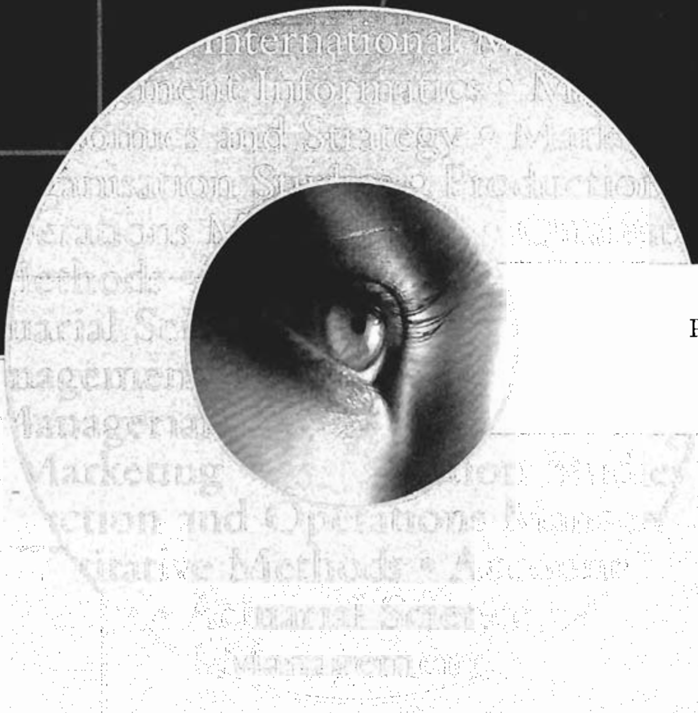
DEPARTMENT OF APPLIED ECONOMICS



RESEARCH REPORT

RISK MEASURES AND DEPENDENCIES OF RISKS

GRZEGORZ DARKIEWICZ • JAN DHAENE • MARC GOOVAERTS

OR 0435

Risk measures and dependencies of risks*

Grzegorz Darkiewicz †‡ Jan Dhaene ‡ Marc Goovaerts ‡

August 10, 2004

Abstract

In the last few years the properties of risk measures that can be considered as suiting "best practice" rules in insurance have been studied extensively in the actuarial literature. In Artzner (1999) so-called coherency axioms were proposed to be satisfied for risk measures that are used for providing capital requirements. On the other hand Goovaerts *et al.* (2003_a), (2003_b),(2003_c) argue that the choice of appropriate set of axioms should depend on the axiomatic "situation at hand".

In this contribution, we show that so-called concave distortion risk measures are not always consistent with some well-known dependency measures such as Pearson's r, Spearman's ρ and Kendall's τ, i.e. higher dependency between random variables does not necessary lead to higher risk measure of corresponding sums. We also test numerically to what extend risk measures are consistent with certain dependency measures and how stable the consistency level is for different one-parametric families of distortion risk measures.

Keywords: Risk measures, distortion risk measures, premium principles, dependency measures.

*Presented as an invited lecture to the First Brazilian Conference on Statistical Modelling in Insurance and Finance, 2003, Ubatuba

†The corresponding author: grzegorz.darkiewicz@econ.kuleuven.ac.be

‡Faculty of Economics and Applied Economics, K.U.Leuven

1 Introduction

A risk measure is an instrument that summarize a distribution of for instance an insurance risk in one single number. There is no commonly accepted classification of insurance risks. "The Report of the IAA's Working Party on Solvency", 2002, suggests to categorize the insurance risks under six major headings: underwriting risk, credit risk, market risk, operational risk, liquidity risk and event risk. This general map of different insurance risks confirms that determining capital requirements for an insurance company (either for reserving or solvency purposes) is a very complex and non-trivial task. By their nature, capital requirements are numeric, based on quantifiable measures of risks.

In general a risk measure is defined as a mapping from the set of risks at hand to the real numbers. It is difficult to specify desirable properties for risk measures. Depending on where it is used for, a risk measure should take into account basic probabilistic quantities such as central tendency, variability, tail behavior or skewness. In many applications it is particularly important to apply risk measures to sums of random variables. In Section 3 we show that the general intuition "the more dependent summands - the more risky sum" is not always the case for the class of so-called distortion risk measures.

Different risk measures do not put the same emphasis on each of the probabilistic quantities and thus the specification of appropriate risk measures must heavily rely on the economic context. In insurance industry there are two main applications of risk measures: at the policy level — the premium, which is understood as the monetary value of the risk associated with the policy, and at the company level — determining the capital requirements for reserving and solvency purposes. In the first case one usually deploys two-sided risk measures which aim to measure the distance between the risky situation and the corresponding risk-free situation when both favorable and unfavorable discrepancies are taken into account. The capital requirements have to be determined much more conservatively and thus so-called one-sided risk measures, to which only unfavorable discrepancies contribute, have to be used. The Value-at-Risk at level p (which is equal to the p-th quantile) is a one-sided risk measure obtained by minimizing the costs capital and residual risk.

A lot of research in actuarial science has been devoted to determine desired properties of risk measures. In the actuarial literature some axiomatic approaches to risk measures (or insurance premium principles) have been proposed. Let us remind some of them: the mean value principle (Hardy *et al.* (1952)), the zero-utility premium principle (Bühlmann (1970)), the Swiss premium principle (Gerber (1974)), the Orlicz premium principle (Haezendonck and Goovaerts (1982)), the Wang's (distortion) premium principle (Wang (1996)). All these risk measures can be described in terms of a few axioms reflecting desirable properties — the related discussion can be found in Goovaerts *et al.* (1984) and Goovaerts *et al.* (2003_b). Recently also so-called coherent risk measures introduced in Artzner (1999) (axioms of monotonicity, translation invariance, subadditivity and positive homogeneity) has drawn a lot of attention in mathematical papers. We discuss the topic of choosing appropriate axioms given the specific economic purpose in Section 2 (see also Goovaerts *et. al* . (2003_a), (2003_b), (2003_c)). Section 3 is devoted to the class of so-called distortion risk measures. In this part we examine the behavior for sums of dependent random variables and its relation with some well-known measures of dependencies. A summary concludes the paper.

2 Risk measures and "best practice" rules

2.1 Premium calculation

When one applies risk measures as premium principles, the coherent risk measures become extremely dangerous, especially in the case of catastrophic risks when one encounters very large claims and strongly dependent risks. In this case the most important shortcoming of coherency is ignoring of available risk capital and as a consequence — the corresponding probability of ruin. In these cases one should be very cautious with risk measures which are subadditive for comonotonic risks (in the extreme case - additive) and/or positively homogeneous.

Obviously there are cases when subadditivity for comonotonic risks reflects the economical reality properly. The so-called **subdecomposability** may be for example useful (see Goovaerts *et al* (1984)):

$$\pi[X] \leq \pi[\alpha X] + \pi[(1-\alpha)X], \text{ where } 0 \leq \alpha \leq 1,$$

(splitting the risk into two separate risks may be more expensive for the company to menage).

This problem can be however solved by the following decomposition of the premium:

$$\pi[X] = \pi'[X] + c[X],$$

where $\pi'[\cdot]$ denotes a pure risk measure and $c[\cdot]$ is the provision for the costs of governing the policy. Then it is reasonable to require $c[\cdot]$ to be subadditive. For some minor policies this subadditivity property may dominate the property of superadditivity for comonotonic risks of pure premium $\pi'[\cdot]$, however for large risks this additional cost premium will become negligible and in this case we can assume that

$$\pi[X] \simeq \pi'[X].$$

Obviously in the case of large claims it often happens that the risk X is much too dangerous for the insurance company to bear as a whole and then splitting the risk between n companies will be advantageous. In such a case the **superdecomposability** of the premium will be a desirable property:

$$\pi[X] \geq \pi[p_1 X] + \ldots + \pi[p_n X], \text{ where } p_i \geq 0 \text{ and } \sum_{i=1}^{n} p_i = 1.$$

Further in this section we will concentrate only on the properties of pure risk premium $\pi[\cdot] = \pi'[\cdot]$, without taking into account provision $c[\cdot]$.

2.1.1 The properties of premium principles

It is reasonable to assume that the following properties should always hold for premium principles:

- $\pi[c] = c$, i.e. when there is no uncertainty, there is no safety loading;
- $\Pr[X \leq Y] = 1 \Rightarrow \pi[X] \leq \pi[Y]$. This condition states that the price of the larger risk must be higher.
- $X \leq_{cx} Y \Rightarrow \pi[X] \leq \pi[Y]$, where $\leq_{cx}$ denotes inequality in the convex order sense. It is the weakest possible condition for risk aversion as follows from utility theory — the risks X and Y are ordered in the convex order sense if all risk averse decision makers prefer

risk X over Y. It is reasonable to assume that in the case of insurance both insurers and insureds are risk averse decisions makers, so the third condition for premiums arises very natural;

Note that two risk measures widely used in practice:

$$\pi_\alpha[X] = E[X] + \alpha\sigma[X])] \text{ and}$$

$$\pi_\beta[X] = E[X] + \beta\mathrm{Var}[X] \tag{1}$$

do not preserve stochastic dominance so generally they should not be used as premium principles. Apart from these general conditions, reasonable premium principles should also satisfy some additional properties for sums of random variables, however they must heavily rely on the dependence structure between the summands. Below we provide some examples in the two extreme cases, namely when random variables are independent and comonotonic.

