Solvency capital, risk measures and comonotonicity: a review

Jan Dhaene Steven Vanduffel Qihe Tang Marc J. Goovaerts
Rob Kaas David Vyncke

July 14, 2004

Abstract

In this paper we examine and summarize properties of several well-known risk measures that can be used in the framework of setting solvency capital requirements for a risky business. Special attention is given to the class of (concave) distortion risk measures. We investigate the relationship between these risk measures and theories of choice under risk. Furthermore we consider the problem of how to evaluate risk measures for sums of non-independent random variables. Approximations for such sums, based on the concept of comonotonicity, are proposed. Several examples are provided to illustrate properties or to prove that certain properties do not hold. Although the paper contains several new results, it is written as an overview and pedagogical introduction to the subject of risk measurement. The paper is an extended version of Dhaene et al. (2003).

1 Introduction

Insurance company risks can be classified in a number of ways, see for instance the "Report of the IAA's Working Party on Solvency"[1]. One possible way of classification is to distinguish between financial risks (asset risks and liability risks) and operational risks, see e.g. Nakada, Shah, Koyluogo & Collignon (1999).

Insurance operations are liability driven. In exchange for a fixed premium, the insurance company accepts the risk to pay the claim amounts related to the insured events. *Liability risks* (also called technical risks) focus on the nature of the risk that the insurance company is assuming by selling insurance contracts. They can be subdivided into non-catastrophic risks (like claims volatility) and catastrophic risks (like September 11)

[1] "Report of the IAA's Working Party on Solvency", 2002, available at www.actuaries.org under "IAA Documents", "Papers".

The insurance company will hold assets to meet its future liabilities. *Asset risks* (or investment risks) are associated with insurers' asset management. They are often subdivided in credit risks (like the issuer of a bond gets ruined) and market risks (like depreciation risk).

Risks that cannot be classified as either asset or liability risks are called *operational risks* and are subdivided in business risks (like lower production than expected) and event risks (like system failure).

A risk measure is defined as a mapping from the set of random variables representing the risks at hand to the real numbers. We will always consider random variables as losses, or payments that have to be made. A negative outcome for the loss variable means that a gain has occurred. The real number denoting a general risk measure associated with the loss random variable Y will be denoted by $\rho[Y]$. Common risk measures in actuarial science are premium principles, see for instance Goovaerts, De Vijlder & Haezendonck (1984), or also chapter 5 in Kaas, Goovaerts, Dhaene & Denuit (2001). Other risk measures are used for determining provisions and capital requirements of an insurer, in order to avoid insolvency. Then they measure the upper tails of distribution functions. Such measures of risk are considered in Artzner, Delbaen, Eber & Heath (1999), Wirch & Hardy (2000), Panjer (2002), Dhaene, Goovaerts & Kaas (2003), Tsanakas & Desli (2003), among others. In this paper, we will concentrate on risk measures that can be used for reserving and solvency purposes.

Let X be the random variable representing the insurance company's risks related to a particular policy, a particular line-of-business or to the entire insurance portfolio over a specified time horizon. We do not specify what kind of risk X is. It could be one specific risk type, such as credit risk for all assets. Or it could be a sum of dependent risks $X_1 + \cdots + X_n$, where the X_i represent the different risk types such as market risk, event risk and so on, or where the X_i represent the claims related to the different policies of the portfolio.

Ensuring that insurers have the financial means to meet their obligations to pay the present and future claims related to policyholders is the purpose of solvency[2]. In order to avoid insolvency over the specified time horizon at some given level of risk tolerance, the insurer should hold assets of value $\rho[X]$ or more. Essentially, $\rho[X]$ should be such that $\Pr[X > \rho[X]]$ is 'small enough'. Note that $\rho[X]$ is a risk measure expressed in monetary terms. It could be defined for instance as the 99-th percentile of the distribution function of X.

A portion of the assets held by the company finds its counterpart on the right hand

[2] "On Solvency, Solvency Assessments and Actuarial Issues, An IAIS Issues Paper", 2000, available at www.iaisweb.org under "Publications".

side of the balance sheet as liabilities (technical provisions or actuarial reserves). The value of these liabilities will be denoted by $P[X]$. The rest of the assets match the equity of the company. Alternative names for equity are surplus or capital.

The 'required capital' will be denoted by $K[X]$. It is defined as the excess of the insurer's required assets over its liabilities: $K[X] = \rho[X] - P[X]$. Alternative names for required capital are minimum capital, minimum surplus, required surplus, capital adequacy reserve, risk-based capital or solvency margin, see e.g. Atkinson & Dallas (2000).

In order to determine the required capital $K[X]$, the value of the liabilities $P[X]$ has to be determined. Since liabilities of insurance companies can in general not be traded efficiently in open markets, they cannot be 'marked to market', but have to be determined by a 'mark to model' approach. Hence, $P[X]$ could be defined as a 'fair value' of the liabilities. The liabilities $P[X]$ could be defined as the 75-th percentile of the distribution of X, or they could be defined as the expected value $E[X]$ increased by some additional prudency margin, or they could be evaluated using a 'replicating portfolio' approach .

The definition of 'required capital' is general in the sense that it can be used to define 'regulatory capital', 'rating agency capital' as well as 'economic capital', depending on the risk measure that is used and the way how the liabilities are evaluated. Regulatory and rating agency capital requirements are often determined using aggregate industry averages. In this case, they may not sufficiently reflect the risks of the particular company under consideration. On the other hand, if they are based on customized internal models, which is an emerging trend, they will reflect the individual company's risk more accurately.

The reference period over which insolvency has to be avoided has to be chosen carefully, taking into account the long-term commitments inherent in insurance products. It might be the time needed to run-off the whole portfolio, or it may be a fixed time period such as one year, in which case X also includes provisions to be set up at the end of the period.

The optimal level of risk tolerance will depend on several considerations such as the length of the reference period, as well as policyholders' concerns and owners' interests. A longer reference period will allow a lower level of risk tolerance. Regulatory authorities and rating agencies want sufficiently high levels of capital because holding more capital increases the capacity of the company to meet its obligations. Tax authorities, on the other hand, will not allow insurance companies to avoid taxes on profits by using these profits to increase the level of the capital. Furthermore, the more capital held, the lower the return on equity. Therefore, the shareholders of the company will only be willing to provide a sufficiently large capital $K[X]$ if they are sufficiently rewarded for it. This 'cost of capital' is covered by the policyholders who will have to pay an extra premium for it, see e.g. Bühlmann (1985).

In order to verify if the actual available capital is in accordance with the desired risk tolerance level, the insurer has to compare the computed monetary value $\rho[X]$ with the value of the assets. It seems obvious to valuate the assets by their market value.

Our definition of 'required capital' is related to one of the definitions of economic capital in the "SOA Specialty Guide on Economic Capital"[3]: Economic capital is 'the excess of the market value of the assets over the fair value of the liabilities required to ensure that obligations can be satisfied at a given level of risk tolerance, over a specified time horizon'.

As pointed out in the "Issues paper on solvency, solvency assessments and actuarial issues"[4] an insurance company's solvency position is not fully determined by its solvency margin alone. In general an insurer's solvency relies on a prudent evaluation of the technical provisions, on the investment of the assets corresponding to these technical provisions in accordance with quantitative and qualitative rules and finally also on the existence of an adequate solvency margin.

In this paper, we will concentrate on risk measures $\rho[X]$ that can be used in determining the 'total balance sheet capital requirement' which is the sum of both liabilities and solvency capital requirement: $\rho[X] = P[X] + K[X]$.

As mentioned above, the risk X will often be a sum of non-independent risks. Hence, we will consider the general problem of determining approximations for risk measures of sums of random variables of which the dependency structure is unknown or too cumbersome to work with.

In Section 2 we introduce several well-known risk measures and the relations that hold between them. Characterizations for ordering concepts in terms of risk measures are explored in Section 3. The concept of comonotonicity is introduced in Section 4. The class of distortion risk measures is examined in Section 5. Approximations for distortion risk measures of sums of non-independent random variables, as well as the relationship between theories of choice under risk and distortion risk measures are considered. Section 6 concludes the paper.

2 Some well-known risk measures

As a first example of a risk measure, consider the p-quantile risk measure, often called the 'VaR' (Value-at-Risk) at level p in the financial and actuarial literature. For any p

[3] "Specialty Guide on Economic Capital", SOA 2003, available at www.soa.org under "Sections, RMTF, Subgroups, Economic Capital and Allocation".

[4] "Issues Paper on Solvency, Solvency Assessment and Actuarial Issues", IAIS 2000, available at www.iaisweb.org under "Publications".

in $(0,1)$, the p-quantile risk measure for a random variable X, which will be denoted by $Q_p(X)$, is defined by

$$Q_p[X] = \inf\{x \in \mathbb{R} \mid F_X(x) \geq p\}, \qquad p \in (0,1), \tag{1}$$

where $F_X(x) = \Pr[X \leq x]$. We also introduce the risk measure $Q_p^+[X]$ which is defined by

$$Q_p^+[X] = \sup\{x \in \mathbb{R} \mid F_X(x) \leq p\}, \qquad p \in (0,1). \tag{2}$$

Note that only values of p corresponding to a horizontal segment of F_X lead to different values of $Q_p[X]$ and $Q_p^+[X]$.

Let X denote the aggregate claims of an insurance portfolio. The liabilities (provisions) for this portfolio are given by P. Assume the insurer establishes a solvency capital $K = Q_p[X] - P$ with p sufficiently large, e.g. $p = 0.99$. In this case, the capital can be interpreted as the 'smallest' capital such that the insurer becomes technically insolvent, i.e.claims exceed provisons and capital, with a (small) probability of at most $1-p$:

$$K = \inf\{L \mid \Pr[X > P + L] \leq 1 - p\} \tag{3}$$

Using the p-quantile risk measure for determining a solvency capital is meaningful in situations where the default event should be avoided, but the size of the shortfall is less important. For shareholders or management e.g., the quantile risk measure gives useful information since avoiding default is the primary concern, whereas the size of the shortfall is only secondary.

Expression (1) can also be used to define $Q_0[X]$ and $Q_1[X]$. For the latter quantile, we take the convention $\inf \emptyset = +\infty$. We find that $Q_0(X) = -\infty$. For a bounded random variable X, we have that $Q_1[X] = \max(X)$. Note that $Q_p[X]$ is often denoted by $F_X^{-1}(p)$. The quantile function $Q_p[X]$ is a non-decreasing and left-continuous function of p. In the sequel, we will often use the following equivalence relation which holds for all $x \in \mathbb{R}$ and $p \in [0,1]$:

$$Q_p[X] \leq x \Leftrightarrow p \leq F_X(x). \tag{4}$$

Note that the equivalence relation (4) holds with equalities if F_X is continuous at this particular x.

