Economic capital allocation derived from risk measures

M.J. Goovaerts*
R. Kaas†
J. Dhaene‡

June 13, 2002

Abstract

We examine properties of risk measures that can be considered to be in line with some 'best practice' rules in insurance, based on solvency margins. We give ample motivation that all economic aspects related to an insurance portfolio should be considered in the definition of a risk measure. As a consequence, conditions arise for comparison as well as for addition of risk measures. We demonstrate that imposing properties that are generally valid for risk measures, in all possible dependency structures, based on the difference of the risk and the solvency margin, though providing opportunities to derive nice mathematical results, violate best practice rules. We show that so-called coherent risk measures lead to problems. In particular we consider an exponential risk measure related to a discrete ruin model, depending on the initial surplus, the desired ruin probability and the risk distribution.

Keywords: Risk measures, Capital allocation, Insurance premiums, coherence.

*K.U. Leuven, University of Amsterdam
†University of Amsterdam, K.U. Leuven
‡K.U. Leuven, University of Amsterdam

1 Introduction

To evaluate risks, several risk measures can be constructed. Classical examples are insurance premiums serving as risk measures. An advantage of such risk measures is that they are expressed in monetary units. This type of risk measures can often be characterized by requiring a limited number of desirable properties. The interested reader is referred to Goovaerts *et al.* (1984). The properties to be deemed relevant however depend on the special type of risk considered. Indeed a bulk risk such as automobile liability or life insurance requires some particular desirable properties, e.g. concerning the risk measure associated to the sum of random variables. In this case, subadditivity of risk measures for sums of independent risks is required, since due to the law of large numbers, the 'risk' diminishes. But the situation changes drastically in case the risks concerned are explicitly allowed to be dependent. Indeed in pure insurance as well as when constructing a risk-free financial portfolio, the aim is to transform risky situations into risk-free equivalent ones. Hence a pure risk measure necessarily aims at measuring the 'distance' between the risky situation and the corresponding risk free situation. For pure risk measures, a distinction has to be made between one-sided and two-sided distances. Indeed if the realization of the risky situation represents a gain to the insurer compared with the risk-free situation, the question can be posed whether or not this situation should contribute to the distance and hence to the risk measure. In case one has to construct a risk measure that aims at measuring the stability (two-sided distance), the above situations should also incorporate the situations that provide a gain from the point of view of the 'insurer'. Hence a distinction should be made between a two-sided pure risk measure and a one-sided pure risk measure. It is clear that a risk measure should reflect the economic elements to measure risk. Indeed for two identical risk portfolios, the riskiness of the situation depends heavily on the available risk capital. The quantity to be measured clearly is the remaining risk in view of the economic capital, and this has no influence on the expected value of the gains or losses of each of the portfolios other than the cost of the available risk capital. In measuring the risk by means of a risk measure we should think of basic probabilistic quantities such as central tendency, variability, tail behavior and skewness. Some of the so-called desirable properties were studied recently by Artzner (1999), and before that comprehensively in Goovaerts *et al.* (1984) and many other sources, by means of economic arguments, but the economy is not reflected into the acceptable risk measures. An expectation is a suitable risk measure and such is the skewness, but they measure different things.
Let us introduce the following concepts:

Two sided risk measure (TRM): A two-sided risk measure measures the 'distance' between the risky situation and the corresponding risk-free situation when both favorable and unfavorable discrepancies are taken into account.

One sided risk measure (ORM) : A one-sided risk measure measures the 'distance' between the risky situation and the corresponding risk-free situation when only unfavorable discrepancies contribute to the 'risk'.

In case one aims at pricing risks, perhaps TRM's are more relevant, while in risk based capital calculations, ORM's might be more suitable. Some examples of thoroughly studied risk measures are for the ORM-case the classical ruin probability, the stop-loss premiums, the conditional tail expectation and the so-called Esscher premiums, which assign to risk X a loaded premium

$E[Se^{hS}]/E[e^{hS}]$ for some $h > 0$. Examples of TRM's include the variance and the standard deviation, and in general risk measures that respect convexity order. When a risk measure is defined as a probability (an example is the ruin probability) then of course no meaning can be attached to addition. Use of a risk measure without outlining the risk measured makes no sense, indeed, if we are interested in the price of a risk, a central tendency should be measured, to which a loading is added. If on the other hand we want to measure a safety loading, some skewness related property has to be examined. A median is a measure of central tendency, but adding medians only makes sense in special cases. A risk premium (risk measure) calculated solely on the basis of the risk can be equal for two companies pricing a risk for some portfolio of risks but from the view of riskiness, the two situations might be completely different in case the ruin probabilities of the two companies are not comparable.