2.1.2 Additivity properties for independent risks

In most calculations in insurance the assumption of independence of risks reasonably well corresponds to the reality. In the case of a balanced risk such as life insurance or automobile third party liability, the claims may be assumed to be independent or at least conditionally independent given some additional information about the mortality (for example calendar year), interest rates, investment opportunities, the skill and experience of the driver, etc. From the law of large numbers it is known that accumulating such risks will be always beneficial for the company. As a conclusion we state that an insurance premium should satisfy the condition of **subadditivity** for independent risks. Thus for example the group insurance policy purchased by the employer for all employees should be always relatively cheaper than policies purchased individually (in this case risks seem to reveal even a slight positive dependence).

In practical applications however it is often convenient to assume additivity for independent risks. It is for example the case when a so-called top-down calculation of insurance premiums is required, i.e. when the premium is determined at the level of a whole portfolio (for example by considering the ruin probability model) and then distributed to the policyholders (see Bühlmann (1970), Gerber (1979, 1985)). From the characterization of Gerber it follows that any premium which is additive for independent risks and preserves the first and the second stochastic dominance, can be expressed as

$$\pi[X] = \frac{1}{R}\log \mathrm{E}\big[e^{RX}\big].$$

This risk measure is known in the literature as exponential premium principle and can be derived also for example from the utility theory (in this case R represents "the risk aversion coefficient") or ruin theory (then $R = \frac{|\log \varepsilon|}{u}$, where ε denotes the imposed probability of ruin and u is the initial capital).

2.1.3 Additivity properties for comonotonic risks

The case of comonotonic risks corresponds to the extreme positive dependency. Formally the vector (X, Y) is said to be comonotonic if

$$(X, Y) =^D (F_X^{-1}(U), F_Y^{-1}(U)), \text{ where } U \sim U(0,1).$$

In this definition we use the generalized inverse function, namely

$$F^{-1}(p) = \inf\{t|F(t) \geq p\}.$$

Clearly from this definition it follows that accumulating comonotonic risks may not be advantageous for an insurer — in this case risks do not hedge against each other and accumulating comonotonic risks substantially increases the probability of ruin. Therefore risk measures which allow strict subadditivity for comonotonic risks do not find any reasonable applications as premium principles. There are some cases when it is convenient and advantageous to use risk measures which are **additive** for comonotonic risks, but we will demonstrate that the additivity may be also very dangerous. In general in the case of insurance premiums one should impose the condition of **superadditivity** for all possible pairs of comonotonic risks.

Example 1 *Suppose that $\pi[\cdot]$ denotes an arbitrary premium principle additive for comonotonic risks and thus also translation invariant. We assume for simplicity that $X_1, \ldots, X_n$ are binomially distributed with parameter $q = 0.1$ and represent comonotonic risks. Suppose also that there is an initial capital u and that we want to ensure that the probability of ruin is smaller than 5%.*

Obviously it is reasonable to assume that $\pi[X] < 1$ *because otherwise nobody would purchase the policy. However then for* n *large enough we get*

$$\Pr\left[\{\mathit{ruin\ will\ occur}\}\right] = \Pr[u + n\pi(X) - \sum_{i=1}^{n} X_i < 0] = 0.1 > 0.05$$

for sufficiently large n*. Thus in this example the strict superadditivity for comonotonic risks is essential.*

Although mathematics hidden behind this example is very simplified, similar situations are well-known from insurance practice. Obviously there is no insurance company which would insure all buildings on the same seismic area or all floors in skyscraper at Manhattan (in both examples the considered risks are close to comonotonicity), unless insureds would pay the premium close to the maximal possible damage. It is not easy to find anybody who would agree to pay such a premium. However after disaggregation such risks are successfully insured and corresponding premiums remain at reasonable high levels. In this particular case the premium principle used by companies satisfy the strict superadditivity condition:

$$\pi[X_1 + \ldots + X_n] > \pi[X_1] + \ldots + \pi[X_n].$$

Note that the exponential premium principle introduced in Section 2.2.2 is superadditive for comonotonic risks (in fact it is superadditive even for the sums of positive quadrant dependent (PQD) couples — see Kaas *et al.* (2001)).

2.1.4 Some comments on positive homogeneity of premium principles

In the actuarial literature it is often argued that premium principles should be positively homogeneous because only such risk measures can be expressed in monetary units and are independent of the actual currency. It is only partially true — indeed, when a risk measure is positive homogeneous then it satisfies these conditions. The opposite implication however does not hold.

Example 2 *Once again we consider the exponential premium principle:*

$$\pi[X] = \frac{1}{R} \log \mathrm{E}\left[e^{RX}\right].$$

It is straightforward to verify from Jensen's inequality that

$$\pi[aX] \begin{cases} \leq a\pi[X] & \textit{for } 0 < a \leq 1 \\ \geq a\pi[X] & \textit{for } a \geq 1 \end{cases} . \tag{2}$$

Does it mean that after exchanging Belgian Francs to Euro we will pay less for the premium if the rules remain unchanged? Not necessary. In Section 2.2.2 we recalled that the exponential premium principle may be derived from the ruin theory and then

$$\frac{1}{R} = \frac{u}{\log \varepsilon},$$

where u is the initial capital and ε denotes the imposed probability of ruin. Thus in this example not only X is expressed in monetary units but also $\frac{1}{R}$, and thus when one changes the currency and adjusts the coefficient R properly — the premium principle turns up to be independent from the currency.

Obviously in other cases one has no such clear interpretation as the ruin theory. However in many cases coefficients in formulae for the corresponding risk measures cannot be interpreted as dimension-free. Let us consider another example.

Example 3 *Recall the risk measure given by (1). In this case parameter β cannot be interpreted as dimension-free because otherwise the first summand will be expressed in Euro while the second — in Euro squared. Thus β must be expressed in $\frac{1}{Euro}$ to give risk measure $\pi_\beta(\cdot)$ in monetary units. Therefore formula (1) can be rewritten for example as follows:*

$$\pi_\beta(X) = E(X) + \beta' \mathrm{E}\Big[\frac{(X - E(X))^2}{u}\Big],$$

where u denotes e.g. the initial capital and β' is a dimension-free constant.

Summarizing, in many cases the positive homogeneity may be a useful and convenient property. However it has nothing to do with the independence of currency. Moreover we are reluctant to require this property for all risk measures used in practice, because it causes very similar problems to those illustrated in Section 2.2.3 for the property of additivity for comonotonic risks (in fact the positive homogeneity and the additivity for comonotonic risks are closely related to each other) — multiplying the risk by a large constant a increases substantially the probability of ruin. We think that more general condition (2) reflects the desirable properties of premium principles much better.

2.2 Risk sharing schemes

In practice we encounter sharing of risks for example when an insurer cedes part of his risk to a reinsurer. Suppose that an insurance company is facing the risk X. Assume that the reinsurer is obliged to cover a part equal to $\phi(X)$ while $X - \phi(X)$ is retained by the insurer. It is reasonable to assume that function ϕ satisfies the following conditions:
a) $0 \leq \phi(x) \leq x$;
b) both $\phi(x)$ and $x - \phi(x)$ are non-decreasing functions of x.
One can easily verify that functions given below which define widely used in practice risk sharing schemes, satisfy conditions a) and b):

- *A stop-loss coverage*: for $d > 0$, $\phi(x) = (x - d)_+$;
- *A quota-share coverage*: for $0 \leq \alpha \leq 1$, $\phi(x) = \alpha x$;
- *A coverage with a maximal limit*: for $d > 0$, $\phi(x) = \min(x, d)$.

Clearly under conditions a) and b) both parts of the vector $(\phi(X), X - \phi(X))$ are comonotonic, thus if one has to distribute the premium between the two parties involved, the property of additivity for comonotonic risks will be desirable, i.e.

$$\pi[X^{(c)} + Y^{(c)}] = \pi[X] + \pi[Y].$$

It is also worth to mention that all risk measures which are additive for comonotonic risks and additionally satisfy the three conditions from Section 2.1.1 can be represented as concave distortion risk measures (at least for bounded random variables). A related discussion can be found in e.g. Wang (1996) or Goovaerts & Dhaene (1998).

2.3 A solvency margin

A calculation of a solvency margin is another typical application of risk measures. However it requires completely different properties of the corresponding risk measures than for example premium calculation (at the policy level) or determination of reserves (at the company level). The solvency margin is interpreted as a provision for the adverse outcome and as a matter of fact it should be equal to zero for all situations where there is no uncertainty involved. In particular it does not make any sense to require the property of monotonicity for the corresponding risk measures.

Example 4 *Consider a Bernoulli risk B_q with parameter $q \in [0,1]$. Then obviously premium principle $\pi[B_q]$ should be increasing in q (monotonicity). On the contrary, consider a risk measure $\rho[\cdot]$ to compute the solvency margin. It is clear that $\rho[B_0] = \rho[B_1] = 0$ because in both situations there is no uncertainty involved. Moreover one can assume that $\rho[B_q] = \rho[B_{1-q}]$ because $B_q =^D 1 - B_{1-q}$ and thus one can think that in these two cases uncertainties are "equal" (note that we put here the same weight to positive and negative discrepancies). Consider a function $f(q)$ for $q \in [0, \frac{1}{2}]$ such that $f(0) = 0$, $f \geq 0$ and $f'(\frac{1}{2}) = 0$. Then risk measure $\rho[\cdot]$ for determining a solvency margin can be defined as*

$$\rho(B_q) = \begin{cases} f(q) & \text{for } 0 \leq q \leq \frac{1}{2} \\ f(1-q) & \text{for } \frac{1}{2} \leq q \leq 1 \end{cases}$$

and the corresponding premium principle as

$$\pi(B_q) = q + \rho(B_q).$$

Recall that $\pi(B_q)$ should be increasing in q, what leads to the following additional condition for f:

$$-1 \leq f'(q) \leq 1. \tag{3}$$

Now consider two specific functions: $f_1(q) = \alpha\sqrt{q(1-q)}$ and $f_2(q) = \beta q(1-q)$. One can easily verify that for any $\alpha > 0$ $f'_{1+}(0) = -\infty$ and that for any $\beta \leq 1$ (3) is satisfied by f_2. Thus in the situation "at hand" f_2 is an example of a consistent risk measure for calculating solvency margin while f_1 not (because it leads to a premium which is not monotonic).