A single quantile risk measure of a predetermined level p does not give any information about the thickness of the upper tail of the distribution function from $Q_p[X]$ on. A regulator for instance is not only concerned with the frequency of default, but also about the severity of default. Also shareholders and management should be concerned with the question "how bad is bad?" when they want to evaluate the risks at hand in a consistent way. Therefore, one often uses another risk measure which is called the Tail Value-at-Risk (TVaR) at level p. It is denoted by $\mathrm{TVaR}_p[X]$, and defined by

$$\mathrm{TVaR}_p[X] = \frac{1}{1-p}\int_p^1 Q_q[X]\ \mathrm{d}q, \qquad p \in (0,1). \tag{5}$$

It is the arithmetic average of the quantiles of X, from p on. Note that the TVaR is always larger than the corresponding quantile. From (5) it follows immediately that the Tail Value-at-Risk is a non-decreasing function of p.

Let X again denote the aggregate claims of an insurance portfolio over a given reference period and P the provision for this portfolio. Setting the capital equal to $\mathrm{TVaR}_p[X] - P$, we could define 'bad times' as those where X takes a value in the interval $[Q_p[X],\ \mathrm{TVaR}_p[X]]$. Hence, 'bad times' are those where the aggregate claims exceed the threshold $Q_p[X]$, but not using up all available capital. The width of the interval is a 'cushion' that is used in case of 'bad times'. For more details, see Overbeck (2000).

The Conditional Tail Expectation (CTE) at level p will be denoted by $\mathrm{CTE}_p[X]$. It is defined as

$$\mathrm{CTE}_p[X] = \mathrm{E}[X \mid X > Q_p[X]], \qquad p \in (0,1). \tag{6}$$

Loosely speaking, the conditional tail expectation at level p is equal to the mean of the top $(1-p)\%$ losses. It can also be interpreted as the VaR at level p augmented by the average exceedance of the claims X over that quantile, given that such exceedance occurs.

The Expected Shortfall (ESF) at level p will be denoted by $\mathrm{ESF}_p[X]$, and is defined as

$$\mathrm{ESF}_p[X] = \mathrm{E}\left[(X - Q_p[X])_+\right], \qquad p \in (0,1). \tag{7}$$

This risk measure can be interpreted as the expected value of the shortfall in case the capital is set equal to $Q_p[X] - P$.

The following relations hold between the four risk measures defined above.

Theorem 1 (Relation between Quantiles, TVaR, CTE and ESF). *For $p \in (0,1)$, we have that*

$$TVaR_p[X] = Q_p[X] + \frac{1}{1-p} ESF_p[X], \tag{8}$$

$$CTE_p[X] = Q_p[X] + \frac{1}{1-F_X(Q_p[X])} ESF_p[X], \tag{9}$$

$$CTE_p[X] = TVaR_{F_X(Q_p[X])}[X]. \tag{10}$$

Proof. Expression (8) follows from

$$\begin{aligned}\mathrm{ESF}_p[X] &= \int_0^1 \left(Q_q[X] - Q_p[X]\right)_+ \mathrm{d}q \\ &= \int_p^1 Q_q[X]\mathrm{d}q - Q_p[X]\,[1-p]\,.\end{aligned}$$

Expression (9) follows from

$$\mathrm{ESF}_p[X] = \mathrm{E}\left[X - Q_p[X] \mid X > Q_p[X]\right]\left(1 - F_X\left(Q_p[X]\right)\right). \tag{11}$$

Expression (10) follows immediately from (8) and (9). ∎

About the Tail Value-at-Risk, from Definition (5) we have the following elementary result, which will be applied later: if X has a finite expectation $\mathrm{E}[X]$, then

$$\lim_{p \searrow 0} \mathrm{TVaR}_p\,[X] = \mathrm{E}[X]. \tag{12}$$

Note that if F_X is continuous then

$$\mathrm{CTE}_p\,[X] = \mathrm{TVaR}_p\,[X]\,, \qquad p \in (0,1)\,. \tag{13}$$

In the sequel, we will often use the following lemma, which expresses the quantiles of a function of a random variable in terms of the quantiles of the random variable.

Lemma 1 (Quantiles of transformed random variables). *Let X be a real-valued random variable, and $0 < p < 1$. For any non-decreasing and left continuous function g, it holds that*

$$Q_p\,[g(X)] = g\left(Q_p\,[X]\right). \tag{14}$$

On the other hand, for any non-increasing and right continuous function g, one has

$$Q_p\,[g(X)] = g(Q_{1-p}^{+}\,[X]). \tag{15}$$

A proof of this result can be found e.g. in Dhaene, Denuit, Goovaerts, Kaas & Vyncke (2002a). As an application of Lemma 1, we immediately find that

$$\mathrm{E}\left[X \mid X < Q_p^{+}\,[X]\right]) = -\mathrm{CTE}_{1-p}\,[-X] \tag{16}$$

holds for any $p \in (0,1)$.

Example 1 (Normal losses).

Consider a random variable $X \sim N(\mu, \sigma^2)$ which is normally distributed with mean μ and variance σ^2. From Lemma 1, it follows immediately that the quantiles of X are given by

$$Q_p[X] = \mu + \sigma\ \Phi^{-1}(p), \qquad p \in (0,1), \tag{17}$$

where Φ denotes the standard normal cumulative distribution function. The stop-loss premiums of X are given by

$$\mathrm{E}\left[(X-d)_+\right] = \sigma\ \phi\left(\frac{d-\mu}{\sigma}\right) - (d-\mu)\left[1-\Phi\left(\frac{d-\mu}{\sigma}\right)\right], \qquad -\infty < d < +\infty, \tag{18}$$

where $\phi(x) = \Phi'(x)$ denotes the density function of the standard normal distribution. For a proof, see e.g. Example 3.9.1 in Kaas, Goovaerts, Dhaene & Denuit (2001). From (18) we find the following expression for the Expected Shortfall:

$$\mathrm{ESF}_p[X] = \sigma\ \phi\left(\Phi^{-1}(p)\right) - \sigma\ \Phi^{-1}(p)\ (1-p), \qquad p \in (0,1). \tag{19}$$

Using (9), we find that the Conditional Tail Expectation is given by

$$\mathrm{CTE}_p[X] = \mu + \sigma\frac{\phi\left(\Phi^{-1}(p)\right)}{1-p}, \qquad p \in (0,1). \tag{20}$$

∇

Example 2 (Lognormal losses).

Consider a random variable X that is lognormally distributed. Hence, $\ln X \sim N(\mu,\ \sigma^2)$. The quantiles of X follow from Lemma 1:

$$Q_p[X] = e^{\mu+\sigma\ \Phi^{-1}(p)}, \qquad p \in (0,1). \tag{21}$$

It is well-known that the stop-loss premiums of X are given by

$$\mathrm{E}\left[(X-d)_+\right] = e^{\mu+\sigma^2/2}\Phi(d_1) - d\ \Phi(d_2), \qquad d > 0, \tag{22}$$

where $d_1 = \sigma + (\mu - \ln d)/\sigma$ and $d_2 = d_1 - \sigma$. The Black & Scholes (1973) call-option pricing formula is based on this expression for the stop-loss premium of a lognormal random variable. The Expected Shortfall is then given by

$$\mathrm{ESF}_p[X] = e^{\mu+\sigma^2/2}\ \Phi\left(\sigma - \Phi^{-1}(p)\right) - e^{\mu+\sigma\ \Phi^{-1}(p)}\ (1-p), \qquad p \in (0,1). \tag{23}$$

The Conditional Tail Expectation is given by

$$\mathrm{CTE}_p[X] = e^{\mu+\sigma^2/2}\frac{\Phi\left(\sigma - \Phi^{-1}(p)\right)}{1-p}, \qquad p \in (0,1). \tag{24}$$

Finally, we also find

$$\mathrm{E}\left[X \mid X < Q_p[X]\right] = e^{\mu+\sigma^2/2}\frac{1-\Phi\left(\sigma - \Phi^{-1}(p)\right)}{p}, \qquad p \in (0,1). \tag{25}$$

∇

3 Risk measures and ordering of risks

Comparing random variables is the essence of the actuarial profession. Several ordering concepts, such as stochastic dominance and stop-loss order, have been introduced for that purpose in the actuarial literature, see e.g. Goovaerts, Kaas, Van Heerwaarden & Bauwelinckx (1990). Other applications of stochastic orders can be found in Shaked & Shanthikumar (1994).

Definition 1 (Stochastic dominance, stop-loss and convex order). *Consider two loss random variables X and Y. X is said to precede Y in the stochastic dominance sense, notation $X \leq_{st} Y$, if and only if the distribution function of X always exceeds that of Y:*

$$F_X(x) \geq F_Y(x), \qquad -\infty < x < +\infty; \tag{26}$$

X is said to precede Y in the stop-loss order sense, notation $X \leq_{sl} Y$, if and only if X has lower stop-loss premiums than Y:

$$E[(X-d)_+] \leq E[(Y-d)_+], \qquad -\infty < d < +\infty; \tag{27}$$

X is said to precede Y in the convex order sense, notation $X \leq_{cx} Y$, if and only if $X \leq_{sl} Y$ and in addition $E[X] = E[Y]$.

In the definitions of stop-loss order and convex order above, we tacitly assume that the expectations exist. In the following theorem it is stated that stochastic dominance can be characterized in terms of ordered quantiles. The proof is straightforward.

Theorem 2 (Stochastic dominance vs. ordered quantiles). *For any random pair (X,Y) we have that X is smaller than Y in stochastic dominance sense if and only if their respective quantiles are ordered:*

$$X \leq_{st} Y \Leftrightarrow Q_p[X] \leq Q_p[Y] \ \textit{for all } p \in (0,1). \tag{28}$$

In the following theorem, we prove that stop-loss order can be characterized in terms of ordered TVaR's.

