, for every random variable $Y \in L$ there exists a function h such that equality holds in (42).

5 Composite 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). 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 and Densing (2014). Shapiro (2015) elaborates properties of these risk measures in the context of convex analysis, while Dentcheva et al. (2015) 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.

5.1 Higher order measures

For the definition of higher order risk measures we refer to Ruszczyński (2010, Example 3) and the references given therein.

Definition 27. The conditional higher order measure is

$$\rho_{c,p}^{ho}(Y, \mathcal{F}) := \operatorname*{ess\,inf}_{x \lhd \mathcal{F}}\; x + c \cdot \mathbb{E}\big((Y-x)_+^p \,|\, \mathcal{F}\big)^{1/p}, \tag{43}$$

where $c \lhd \mathcal{F}$; $x \lhd \mathcal{F}$ ($c \lhd \mathcal{F}$, resp.) is shorthand for x (c, resp.) being measurable with respect to the sigma algebra $\mathcal{F}$. For convenience it is accepted in the literature to write also $\rho_{c,p}^{ho}(Y|\,\mathcal{F}) := \rho_{c,p}^{ho}(Y, \mathcal{F})$.

Remark 28. With $\mathcal{F}_0 := \{\emptyset, \Omega\}$ we denote the trivial sigma algebra. Then $\rho_{c,p}^{ho}(Y) = \rho_{c,p}^{ho}(Y, \mathcal{F}_0)$, the higher order measure introduced in (13).

Lemma 29 (Composition of risk measures)**.** *The composite risk measure* $Y \mapsto \rho^{ho}_{c,p}\left(\rho^{ho}_{c,p}(Y, \mathcal{F})\right)$ *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 30. *For the composite risk measure it holds that*

$$\|Y\|_p \le \rho^{ho}_{c_0,p}\left(\rho^{ho}_{c_1,p}(|Y|,\, \mathcal{F})\right) \le c_0 \cdot \|c_1\|_q \cdot \|Y\|_p\,,$$

where $c_0 \lhd \mathcal{F}_0$ *is deterministic,* $c_1 \lhd \mathcal{F}$ *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 10 conditioned on $\mathcal{F}$ to see that

$$\mathbb{E}\left(|Y|^p|\, \mathcal{F}\right)^{1/p} \le \rho^{ho}_{c_1,p}(|Y|,\, \mathcal{F}) \le c_1 \cdot \mathbb{E}\left(|Y|^p\, |\mathcal{F}\right)^{1/p}. \tag{44}$$

It follows that

$$\begin{aligned} \|Y\|_p^p &= \mathbb{E}\mathbb{E}\left(|Y|^p\,|\,\mathcal{F}\right) = \left\|\mathbb{E}\left(|Y|^p\,|\,\mathcal{F}\right)^{1/p}\right\|_p^p \\ &\le \left\|\rho^{ho}_{c_1,p}(|Y|\,,\mathcal{F})\right\|_p^p \le \rho^{ho}_{c_0,p}\left(\rho^{ho}_{c_1,p}(|Y|\,,\mathcal{F})\right)^p \end{aligned}$$

by (44) and (14), and this is the first claim. As for the second observe that

$$\begin{aligned} \rho^{ho}_{c_0,p}\left(\rho^{ho}_{c_1,p}(|Y|,\mathcal{F})\right) \le c_0 \left\|\rho^{ho}_{c_1,p}(|Y|,\mathcal{F})\right\|_p \le c_0 \left\|c_1 \mathbb{E}\left(|Y|^p\,|\mathcal{F}\right)^{1/p}\right\|_p \\ \le c_0 \cdot \|c_1\|_q \cdot \left(\mathbb{E}\,\mathbb{E}\left(|Y|^p\,|\mathcal{F}\right)\right)^{1/p} = c_0 \cdot \|c_1\|_q \cdot \|Y\|_p \end{aligned}$$

by (15) and (44) again and Hölder's inequality. This proves the second assertion. ☐

Example 31 (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 de Matos (2012)), is the conditional Average Value-at-Risk,

$$\mathsf{AV@R}_\alpha(Y, \mathcal{F}) := \rho^{ho}_{c_1,1}(Y, \mathcal{F}), \tag{45}$$

where c_1 is the constant function $c_1(\cdot) = \frac{1}{1-\alpha}$. It is immediate from Theorem 30 that

$$\|Y\|_1 \le \mathsf{AV@R}_\alpha\big(\mathsf{AV@R}_\alpha(Y, \mathcal{F})\big) \le \frac{1}{(1-\alpha)^2}\,\|Y\|_1$$

and

$$\|Y\|_1 \le \mathbb{E}\,\mathsf{AV@R}_\alpha(Y, \mathcal{F})\big) \le \frac{1}{1-\alpha}\,\|Y\|_1$$

whenever $Y \ge 0$ (cf. Xin and Shapiro (2012) for a related result).