2.4 An allocation of an economic capital

There must be a substantial difference between risk measures applicable as premium principles and those used to allocate economic capital. The capital allocation problem is somehow dual — in this case the risk (at the level of a company) is given and one has to determine the required capital sufficiently large too make the ruin unlikely. We will demonstrate that also in this case risk measures which have to be used exhibit very complex behavior. In particular coherent risk measures do not always lead to optimal solutions.

Example 5 *(A capital allocation based on the cost minimization) Consider the following problem. Suppose that an insurance company faces a risk X and that the shareholders have to provide the capital u to let the business run. However when at the end of the year the shortfall occurs, they are also obliged to cover the deficit. On the other hand it is not allowed to withdraw the capital if the shortfall does not occur. Suppose that the capital will be provided at the price i per unit and that a risk-free interest rate is equal to r. Under these assumptions the shareholders will aim to solve the following minimization problem of their expected cost:*

$$\min_u (i-r)u + \mathrm{E}\big[(X-u)_+\big],$$

which has the unique solution equal to $F_X^{-1}\big(1-\frac{i-r}{1+r}\big)$ (see Goovaerts et al., 2003$_a$). Thus in this case a very natural optimization problem leads to the Value-at-Risk which is a non-coherent risk measure.

Example 6 *(An allocation of an available economic capital between the subsidiaries) Now consider the following problem. Suppose that a company faces a risk X and that the capital u to cover this risk has been allocated already. Now suppose that risk X has to be split into two risks $X = X_1 + X_2$. Then one faces the problem of finding the optimal division of economical capital u into $u = u_1 + u_2$ where u_1 is allocated to risk X_1 and u_2 to X_2. The optimal solution will be given by solving the following minimization problem:*

$$\min_{u=u_1+u_2} \big\{\rho[X_1-u_1] + \rho[X_2-u_2]\big\}, \tag{4}$$

where ρ is a risk measure which has to be used in this situation. In this case also a non-coherent risk measure has to be used. Otherwise, because of the property of translation invariance, (4) *simplifies to*

$$\min_{u=u_1+u_2} \rho[X_1] + \rho[X_2] - u,$$

which does not lead to any solution.

Example 7 *(An allocation of an economic capital for sums of risks) In this example we consider a risk measure $\rho[\cdot]$ which has to be used as a rule of determining an economic capital, i.e. the amount $u = \rho(X)$ to be allocated to the risk X. Now suppose that two companies represented by risks X_1 and X_2 merge to $X = X_1 + X_2$. From the regulatory's point of view the merger should be efficient in the following sense:*

$$\pi\big[(X_1 + X_2 - u)_+\big] \leq \pi\big[(X_1 - u_1)_+ + (X_2 - u_2)_+\big] \tag{5}$$

(both sides of the inequality represent the cost to the society). Note that under a mild and natural assumption that risk measure $\pi[\cdot]$ has to preserve the stochastic dominance, subadditive risk measure ρ may lead to problems for (5) to be satisfied. On the other hand note that one has with probability one an inequality:

$$(X_1 + X_2 - u_1 - u_2)_+ \leq (X_1 - u_1)_+ + (X_2 - u_2)_+$$

Thus the residual risk of the merged company is always smaller than the risk of the split company. This fact will hold in general for risk measures $\rho[\cdot]$ which are superadditive.

We are far from requiring superadditivity for risk measures used for economic capital purposes. Example 7 aims only to illustrate that risk measures which are subadditive for all possible dependence structures of the vector (X_1, X_2) do not reflect properly the dependency between $(X_1 - u_1)_+$ and $(X_2 - u_2)_+$. Taking this dependency into account, the risk measure providing capitals u, u_1 and u_2 will not always be subadditive nor always superadditive, but may instead exhibit behavior similar to the Value-at-Risk (see Embrechts *et al.* (2002)). From this perspective the fact that the Value-at-Risk is neither sub- nor superadditive is a desirable property rather than a pitfall!

2.5 Consistent risk measures

In this section we provided several examples to demonstrate that "best practice" rules in insurance require sometimes much more complex properties of risk measures than those following from coherency axioms. It does not seem to be reasonable to require one particular set of axioms to hold in all risky situations, without taking into consideration the available economic capital or the dependency structure between random variables. In Goovaerts *et al.* (2003_b) and (2003_c) it was argued that in any realistic situation at hand, a specific set of axioms $\mathbb{S}$ "consistent" with the given situation has to be considered. More precisely, they considered the following definition.

Definition 1 *Let $\mathbb{S}$ be a set of axioms for risk measures and α denotes an arbitrary number number from the interval $(0,1)$. A risk measure $\pi[\cdot] = \pi_{(\mathbb{S},\alpha)}[\cdot] = \pi_\alpha$ is called $(\mathbb{S},\alpha)$-consistent if $\pi[\cdot]$ satisfies the set of axioms $\mathbb{S}$ and inequality $\pi[X] > F_X^{-1}(\alpha)$ for any risk X, where $F_X^{-1}(\alpha)$ denotes α-quantile.*

The condition on α ensures that the risk measure is acceptable for regulators who impose the Value-at-Risk at level α. In Goovaerts *et al.* (2003_b) some universal procedures based on the Markov inequality were provided to generate $(\mathbb{S},\alpha)$-consistent risk measures.

3 Distortion risk measures and dependency measures

3.1 Introduction

Distortion risk measures were introduced in Wang (1996). For a given non-decreasing function $g : [0,1] \to [0,1]$ such that $g(0) = 0$ and $g(1) = 1$ for every risk the corresponding risk measure is defined as follows:

$$H_g[X] = \int_0^\infty g(1 - F_X(t))dt = \int_0^1 F_X^{-1}(1-q)dg(q), \tag{6}$$

where $F_X(t)$ denotes the distribution function of X. We will call g a distortion function. Distortion risk measures have many properties discussed in the previous section: positive homogeneity, translation invariance, additivity for comonotonic risks, preservation of first order stochastic dominance. Moreover if we additionally assume concavity of distortion function g than the corresponding risk measure is also subadditive, and thus is Artzner-coherent.

These properties of distortion risk measures were comprehensively studied in many works (see e.g. Wang (1996), Wang *et. al* (1997), Dhaene and Wang (1998), Wang and Young (1998), Wirch and Hardy (2000), Dhaene *et. al* (2004)). In this section we investigate the relation between distortion risk measures applied to sums of random variables and some well-known dependency measures between summands (throughout this section we assume that marginal distributions are fixed). The theorem we cite below says that when the dependency level differs strongly (which is expressed in the terms of the so-called correlation order of pairs of random variables) then all concave distortion risk measures behave intuitively, i.e. the more dependent summands - the more risky sums.

Definition 2 *Let* (X_1, Y_1) *and* (X_2, Y_2) *be elements of* $R(F_X, F_Y)$ *(i.e. have the same marginal distributions equal to* F_X *and* F_Y*). Then we say that* (X_1, Y_1) *precede* (X_2, Y_2) *in correlation order sense when either of the two equivalent conditions holds:*
(a) for all non-decreasing functions f, g *one has that* $\mathrm{Cov}(f(X_1), g(Y_1)) \leq \mathrm{Cov}(f(X_2), g(Y_2))$, *provided that the respective covariance functions exist;*
(b) $F_{(X_1,Y_1)}(x, y) \leq F_{(X_2,Y_2)}(x, y)$ *for any non-negative pair* (x, y).

We denote the correlation order by $\leq_{corr}$.

Theorem 1 *Suppose that* g *is a concave distortion function. Assume* $(X_1, Y_1), (X_2, Y_2) \in R(F_X, F_Y)$ *are such that* $(X_1, Y_1) \leq_{corr} (X_2, Y_2)$. *Then* $H_g(X_1 + Y_1) \leq H_g(X_2 + Y_2)$.

PROOF. See Wang, Dhaene (1998). ■

However the correlation order is only a partial order and recognizes only very strong differences. In this section we investigate how distortion risk measures are related to some more elastic measures of dependency, namely:

- Pearson's correlation coefficient

$$r(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma(X)\sigma(Y)};$$

- Spearman's rank correlation coefficient

$$\rho(X,Y) = \frac{E[F_X(X)F_Y(Y)] - E[F_X(X)]E[F_Y(Y)]}{\sigma\big(F_X(X)\big)\sigma\big(F_Y(Y)\big)}; \tag{7}$$

- Kendall's rank correlation coefficient

$$\tau(X,Y) = \Pr\big((X-X')(Y-Y')>0\big) - \Pr\big((X-X')(Y-Y')<0\big), \tag{8}$$

  where (X,Y) and (X',Y') are two independent copies from the considered bivariate distribution.