Consider next a set of $n > 1$ identically distributed but possibly dependent risks with sum $X = X_1 + X_2 + \ldots + X_n$, where each X_i has mean $E[X_i] = \mu$ and variance σ^2. We look at the variance as a TRM. For independent risks $X_i^{\perp}, i = 1, ..., n$ with the same marginal distribution as the X_i, one has $Var[\sum X_i^{\perp}] = n\,Var[X_i]$ and one has additivity. But in case of dependence things are different. For the comonotonic copies X_i^c one gets $Var[\sum X_i^c] = n^2\,Var[X_i]$, and hence $Var[\sum X_i^c] = n\,Var[\sum X_i^{\perp}]$. So the comonotonic sum has a larger variance than the independent sum. It is clear that for the present situation the standard deviation is inadequate if addition is allowed, because $\sigma[\sum X_i] \leq \sigma[\sum X_i^c] = \sum \sigma[X_i]$, the inequality holding with equality only if the X_i are comonotonic to begin with. The equality holds because all X_i are identically distributed, so their comonotonic equivalents X_i^c are in fact identical and have correlation 1. The number $\sigma[\sum X_i^c]$ corresponds to the situation of insured lives with an extremely dependent mortality pattern, while $\sigma[\sum X_i]$ is the standard deviation of the total risk when the lives are independent, e.g., spread all over the world. Insurers and reinsurers who prefer the second situation to the first have yet to be found. But a reinsurer with an enormous free capital (obtained by covering independent risks) could see this subadditivity as a challenge for commercial reasons. Both the variance and the standard deviation rank risks with the same mean in the same way, but they have quite different additivity properties. While the variance exhibits superadditivity for two positively correlated risks, hence $Var[X+Y] > Var[X] + Var[Y]$, the standard deviation is always subadditive, since $\sigma[X+Y] \leq \sigma[X] + \sigma[Y]$ always holds.

The need to add risk measures stems from a quite different problem. Indeed suppose one has to construct a system for dividing the economic capital available for the entire conglomerate between the daughter companies. One wants to allocate the capital in such a way that the sum of the risk measures of each of the daughter companies does not exceed the total risk measure. The main motivation here is the fear of increasing the riskiness of the business by splitting it into different entities because to some extent, one loses the advantage of economies of scale. The measure of the total risk, with its unknown dependencies, should be bounded by the sum of the risks as measured for each daughter company, hence based on marginal data of the different companies. While in reality it is clear that because of dependencies the global risk measure a priori might be larger than the sum of the risk measures for the subcompanies, it seems unrealistic to require subadditivity for risk measures in this content. Restricting oneself to only subadditive pure risk measures is acting with an ostrich attitude.

We might consider a risk measure for the remaining risk, given an attributed economic capital. An expectation is something like a price, measuring the central tendency of a risk, while a variance is a risk measure reflecting the spread. The first one can be derived from the other by an optimization procedure. Indeed minimizing the quadratic distance $E[(X-p)^2]$ between