5.2 Higher order semideviation

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

Definition 32. The conditional higher order semideviation measure is

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

Remark 33. For the trivial sigma algebra $\mathcal{F}_0 := \{\emptyset,\,\Omega\}$, $\rho_{\lambda,p}^{sd}(Y) = \rho_{\lambda,p}^{sd}(Y,\,\mathcal{F}_0)$, the higher order semideviation risk measure introduced in Theorem 11.

Lemma 34. *$\rho(Y) := \rho_{\lambda,p}^{sd}\left(\rho_{\lambda,p}^{sd}(Y,\mathcal{F})\right)$ 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 14.

Theorem 35. *For $\lambda_0 \lhd \mathcal{F}_0$ deterministic and $\lambda_1 \lhd \mathcal{F}$ it holds that*

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

and

$$\lambda_0\cdot(\operatorname{ess\,inf}\lambda_1)\cdot\|Y\|_p \le \rho_{\lambda_0,p}^{sd}\left(\rho_{\lambda_1,p}^{sd}(Y,\mathcal{F})\right) \le k_{\lambda_0,p}\cdot\|k_{\lambda_1,p}\|_q\cdot\|Y\|_p\,,$$

where the constants $k_{\lambda,p}$ are as in Theorem 14.

Proof. The first claim is analoguously to the second, so we pdove only the second claim.

One may repeat the computations from the proof of Theorem 14 conditioned on $\mathcal{F}$ to see that

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

and

$$\lambda_1\cdot\mathbb{E}(|Y|^p|\,\mathcal{F})^{1/p} \le \rho_{\lambda_1,p}^{sd}(|Y|,\,\mathcal{F}) \le (1+\lambda_1^q)^{1/q}\cdot\mathbb{E}(|Y|^p\,|\mathcal{F})^{1/p}. \tag{46}$$

It follows that

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

by (46) and (20), and this is the first inequality. As for remaining inequality observe that

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

by (21) and (46), the second assertion. □

Aggregation of risk – independent sigma algebras. The inequalities in Theorems 30 and 35 are given for a general random variable Y and a sigma algebra $\mathcal{F}$. The inequalities improve significantly, if the random variable Y is independent from the sigma algebra $\mathcal{F}$ (i.e., $P(A\cap B)=P(A)\cdot P(B)$ for every $A\in\mathcal{F}$ and $B\in\sigma(Y)$, the sigma algebra generated by Y) and $\lambda\lhd\mathcal{F}_0$ is constant. In this situation it holds that

$$\begin{aligned}\rho^{sd}_{\lambda,p}\left(Y,\mathcal{F}\right) &= \mathbb{E}(Y|\,\mathcal{F})+\lambda\cdot\mathbb{E}\big((Y-\mathbb{E}(Y|\,\mathcal{F}))_+^p\,|\,\mathcal{F}\big)^{1/p}\\ &\equiv \mathbb{E}Y+\lambda\cdot\mathbb{E}\big((Y-\mathbb{E}Y)_+^p\big)^{1/p}=\rho^{sd}_{\lambda,p}(Y)\end{aligned}$$

and

$$\rho^{ho}_{c,p}\left(Y,\mathcal{F}\right)\equiv\rho^{ho}_{c,p}(Y)$$

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

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

$$\rho^{ho}_{c_1,p}\left(Y_1+\rho^{ho}_{c_2,p}\left(Y_2|\,\mathcal{F}\right)\right)=\rho^{ho}_{c_1,p}\left(Y_1\right)+\rho^{ho}_{c_2,p}\left(Y_2\right) \tag{47}$$

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 $\mathcal{F}$ ($Y_1\lhd\mathcal{F}$) and c_1, $c_2\lhd\mathcal{F}_0$, then the equalities extend to

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

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

5.3 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. (45)). 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 (2015), 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 useful in providing estimates.