Theorem 3 (Stop-loss order vs. ordered TVaR's). *For any random pair (X,Y) we have that $X \leq_{sl} Y$ if and only if their respective TVaR's are ordered:*

$$X \leq_{sl} Y \Leftrightarrow TVaR_p[X] \leq TVaR_p[Y] \ \textit{for all } p \in (0,1). \tag{29}$$

Proof. First we assume $X \leq_{sl} Y$ and let $p \in (0,1)$. Consider the function $f(d)$ defined by

$$f(d) = (1-p)\,d + \mathrm{E}[(X-d)_+] = (1-p)\,d + \int_d^\infty \overline{F}_X(x)\mathrm{d}x,$$

where $\overline{F}_X(x) = 1 - F_X(x)$ is the decumulative distribution function of X. Observe that $\overline{F}_X(Q_p[X]) \leq 1-p \leq \overline{F}_X(Q_p[X]-0)$. So by the monotonicity of the function $\overline{F}_X(x)$, one easily sees that the function $f(d)$, and hence also the function $f(d)/(1-p)$, is minimized for d equal to $Q_p[X]$. Hence, by choosing $d = Q_p[Y]$, we find

$$\begin{aligned}
\mathrm{TVaR}_p[X] &= Q_p[X] + \frac{1}{1-p}\mathrm{E}\left[(X - Q_p[X])_+\right] \\
&= \frac{f(Q_p[X])}{1-p} \\
&\leq \frac{f(Q_p[Y])}{1-p} \\
&= Q_p[Y] + \frac{1}{1-p}\mathrm{E}\left[(X - Q_p[Y])_+\right] \\
&\leq \mathrm{TVaR}_p[Y].
\end{aligned}$$

To prove the other implication, we assume that the TVaR's are ordered for all $p \in (0,1)$. Note that for any random variable X, we have that

$$\begin{aligned}
E[(X-d)_+] &= E[(F_X^{-1}(U) - d)_+] \\
&= \int_{F_X(d)}^1 Q_q[X]\mathrm{d}q - d\,(1 - F_X(d))\,.
\end{aligned}$$

Hence, for d such that $0 < F_X(d) < 1$, we find

$$\begin{aligned}
\mathrm{E}[(X-d)_+] &= \big(\mathrm{TVaR}_{F_X(d)}\,[X] - d\big)\,(1 - F_X(d)) \\
&\leq \big(\mathrm{TVaR}_{F_X(d)}\,[Y] - d\big)\,(1 - F_X(d)) \\
&= \int_{F_X(d)}^1 Q_q[Y]\mathrm{d}q - d\,(1 - F_X(d)) \\
&= \int_{F_Y(d)}^1 Q_q[Y]\mathrm{d}q - d\,(1 - F_Y(d)) \\
&\quad + \int_{F_X(d)}^{F_Y(d)} Q_q[Y]\mathrm{d}q - d\,(F_Y(d) - F_X(d)) \\
&= \mathrm{E}[(Y-d)_+] + \int_{F_X(d)}^{F_Y(d)} (Q_q[Y] - d)\ \mathrm{d}q.
\end{aligned}$$

Using the equivalence $q \leq F_Y(d) \Leftrightarrow d \geq Q_q[Y]$, it is straightforward to prove that

$$\int_{F_X(d)}^{F_Y(d)} (Q_q[Y] - d)\,\mathrm{d}q \leq 0.$$

This proves that the stop-loss premiums of X are smaller than that of Y for any retention d such that $0 < F_X(d) < 1$. If $F_X(d) = 1$, we find $\mathrm{E}[(X-d)_+] = 0 \leq \mathrm{E}[(Y-d)_+]$. Recalling (12), the assumption that $\mathrm{TVaR}_p[X] \leq \mathrm{TVaR}_p[Y]$ for all $p \in (0,1)$ immediately implies that $\mathrm{E}[X] \leq \mathrm{E}[Y]$. Thus $\mathrm{E}[(X-d)_+] \leq \mathrm{E}[(X-d)_+]$ also holds for d such that $F_X(d) = 0$. Hence, we have proven that $X \leq_{sl} Y$. ∎

Remark 1 (CTE does not preserve convex order).

Recall the third item of Theorem 1. The identity $\mathrm{TVaR}_{F_X(d)}[X] = \mathrm{CTE}_{F_X(d)}[X]$ holds for any d such that $0 < F_X(d) < 1$. Hence along the same line as the proof of (b) above, we can obtain the implication that

$$X \leq_{sl} Y \Leftarrow \mathrm{CTE}_p[X] \leq CTE_p[Y] \text{ for all } p \in (0,1).$$

However, the other implication is not true, in general. Actually, we make a somewhat stronger statement below:

$$X \leq_{cx} Y \nRightarrow \mathrm{CTE}_p[X] \leq \mathrm{CTE}_p[Y] \text{ for all } p \in (0,1). \tag{30}$$

A simple illustration for (30) is as follows: Let X and Y be two random variables where F_Y is uniform over $[0,1]$, and F_X is given by

$$F_X(x) = \begin{cases} x & \text{if } 0 \leq x < 0.85, \\ 0.85 & \text{if } 0.85 \leq x < 0.9, \\ 0.95 & \text{if } 0.9 \leq x < 0.95, \\ x & \text{if } 0.95 \leq x \leq 1. \end{cases} \tag{31}$$

Clearly, $F_X(x) \leq F_Y(x)$ for $x < 0.9$, and $F_X(x) \geq F_Y(x)$ for $x \geq 0.9$. We have that $\mathrm{E}[X] = \mathrm{E}[Y] = 0.5$ and $X \leq_{sl} Y$, hence that $X \leq_{cx} Y$. However, we easily check that $\mathrm{CTE}_{0.9}[X] > \mathrm{CTE}_{0.9}[Y]$ since $\mathrm{CTE}_{0.9}[X] = 0.975$ and $\mathrm{CTE}_{0.9}[Y] = 0.95$. ∇

4 Comonotonicity

4.1 Comonotonic bounds for sums of dependent random variables

A set S in R^n is said to be comonotonic, if, for all $(y_1, y_2, \ldots, y_n)$ and $(z_1, z_2, \ldots, z_n)$ in this set, $y_i < z_i$ for some i implies $y_j \leq z_j$ for all j. Notice that a comonotonic set is a 'thin' set, in the sense that it is contained in a one-dimensional subset of R^n. When the support of a random vector is a comonotonic set, the random vector itself and its joint distribution are called comonotonic.

It can be proven that an n-dimensional random vector $\underline{Y} = (Y_1, Y_2, \ldots, Y_n)$ is comonotonic if and only if

$$\underline{Y} \stackrel{d}{=} (F_{Y_1}^{-1}(U), F_{Y_2}^{-1}(U), \ldots, F_{Y_n}^{-1}(U)), \tag{32}$$

where $\stackrel{d}{=}$ stands for 'equality in distribution', and U is a random variable that is uniformly distributed over the unit interval (0,1). In the remainder of this paper, the notation U will only be used to denote such a uniformly distributed random variable.

For any random vector $\underline{X} = (X_1, X_2, \ldots, X_n)$, not necessarily comonotonic, we will call its comonotonic counterpart any random vector with the same marginal distributions and with the comonotonic dependency structure. The comonotonic counterpart of $\underline{X} = (X_1, X_2, \ldots, X_n)$ will be denoted by $\underline{X}^c = (X_1^c, X_2^c, \ldots, X_n^c)$. Note that

$$(X_1^c, X_2^c, \ldots, X_n^c) \stackrel{d}{=} (F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), \ldots, F_{X_n}^{-1}(U)).$$

It can be proven that a random vector is comonotonic if and only if all its marginal distribution functions are non-decreasing functions (or all are non-increasing functions) of the same random variable. For other characterizations and more details about the concept of comonotonicity and its applications in actuarial science and finance, we refer to the overview papers by Dhaene, Denuit, Goovaerts, Kaas & Vyncke (2002a,b).

A proof for the following theorem concerning convex order bounds for sums of dependent random variables is presented in Kaas, Dhaene & Goovaerts (2000).

Theorem 4 (Convex bounds for sums of random variables). *For any random vector $(X_1, X_2, \ldots, X_n)$ and any random variable Λ, we have that*

$$\sum_{i=1}^{n} E[X_i \mid \Lambda] \leq_{cx} \sum_{i=1}^{n} X_i \leq_{cx} \sum_{i=1}^{n} F_{X_i}^{-1}(U). \tag{33}$$

The theorem above states that the least attractive random vector $(X_1, \ldots, X_n)$ with given marginal distribution functions F_{X_i}, in the sense that the sum of its components is largest in the convex order, has the *comonotonic* joint distribution, which means that it has the joint distribution of $\left(F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), \ldots, F_{X_n}^{-1}(U)\right)$. The components of this random vector are maximally dependent, all components being non-decreasing functions of the same random variable. Several proofs have been given for this result, see e.g. Denneberg (1994), Dhaene & Goovaerts (1996), Müller (1997) or Dhaene, Wang, Young & Goovaerts (2000).

The random vector $(\mathrm{E}[X_1 \mid \Lambda], \mathrm{E}[X_2 \mid \Lambda], \ldots, \mathrm{E}[X_n \mid \Lambda])$ will in general not have the same marginal distributions as $(X_1, X_2, \ldots, X_n)$. If one can find a conditioning random variable Λ with the property that all random variables $\mathrm{E}[X_i \mid \Lambda]$ are non-increasing functions of Λ (or all are non-decreasing functions of Λ), the lower bound $S^l = \sum_{i=1}^{n} \mathrm{E}[X_i \mid \Lambda]$ is a sum of n comonotonic random variables.

4.2 Risk measures and comonotonicity

In the following theorem, we prove that the quantile risk measure, the Tail Value-at-Risk and the expected shortfall are additive for a sum of comonotonic random variables.

Theorem 5 (Additivity of risk measures for sums of comonotonic risks). *Consider a comonotonic random vector* $(X_1^c, X_2^c, \ldots, X_n^c)$*, and let* $S^c = X_1^c + X_2^c + \cdots + X_n^c$*. Then we have for all* $p \in (0,1)$ *that*

$$Q_p\left[S^c\right] = \sum_{i=1}^{n} Q_p\left[X_i\right], \tag{34}$$

$$TVaR_p\left[S^c\right] = \sum_{i=1}^{n} TVaR_p\left[X_i\right], \tag{35}$$

$$ESF_p\left[S^c\right] = \sum_{i=1}^{n} ESF_p\left[X_i\right]. \tag{36}$$

Proof. (a) We have that

$$S^c \overset{d}{=} F_{X_1}^{-1}(U) + F_{X_2}^{-1}(U) + \cdots + F_{X_n}^{-1}(U) = g\left(U\right),$$

with g a non-decreasing and left-continuous function. Hence, (34) follows from Lemma 1.
(b) Relation (35) follows immediately from (5) and (34).
(c) Relation (36) follows from (8), (34) and (35). ∎

From the theorem above, we can conclude that the quantile risk measure, TVaR and ESF risk measure for a comonotonic sum can easily be obtained by summing the corresponding risk measures of the marginal distributions involved. Specifically, if all the random variables X_i above have the same distribution as that of X, then we find $Q_p\left[nX\right] = n\ Q_p\left[X\right]$, $\mathrm{TVaR}_p\left[nX\right] = n\ \mathrm{TVaR}_p\left[X\right]$ and $\mathrm{E}\big[(nX - Q_p\left[nX\right])_+\big] = n\ \mathrm{E}\big[(X - Q_p\left[X\right])_+\big]$. As we will see, the CTE risk measure is in general not additive for sums of comonotonic risks. Nevertheless, we immediately find that $\mathrm{CTE}_p\left[nX\right] = n\ \mathrm{CTE}_p\left[X\right]$. Another case where the additivity property does hold for CTE is given in the following remark.