We show that in general there is no strict relation between distortion risk measures and those measures of dependencies. In the following subsection we show that for Tail Value-at-Risk it is possible to find random pairs with fixed marginals (X_1,Y_1) and (X_2,Y_2) such that

$$TVaR_p[X_1+Y_1] > TVaR_p[X_2+Y_2]$$

despite the ordering of all corresponding correlation coefficients is the opposite, i.e.:

$$r(X_1,Y_1) < r(X_2,Y_2), \quad \rho(X_1,Y_1) < \rho(X_2,Y_2) \quad \text{and} \quad \tau(X_1,Y_1) < \tau(X_2,Y_2). \tag{9}$$

Next, we show that for any distortion risk measure $H_g[\cdot]$ it is possible to construct such random pairs $(X_1,Y_1),(X_2,Y_2) \in R(F_X,F_Y)$ that $TVaR_p[X_1+Y_1] > TVaR_p[X_2+Y_2]$ and $r(X_1,Y_1) < r(X_2,Y_2)$. Moreover it turns out that under a special condition which ensures that distortion function g is not "too concave", all three inequalities (9) hold.
Finally, we propose an experimental test which aims to indicate how strong is the relationship between the riskiness of sums of random variables generated by distortion risk measures and the measures of dependency between appropriate summands.

3.2 A counterexample for the Tail Value-at-Risk

The Tail Value-at-Risk (further we will call it TVaR) has been recognized as a very important risk measure which can be used for solvency purposes. Artzner (1999) recommended this risk measure to determine solvency capital requirements, in Panjer (2002) it was used to allocate solvency economic capital between subsidiaries for normally distributed risks. The practical importance of the TVaR is intuitively clear — for continuous distributions it can be interpreted as an expected loss when a specified threshold (defined here as an appropriate quantile) is exceeded. The TVaR at level p is also the smallest concave distortion risk measure exceeding VaR at level p and thus is acceptable by regulators, see Dhaene *et al.* (2004).

Formally the TVaR at level p is defined as follows:

$$TVaR_p[X] = \frac{1}{1-p}\int_p^1 Q_q(X)dq,$$

and it is straightforward to prove that $TVaR_p$ is determined by the concave distortion function:

$$g_p(x) = \begin{cases} \frac{1}{p}x & \text{for } 0 \le t \le p \\ 1 & \text{for } p < t \le 1 \end{cases} \quad \text{where } 0 \le p \le 1.$$

Remark 1 *In the actuarial literature the* $TVaR$ *is often confused with the so-called Conditional Tail Expectation (CTE)defined below:*

$$CTE_p(X) = \mathrm{E}\big[X|X > Q_p(X)\big],$$

where $Q_p(X)$ *denotes p-th quantile of* X*. Indeed, in the case of continuous random variables* $TVaR$ *and* CTE *do coincide, however they are not necessary the same in the discrete case and in general* CTE_p *cannot be expressed as a distortion risk measure. The subtle differences between those two risk measures were investigated in Dhaene et. al (2004).*

The following example shows that for sums of random variables with fixed marginal distributions, $TVaR$ does not preserve in general neither of the three well-known dependency measures: Pearson's r, Spearman's ρ and Kendall's τ.

Example 8 Let X and Y be two random variables with probabilities $\Pr(X = i) = p_i$ and $\Pr(Y = i) = q_i$ given by:

$$p_0 = p_1 = \frac{1-\sqrt{p}}{2}, \quad p_2 = \sqrt{p} \tag{10}$$

and

$$q_0 = 1 - \sqrt{p}, \quad q_1 = \sqrt{p}.$$

Let (X_1, Y_1) and (X_2, Y_2) be two elements of $R(F_X, F_Y)$. Concerning the dependency structure of the couples, we assume that X_1 and Y_1 are mutually independent, while the distribution of (X_2, Y_2) is given in the following table:

	X_2		
Y_2	0	1	2
0	$p_0q_0 + x\varepsilon$	$p_1q_0 - \varepsilon$	$p_2q_0 + (1-x)\varepsilon$
1	$p_0q_1 - x\varepsilon$	$p_1q_1 + \varepsilon$	$p_2q_1 - (1-x)\varepsilon$

In this definition x denotes a positive number satisfying the following inequalities

$$1 \geq x \geq \max\left(\frac{1}{2}, \frac{2\sqrt{p}}{1+\sqrt{p}}, \frac{1+\sqrt{p}}{3-\sqrt{p}}\right) \tag{11}$$

and ε is an arbitrary positive number such that:

$$\varepsilon \leq \min\left(\frac{p_0q_1}{x}, p_1q_0, \frac{p_2q_1}{1-x}\right).$$

One can immediately verify that $(X_2, Y_2) \in R(F_X, F_Y)$. Note also that for the first independent pair one has $r(X_1, Y_1) = \rho(X_1, Y_1) = \tau(X_1, Y_1) = 0$.

All correlation coefficients for the second pair are positive, which can be verified as follows:

- $\mathrm{Cov}(X_2, Y_2) = (2x-1)\varepsilon > 0$ because $x > \frac{1}{2}$ and thus also $r(X_2, Y_2) > 0$.
- From (7) we have that

$$\rho(X_2, Y_2) = \frac{\varepsilon(1-q_0)\big((1-x)p_0 + p_1 - (1-x)\big)}{\sigma\big(F_X(X)\big)\sigma\big(F_Y(Y)\big)},$$

  which is positive when $x > \frac{1-p_0-p_1}{1+p_0}$. Combining this with (10) we get that $x > \frac{2\sqrt{p}}{1+\sqrt{p}}$ which is always true in view of (17).
- A straightforward manipulation of (8) leads to the formula:

$$\begin{aligned}\tau(X_2, Y_2) &= 2\big((p_0q_0 + x\varepsilon)(p_2q_1 - (1-x)\varepsilon) + (p_0q_0 + x\varepsilon)(p_1q_1 + \varepsilon) \\ &+ (p_1q_0 - \varepsilon)(p_2q_1 - (1-x)\varepsilon)\big) - 2\big((p_0q_0 + x\varepsilon)(p_2q_1 - (1-x)\varepsilon) \\ &+ (p_0q_0 + x\varepsilon)(p_1q_1 + \varepsilon) + (p_1q_0 - \varepsilon)(p_2q_1 - (1-x)\varepsilon)\big).\end{aligned}$$

Note that all expressions without ε sum up to 0 as well as all expressions with ε^2 and thus (after some calculations) the condition for $\tau(X_2, Y_2)$ to be positive is equivalent to the inequality

$$xp_0 + (2x-1)p_1 + xp_1 - (1-x)p_2 > 0,$$

what — after taking into account (10) — gives $x > \frac{1+\sqrt{p}}{3-\sqrt{p}}$, which holds because of (17).

Now let us return to the TVaR. For the decumulative distribution functions of the sums $S_i = X_i + Y_i$ we find:

$$\overline{F}_{S_1}(t) = \begin{cases} 1 & \text{for } t < 0, \\ p + \upsilon + \vartheta & \text{for } 0 \le t < 1, \\ p + \upsilon & \text{for } 1 \le t < 2, \\ p & \text{for } 2 \le t < 3, \\ 0 & \text{for } t \ge 3. \end{cases}$$

and

$$\overline{F}_{S_2}(t) = \begin{cases} 1 & \text{for } t < 0, \\ p + \upsilon + \vartheta - x\varepsilon & \text{for } 0 \le t < 1, \\ p + \upsilon + \varepsilon & \text{for } 1 \le t < 2, \\ p - (1-x)\varepsilon & \text{for } 2 \le t < 3, \\ 0 & \text{for } t < 0. \end{cases}$$

(for simplicity of notation we denote $\Pr[S_1 = 2]$ by υ and $\Pr[S_1 = 1]$ by ϑ).

The computation of the first integral in formula (6) is now straightforward:

$$H_{g_p}[S_1] = g_p(p + \upsilon + \vartheta) + g_p(p + \upsilon) + g_p(p) = 1 + 1 + 1 = 3,$$

$$\begin{aligned} H_{g_p}[S_2] &= g_p(p + \upsilon + \vartheta - x\varepsilon) + g_p(p + \upsilon + \varepsilon) + g_p(p - (1-x)\varepsilon) = \\ &= 1 + 1 + \frac{p - (1-x)\varepsilon}{p} < 3 = H_{g_p}[S_1]. \end{aligned}$$

Thus $TVaR_p(X_1 + Y_1) > TVaR_p(X_2 + Y_2)$ despite $r(X_1, Y_1) < r(X_2, Y_2)$, $\rho(X_1, Y_1) < \rho(X_2, Y_2)$ and $\tau(X_1, Y_1) < \tau(X_2, Y_2)$.

■

3.3 A construction of a general counterexample

We split the construction into two cases: the critical case when g is concave and the easy case of non-concave distortion functions.

3.3.1 The case of concave distortion functions

We restrict ourselves only to the case when a distortion function $g : [0,1] \to [0,1]$ satisfies some additional smoothness conditions. More precisely we will assume the following:

(i) $g(0) = 0$ and $g(1) = 1$;

(ii) g is piecewise twice continuously differentiable;

(iii) for all x $g'(x) \geq 0$ (thus g is nondecreasing) and $g''(x) \leq 0$;

(iv) g differs from the identity function.

Condition (iv) excludes the trivial case of the expectation. Note that assumption (ii) allows for example piecewise linear distortion functions. In fact in our prove we use only left continuity of first derivative at 1 and right continuity at 0.
We start with a helpful technical lemma.

Lemma 1 *Let g be an arbitrary function satisfying conditions (i)-(iv). Then there exist real numbers $\alpha_1 < \alpha_2$ in $(0,1)$ such that $g'(\alpha_1) > g'(\alpha_2)$ and*

$$(1-x)g'(\alpha_1) + xg'_-(1) > g'(\alpha_2), \tag{12}$$

where x is an arbitrary number from the interval $(\frac{1}{2}, 1)$.
If we additionally assume that $-4g''_-(1) < g'_+(0) - g'_-(1)$ then for (12) to hold true it may be assumed that $x = \frac{2}{3-\sqrt{\alpha_2}}$.