The Average Value-at-Risk, as introduced in Example 31, is well-defined for α a $\mathcal{F}_1$-measurable random variable. By noting that the norm corresponding to the $\mathsf{AV@R}$ is $\|\cdot\|_1$ with dual $\|\cdot\|_1^*=\|\cdot\|_\infty$ one may repeat the arguments in Theorem 3 and obtain the alternative formulation

$$\mathsf{AV@R}_\alpha(Y|\,\mathcal{F})=\operatorname{ess\,sup}\left\{\mathbb{E}(YZ|\,\mathcal{F}):\;Z\geq 0,\;\mathbb{E}(Z|\,\mathcal{F})=1\text{ and }Z\leq\frac{1}{1-\alpha}\right\}. \tag{48}$$

The dual representation of a general coherent risk measure,

$$\rho_S(Y)=\sup_{\sigma\in S}\left\{\mathbb{E}\,YZ:\;Z\geq 0,\;\mathbb{E}Z=1\text{ and }\mathsf{AV@R}_\alpha(Z)\leq\frac{1}{1-\alpha}\int_\alpha^1\sigma(u)\mathrm{d}u,\;0\leq\alpha<1\right\},$$

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

Definition 36. The *conditional risk measure* $\rho_S(\cdot|\,\mathcal{F})$ corresponding to the risk measure $\rho_S(Y) := \sup_{\sigma\in S}\int_0^1 \sigma(u)F_Y^{-1}(u)\mathrm{d}u$ and the sigma algebra $\mathcal{F}$ is

$$\rho_S(Y|\,\mathcal{F}) := \operatorname*{ess\,sup}_{\sigma\in S}\left\{\mathbb{E}(YZ|\,\mathcal{F}) \;\middle|\; \begin{array}{l} Z\geq 0,\ \mathbb{E}(Z|\,\mathcal{F})=1 \text{ and} \\ \mathsf{AV@R}_\alpha(Z|\,\mathcal{F}) \leq \frac{1}{1-\alpha}\int_\alpha^1 \sigma(u)\mathrm{d}u \text{ for all } 0\leq\alpha<1 \end{array}\right\}.$$

The conditional norm, associated with the conditional risk measure, is $\|(Y|\,\mathcal{F})\|_\sigma := \rho_\sigma(|Y|,\mathcal{F})$.

The particular advantage of this definition is the intuition, as the profile σ is repeated on every atom of $\mathcal{F}$. With this definition the results of Section 4 extend analogously and naturally, in line with the results already mentioned for the higher order measure (Subsection 5.1) and the higher order semideviation (Subsection 5.2).

We finally have the following estimate for compositions of risk measures.

Theorem 37. *It holds that*

$$\mathbb{E}\,\mathsf{AV@R}_\alpha\,(Y|\,\mathcal{F}) \leq \mathsf{AV@R}_\alpha(Y),$$
$$\mathsf{AV@R}_\alpha\big(\mathbb{E}\,(Y|\,\mathcal{F})\big) \leq \mathsf{AV@R}_\alpha(Y)$$

and[6]

$$\operatorname{ess\,sup}\mathsf{AV@R}_\alpha\,(Y|\,\mathcal{F}) \leq \|Y\|_\infty\,.$$

$$\mathbb{E}\rho_S(Y|\,\mathcal{F}) \leq \rho_S(Y).$$

Proof. We prove the statement for the Average Value-at-Risk first. For $\alpha \lhd \mathcal{F}_0$ a fixed constant it follows with (43) that

$$\begin{aligned}\mathsf{AV@R}_\alpha\,(Y|\,\mathcal{F}) &= \operatorname*{ess\,inf}_{x\lhd\mathcal{F}}\ x + \frac{1}{1-\alpha}\cdot\mathbb{E}\big((Y-x)_+\,|\,\mathcal{F}\big)\\ &\leq \operatorname*{ess\,inf}_{x\in\mathbb{R}}\ x + \frac{1}{1-\alpha}\cdot\mathbb{E}\big((Y-x)_+\,|\,\mathcal{F}\big),\end{aligned}$$

and by taking expectations that

$$\begin{aligned}\mathbb{E}\,\mathsf{AV@R}_\alpha\,(Y|\,\mathcal{F}) &\leq \mathbb{E}\operatorname*{ess\,inf}_{x\in\mathbb{R}}\ x + \frac{1}{1-\alpha}\cdot\mathbb{E}\big((Y-x)_+\,|\,\mathcal{F}\big)\\ &\leq \inf_{x\in\mathbb{R}}\mathbb{E}\left(x + \frac{1}{1-\alpha}\cdot\mathbb{E}\big((Y-x)_+\,|\,\mathcal{F}\big)\right) = \mathsf{AV@R}_\alpha\,(Y)\,.\end{aligned}$$