Remark 2 (Additivity of CTE for sums of comonotonic continuous risks).

Consider a comonotonic random vector $(X_1^c, X_2^c, \ldots, X_n^c)$ with continuous marginal distributions. For any random variable X, we have that $F_X(x)$ is continuous in $x \in (-\infty, \infty)$ if and only if $Q_p\left[X\right]$ is strictly increasing in $p \in (0,1)$. This implies that the sum S^c is continuously distributed. Furthermore, the continuity of the distribution function of S^c

implies that $\mathrm{CTE}_p[S^c] = \mathrm{TVaR}_p[S^c]$ for each $p \in (0,1)$. Therefore it follows from (35) that

$$\mathrm{CTE}_p[S^c] = \mathrm{TVaR}_p[S^c] = \sum_{i=1}^{n} \mathrm{TVaR}_p[X_i] = \sum_{i=1}^{n} \mathrm{CTE}_p[X_i].$$

∇

For the case where the marginal distributions are not continuous and not the same, however, the CTE is, in general, not additive for comonotonic risks. Here we propose an illustration for this case.

Remark 3 (CTE is not additive for sums of comonotonic risks).

Consider the comonotonic random vector (X^c, Y^c), where X has a distribution F_X given by

$$F_X(x) = \begin{cases} x & 0 \leq x < 0.85 \\ 0.85 & 0.85 \leq x < 0.9 \\ 0.95 & 0.9 \leq x < 0.95 \\ x & 0.95 \leq x \leq 1 \end{cases},$$

and Y is uniformly distributed in $(0,1)$. We write

$$S^c = X^c + Y^c \stackrel{d}{=} F_X^{-1}(U) + U = g(U).$$

Since $F_X^{-1}(y)$ and $F_Y^{-1}(y)$ are non-decreasing in $y \in (0,1)$ and the monotonicity of $F_Y^{-1}(y)$ is strict, the function $g(y)$ is strictly increasing. Hence the sum $S^c = X^c + Y^c$ has a continuous distribution. Because of the additivity of the risk measures $Q_{0.9}(\cdot)$ and $\mathrm{ESF}_{0.9}(\cdot)$, we find

$$\begin{aligned}
\mathrm{CTE}_{0.9}[S^c] &= Q_{0.9}[S^c] + \frac{1}{1-0.9}\mathrm{ESF}_{0.9}[S^c] \\
&= Q_{0.9}[X^c] + Q_{0.9}[Y^c] + \frac{1}{1-0.9}\left(\mathrm{ESF}_{0.9}[X^c] + \mathrm{ESF}_{0.9}[Y^c]\right) \\
&= \left(Q_{0.9}[X^c] + \frac{1}{1-0.95}\mathrm{ESF}_{0.9}[X^c]\right) + \left(Q_{0.9}[Y^c] + \frac{1}{1-0.9}\mathrm{ESF}_{0.9}[Y^c]\right) \\
&\qquad - \left(\frac{1}{1-0.95} - \frac{1}{1-0.9}\right)\mathrm{ESF}_{0.9}[X^c] \\
&= \mathrm{CTE}_{0.9}[X^c] + \mathrm{CTE}_{0.9}[Y^c] - \left(\frac{1}{1-0.95} - \frac{1}{1-0.9}\right)\mathrm{ESF}_{0.9}[X^c] \\
&< \mathrm{CTE}_{0.9}[X^c] + \mathrm{CTE}_{0.9}[Y^c].
\end{aligned}$$

∇

A risk measure ρ is said to be sub-additive if for any random variables X and Y, one has $\rho(X+Y) \leq \rho(X)+\rho(Y)$. Sub-additivity of a risk measure ρ immediately implies

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

A risk measure is said to preserve stop-loss order if for any X and Y, one has that $X \leq_{sl} Y$ implies $\rho\left[X\right] \leq \rho\left[Y\right]$.

Theorem 6 (Sub-additivity of risk measures). *Any risk measure that preserves stop-loss order and that is additive for comonotonic risks is sub-additive.*

Proof. From Theorem 4, we have that a sum of random variables with given marginal distributions is largest in the convex order sense if these random variables are comonotonic:

$$X+Y \leq_{sl} X^c + Y^c.$$

If the risk measure ρ preserves stop-loss order and is additive for comonotonic risks, then

$$\rho\left[X+Y\right] \leq \rho\left[X^c+Y^c\right] = \rho\left[X\right]+\rho\left[Y\right],$$

which proves the stated result. ∎

Recall Theorems 3 and 5. As a special case of Theorem 6, we find that TVaR is sub-additive:

$$\mathrm{TVaR}_p\left[X+Y\right] \leq \mathrm{TVaR}_p\left[X\right]+\mathrm{TVaR}_p\left[Y\right], \qquad p \in (0,1). \tag{37}$$

In the following remark we show that CTE is not sub-additive.

Remark 4 (CTE is not sub-additive).

Let X be a random variable uniformly distributed in $(0,1)$, and let Y be another random variable defined by

$$Y = (0.95-X)I_{(0<X\leq 0.95)} + (1.95-X)I_{(0.95<X<1)},$$

where I_A denotes the indicator function which equals 1 if condition A holds and 0 otherwise. It is easy to see that Y is also uniformly distributed on $(0,1)$ and

$$X+Y = 0.95\ I_{(0<X\leq 0.95)} + 1.95\ I_{(0.95<X<1)}. \tag{38}$$

Equation (38) indicates that $X+Y$ follows a discrete law with only two jumps:

$$\Pr(X+Y=0.95) = 1-\Pr(X+Y=1.95) = 0.95.$$

For $p = 0.90$, by formula (9) one easily checks that

$$\mathrm{CTE}_p[X+Y] = 1.95, \quad \mathrm{CTE}_p[X] = \mathrm{CTE}_p[Y] = 0.95.$$

Hence

$$\mathrm{CTE}_p[X+Y] > \mathrm{CTE}_p[X] + \mathrm{CTE}_p[Y].$$

Notice that also the risk measure $\mathrm{E}[X+Y \mid X+Y \geq Q_p[X+Y]]$, $p \in (0,1)$, is not subadditive. Indeed, consider two random variables X and Y which are i.i.d. Bernoulli (0.01) distributed. We immediately find that

$$E[X+Y \mid X+Y \geq Q_{0.985}[X+Y]] > E[X \mid X \geq Q_{0.985}[X]] + E[Y \mid Y \geq Q_{0.985}[Y]] \tag{39}$$

∇

In the following remarks we show that both the quantile risk measure and ESF are not sub-additive.

Remark 5 (VaR is not sub-additive).

Let X and Y be i.i.d. random variables which are Bernoulli (0.02) distributed. We immediately find that $Q_{0.975}[X] = Q_{0.975}[Y] = 0$. On the other hand, $\Pr(X+Y=0) = 0.9604$, which implies that $Q_{0.975}[X+Y] > 0$. As another illustration of the fact that the quantile risk measure is not sub-additive, consider a bivariate normal random vector (X, Y). One can easily prove that the distribution functions of $X+Y$ and X^c+Y^c only cross once, in $(\mu_X+\mu_Y,\ 0.5)$. This implies that $Q_p[X+Y] > Q_p[X]+Q_p[Y]$ if $p < 0.5$, whereas $Q_p[X+Y] < Q_p[X]+Q_p[Y]$ if $p > 0.5$. ∇

Remark 6 (ESF is not sub-additve).

Let X and Y be i.i.d. random variables which are Bernoulli (0.02) distributed. It is straightforward to prove that $\mathrm{ESF}_{0.99}[X] = 0$, while $\mathrm{ESF}_{0.99}[X+Y] > 0$. ∇

Remark 7 (Translation-scale invariant distributions).

The distribution functions of the risks $X_1, X_2, \ldots, X_n$ are said to belong to the same translation-scale invariant family of distributions if there exist a random variable Y, positive real constants a_i and real constants b_i such that X_i has the same distribution as $a_iY + b_i$ for each $i = 1, 2, \ldots, n$. Examples of translation-scale invariant families of distributions are normal distributions, or more generally, elliptical distributions with the same characteristic generator, see e.g. Valdez & Dhaene (2003). Now assume that the risk measure ρ preserves stop-loss order and that $\rho[aX+b] = a\,\rho[X] + b$ for any positive real number a and any real number b. It is easy to prove that if the set of risks is restricted to a translation-scale invariant family, then the risk measure ρ is sub-additive in this family. ∇

5 Distortion risk measures

5.1 Definition, examples and properties

In this section we will consider the class of distortion risk measures, introduced by Wang (1996). The quantile risk measure and TVaR belong to this class. A number of the properties of these risk measures can be generalized to the class of distortion risk measures.

The expectation of X, if it exists, can be written as

$$\mathrm{E}[X] = -\int_{-\infty}^{0}\left[1-\overline{F}_X(x)\right]\mathrm{d}x + \int_{0}^{\infty}\overline{F}_X(x)\mathrm{d}x. \tag{40}$$

Wang (1996) defines a family of risk measures by using the concept of distortion function as introduced in Yaari's dual theory of choice under risk, see also Wang & Young (1998). A distortion function is defined as a non-decreasing function $g:[0,1]\to[0,1]$ such that $g(0)=0$ and $g(1)=1$. The distortion risk measure associated with distortion function g is denoted by $\rho_g[\cdot]$ and is defined by

$$\rho_g[X] = -\int_{-\infty}^{0}\left[1-g\left(\overline{F}_X(x)\right)\right]\mathrm{d}x + \int_{0}^{\infty} g\left(\overline{F}_X(x)\right)\mathrm{d}x, \tag{41}$$

for any random variable X. Note that the distortion function g is assumed to be independent of the distribution function of the random variable X. The distortion function $g(q)=q$ corresponds to $\mathrm{E}[X]$. Note that if $g(q)\geq q$ for all $q\in[0,1]$, then $\rho_g[X] \geq \mathrm{E}[X]$. In particular this result holds in case g is a concave distortion function. Also note that $g_1(q)\leq g_2(q)$ for all $q\in[0,1]$ implies that $\rho_{g_1}[X]\leq\rho_{g_2}[X]$.