random variable X and price p leads to a minimal distance equal to $Var[X]$, which is obtained when the price is $p = E[X]$. Our starting point is that a risk measure should take into account the economic situation. Indeed after having distributed the available economic capital there still will be a residual risk which is measured by a risk measure depending on the risk X and the economic capital at hand. The only stringent restriction that is needed is that the total available economic capital has to be equal to the sum of the economic capitals. This is different from the additivity property of the measures themselves. In case addition of risk measures is possible (even without properties in the framework of additivity), the optimal choice can be formulated in a mathematical way as a minimization problem. In the approach of Panjer (2001), the risk measure coincides with the economic capital, which ensures that additivity evolves as a consequence of the defined expectations. The unrealistic ideal might be to have a risk measure that is additive and moreover, invariant for all possible hierarchical subdivisions. Then expectations will be the only answer to the problem. We are a bit reluctant to require this additivity property for the following reason. The problem is related to an heritage problem. As a grandparent, does one want to be fair towards one's childeren or towards the grandchildren? This is a decision problem but it is clear that both criteria give distinct solutions. Translated to the problem of allocation of economic or risk capital the problem is also to some extent strategic. How does one distribute the liabilities in e.g. juridical entities considered to have comparable risk measures? The question is whether at the moment of the further splitting of a subcompany the liability remains with the subcompany, or with the splitted companies. If the liability remains on the level of the original subcompany consisting of two newly formed companies, then of course the risk capital for the other subcompanies at the higher level is unchanged. When the risk attitude towards the former subcompanies and the newly splitted companies, hence supposed at the same hierarchical level, is the same, a reallocation is needed. If one buys a new company, must the economic capital be reallocated? We do not exclude that the risk of the mother company is calculated by the same principle, e.g. an Esscher transform or a stop loss premium. In conclusion: we think that a distinction has to be made between how the capital is allocated and the risk measure on which this allocation is based, and that the risk measure has to be minimized. The optimal solution is the desired capital allocation.
In this paper we will derive an answer to the real problem concerned with the addition of risks. The situation of having bounds for the total risk based on the risk measures for the individual risks is different if one considers a situation with incomplete dependence information or the case of known dependence. This is also related to the level on which risk measures are compared. Indeed the regulators will only use marginal data of the different juridical companies in adding risk measures, while within a conglomerate use can be made of some dependence measures (say correlations). Hence the realistic question is what can be said about the comparison of the sum of risk measures and the risk measure of the sum and how it can provide information for decision makers. In what sense has addition of risks meaning? In contrast with a pure risk measure (compare with measuring a central tendency combined with the spread) which is a TRM one also has to consider a solvency risk measure (compare with measuring a tail property) where a ORM is needed.
In conclusion, we think that a viable risk measure has to be defined for each of the daughter companies, that can be added. This measure of course has to take into account all the risk characteristics, the total claim size X, the premium income π, the economic capital u as well as the corresponding ruin probability ε. This risk measure $\rho(X, \pi, u, \varepsilon)$ might even have an appropriate definition for each daughter company, depending e.g. on the branches and the lines

of business. If the risk measures can be added, the optimal allocation of economic capital is obtained by solving

$$\text{Minimize} \sum_j \rho_j(X_j, \pi_j, u_j, \varepsilon) \quad \text{over} \quad u_1, ..., u_n \text{ with } \sum u_i = u.$$

This result should be compared with $\rho_m(X_1 + ... + X_n, \pi_1 + ... + \pi_n, u_1 + ... + u_n, \varepsilon)$, the risk measure for the mother company.

2 Risk measures

Let us consider first the case of ORM. Because the order $\leq_{st}$is in essence the same as order with probability one, see e.g. Kaas *et al.* (2001, Ch. 10.2), one might impose the requirement that 'good' risk measures should satisfy the following property:

$$X \leq_{st} Y \Longrightarrow \rho(X) \leq \rho(Y) \tag{1}$$

Even if X and Y are regarded as 'net gains', i.e. payments minus premiums, $X \leq_{st} Y$ still does not imply $Var[X] \leq Var[Y]$, nor $\sigma(X) \leq \sigma(Y)$. From this it follows that neither $Var[X]$ nor $\sigma(X)$ is acceptable as a risk measure. This might be true for the ORM case, because a variance is indeed a two-sided measure. But subadditivity is often substantiated by pointing at properties of the standard deviation, because $\sigma(X_1 + X_2) \leq \sigma(X_1) + \sigma(X_2)$, with equality only if the correlation equals $+1$.
Note that $X \leq_{st} Y$ together with $E[X] = E[Y]$ means that $X \sim Y$. Stop-loss order, hence $E[(X-t)_+] \leq E[(Y-t)_+]$ for all real t, together with equal means $E[X] = E[Y]$ is equivalent to $X \leq_{cx} Y$ (convex order), see e.g. Kaas *et al.* (2001, Ch. 10.6). Consequently certainly for ORM, respecting convex order must be an important criterion. In what follows, we will develop arguments indicating that imposing general axioms valid for all risky situations conflicts with generally accepted properties for dealing with particular sets of risks, based on what could be called as 'best practice' rules. We will show that pure risk measures should possess other properties than measures developed for solvency purposes.
Some examples are examined to support this assertion.