PROOF. To prove the first part, we start with choosing any $\alpha_1 \in (0,1)$ such that

$$g'(\alpha_1) > g'_-(1) > 0$$

(this is always possible in view of conditions (i)-(iv)). Define $\varepsilon = (1-x)\big(g'(\alpha_1) - g'_-(1)\big) > 0$. Left continuity of g' in 1 implies that it is possible to choose a point α_2 such that

$$g'(\alpha_2) - g'_-(1) < \varepsilon.$$

Then one gets

$$g'(\alpha_2) < g'_-(1) + \varepsilon = (1-x)g'(\alpha_1) + xg'(1).$$

Moreover,

$$g'(\alpha_2) - g'_-(1) < \varepsilon = \frac{g'(\alpha_1) - g'_-(1)}{3} < g'(\alpha_1) - g'_-(1),$$

and hence

$$g'(\alpha_1) > g'(\alpha_2),$$

which completes the proof of the first part.

The proof of the second part is a bit more subtle, because α_2 cannot be chosen as a function of x. Recall that we assume here additionally that

$$-4g''_-(1) < g'_+(0) - g'_-(1).$$

From continuity of the first derivative it immediately follows that it can be chosen such number $\alpha_1 > 0$ that

$$-4g''_-(1) < g'(\alpha_1) - g'_-(1).$$

Note that inequality (12) which has to be proven can be rewritten as

$$g'(\alpha_1) - g'(\alpha_2) < \frac{2}{3-\sqrt{\alpha_2}}\big(g'(\alpha_1) - g'_{(-)}(1)\big).$$

Consider an auxiliary function f defined as follows:

$$f(p) = g'(\alpha_1) - g'(p) - \frac{2}{3-\sqrt{p}}\big(g'(\alpha_1) - g'_{(-)}(1)\big).$$

One can easily check that $f(1)=0$ and

$$f'_-(1) = -f''_-(1) - \frac{1}{4}\big(g'(\alpha_1) - g'_-(1)\big) < 0.$$

Thus it is possible to choose $\alpha_1 < \alpha_2 < 1$ such that $f(\alpha_2) > 0$. Moreover from the identity

$$g'(\alpha_1) = g'(t) \Rightarrow f(t) < 0$$

we conclude that $g'(\alpha_1) > g'(\alpha_2)$ what completes the proof of Lemma 1. ∎

Theorem 2 *Let g be an arbitrary function satisfying conditions (i)-(iv). Then there exist univariate distributions $F_{X^{(g)}}$, $F_{Y^{(g)}}$ and random couples $\big(X_1^{(g)}, Y_1^{(g)}\big)$, $\big(X_2^{(g)}, Y_2^{(g)}\big)$ belonging to $R(F_{X^{(g)}}, F_{Y^{(g)}})$ such that*
(i) $r\big(X_1^{(g)}, Y_1^{(g)}\big) < r\big(X_2^{(g)}, Y_2^{(g)}\big)$,
(ii) $H_g\big[X_1^{(g)} + Y_1^{(g)}\big] > H_g\big[X_2^{(g)} + Y_2^{(g)}\big]$.
Moreover under additional assumption that

$$-4g''_-(1) < g'_+(0) - g'_-(1), \tag{13}$$

the random couples can be chosen such that also $\rho\big(X_1^{(g)}, Y_1^{(g)}\big) < \rho\big(X_2^{(g)}, Y_2^{(g)}\big)$ and $\tau\big(X_1^{(g)}, Y_1^{(g)}\big) < \tau\big(X_2^{(g)}, Y_2^{(g)}\big)$.

PROOF. Consider two points $0 < \alpha_1 < \alpha_2 < 1$ satisfying the conditions of Lemma 1. Consider the random variables $X^{(g)}$ and $Y^{(g)}$ for which $\Pr\big(X^{(g)} = i\big) = p_i$ and $\Pr\big(Y^{(g)} = j\big) = q_j$ are given below:

$$p_0 = p_1 = \frac{1-\sqrt{\alpha_2}}{2}, \quad p_2 = \sqrt{\alpha_2} \tag{14}$$

and

$$q_0 = 1 - \frac{\alpha_1}{\sqrt{\alpha_2}}, \quad q_1 = \frac{\alpha_1}{\sqrt{\alpha_2}}. \tag{15}$$

Furthermore, let $\big(X_1^{(g)}, Y_1^{(g)}\big)$ be an independent pair with marginal distributions as defined in (14) and (15), i.e.:

$$\Pr\big[X_1^{(g)} = i, Y_1^{(g)} = j\big] = p_i q_j. \tag{16}$$

The joint distribution of $\big(X_2^{(g)}, Y_2^{(g)}\big)$ is defined in Table 1, where x denotes

(i) any fixed number in the interval $(\frac{1}{2}, 1)$ if (13) is not satisfied;

(ii) $x = \frac{2}{3-\sqrt{\alpha_2}}$ if (13) is satisfied

and ε is chosen as an arbitrary positive number such that

$$\varepsilon \leq \min\Big(\frac{p_0 q_1}{x}, p_1 q_0, \frac{p_2 q_1}{1-x}\Big).$$

	$X_2^{(g)}$		
$Y_2^{(g)}$	0	1	2
0	$p_0q_0 + x\varepsilon$	$p_1q_0 - \varepsilon$	$p_2q_0 + (1-x)\varepsilon$
1	$p_0q_1 - x\varepsilon$	$p_1q_1 + \varepsilon$	$p_2q_1 - (1-x)\varepsilon$

Table 1: The distribution of $\big(X_2^{(g)}, Y_2^{(g)}\big)$.

Note that in the case when (13) is satisfied the follwing inequalities hold:

$$1 \geq x \geq \max\Big(\frac{1}{2}, \frac{2\sqrt{\alpha_2}}{1+\sqrt{\alpha_2}}, \frac{1+\sqrt{\alpha_2}}{3-\sqrt{\alpha_2}}\Big) \tag{17}$$

One can immediately verify that $(X_2^{(g)}, Y_2^{(g)}) \in R(F_X, F_Y)$. Note also that for the first independent pair one has $r(X_1, Y_1) = \rho(X_1, Y_1) = \tau(X_1, Y_1) = 0$, which have to be compared to the correlation coeficients of the second pair calculated as follows:

(i) $\mathrm{Cov}(X_2, Y_2) = (2x-1)\varepsilon > 0$ and thus also $r(X_2, Y_2) > 0$;

(ii) From (7) we have that

$$\rho(X_2, Y_2) = \frac{\varepsilon(1-q_0)\big((1-x)p_0 + p_1 - (1-x)\big)}{\sigma\big(F_X(X)\big)\sigma\big(F_Y(Y)\big)},$$

which is positive when $x > \frac{1-p_0-p_1}{1+p_0}$. Combining this with (14) we get that $x > \frac{2\sqrt{\alpha_2}}{1+\sqrt{\alpha_2}}$ which is in view of (17) true in the case when (13) holds.

(iii) A straightforward manipulation on (8) leads to the formula:

$$\begin{aligned}\tau(X_2, Y_2) &= 2\big((p_0q_0 + x\varepsilon)(p_2q_1 - (1-x)\varepsilon) + (p_0q_0 + x\varepsilon)(p_1q_1 + \varepsilon) \\ &+ (p_1q_0 - \varepsilon)(p_2q_1 - (1-x)\varepsilon)\big) - 2\big((p_0q_1 - x\varepsilon)(p_2q_0 + (1-x)\varepsilon) \\ &+ (p_0q_1 - x\varepsilon)(p_1q_0 - \varepsilon) + (p_1q_1 + \varepsilon)(p_2q_0 + (1-x)\varepsilon)\big).\end{aligned}$$

Note that all expressions without ε sum up to 0 as well as all expressions with ε^2 and thus (after some calculations) the condition for $\tau(X_2, Y_2)$ to be positive is equivalent to the inequality

$$xp_0 + (2x-1)p_1 - (1-x)p_2 > 0,$$

which — after taking into account (14) — gives $x > \frac{1+\sqrt{\alpha_2}}{3-\sqrt{\alpha_2}}$, which is true in the case when (13) holds.