One immediately finds that $g\left(\overline{F}_X(x)\right)$ is a non-increasing function of x with values in the interval $[0,1]$. However $\rho_g[X]$ cannot always be considered as the expectation of X under a new probability measure, because $g\left(\overline{F}_X(x)\right)$ will not necessarily be right-continuous. For a general distortion function g, the risk measure $\rho_g[X]$ can be interpreted as a "distorted expectation" of X, evaluated with a "distorted probability measure" in the sense of a Choquet-integral, see Denneberg (1994). Substituting $g\left(\overline{F}_X(x)\right)$ by $\int_0^{\overline{F}_X(x)}\mathrm{d}g(q)$ in (41) and reverting the order of the integrations, one finds that any distortion risk measure

$\rho_g[X]$ can be written as

$$\rho_g[X] = \int_0^1 Q_{1-q}[X] \mathrm{d}g(q). \tag{42}$$

Notice that when the distortion function g is differentiable, (42) can be rewritten as

$$\rho_g[X] = E\left[Q_{1-U}[X]\ g'(U)\right]. \tag{43}$$

From (42), one can easily verify that the quantile $Q_p[X]$, $p \in (0,1)$, corresponds to the distortion function

$$g(x) = I_{(x>1-p)}, \qquad 0 \leq x \leq 1. \tag{44}$$

On the other hand, $\mathrm{TVaR}_p[X]$, $p \in (0,1)$, corresponds to the distortion function

$$g(x) = \min\left(\frac{x}{1-p}, 1\right), \qquad 0 \leq x \leq 1. \tag{45}$$

Remark 8 (ESF is not a distortion risk measure).

Let us assume that the risk measure $\mathrm{ESF}_p[X]$ can be expressed as (41) for some distortion function g. We first substitute to (41) a risk variable X with a uniform distribution on $(0,1)$. Hence, for the given $p \in (0,1)$, we have

$$\frac{1}{2}(1-p)^2 = \int_0^1 g(s)ds. \tag{46}$$

We then substitute to (41) a Bernoulli variable Y with

$$\Pr(Y=0) = 1 - \Pr(Y=1) = 1-r,$$

for some arbitrarily but fixed $0 < r \leq 1-p$. We easily obtain that $g(r) = r$ for $0 < r \leq 1-p$. From (46) we find that

$$\frac{1}{2}(1-p)^2 = \int_0^{1-p} s ds + \int_{1-p}^1 g(s)ds \geq \frac{1}{2}(1-p)^2 + p(1-p),$$

which is obviously a contradiction since $0 < p < 1$. This illustrates that ESF is not a distortion risk measure. ∇

From (10) and the fact that $\mathrm{TVaR}_p[X]$, $p \in (0,1)$, corresponds to the distortion function given in (45), we find that $\mathrm{CTE}_p[X]$, $p \in (0,1)$, can be written in the form $\rho_g[X]$ with g given by

$$g(x) = \min\left(\frac{x}{1-F_X(Q_p[X])}, 1\right), \qquad 0 \leq x \leq 1. \tag{47}$$

This function g, however, depends on the distribution function of X; hence we cannot infer that $\mathrm{CTE}_p[\cdot]$ is a distortion risk measure. Actually, we can obtain the following result.

Remark 9 (CTE is not a distortion risk measure).

Along the same approach as in Remark 8, we assume by contradiction that the risk measure $\mathrm{CTE}_p[X]$ can be expressed as (41) for some distortion function g. We first substitute to (41) a risk variable X with a uniform distribution on $(0,1)$. Hence, for the given $p \in (0,1)$, recalling (9), we find

$$p + \frac{1}{2}(1-p) = \int_0^1 g(1-x)\,dx.$$

Simplification on the above equation leads to

$$\int_0^1 g(x)\,dx = \frac{1}{2}(1+p). \tag{48}$$

We then substitute to (41) some other risk variable. To this end we choose a Bernoulli variable Y with

$$\Pr(Y=0) = 1 - \Pr(Y=1) = 1 - r,$$

for some arbitrarily but fixed $0 < r \leq 1-p$. Again applying (9) we obtain $\mathrm{CTE}_p[Y] = 1$. Hence by (41) it should hold that $g(r) = 1$. By virtue of the monotonicity of the distortion function g and the arbitrariness of $0 < r \leq 1-p$ we conclude that $g(\cdot) \equiv 1$ on $(0,1]$, which contradicts the equation (48). This illustrates that $\mathrm{CTE}_p[X]$ is not a distortion risk measure. ∇

Example 3 (The Wang transform risk measure).

From (5), we see that the TVaR_p risk measure uses only the upper tail of the distribution. Hence, this risk measure does not create incentive for taking actions that increase the distribution function for outcomes smaller than Q_p. Also, from (8) we see that TVaR_p only accounts for the expected shortfall and hence, does not properly adjust for extreme low-frequency and high severity losses. The Wang Transform risk measure was introduced by Wang (2000) as an example of a risk measure that could give a solution to these problems. For any $0 < p < 1$, define the distortion function

$$g_p(x) = \Phi\left[\Phi^{-1}(x) + \Phi^{-1}(p)\right], \qquad 0 \leq x \leq 1,\ 0 < p < 1, \tag{49}$$

which is called the 'Wang Transform at level p'. The corresponding distortion risk measure is called the Wang Transform risk measure and is denoted by $\mathrm{WT}_p[X]$.

For a normally distributed random variable X, we find

$$1 - g_p\left(\overline{F}_X(x)\right) = \Phi\left[\frac{x - Q_p[X]}{\sigma}\right]$$

which implies that the Wang Transform risk measure is identical to the quantile risk measure at the same probability level in case of a normal random variable:

$$\mathrm{WT}_p[X] = Q_p[X]. \tag{50}$$

For a lognormal distributed random variable Y with parameters μ and σ^2, we find

$$\mathrm{WT}_p(Y) = Q_{\Phi[\Phi^{-1}(p)+\sigma/2]}(Y), \tag{51}$$

which is larger than $Q_p[Y]$.

Examples illustrating the fact that the WT risk measure uses the whole distribution and that it accounts for extreme low-frequency and high severity losses can be found in Wang (2001). ∇

It is easy to prove that any distortion risk measure ρ_g obeys the following properties, see also Wang (1996):

- **Additivity for comonotonic risks**:
  For any distortion function g and all random variables X_i,

$$\rho_g\left[X_1^c + X_2^c + \cdots + X_n^c\right] = \sum_{i=1}^{n} \rho_g\left[X_i\right]. \tag{52}$$

- **Positive homogeneity**:
  For any distortion function g, any random variable X and any non-negative constant a, we have

$$\rho_g[aX] = a\rho_g[X]. \tag{53}$$

- **Translation invariance**:
  For any distortion function g, any random variable X and any constant b, we have

$$\rho_g[X + b] = \rho_g[X] + b. \tag{54}$$

- **Monotonicity**:
  For any distortion function g and any two random variables's X and Y where $X \leq Y$ with probability 1, we have
$$\rho_g [\, X] \leq \rho_g [Y] \tag{55}$$

The first property follows immediately from (42) and the additivity property of quantiles for comonotonic risks. The second and the third properties follow from (42) and Lemma 1. The fourth property follows from (42) and the fact that $X \leq Y$ with probability 1 implies that each quantile of Y exceeds the corresponding quantile of X. Note that in the literature the property of positive homogeneity is often wrongly explained as 'currency independence'. Take as an example the risk measure

$$\rho [X] = E \left[(X - d)_+ \right] , \tag{56}$$

where clearly X and d have to be expressed in the same monetary unit. This risk measure is not positive homogeneous but it is currency independent, see also Remark 3.5 in Goovaerts, Kaas, Dhaene & Tang (2003).

In the following theorem, stochastic dominance is characterized in terms of ordered distortion risk measures.

Theorem 7 (Stochastic dominance vs. ordered distortion risk measures). *For any random pair (X, Y) we have that X is smaller than Y in stochastic dominance sense if and only if their respective distortion risk measures are ordered:*

$$X \leq_{st} Y \Leftrightarrow \rho_g [X] \leq \rho_g [Y] \text{ for all distortion functions } g. \tag{57}$$

Proof. This follows immediately from (42) and Theorem 2. ∎

5.2 Concave distortion risk measures

A subclass of distortion functions that is often considered in the literature is the class of concave distortion functions. A distortion function g is said to be concave if for each q in $(0, 1]$, there exist real numbers a_y and b_y and a line $l(x) = a_y x + b_y$, such that $l(q) = g(q)$ and $l(q) \geq g(q)$ for all q in $(0, 1]$. A concave distortion function is necessarily continuous in $(0, 1]$. For convenience, we will always tacitly assume that a concave distortion function is also continuous at 0. A risk measure with a concave distortion function is then called a 'concave distortion risk measure'.

For any concave distortion function g, we have that $g\left(\overline{F}_X(x)\right)$ is right-continuous, so that in this case the risk measure $\rho_g [X]$ can be interpreted as the expectation of X under a 'distorted probability measure'. Note that the quantile risk measure is not a concave distortion risk measure whereas TVaR is a concave distortion risk measure.

In the following theorem, we show that stop-loss order can be characterized in terms of ordered concave distortion risk measures.

Theorem 8 (SL-order vs. ordered concave distortion risk measures). *For any random pair (X, Y) we have that $X \leq_{sl} Y$ if and only if their respective concave distortion risk measures are ordered:*

$$X \leq_{sl} Y \Leftrightarrow \rho_g[X] \leq \rho_g[Y] \text{ for all concave distortion functions } g. \tag{58}$$

Proof. The " $\Leftarrow$ " implication follows immediately from Theorem 3. To complete the proof of Theorem 8, we first prove the " $\Rightarrow$ " implication for concave piecewise linear distortion functions g. Any such distortion function can be written in the form

$$g(x) = \sum_{i=1}^{n} a_i \left(\beta_i - \beta_{i+1}\right) \min\left(x/a_i, 1\right)$$

where $0 = a_0 < a_1 < \cdots < a_{n-1} < a_n = 1$. Further, β_i is the derivative of g in the interval (a_{i-1}, a_i) and $\beta_{n+1} = 0$. Because of the concavity of g, we have that β_i is a decreasing function of i. From (42) and (45), it follows that the related risk measure $\rho_g[X]$ can be written as

$$\rho_g[X] = \sum_{i=1}^{n} a_i \left(\beta_i - \beta_{i+1}\right) \text{TVaR}_{1-a_i}[X].$$

In view of Theorem 3, we find that $X \leq_{sl} Y$ implies $\rho_g[X] \leq \rho_g[Y]$ for all concave piecewise linear distortion functions g.
Now we are able to prove the " $\Rightarrow$ " for general concave distortion functions g. If $\rho_g[Y] = \infty$, the result is obvious. Let us now assume that $\rho_g[Y] < \infty$. The concave distortion function g can be approximated from below by concave piecewise linear distortion functions g_n such that for any $x \in [0, 1]$, we have that $g_1(x) \leq g_2(x) \leq \cdots \leq g_n(x) \leq \cdots \leq g(x)$ and $\lim_{n\to\infty} g_n(x) = g(x)$. As we have just proven, the inequality $X \leq_{sl} Y$ implies $\rho_{g_n}[X] \leq \rho_{g_n}[Y]$ for all n. Further, $g_n(x) \leq g(x)$ implies $\rho_{g_n}[Y] \leq \rho_g[Y] < \infty$. From the monotone convergence theorem we find that $\lim_{n\to\infty} \rho_{g_n}[X] = \rho_g[X]$, so that we can conclude that $\rho_g[X] \leq \rho_g[Y]$. ∎

Proofs for the theorem above can also be found in Yaari (1987), Wang & Young (1998) or Dhaene, Wang, Young & Goovaerts (2000).