Suppose further that Z is feasible for (48), then $Z\geq 0$, $\mathbb{E}Z=1$ and $Z\leq\frac{1}{1-\alpha}$, that is, Z is feasible for the Average Value-at-Risk and thus □

[6] It should be noted that $\mathsf{AV@R}_\alpha(Y) \not\leq \|\mathsf{AV@R}_\alpha\,(Y|\,\mathcal{F})\|_\infty$. As a counterexample consider $P(Y=20)=5\,\%$, $P(Y=0)=75\,\%$ and $P(Y=10)=20\,\%$, such that $\mathsf{AV@R}_{75\%}(Y)=12$. But with $\mathcal{F}=\sigma\left(\{Y\neq 10\},\{Y=10\}\right)$ it holds that $\mathsf{AV@R}_{75\%}(Y|\,Y\neq 10)=5$ and $\mathsf{AV@R}_{75\%}(Y|\,Y=10)=10$.

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 every appropriate norm specifies a risk measure. We use these observations to establish continuity relations between risk functionals and norms. Many examples of important risk measures are included, for which precise upper and lower bounds are given, if available. In this way we get a comprehensive collection of (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 are simpler in implementations.

The second part of the paper addresses composite risk functionals. We elaborate multiplicative bounds for composite risk measures, based on bounds of its components.

7 Acknowledgment

The author gratefully acknowledges support of the Research Council of Norway (grant 207690/E20).

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

References

C. Acerbi. Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26:1505–1518, 2002.

A. Ahmadi-Javid. Entropic value-at-risk: A new coherent risk measure. *Journal of Optimization Theory and Applications*, 155(3):1105–1123, Dec 2011.

P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, and H. Ku. Coherent multiperiod risk adjusted values and Bellman's principle. *Annals of Operations Research*, 152:5–22, 2007.

T. Asamov and A. Ruszczyński. Time-consistent approximations of risk-averse multistage stochastic optimization problems. *Mathematical Programming*, pages 1–35, 2014.

F. Bellini and C. Caperdoni. Coherent distortion risk measures and higher-order stochastic dominances. *North American Actuarial Journal*, 11(2):35–42, 2007.

F. Bellini and E. R. Gianin. On Haezendonck risk measures. *Journal of Banking & Finance*, 32(6): 986–994, 2008.

F. Bellini and E. R. Gianin. Haezendonck–Goovaerts risk measures and Orlicz quantiles. *Insurance: Mathematics and Economics*, 51(1):107–114, 2012.

P. Cheridito and M. Kupper. Composition of time-consistent dynamic monetary risk measures in discrete time. *International Journal of Theoretical and Applied Finance*, 14(1):137–162, 2011.

R. A. Collado, D. Papp, and A. Ruszczyński. Scenario decomposition of risk-averse multistage stochastic programming problems. *Annals of Operations Research*, 200(1):147–170, 2012.

M. De Lara and V. Leclère. Building up time-consistency for risk measures and dynamic optimization. *European Journal of Operational Research*, (0):–, 2015.

F. Delbaen. Remark on the paper "entropic value-at-risk: A new coherent risk measure" by Amir Ahmadi-Javid, j. opt. theory and appl., 155 (2001),1105–1123. arXiv:1504.00640, April 2015. arXiv.org.

M. Densing. Stochastic progamming of time-consistent extensions of AVaR. *SIAM Journal on Optimization*, 24(3):993–1010, 2014.

D. Dentcheva, S. Penev, and A. Ruszczyński. Kusuoka representation of higher order dual risk measures. *Annals of Operations Research*, 181:325–335, 2010.

D. Dentcheva, S. Penev, and A. Ruszczyński. Statistical estimation of composite risk functionals and risk optimization problems. Number 3 in Stochastic Programming E-Print Series. Institut für Mathematik, 2015. Unpublished.