Let us define $S_1^{(g)} = X_1^{(g)} + X_1^{(g)}$ and $S_2^{(g)} = X_2^{(g)} + Y_2^{(g)}$. To complete the proof of Theorem 2, it suffices to prove that

$$H_g\big[S_1^{(g)}\big] > H_g\big[S_2^{(g)}\big]. \tag{18}$$

We compute the distribution of $S_1^{(g)}$ as follows:

$$f_1(2) = \Pr\big[S_1^{(g)} > 2\big] = p_2 q_1 = \sqrt{\alpha_2}\frac{\alpha_1}{\sqrt{\alpha_2}} = \alpha_1; \tag{19}$$

$$\begin{aligned} f_1(1) &= \Pr\big[S_1^{(g)} > 1\big] = p_2q_1 + p_1q_1 + p_2q_0 = \alpha_1 + \frac{1-\sqrt{\alpha_2}}{2} \\ &= \frac{\alpha_1}{\sqrt{\alpha_2}} + \sqrt{\alpha_2}\Big(1 - \frac{\alpha_1}{\sqrt{\alpha_2}}\Big) > \sqrt{\alpha_2} > \alpha_2; \end{aligned} \tag{20}$$

$$f_1(0) = \Pr\big[S_1^{(g)} > 0\big] = 1 - p_0q_0 < 1. \tag{21}$$

One finds the following expression for the decumulative distribution function:

$$\overline{F}_{S_1^{(g)}}(t) = \begin{cases} 1 & \text{for } t < 0 \\ f_1(k) & \text{for } k \le t < k+1 \text{ and } k = 0,1,2 \\ 0 & \text{for } t \ge 3 \end{cases}$$

Now using formula (6), we find

$$H_g\big[S_1^{(g)}\big] = g\big(f_1(0)\big) + g\big(f_1(1)\big) + g\big(f_1(2)\big). \tag{22}$$

Analogously, we define values $f_2(k) = \Pr\big[S_2^{(g)} > k\big]$ for $k = 0,1,2$. We get the following identities:

$$f_2(2) = f_1(2) - (1-x)\varepsilon,$$

$$f_2(1) = f_1(1) + \varepsilon,$$

$$f_2(0) = f_1(0) - x\varepsilon.$$

Thus

$$H_g\big[S_2^{(g)}\big] = g\big(f_1(0) - x\varepsilon\big) + g\big(f_1(1) + \varepsilon\big) + g\big(f_1(2) - (1-x)\varepsilon\big). \tag{23}$$

After combining (22) with (23) we see that in order to complete the proof of inequality (18) it suffices to prove that

$$g\big(f_1(2)\big) - g\big(f_1(2) - (1-x)\varepsilon\big) + g\big(f_1(0)\big) - g\big(f_1(0) - x\varepsilon\big) > g\big(f_1(1) + \varepsilon\big) - g\big(f_1(1)\big). \tag{24}$$

Now let us take a closer insight in differences occurring in inequality (24). From the Lagrange Theorem it follows that there exist $0 < \varepsilon_0, \varepsilon_1, \varepsilon_2 < \varepsilon$ such that the following identities hold:

$$g\big(f_1(0)\big) - g\big(f_1(0) - x\varepsilon\big) = x\varepsilon g'\big(f_1(0) - x\varepsilon_0\big) > xg'_-(1)\varepsilon, \tag{25}$$

$$g\big(f_1(1) + \varepsilon\big) - g\big(f_1(0)\big) = \varepsilon g'\big(f_1(1) + \varepsilon_1\big) < g'(\alpha_2)\varepsilon, \tag{26}$$

$$g\big(f_1(2)\big) - g\big(f_1(2) - (1-x)\varepsilon\big) = (1-x)\varepsilon g'\big(f_1(2) - (1-x)\varepsilon_2\big) > (1-x)g'(p_1)\varepsilon. \tag{27}$$

However, from Lemma 1 we find that

$$(1-x)g'(p_1) + xg'_-(1) > g'(p_2). \tag{28}$$

Multiplying both sides of (28) by ε and combining with inequalities (25), (26) and (27), we get the sequence of inequalities:

$$\begin{aligned} g\big(f_1(2)\big) - g\big(f_1(2) - (1-x)\varepsilon\big) \quad &+ \quad g\big(f_1(0)\big) - g\big(f_1(0) - x\varepsilon\big) > \\ &> \quad (1-x)\varepsilon\, g'(\alpha_1) + x\varepsilon\, g'_-(1) > \varepsilon g'(\alpha_2) > g\big(f_1(1) + \varepsilon\big) - g\big(f_1(1)\big), \end{aligned}$$

what completes the proof. ■

Remark 2 *Condition (13) requires an additional comment. We believe that this assumption can be somehow released (compare Darkiewicz et al. (2004)), however for our construction this kind of restriction seems to be necessary. Fortunately a lot of distortion functions encountered* *in practice satisfy this additional limitation. In particular the theorem holds true for all concave piecewise linear functions (e.g. Tail Value-at-Risk admits such representation), because then* $g''(1) = 0$. *At the second extreme we have distortion functions for which the first derivative at* 0 *is infinite and also in this case condition (13) follows automatically. The latter case contains other favorite distortion risk measures, like Proportional Hazard Transform (Wang (1995) and Wang (1996)) or its generalization — a Beta distortion risk measure (Wirch and Hardy (2000)).*

3.3.2 The case of non-concave distortion functions

Intuitively, it is clear that the assumption of concavity of g is somehow critical. However in the proof we use this assumption explicitly. In fact, when one releases the assumption of concavity, the construction follows easily from a general theorem proved by Greco and later also by Schmeidler.

Theorem 3 *Let* BV *be a set of bounded random variables. Suppose that a functional* $H : BV \longrightarrow [0, \infty)$

(i) is additive for comonotonic risks;

(ii) preserves the first order stochastic dominance (i.e. $\forall_t\ F_X(t) \le F_Y(t) \Rightarrow H[X] \le H[Y]$*);*

(iii) satisfies $H[1] = 1$.

Then there exists a distortion function h *such that* $H[X] = H_h[X]$ *for all* $X \in BV$*. Moreover* $H[X+Y] \le H[X] + H[Y]$ *holds for all* X, $Y \in BV$ *if and only if* h *is concave.*

PROOF. See e.g. Dennenberg (1994), Wang (1996). ∎

Consider a distortion risk measure H_g generated by the distortion function g which is not concave. Clearly, H_g obeys (i), (ii) and (iii) in the theorem above and therefore we find the following corollary.

Corollary 1 *Let* H_g *denote a distortion risk measure generated by a distortion function* g *which is not concave. Then there exists a bivariate random variable* (X, Y) *such that* $H_g[X+Y] > H_g[X] + H_g[Y]$.

Now it is straightforward to prove the general theorem.

Theorem 4 *Let g be an arbitrary non-concave distortion function. Then there exist univariate distributions $F_{X^{(g)}}$, $F_{Y^{(g)}}$ and bivariate distributions $\left(X_1^{(g)}, Y_1^{(g)}\right)$, $\left(X_2^{(g)}, Y_2^{(g)}\right)$ belonging to $R\left(F_{X^{(g)}}, F_{Y^{(g)}}\right)$ such that*

(i) $r\left(X_1^{(g)}, Y_1^{(g)}\right) < r\left(X_2^{(g)}, Y_2^{(g)}\right)$;

(ii) $\rho\left(X_1^{(g)}, Y_1^{(g)}\right) < \rho\left(X_2^{(g)}, Y_2^{(g)}\right)$;

(iii) $\tau\left(X_1^{(g)}, Y_1^{(g)}\right) < \tau\left(X_2^{(g)}, Y_2^{(g)}\right)$;

(iv) $H_g\left[X_1^{(g)} + Y_1^{(g)}\right] > H_g\left[X_2^{(g)} + Y_2^{(g)}\right]$.

PROOF. If g is not concave, one finds from Corollary 1 that there exists a random couple (X, Y) such that

$$H_g[X+Y] > H_g[X] + H_g[Y]. \tag{29}$$

On the other hand, for the couple (X^c, Y^c) with the same marginal distributions as the couple (X, Y), but with the comonotonic dependency structure, one has that

$$H_g\left[X^c + Y^c\right] = H_g[X] + H_g[Y]. \tag{30}$$

Combining (29) with (30), one gets

$$H_g[X+Y] > H_g\left[X^c + Y^c\right].$$

However we have that $\mathrm{Var}[X+Y] < \mathrm{Var}\left[X^c + Y^c\right]$ and thus $r(X,Y) < r\left(X^c, Y^c\right)$ (see Dhaene *et al.* (2002a)). The same is true also for Spearman's ρ and Kendall's τ because $\rho = \tau = 1$ holds true only in the comonotonic case.

Hence, taking $\left(X_1^{(g)}, Y_1^{(g)}\right) = (X, Y)$ and $\left(X_2^{(g)}, Y_2^{(g)}\right) = (X^c, Y^c)$ leads to the desired result. ■

3.4 The consistency between distortion risk measures and dependency measures

In this subsection we provide a simple methodology to test the consistency of distortion risk measures of sums of random variables with the order induced by different dependency measures between the summands (in all cases we keep the marginal distributions fixed). We want to emphasize that the test presented here is just a first attempt to test this form of consistency. Our conclusions cannot be interpreted formally because there are no accepted procedures of generating samples from the population of all random distributions. Our methodology is rather subjective and takes into account computational convenience. However it seems to provide quite realistic intuition of the problem.