Example 4 (The Beta distortion risk measure).

We will write $X <_{sl} Y$ if $X \leq_{sl} Y$ and $\mathrm{E}\left[(X-d)_+\right] < \mathrm{E}\left[(Y-d)_+\right]$ for at least one retention d. Wirch & Hardy (2000) give the following example that illustrates that $\mathrm{TVaR}_{0.95}$ does not strongly preserve stop-loss order.
Let $\Pr[X = (0,1,2)] = (0.95, 0.025, 0.025)$ and $\Pr[Y = (1,2)] = (0.975, 0.025)$. It is easy to verify that $X <_{sl} Y$ and $\mathrm{TVaR}_{0.95}[X] = \mathrm{TVaR}_{0.95}[Y] = 1.5$. This means that there exist random variables X and Y such that $X <_{sl} Y$ but $\mathrm{TVaR}_{0.95}[X] = \mathrm{TVaR}_{0.95}[Y]$.
More generally, they prove that any distortion risk measure derived from a distortion function g which is concave but not strictly concave (i.e. g has a linear part) does not strongly preserve stop-loss order. They also prove that for any distortion function g which is strictly concave one has that $X <_{sl} Y$ implies $\rho_g[X] < \rho_g[Y]$. Wirch & Hardy (2000) start from the Beta distribution function

$$F_\beta(x) = \frac{1}{\beta(a,b)} \int_0^x t^{a-1}(1-t)^{b-1}\,\mathrm{d}t, \qquad 0 \leq x \leq 1, \tag{59}$$

where $\beta(a,b)$ is the Beta function with parameters $a > 0, b > 0$, i.e.

$$\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \tag{60}$$

to define the Beta distortion function

$$g(x) = F_\beta(x), \qquad 0 \leq x \leq 1. \tag{61}$$

The Beta distortion function is strictly concave for any parameters $0 < a \leq 1$ and $b \geq 1$, provided a and b are not both equal to 1. This implies that the risk measure derived from the distortion function $F_\beta(q)$ strictly preserves stop-loss order. For random variables X and Y as defined above, and parameters $a = 0.1$ and $b = 1$, we find

$$\rho_{F_\beta}[X] = 1.4326 < \rho_{F_\beta}[Y] = 1.6915.$$

Note that the Beta distortion risk measure with the parameters $a = 0.1$ and $b = 1$ reduces to the PH-transform risk measure, which is considered in Wang (1995). ∇

Concave distortion risk measures are subadditive, which means that the risk measure for a sum of random variables is smaller than or equivalent to the sum of the risk measures.

- **Subadditivity**:
  For any concave distortion function g, and any two random variables X and Y, we have

  $$\rho_g[X+Y] \leq \rho_g[X] + \rho_g[Y]. \tag{62}$$

  The proof follows immediately from Theorem 6, see also Wang & Dhaene (1998).

In Artzner (1999) and Artzner, Delbaen, Eber & Heath (1999) a risk measure satisfying the four axioms of subadditivity, monotonicity, positive homogeneity and translation invariance is called "coherent". As we have proven, any concave distortion risk measure is coherent. As the quantile risk measure is not subadditive, it is not a "coherent" risk measure.

Note that the class of concave distortion risk measures is only a subset of the class of "coherent" risk measures, as is shown by the following example.

Example 5 (The Dutch risk measure).

For any random variable X, consider the risk measure

$$\rho\left[X\right] = \mathrm{E}\left[X\right] + \theta\ \mathrm{E}\left[\left(X - \alpha\ \mathrm{E}\left[X\right]\right)_+\right], \qquad \alpha \geq 1,\ 0 \leq \theta \leq 1. \tag{63}$$

We will call this risk measure the "Dutch risk measure", because for non-negative random variables, it is called the "Dutch premium principle", see Kaas, van Heerwaarden & Goovaerts (1994).
In the sequel of this example we assume that the parameters α and θ are both equal to 1. In this case the Dutch risk measure is coherent. Indeed, the verifications of the properties of positive homogeneity, translation invariance and subadditivity are immediate. Finally, if $X \leq Y$ with probability 1, then $\mathrm{E}[X] \leq \mathrm{E}[Y]$, so that the property of monotonicity follows from

$$\rho\left[X\right] = \mathrm{E}\left[\max\left(\mathrm{E}\left[X\right],\ X\right)\right] \leq \mathrm{E}\left[\max\left(\mathrm{E}\left[Y\right],\ Y\right)\right] = \rho\left[Y\right].$$

Next, we will prove that the Dutch risk measure $\rho(\cdot)$ is in general not additive for comonotonic risks. Let (X_1^c, X_2^c) be a comonotonic random couple with Bernoulli marginal distributions: $\Pr\left[X_i = 1\right] = q_i$ with $0 < q_1 < q_2 < 1$ and $q_1 + q_2 > 1$. After some straightforward computations, we find

$$\rho\left[X_i\right] = q_i\left(2 - q_i\right), \qquad i = 1, 2,$$

and

$$\rho\left[X_1^c + \ X_2^c\right] = 2\ q_1 + \left(1 - q_1\right)\left(q_1 + q_2\right),$$

from which we can conclude that the Dutch premium principle is in general not additive for comonotonic risks. Hence, the Dutch risk measure (with parameters equal to 1) is an example of a risk measure that is coherent, although it is not a distortion risk measure. The example also illustrates the fact that coherent risk measures are not necessarily additive for comonotonic risks. ∇

As we have seen, the quantile risk measure Q_p is not a concave distortion risk measure. The following theorem states that in the class of concave distortion risk measures, the one that leads to the minimal extra-capital compared to the quantile risk measure at probability level p is the TVaR risk measure at the same level p.

Theorem 9 (Characterization of TVaR). *For any $0 < p < 1$ and for any random variable X one has*

$$TVaR_p[X] = \min\left\{\rho_g[X] \mid g \text{ is concave and } \rho_g \geq Q_p\right\}. \tag{64}$$

Proof. The distortion risk measure TVaR_p has a concave distortion function $\min\left(\frac{x}{1-p}, 1\right)$. Further, $\mathrm{TVaR}_p \geq Q_p$. This implies that

$$\mathrm{TVaR}_p(X) \geq \inf\left\{\rho_g(X) \mid g \text{ is concave and } \rho_g \geq Q_p\right\}.$$

In order to prove the opposite inequality, consider a concave distortion function g such that $\rho_g(Y) \geq Q_p(Y)$ holds for all random variables Y. For any q with $1-p < q < 1$, we define the Bernoulli random variable Y_q with

$$\Pr(Y_q = 1) = q.$$

It is easy to verify that $Q_p(Y_q) = 1$, and also $\rho_g(Y_q) = g(q)$. As $g(x) \leq 1$, we find that the condition $\rho_g(Y_q) \geq Q_p(Y_q)$ can be rewritten as $g(q) = 1$. This means that g is equal to 1 on the interval $(1-p, 1]$. As g is concave, it follows immediately that

$$g(x) \geq \min\left(\frac{x}{1-p}, 1\right), \qquad 0 < x < 1.$$

Hence,

$$\rho_g(X) \geq \mathrm{TVaR}_p(X)$$

holds for all concave distortion risk measures g for which $\rho_g \geq Q_p$. This implies

$$\mathrm{TVaR}_p(X) = \inf\left\{\rho_g(X) \mid g \text{ is concave and } \rho_g \geq Q_p\right\}.$$

∎

A result with a taste similar to our Theorem 9 is Proposition 5.2 in Artzner, Delbaen, Eber & Heath (1999), which says that

$$\mathrm{VaR}_p[X] = \inf\{\rho[X] \mid \rho \text{ coherent and } \rho \geq Q_p\}$$

holds for each risk variable X, see also Proposition 3.3 in Artzner (1999).

5.3 Risk measures for sums of dependent random variables

In this subsection, we will consider the problem of finding approximations for distorted expectations (such as quantiles and TVaR's) of a sum $S = \sum_{i=1}^{n} X_i$ of which the marginal distributions of the random variables X_i are given, but the dependency structure between the X_i is unknown or too cumbersome to work with. In view of Theorem 4, we propose to approximate (the d.f.) of S by (the d.f. of) $S^c = \sum_{i=1}^{n} F_{X_i}^{-1}(U)$ or (the d.f.) of $S^l = \sum_{i=1}^{n} \mathrm{E}[X_i \mid \Lambda]$, and approximate $\rho_g[S]$ by $\rho_g[S^c]$ or by $\rho_g\left[S^l\right]$. Note that S^c is a comonotonic sum, hence from the additivity property for comonotonic risks we find

$$\rho_g[S^c] = \sum_{i=1}^{n} \rho_g[X_i]. \tag{65}$$

On the other hand, if the conditioning random variable Λ is such that all $\mathrm{E}[X_i \mid \Lambda]$ are non-decreasing functions of Λ (or all are non-increasing functions of Λ), then S^l is a comonotonic sum too. Hence, in this case

$$\rho_g\left[S^l\right] = \sum_{i=1}^{n} \rho_g\left[\mathrm{E}\left[X_i \mid \Lambda\right]\right]. \tag{66}$$

In case of a concave distortion function g, we find from Theorem 4 that $\rho_g\left[S^l\right]$ is a lower bound whereas $\rho_g[S^c]$ is an upper bound for $\rho_g[S]$:

$$\rho_g\left[S^l\right] \leq \rho_g[S] \leq \rho_g[S^c]. \tag{67}$$

In particular, we have that

$$\mathrm{TVaR}_p\left[S^l\right] \leq \mathrm{TVaR}_p[S] \leq \mathrm{TVaR}_p[S^c]. \tag{68}$$

Note that the quantiles of S^l, S and S^c are not necessarily ordered in the same way.