D. A. Iancu, M. Petrik, and D. Subramanian. Tight approximations of dynamic risk measures. *Mathematics of Operations Research*, 2013.

P. A. Krokhmal. Higher moment coherent risk measures. *Quantitative Finance*, 7(4):373–387, 2007.

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

M. López-Díaz, M. A. Sordo, and A. Suárez-Llorens. On the L_p-metric between a probability distribution and its distortion. *Insurance: Mathematics and Economics*, 51:257–264, 2012.

J. P. Luna, C. Sagastizábal, and M. Solodov. An approximatioin scheme for a class of risk-averse stochastic equilibrium problems. 2015.

N. Miller and A. Ruszczyński. Risk-averse two-stage stochastic linear programming: Modeling and decomposition. *Operations Research*, 59:125–132, 2011.

G. Ch. Pflug. Some remarks on the Value-at-Risk and the Conditional Value-at-Risk. In S. Uryasev, editor, *Probabilistic Constrained Optimization*, volume 49 of *Nonconvex Optimization and its Application*, pages 272–281. Springer US, 2000.

G. Ch. Pflug and A. Pichler. Time consistent decisions and temporal decomposition of coherent risk functionals. *Mathematics of Operations Research*, 2015.

G. Ch. Pflug and A. Ruszczyński. Measuring risk for income streams. *Computational Optimization and Applications*, 32(1-2):161–178, 2005.

A. B. Philpott and V. L. de Matos. Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. *European Journal of Operational Research*, 218(2):470–483, 2012.

A. B. Philpott, V. L. de Matos, and E. Finardi. On solving multistage stochastic programs with coherent risk measures. *Operations Research*, 61(4):957–970, 2013.

A. Pichler. The natural Banach space for version independent risk measures. *Insurance: Mathematics and Economics*, 53(2):405–415, 2013.

A. Pichler. Premiums and reserves, adjusted by distortions. *Scandinavian Actuarial Journal*, 2015 (4):332–351, 2015.

A. Pichler and A. Shapiro. Minimal representations of insurance prices. *Insurance: Mathematics and Economics*, 62:184–193, 2015.

R. T. Rockafellar and S. Uryasev. Optimization of Conditional Value-at-Risk. *Journal of Risk*, 2: 21–41, 2000.

D. Roman, K. Darby-Dowman, and G. Mitra. Mean-risk models using tow risk measures: a multi-objective approach. *Quantitative Finance*, 7(4):443–458, August 2007.

A. Ruszczyński. Risk-averse dynamic programming for Markov decision processes. *Math. Program., Ser. B*, 125:235–261, 2010.

A. Ruszczyński and A. Shapiro. Conditional risk mappings. *Mathematics of Operations Research*, 31(3):544–561, 2006.

A. Ruszczyński and J. Yao. *A Risk-Averse Analog of the Hamilton–Jacobi–Bellman Equation*, chapter 62, pages 462–468.

A. Shapiro. Analysis of stochastic dual dynamic programming method. *European Journal of Operational Research*, 209:63–72, 2010.

A. Shapiro. On Kusuoka representation of law invariant risk measures. *Mathematics of Operations Research*, 38(1):142–152, 2013.

A. Shapiro. Rectangular sets of probability measures. Optimization Online, 2015.

A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming.* MOS-SIAM Series on Optimization. 2009.

K. Sundaresan. Extreme points of the unit cell in Lebesgue–Bochner function spaces. *Proceedings of the American Mathematical Society*, 23(1):179–184, 1969.

A. E. van Heerwaarden and R. Kaas. The Dutch premium principle. *Insurance: Mathematics and Economics*, 11:223–230, 1992.

S. S. Wang. Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics*, (17):43–54, 1995.

P. Wojtaszczyk. *Banach Spaces for Analysts.* Cambridge University Press, Cambridge, UK, 1991.

D. Wozabal. Robustifying convex risk measures for linear portfolios: A nonparametric approach. *Operations Research*, 62(6):1302–1315, 2014.

L. Xin and A. Shapiro. Bounds for nested law invariant coherent risk measures. *Operations Research Letters*, 40:431–435, 2012.

V. R. Young. *Premium Principles.* Encyclopedia of Actuarial Science. John Wiley & Sons, Ltd, Pennsylvania State University, 2006. ISBN 9780470012505.