3.4.1 Description of the methodology

First, we will select 100 000 couples $(X_{1,k}, Y_{1,k})$ in the class of bivariate random variables with support $\{(i, j) \mid i, j = 0, \ldots, 9\}$. For each of the selected couples, we will also consider a random couple $(X_{2,k}, Y_{2,k})$ with the same marginal distributions as $(X_{1,k}, Y_{1,k})$, but of which $X_{2,k}$ and $Y_{2,k}$ are mutually independent. Finally, we will check how many of these couples $(X_{1,k}, Y_{1,k})$ and $(X_{2,k}, Y_{2,k})$ satisfy the following relations:

$$\operatorname{sign}\big(r(X_{1,k}, Y_{1,k}) - r(X_{2,k}, Y_{2,k})\big) = \operatorname{sign}\big(H_g[X_{1,k} + Y_{1,k}] - H_g[X_{2,k} + Y_{2,k}]\big), \tag{31}$$

$$\operatorname{sign}\big(\rho(X_{1,k}, Y_{1,k}) - \rho(X_{2,k}, Y_{2,k})\big) = \operatorname{sign}\big(H_g[X_{1,k} + Y_{1,k}] - H_g[X_{2,k} + Y_{2,k}]\big), \tag{32}$$

$$\operatorname{sign}\big(\tau(X_{1,k}, Y_{1,k}) - \tau(X_{2,k}, Y_{2,k})\big) = \operatorname{sign}\big(H_g[X_{1,k} + Y_{1,k}] - H_g[X_{2,k} + Y_{2,k}]\big). \tag{33}$$

In order to select (the distribution function of) the couple $(X_{1,k}, Y_{1,k})$, we start by generating 99 random numbers $U_{i,k}$ in the interval $(0, 1)$. Let

$$V_{0,k} = 0,$$

$$V_{i,k} = U'_{i,k} \text{ for } i = 1, \ldots, 99,$$

$$V_{100,k} = 1,$$

where $U'_{i,k}$ denotes the i-th order statistic of the sequence $\{U_{i,k}\}$. We consider the differences

$$a_{i,k} = V_{i,k} - V_{i-1,k}$$

for $i = 1, \ldots, 100$. In this way, we get 100 identically distributed random numbers such that

$$a_{1,k} + \ldots + a_{100,k} = 1.$$

Now we define the probability distribution of $(X_{1,k}, Y_{1,k})$ as follows:

$$\Pr[X_{1,k} = i, Y_{1,k} = j] = a_{i+1+10j,k}.$$

Then the marginal distributions of $X_{1,k}$ and $Y_{1,k}$ are given by $\Pr[X_{1,k} = i] = \sum_{j=0}^{9} a_{i+1+10j,k}$ and $\Pr[Y_{1,k} = j] = \sum_{i=0}^{9} a_{i+1+10j,k}$.
The related random couple $(X_{2,k}, Y_{2,k})$ is defined as the independent counterpart of $(X_{1,k}, Y_{1,k})$, hence

$$\Pr[X_{2,k} = i, Y_{2,k} = j] = \Pr[X_{1,k} = i] \Pr[Y_{1,k} = j].$$

Next, we compute Pearson's $r(X_{1,k}, Y_{1,k})$, Spearman's $\rho(X_{1,k}, Y_{1,k})$, Kendall's $\tau(X_{1,k}, Y_{1,k})$ and the considered risk measure of appropriate sums $\left(H_g[X_{1,k} + Y_{1,k}]\ ,\ H_g[X_{2,k} + Y_{2,k}]\right)$. Finally we verify whether the equations (31), (32) and (33) are satisfied (note that all the correlation coefficients for the second independent pair are always equal to 0).
This procedure is repeated for every $k = 1, \ldots, 100000$.
Then, for any particular choice of a distortion risk measure $H_g[\cdot]$ we determine the frequencies

$$r_{g,r} = \frac{N_{g,r}}{100,000}, \quad r_{g,\rho} = \frac{N_{g,\rho}}{100,000}, \quad r_{g,\tau} = \frac{N_{g,\tau}}{100,000},$$

with $N_{g,r}$, $N_{g,\tau}$ and $N_{g,\tau}$ defined as

$$N_{g,r} = \#\Big\{ \big((X_{1k}, Y_{1k}), (X_{2k}, Y_{2k})\big) \mid (31) \text{ holds}\Big),$$

$$N_{g,\rho} = \#\Big\{ \big((X_{1k}, Y_{1k}), (X_{2k}, Y_{2k})\big) \mid (32) \text{ holds}\Big),$$

$$N_{g,\tau} = \#\Big\{ \big((X_{1k}, Y_{1k}), (X_{2k}, Y_{2k})\big) \mid (33) \text{ holds}\Big).$$

We will call $r_{g,\cdot}$ the (Pearson's, Spearman's, Kendall's) correlation consistency coefficient of the risk measure H_g for the particular set of constructed bivariate distributions.

3.4.2 The risk measures under consideration

We performed the procedure described above for the following one-parameter families of distortion functions. Most of these distortion risk measures were introduced in Wang (1996). For each family the parameter p comes from the interval $(0,1)$.

- **The Value at Risk:**

$$g_p(x) = \mathbf{1}_{(\mathbf{p},\mathbf{1}]}(\mathbf{x});$$

- **The Tail Value at Risk:**

$$g_p(x) = \min\Big(\frac{x}{p}, 1\Big);$$

- **The proportional hazard transform:**

$$g_p(x) = x^p; \tag{34}$$

- **The dual-power transform:**

$$g_p(x) = 1 - (1-x)^{\frac{1}{p}};$$

- **Dennensberg's absolute deviation principle:**

$$g_p(x) = \begin{cases} (1+p)x & \text{for } 0 \leq x \leq \frac{1}{2} \\ p + (1-p)x & \text{for } \frac{1}{2} \leq x \leq 1 \end{cases}; \tag{35}$$

- **Gini's principle:**

$$g_p(x) = (1+p)x - px^2; \tag{36}$$

- **The square-root transform:**

$$g_p(x) = \frac{\sqrt{1-\ln(p)x} - 1}{\sqrt{1-\ln(p)} - 1};$$

- **The exponential transform:**

$$g_p(x) = \frac{1-p^x}{1-p};$$

- **The logarithmic transform:**

$$g_p(x) = \frac{\ln\big(1-\ln(p)x\big)}{\ln(1-\ln(p))}.$$

3.4.3 Results and conclusions

In Table 2, Table 3 and Table 4 we present the results respectively for the Pearson's, Spearman's and Kendall's correlation consistency coefficient for different distortion functions g.

From Table 2 we can draw the overall conclusion, that the correlation coefficient is preserved in the majority of cases, for many tested distortion risk measures more frequently than nine times out of ten, for some of them even more than nineteen times out of twenty. Favorite risk measures, such as the Value-at-Risk, the Tail Value-at-Risk and the Proportional Hazard do not perform very well. We also observe that the correlation consistency differs not only between different families of distortion risk measures, but also between different parameters within the same family. In this respect, the dispersion of the correlation consistency seems to be the worst for the Dual-power transform.

Risk measures such as the square root transform, the exponential transform, the logarithmic transform and Gini's principle perform very well. For these distortion risk measures, the Pearson's correlation consistency coefficient does not seem to be very dispersed and tends to increase monotonically together with the parameter p.

The results for Spearman's coefficient differ significantly from the ones obtained for Pearson's

Table 2: The results for Pearson's correlation consistency $r_{\cdot,r}$

	Parameter p						
Risk measure	0.01	0.1	0.25	0.5	0.75	0.9	0.99
Value at Risk	84.25%	93.01%	94.26%	89.00%	75.31%	69.01%	74.45%
Tail Value at Risk	66.98%	71.33%	82.35%	89.58%	82.06%	70.99%	59.02%
PH transform	70.09%	71.69%	74.80%	80.51%	85.56%	88.04%	89.40%
Dual-power	60.05%	77.85%	89.22%	96.86%	93.59%	91.04%	89.72%
Dennenberg	89.58%	89.58%	89.58%	89.58%	89.58%	89.58%	89.58%
Gini	96.86%	96.86%	96.86%	96.86%	96.86%	96.86%	96.86%
Square-root	92.02%	93.98%	95.12%	96.16%	96.73%	96.84%	96.86%
Exponential	86.96%	92.49%	94.80%	96.28%	96.78%	96.84%	96.86%
Logarithmical	89.49%	92.24%	94.01%	95.63%	96.57%	96.84%	96.86%

Table 3: The results for Spearman's correlation consistency $r_{\cdot,\rho}$

	Parameter p						
Risk measure	0.01	0.1	0.25	0.5	0.75	0.9	0.99
Value at Risk	85.80%	89.63%	91.64%	89.01%	77.94%	72.40%	72.77%
Tail Value at Risk	73.74%	67.15%	71.77%	73.75%	71.79%	67.19%	65.82%
PH transform	70.62%	71.41%	72.90%	74.91%	75.87%	76.13%	76.26%
Dual-power	63.84%	71.15%	74.81%	75.78%	76.23%	76.32%	76.31%
Dennenberg	73.75%	73.75%	73.75%	73.75%	73.75%	73.75%	73.75%
Gini	75.78%	75.78%	75.78%	75.78%	75.78%	75.78%	75.78%
Square-root	75.74%	75.82%	75.87%	75.84%	75.79%	75.82%	75.79%
Exponential	74.50%	75.56%	75.78%	75.80%	75.83%	75.81%	75.79%
Logarithmical	75.48%	75.66%	75.80%	75.87%	75.79%	75.82%	75.78%

coefficient. The values are much smaller but also much more stable — all but only few coefficients fall between 70% and 77%. Surprisingly the largest consistency seems to be obtained by the Value at Risk for low values of parameter p - however these risk measures are useless in practical applications. Once again the most stable and relatively large values were obtained for the square

Table 4: The results for Kendall's correlation consistency $r_{\cdot,\tau}$

	Parameter p						
Risk measure	0.01	0.1	0.25	0.5	0.75	0.9	0.99
Value at Risk	84.17%	92.98%	94.23%	88.98%	75.31%	69.07%	74.52%
Tail Value at Risk	66.89%	71.14%	82.08%	89.31%	81.86%	70.73%	58.83%
PH transform	69.88%	71.45%	74.53%	80.15%	85.12%	87.54%	88.87%
Dual-power	59.92%	77.56%	88.83%	95.69%	92.77%	90.41%	89.13%
Dennenberg	89.31%	89.31%	89.31%	89.31%	89.31%	89.31%	89.31%
Gini	95.69%	95.69%	95.69%	95.69%	95.69%	95.69%	95.69%
Square-root	91.43%	93.21%	94.23%	95.08%	95.51%	95.63%	95.68%
Exponential	86.59%	91.91%	93.99%	95.21%	95.56%	95.65%	95.68%
Logarithmical	89.02%	91.66%	93.26%	94.64%	95.40%	95.64%	95.68%

root transform, the exponential transform, the logarithmic transform and Gini's principle. The coefficients for Kendall's τ in Table 4 are very close to those obtained for Pearson's correlation, so the conclusions are analogical.