Example 6 (Sums of lognormals).

Consider the sum

$$S = \sum_{i=0}^{n} \alpha_i \, e^{Z_i} \tag{69}$$

where the α_i are non-negative constants and the Z_i are linear combinations of the components of the random vector $(Y_1, Y_2, \ldots, Y_n)$ which is assumed to have a multivariate normal distribution:

$$Z_i = \sum_{j=1}^{n} \lambda_{ij} \, Y_j. \tag{70}$$

Let U be uniformly distributed on the unit interval. Then from Lemma 1, we find that the comonotonic upper bound $S^c = \sum_{i=0}^n F^{-1}_{\alpha_i \, e^{Z_i}}(U)$ of S is given by

$$S^c = \sum_{i=0}^n \alpha_i \; e^{\mathrm{E}[Z_i] + \sigma_{Z_i} \;\; \Phi^{-1}(U)}. \tag{71}$$

From Theorem 5 and Example 2, we find the following expressions for the risk measures associated with S^c:

$$Q_p\left[S^c\right] = \sum_{i=0}^n \alpha_i \; e^{\mathrm{E}[Z_i] + \sigma_{Z_i} \;\; \Phi^{-1}(p)}, \tag{72}$$

$$\mathrm{CTE}_p\left[S^c\right] = \sum_{i=0}^n \alpha_i \; e^{\mathrm{E}[Z_i] + \frac{1}{2}\sigma^2_{Z_i}} \frac{\Phi\left(\sigma_{Z_i} - \Phi^{-1}(p)\right)}{1-p}, \qquad p \in (0,1)\,, \tag{73}$$

where in deriving (73) we have used the fact that the CTE is additive for comonotonic risks with continuous marginal distributions; recall Remark 2 for details. From (16), Theorem 5 and Example 2 we also find

$$\mathrm{E}(S^c \mid S^c < Q_p^+\left[S^c\right]) = \sum_{i=0}^n \alpha_i \; e^{\mathrm{E}[Z_i] + \frac{1}{2}\sigma^2_{Z_i}} \frac{1 - \Phi\left(\sigma_{Z_i} - \Phi^{-1}(p)\right)}{p}, \qquad p \in (0,1)\,. \tag{74}$$

From (46) and (47) in Dhaene, Denuit, Goovaerts, Kaas & Vyncke (2002a), one can prove that S^c has a strictly increasing distribution function. This implies that in the expression above $Q_p^+(S^c)$ can be replaced by $Q_p(S^c)$.

In order to define a stochastic lower bound for S, we choose a conditioning random variable Λ which is a linear combination of the Y_j:

$$\Lambda = \sum_{j=1}^n \beta_j \; Y_j. \tag{75}$$

After some computations, we find that the lower bound $S^l = \sum_{i=0}^n \alpha_i \; \mathrm{E}\left[e^{Z_i} \mid \Lambda\right]$ is given by

$$S^l = \sum_{i=0}^n \alpha_i \; e^{\mathrm{E}[Z_i] + \frac{1}{2}\left(1 - r_i^2\right)\sigma^2_{Z_i} + r_i \;\; \sigma_{Z_i} \;\; \Phi^{-1}(U)}, \tag{76}$$

where the uniformly distributed random variable U follows from $\Phi^{-1}(U) \equiv \frac{\Lambda - \mathrm{E}(\Lambda)}{\sigma_\Lambda}$, and r_i is the correlation between Z_i and Λ.

If all r_i are positive, then S^l is a comonotonic sum, which means that quantiles and conditional tail expectations related to S^l can be computed by summing the associated risk measures for the marginal distributions involved. Assuming that all r_i are positive, we find the following expressions for the risk measures associated with S^l:

$$Q_p\left[S^l\right] = \sum_{i=0}^{n} \alpha_i \; e^{\mathrm{E}[Z_i]+\frac{1}{2}\left(1-r_i^2\right)\sigma_{Z_i}^2+r_i\;\sigma_{Z_i}\;\Phi^{-1}(p)}, \tag{77}$$

$$\mathrm{CTE}_p\left[S^l\right] = \sum_{i=0}^{n} \alpha_i \; e^{\mathrm{E}[Z_i]+\frac{1}{2}\sigma_{Z_i}^2} \frac{\Phi\left(r_i\;\sigma_{Z_i} - \Phi^{-1}(p)\right)}{1-p}, \qquad p \in (0,1)\,, \tag{78}$$

and also

$$\mathrm{E}(S^l \mid S^l < Q_p^+\left[S^l\right]) = \sum_{i=0}^{n} \alpha_i \; e^{\mathrm{E}[Z_i]+\frac{1}{2}\sigma_{Z_i}^2} \frac{1-\Phi\left(r_i\;\sigma_{Z_i} - \Phi^{-1}(p)\right)}{p}, \qquad p \in (0,1)\,. \tag{79}$$

Again, one can prove that S^l has a strictly increasing distribution function, which implies that $Q_p^+(S^l) = Q_p(S^l)$. We have that

$$\mathrm{CTE}_p\left[S^l\right] \le \mathrm{CTE}_p\left[S\right] \le \mathrm{CTE}_p\left[S^c\right], \qquad p \in (0,1)\,. \tag{80}$$

Note however that this ordering does not hold in general for $Q_p\left[S\right]$ and its approximations $Q_p\left[S^l\right]$ and $Q_p\left[S^c\right]$.

The correlation coefficients r_i follow from the correlations between the random variables Y_i. In the special case that all Y_i are i.i.d., we find

$$r_i = \frac{\sum_{j=1}^{n} \lambda_{ij}\;\beta_j}{\sqrt{\sum_{j=1}^{n} \lambda_{ij}^2}\sqrt{\sum_{j=1}^{n} \beta_j^2}}, \qquad i = 1,\; 2,\; \ldots,\; n. \tag{81}$$

The optimal choice for the coefficients β_j and the performance of (the quantiles of) S^l and S^c as approximations for (the quantiles of) S are investigated in Dhaene, Denuit, Goovaerts, Kaas & Vyncke (2002b). It turns out that S^l performs very well as an approximation for S, even at very high quantiles. ∇

5.4 Theories of choice under risk

In expected utility theory a decision maker asserts a utility $u(x)$ to each possible wealth-level x, see von Neumann & Morgenstern (1947). This real-valued function $u(\cdot)$ is called his utility function. As a rational decision maker is assumed to prefer more to less, it is assumed that a utility function is non-decreasing. If the decision maker, with initial wealth w, has to choose between random losses X and Y, then he compares $\mathrm{E}[u(w-X)]$ with $\mathrm{E}[u(w-Y)]$ and chooses the loss which gives rise to the highest expected utility. Hence, the decision-maker acts in order to maximize his expected utility.

Yaari (1987) presents a dual theory of choice under risk. In this dual theory, the decision maker has a distortion function f. This "distortion function" can be considered as the parallel to the concept of "utility function" in utility theory. While in utility theory, choosing among risks is performed by comparing expected values of transformed wealth levels (utilities), in Yaari's theory the quantities that are compared are the distorted expectations of wealth levels. Consider a decision maker with initial wealth w, which has to choose between two random losses X and Y. The decision-maker acts in order to maximize his distorted expectation. Hence, he will prefer loss X over loss Y if and only if $\rho_f[w-X] \geq \rho_f[w-Y]$, where ρ_f is the distortion risk measure associated with the distortion function f. Comparing the expression

$$\mathrm{E}[w-X] = \int_0^1 Q_{1-q}[w-X]\,\mathrm{d}q \tag{82}$$

with

$$\mathrm{E}[u(w-X)] = \int_0^1 u\left(Q_{1-q}[w-X]\right)\mathrm{d}q \tag{83}$$

and

$$\rho_f[w-X] = \int_0^1 Q_{1-q}[w-X]\,\mathrm{d}f(q), \tag{84}$$

we see that both the expressions (83) and (84) transform the expected wealth level $\mathrm{E}[w-X]$. Under the expected utility hypothesis, the possible levels-of-wealth are adjusted by a utility function, whereas under the distorted expectation hypothesis, the probabilities are adjusted. It is well-known that stochastic dominance and stop-loss order have a natural interpretation in terms of expected utility theory. The pairs of losses X and Y with $X \leq_{st} Y$ are exactly those pairs of losses about which all decision makers (with a non-decreasing) utility function agree:

$$X \leq_{st} Y \Leftrightarrow \mathrm{E}[u(w-X)] \geq \mathrm{E}[u(w-Y)] \text{ for all utility functions } u. \tag{85}$$

In expected utility theory, a decision maker is said to be risk-averse if his utility function is concave. Stop-loss order represents the common preferences of all risk averse decision makers:

$$X \leq_{sl} Y \Leftrightarrow \mathrm{E}[u(w-X)] \geq \mathrm{E}[u(w-Y)] \text{ for all concave utility functions } u. \tag{86}$$

For more details about actuarial applications of expected utility theory and its relation to ordering of random variables, see e.g. Kaas, Goovaerts, Dhaene & Denuit (2001).

On the other hand, one has that

$$X \leq_{st} Y \Leftrightarrow \rho_f[w-X] \geq \rho_f[w-Y] \text{ for all distortion functions } f. \tag{87}$$

Hence, stochastic dominance of loss Y over loss X holds if and only if all decision makers in Yaari's dual theory of choice under risk prefer loss X over loss Y. This characterization for stochastic dominance follows from Theorem 7 by introducing the 'dual distortion function' $\overline{f}$ for each distortion function f:

$$\overline{f}(x) = 1 - f(1-x), \qquad x \in [0,1]. \tag{88}$$

The dual distortion function is again a distortion function. It is clear that $\overline{\overline{f}} \equiv f$. Furthermore, we have that

$$\rho_f[-X] = -\rho_{\overline{f}}[X], \tag{89}$$

and therefore that

$$\rho_f[w-X] \geq \rho_f[w-Y] \Leftrightarrow \rho_{\overline{f}}[X] \leq \rho_{\overline{f}}[Y]. \tag{90}$$

The former relation (89) can easily be proven from the definition (41). Actually, we have

$$\begin{aligned}
\rho_f[-X] &= -\int_{-\infty}^{0} [1 - f(\Pr(-X > x))] \, \mathrm{d}x + \int_0^{\infty} f(\Pr(-X > x)) \, \mathrm{d}x \\
&= -\int_{-\infty}^{0} \overline{f}(\Pr(-X \leq x)) \, \mathrm{d}x + \int_0^{\infty} \left[1 - \overline{f}(\Pr(-X \leq x))\right] \mathrm{d}x.
\end{aligned}$$

Substituting $s = -x$ gives that

$$\begin{aligned}
\rho_f[-X] &= -\int_0^{\infty} \overline{f}(\Pr(-X \leq -s)) \, \mathrm{d}s + \int_{-\infty}^{0} \left[1 - \overline{f}(\Pr(-X \leq -s))\right] \mathrm{d}s \\
&= \int_{-\infty}^{0} \left[1 - \overline{f}(\Pr(X \geq s))\right] \mathrm{d}s - \int_0^{\infty} \overline{f}(\Pr(X \geq s)) \, \mathrm{d}s \\
&= \int_{-\infty}^{0} \left[1 - \overline{f}(\Pr(X > s))\right] \mathrm{d}s - \int_0^{\infty} \overline{f}(\Pr(X > s)) \, \mathrm{d}s,
\end{aligned}$$

where in the last we used the fact that the Lebesgue measure of the set of all discontinuities of a monotone function is 0. This proves the stated result (89).

In Yaari's dual theory of choice under risk, a decision maker is said to be risk-averse if his distortion function is convex. Here we will tacitly assume that a convex distortion function is continuous on $[0,1]$. This means that a risk averse decision maker systematically underestimates his tail probabilities $g\left(\overline{F}_{w-X}(x)\right)$ related to levels-of-wealth, which is a prudent attitude. One finds that stop-loss order of loss Y over loss X can be characterized as follows:

$$X \leq_{sl} Y \Leftrightarrow \rho_f[w-X] \geq \rho_f[w-Y] \text{ for all convex distortion functions } f. \tag{91}$$

Hence, also in Yaari's dual theory of choice under risk, stop-loss order represents the common preferences of all risk averse decision makers. This characterization for stop-loss order follows from Theorem 8 by noting that a distortion function is convex if and only if its dual distortion function $\overline{f}$ is concave. A proof for these characterizations in case of non-negative random variables can be found e.g. in Wang & Young (1998), see also Dhaene, Wang, Young & Goovaerts (2000). Note that the relation between theories of choice under risk and distortion risk measures is also investigated in Denuit, Dhaene & Van Wouwe (1999) and Tsanakas & Desli (2003).

The zero utility risk measure $\rho(X)$ associated with a utility function u is the solution to the following indifference equation:

$$u(0) = E\left[u(\rho\left[X\right] - X)\right], \tag{92}$$

which has an ituitive interpretation in terms of utility theory, see e.g. Kaas, Goovaerts, Dhaene & Denuit (2001). An interesting risk measure arises when the utility function is of the exponential type,

$$u(x) = \frac{1}{\alpha}\left(1 - e^{-\alpha x}\right). \tag{93}$$

In this case we find

$$\rho(X) = \frac{1}{\alpha} \ln E\left[e^{\alpha X}\right]. \tag{94}$$

The class of distortion risk measures can be considered as Yaari's equivalent of the class of zero-utility risks measures in expected utility theory. Indeed, the solution of the indifference equation

$$\rho_f(0) = \rho_f\left[\rho\left[X\right] - X\right] \tag{95}$$

is given by

$$\rho\left[X\right] = \rho_{\overline{f}}\left[X\right]. \tag{96}$$

Hence, any (concave) distortion risk measure $\rho_g\left[X\right]$ can be considered as the solution of the indifference equation (95) of a (risk-averse) decision maker with distortion function $f(x) = \overline{g}(x)$.

Tsanakas & Desli (2003) introduce a class of risk measures which can be considered as the solutions of the indifference equations in 'Generalised Expected Utility Theory. This theory combines both above mentioned theories of choice under risk, see Quiggin (1993).

6 Final remarks

In this paper we examined and summarized properties of several well-known risk measures that can be used in the framework of setting capital requirements for a risky business.

Special attention was given to the class of (concave) distortion risk measures. We investigated the relationship between these risk measures and theories of choice under risk. We considered the problem of how to evaluate risk measures for sums of non-independent random variables. Approximations for such sums, based on the concept of comonotonicity, were proposed. Several examples were provided to illustrate properties or to prove that certain properties do not hold.

Several of the results presented in this paper for (log)normal random variables can be generalized to the class of (log)elliptical distributions, see Dhaene & Valdez (2003). A problem that we did not consider in this paper is how to determine the optimal threshold for determining the required capital. This problem is considered in Examples 9 and 10 of Dhaene, Goovaerts & Kaas (2003).

Acknowledgements

Jan Dhaene, Marc Goovaerts, Steven Vanduffel and David Vyncke acknowledge the financial support of the Onderzoeksfonds K.U. Leuven (GOA/02: Actuariële, financiële en statistische aspecten van afhankelijkheden in verzekerings- en financiële portefeuilles). Qihe Tang acknowledges the support of the Dutch Organization for Scientific Research (NWO 42511013).

References

[1] Artzner, P. (1999). "Application of coherent capital requirements", *North American Actuarial Journal,* vol. 3(2), 11–25.

[2] Artzner, P.; Delbaen, F.; Eber, J.M.; Heath, D. (1999). "Coherent measures of risk", *Mathematical Finance, 9, 203–228.*

[3] Atkinson, D.B.; Dallas, J.W. (2000). *Life Insurance : Products and Finance*, Society of Actuaries.

[4] Black, F.; Scholes, M. (1973). "The pricing of options and corporate liabilities", *Journal of Political Economy*, 81 (May-June), 637–659.

[5] Bühlmann, H. (1985). "Premium calculation from top down", *ASTIN Bulletin*, 15, 89-101.

[6] Denneberg, D. (1994). "Non-additive measure and integral", Kluwer Academic Publishers, Boston.

[7] Denuit, M.; Dhaene, J. & Van Wouwe, M. (1999). "The economics of insurance: a review and some recent developments", *Mitteilungen der Schweiz. Aktuarvereinigung*, 1999(2), 137-175.

[8] Dhaene, J.; Denuit, M.; Goovaerts, M.J.; Kaas, R.; Vyncke, D. (2002a). "The concept of comonotonicity in actuarial science and finance: Theory", *Insurance: Mathematics & Economics*, vol. 31(1), 3–33.

[9] Dhaene, J.; Denuit, M.; Goovaerts, M.J.; Kaas, R.; Vyncke, D. (2002b). "The concept of comonotonicity in actuarial science and finance: Applications", *Insurance: Mathematics & Economics*, vol. 31(2), 133–161.

[10] Dhaene, J.; Goovaerts, M.J. (1996). "Dependency of risks and stop-loss order", *ASTIN Bulletin,* vol. 26(2), 201–212.

[11] Dhaene, J.; Vanduffel, S.; Tang, Qi He; Goovaerts, M.J.; Kaas, R.; Vyncke, D. (2003). "Capital requirements, risk measures and comonotonicity", to be published.

[12] Dhaene, J.; Wang, S.; Young, V; Goovaerts, M. (2000). "Comonotonicity and maximal stop-loss premiums", *Mitteilungen der Schweiz. Aktuarvereinigung, 2000(2)*, 99–113.

[13] Goovaerts, M.J.; De Vijlder, F.; Haezendonck, J. (1984). "Insurance Premiums", North-Holland, Amsterdam.

[14] Goovaerts, M.J.; Kaas, R.; Dhaene, J.; Tang Q. (2003). "A unified approach to generate risk measures". Submitted.

[15] Goovaerts, M.J.; Kaas, R.; Van Heerwaarden, A.E. & Bauwelinkx, T. (1990). "Effective Actuarial Methods", North-Holland, Amsterdam.

[16] Kaas, R.; Dhaene, J.; Goovaerts; M.J. (2000). "Upper and lower bounds for sums of random variables", *Insurance: Mathematics & Economics*, 23, 151–168.

[17] Kaas, R.; Goovaerts, M.J., Dhaene, J.; Denuit, M. (2001). "Modern Actuarial Risk Theory", Kluwer Academic Publishers, pp. 328.

[18] Kaas, R.; Van Heerwaarden, A.E.; Goovaerts, M.J. (1994). "Ordering of Actuarial Risks", Caire Education Series, Amsterdam.

[19] Nakada, P.; Shah, H.; Koyluogo, H.U. & Collignon, O. (1999). "P&C RAROC: A catalyst for improved capital management in the property and casualty insurance industry", *The Journal of Risk Finance*, Fall 1999, 1-18.

[20] Overbeck, L. (2000). "Allocation of economic capital in loan portfolios", *Measuring risk in complex systems*, Franke, J.; Haerdle, W.; Stahl, G. (eds), Springer.

[21] Panjer, H.H. (2002). “Measurement of risk, solvency requirements and allocation of capital within financial conglomerates”, Institute of Insurance and Pension Research, University of Waterloo, Research Report 01–15.

[22] Quiggin, J. (1993). ”Generalised Expected Utility Theory: The Rank-Dependent Model”, Kluwer Academic Publishers, Boston.

[23] Shaked, M.; Shanthikumar, J.G. (1994). “Stochastic orders and their applications”, Academic Press, pp. 545.

[24] Tsanakas, A.; Desli E. (2003). “Risk measures and theories of choice”, to be published.

[25] Valdez, E.; Dhaene, J. (2003). “Bounds for sums of non-independent log-elliptical random variables”, forthcoming.

[26] Von Neumann, J; Morgenstern, O. (1947). “Theory of games and economic behavior”, Princeton University Press, Princeton.

[27] Wang, S. (1995). “Insurance pricing and increased limits rate-making by proportional hazard transforms”, *Insurance: Mathematics & Economics* 17, 43–54.

[28] Wang, S. (1996). “Premium calculation by transforming the layer premium density”, *ASTIN Bulletin* 26, 71–92.

[29] Wang, S. (2000). “A class of distortion operators for pricing financial and insurance risks”, *Journal of Risk and Insurance* 67(1), 15–36.

[30] Wang, S. (2001). “A risk measure that goes beyond coherence”, *Research Report 01-18, Institute of Insurance and Pension Research, University of Waterloo, Canada*, pp. 11.

[31] Wang, S. and Dhaene, J. (1998). “Comonotonicity, correlation order and premium principles”, *Insurance: Mathematics & Economics* 22, 235–242.

[32] Wang, S. and Young, V.R. (1998). “Ordering risks: expected utility theory versus Yaari’s dual theory of risk”, *Insurance: Mathematics & Economics* 22, 235–242.

[33] Wirch, J.L.; Hardy, M.R. (2000). “Ordering of risk measures for capital adequacy”, Institute of Insurance and Pension Research, University of Waterloo, Research Report 00-03.

[34] Yaari, M.E. (1987). “The dual theory of choice under risk”, Econometrica 55, 95–115.