From the tables it seems that Dennenberg's principle and Gini's principle have very stable correlation consistency coefficients (Pearson's, Spearman's and Kendall's). In our test these coefficients are even identical for all parameters p. This is not accidental, because both risk measures can be expressed as a sum of the expectation and a summand proportional to some dispersion measures independent from the parameter p. We discuss it more comprehensively in Section 3.4.

Interested readers are also referred to Dennenberg (1990).

3.5 Dennenberg's and Gini's principles

In this section we briefly discuss Dennenberg's and Gini's risk measures. They were recommended as premium principles in Dennenberg (1990).

Firstly we take a closer view at Dennenberg's principle. Substituting (35) into (6) we get:

$$\begin{aligned}
H_{g_p}[X] &= \int_0^{\overline{F}_X^{-1}(\frac{1}{2})} \big(p + (1-p)\overline{F}_X(t)\big)dt + \int_{\overline{F}_X^{-1}(\frac{1}{2})}^{\infty} (1+p)\overline{F}_X(t)dt \\
&= \int_0^{\frac{1}{2}} (1+p)\overline{F}_X^{-1}(q)dq + \int_{\frac{1}{2}}^{1} (1-p)\overline{F}_X^{-1}(q)dq = \mathrm{Me}[X] \\
&+ (1+p)\int_0^{\frac{1}{2}} \big(\overline{F}_X^{-1}(q) - \mathrm{Me}[X]\big)dq - (1-p)\int_{\frac{1}{2}}^{1} \big(\mathrm{Me}[X] - \overline{F}_X^{-1}(q)\big)dq \\
&= \mathrm{Me}[X] + \int_0^1 \big(\overline{F}_X^{-1}(q) - \mathrm{Me}[X]\big)dq + p\int_0^1 \big|\overline{F}_X^{-1}(q) - \mathrm{Me}[X]\big|dq = \\
&= \mathrm{E}(X) + p\mathrm{E}|X - \mathrm{Me}[X]|, \qquad (37)
\end{aligned}$$

where $\mathrm{Me}[X]$ denotes the median of a random variable X.

Analogous calculations can be done for Gini's principle. Thus, starting from (36) and (6), we get:

$$
\begin{aligned}
H_{g_p}[X] &= \int_0^\infty \Big((1+p)\overline{F}_X(t) - p\big(\overline{F}_X(t)\big)^2\Big)dt \\
&= \mathrm{E}[X] + p\int_0^\infty \overline{F}_X(t)\big(1-\overline{F}_X(t)\big)dt \\
&= \mathrm{E}[X] + p\int_0^\infty \mathrm{E}\big[(X-t)_+\big]dF_X(t) = \mathrm{E}[X] + p\mathrm{E}\big[(X-Y)_+\big], \qquad (38)
\end{aligned}
$$

where X and Y are independent copies from the same distribution F_X.

Notice that for the special case when $p = 1$, one can write the insurance premium as:

$$
H_{g_1}[X] = \mathrm{E}\big[\max(X, Y)\big],
$$

thus the premium can be understood as the expectation of the greater of the first two claims (assuming independence).

Therefore, both Dennenberg's and Gini's principles can be written in the form of a sum of an expectation and a summand proportional to a specific dispersion measure. It explains why correlation consistencies given in Table 2, Table 3 and Table 4 do not depend on the parameter p for these risk measures.

This representation can be seen as an analogous to the well-known premium principle:

$$
H_\alpha[X] = \mathrm{E}[X] + \alpha\sigma[X],
$$

however the property of preserving stochastic dominance make them much more attractive. Dennenberg's and Gini's risk measures are also computable for a larger class of random variables — one does not need the existence of moments of order higher than one. In some cases also the property of additivity for comonotonic risks which holds for these risk measures may be useful — for premium principles this topic was discussed in Section 2.2.

These risk measures however should not be applied to very heavy tailed distributions. This limitation results from the fact that their respective values are restricted by $2\mathrm{E}[X] + \mathrm{Me}[X]$ and $2\mathrm{E}[X]$, and hence the resulting safety loading may turn out to be too small (sometimes it is even impossible to find a premium which would compensate risk for random variables with very heavy tails). It is however a typical problem for most distortion risk measures. For this reason Wang (1996) postulated to consider one more condition for distortion functions,

namely $g'_+(0) = \infty$. Among all analyzed distortion risk functions, only the Proportional Hazard transform (34) satisfies this additional property.
For risk measures (37) and (38) this problem may be partially solved by extending the range of the parameter p to all positive values. Then Dennenberg's and Gini's premiums will not satisfy the distortion conditions any more (the corresponding function will not be non-decreasing), however all desirable properties will be preserved.

4 Summary

In this paper we investigated how risk measures of sums of risks are related to the level of dependency between the corresponding summands.
In the first part we demonstrated by means of a number of practical examples that it is impossible to find a combination of axioms for risk measures which would hold in all risky situations, no matter what the dependency structure between the risks is. We analyzed different contexts in which risk measures are typically used, such as calculation of premiums, risk sharing schemes, calculation of the solvency margin and an allocation of an economic capital, and related our observations to the coherency axioms.
In the second part we investigated how dependency measures of couples of risks such as Pearson's r, Spearman's ρ and Kendall's τ are related to the ordering generated by distortion risk measures applied to corresponding sums. We found that for distortion risk measures one can construct random couples for which the order is not preserved by neither of the three dependency measures. We also tested the consistency between risk measures generated by some one-parameter families of distortion functions and the coefficients r, ρ and τ. We found that the consistency varies significantly between different risk measures. For Gini's principle for example the level of consistency could be seen as very high and stable.

5 Acknowledgments

Grzegorz Darkiewicz, Jan Dhaene and Marc Goovaerts acknowledge the financial support of the Onderzoeksfonds K.U.Leuven (GOA/02: Actuariële, financiële en statistische aspecten van

afhankelijkheden in verzekerings- en financiële portefeuilles).

References

[1] Artzner, P. (1999). "Application of coherent risk measures to capital requirements in insurance", *North American Actuarial Journal* 3(2), 11-25.

[2] Bühlmann, H. (1970). "Mathematical methods in risk theory". Springer-Verlag, Berlin.

[3] Dennenberg, D. (1990) "Premium Calculation: Why Standard Deviation should be replaced by Absolute Deviation", *ASTIN Bulletin*, vol.20(2), 181-190.

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

[5] 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.

[6] 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.

[7] Dhaene, J.; Vanduffel, S.; Tang, Q.H.; Goovaerts, M.J.; Kaas, R.; Vyncke, D. (2004). "Solvency Capital, Risk Measures and Comonotonicity: a Review", *Journal of Actuarial Practice*, to appear.

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

[9] Embrechts, P.; McNeil, A.K.; Straumann, D. (2002). "Correlation and dependence in risk management: properties and pitfalls", *Risk management: value at risk and beyond* , 176-223, Cambridge University Press, Cambridge.

[10] Gerber, H.U. (1974). "On additive premium calculation principles", *ASTIN Bulletin* 7, 215-222.

[11] Gerber, H.U. (1979). "On additive principles of zero utility". *Insurance: Mathematics & Economics* 4, 249-251.

[12] Gerber, H.U. (1985). "An introduction to mathematical risk theory". *Huebner Foundation Monograph* 8, distributed by Richard D. Erwin, Inc., Homewood, Illinois.

[13] Gerber, H.U.; Goovaerts, M.J. (1981). "On the representation of additive principles of premium calculation", *Scandinavian Actuarial Journal* 4, 221-227.

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

[15] Goovaerts, M.J.; Dhaene, J. (1998). "On the characterization of Wang's class of premium principles", *Transactions of the 26th International Congress of Actuaries* 4, 121-134.

[16] Goovaerts, M.J.; Kaas, R.;Dhaene, J. (2003_a). "Economical capital allocation derived from risk measures",*North American Actuarial Journal*, vol. 7(2).

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

[18] Goovaerts, M.J.; Kaas, R.; Dhaene, J.; Tang, Q. (2003_c). "Some new classes of consistent risk measures". Submitted

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

[20] Hardy, G.H.; Littlewood, J.E.; Polya, G. (1982). "Inequalities", 2nd ed. Cambridge University Press.

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

[22] Kaas, R.; Dhaene, J.; Vyncke, D.; Goovaerts, M.J.; Denuit, M. (2002). "A simple geometric proof that comonotonic risks have the convex-largest sum", *ASTIN Bulletin*, vol.32(1), 71-80.

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

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

[25] 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.

[26] Ramsay, C.M. (1994). "Loading gross premiums for risk without using utility theory, with Discussions", *Transactions of the Society of Actuaries* XLV, 305-349.

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

[28] Schmeidler, D. (1986). "Integral representation without additivity", *Proceedings of the American Mathematical Society* 97, 225-261.

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

[30] Wang, S. (1995). "Insurance pricing and increased limits ratemaking by proportional hazard transforms", *Insurance: Mathematics & Economics* 17, 43-54.

[31] Wang, S. (1996). "Premium calculation by transforming the layer premium density", *ASTIN Bulletin* 26, 71-92.

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

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

[34] Wang, S. and Dhaene, J. (1998). "Comonotonicity, correlation order and premium principles", *Insurance: Mathematics & Economics* 22, 235-242.

[35] 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.

[36] 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.

[37] Yaari, M.E. (1987). "The dual theory of choice under risk", *Econometrica* 55, 95-115.