Example 1: In earthquake risk insurance, it is better, in the sense that a lower total price is possible, to insure two independent risks than two positively dependent risks, like two buildings in the same area. For insuring both buildings, the premium should be more than twice the premium for insuring only a single building. The exchange of portions of life portfolios between different continents such as one sometimes encounters in practice is another example illustrating the importance of a geographical spread of risks in order to make them more independent. As a consequence, we see that imposing subadditivity $\pi[X+Y] \leq \pi[X] + \pi[Y]$ for all risks (including dependent risks) is not in line with what could be called 'best practice'.
In the framework of risk measures, it is also clear that percentiles or related measures do not catch the risky character of a risk in an economically sensible way. This simply means that when the Value At Risk is used to measure risk, it makes for instance no sense to consider subadditivity. This notion arises from the contamination with the problem of supervision, where the supervisor or a rating agency wants to end up with an upper bound

for the integrated risk of several portfolios. In that situation it would be nice to have a measure for insolvency risk that can be obtained by adding the measures for each of the portfolios, or that this procedure provides an upper bound for it.

Example 2: Consider a combined risk (payments) distributed uniformly on $(0, 1)$. The 90% percentile equals 0.9. By the percentile criterion at the level 10%, this random variable is just as dangerous as the one which is uniformly distributed on $(0, 0.901) \cup (9.901, 10)$. So a tail characteristic like the VAR by itself is not a good risk measure and is not in line with best practice rules. By using the VAR as a criterion, one implicitly assumes that the distributions to be compared are of a similar type, for instance a normal distribution.

It should be noted that the conditional expectation, e.g. above the 90% percentile, does make a distinction between the two situations in Example 2. Example 1 indicates that serious problems may arise from assuming subadditivity. Clearly, subadditivity is not desirable in case dependence aspects of the risks are important. Premium principles satisfying the properties of (sub-)additivity were restricted to independent risks. Risk measures should cope with dependencies as well as with tails.

Example 3: In the case of a translation invariant risk measure the problem of allocation of economic capital does not play a role in judging the safety of a financial conglomerate. Indeed suppose the economic capital is $u = u_1 + u_2 + \ldots + u_n$. Then we have to compare the conglomerate risk measured as $\rho_{congl}(X_1 + \ldots + X_n - u)$ with the sum of the company risks $\rho_1(X_1 - u_1) + \ldots + \rho_n(X_n - u_n)$. Coherence of a risk measures implies translation invariance, so this boils down to comparing $\rho_{congl}(X_1 + \ldots + X_n) - u$ with $\rho_1(X_1) + \ldots + \rho_n(X_n) - u$, meaning that the way in which we allocate the economic capital is irrelevant. Apparently all work done nowadays on capital allocation assumes incoherent risk measures.

Example 4: Also the property of positive homogeneity is not always desirable, because it corresponds to linear utility. In this case, a rational decision maker will not accept that risk is linear function of scale. For a risk neutral decision maker this holds, but it certainly is not valid in insurance practice.

A problem not to be confused with the problem of defining a risk measure for a set of risks consists in the determination of a measure for insolvency risk. This problem originates from a very practical situation where within a financial conglomerate one wants one figure to summarize the risks of a set of different (possibly) dependent subcompanies. The same problem arises in case we consider one financial and/or insurance institution with different portfolios or business lines. Here the final aim is related but different from the aim of determining a risk measure. For each of the separate subcompanies (dependent or not) one can derive a measure for the insolvency risk based on the relevant statistical material that comes from within the subcompany (hence only marginal statistical data are used). Here the question arises whether the sum of the measures of insolvency for the individual subcompanies gives an upper bound of the risk measure for the sum of risks contained in the financial conglomerate. This may resemble the concept of subadditivity but in reality it is not the same. It is a problem of finding the best upper bound for the measure of insolvency of the sum of risks for which we know a measure of insolvency for each of the individual companies (marginally). This is directly related to the following question: if a financial conglomerate has a risk based capital available that amounts to u then how can one

distribute this amount over the subcompanies in such a way that the total measure of insolvency is known, only based on the measures of insolvency risk of each of the separate companies.
We consider some other examples indicating the danger of imposing general properties for measures of insolvency risk.

Example 5: Sometimes bankruptcy risk, where some argue that society requires less capital from a group without firewalls, is introduced to motivate subadditivity of pure risk measures. One should however not confuse risk-free reserves with the notion of a risk capital. Indeed large companies generally have larger free reserves, however large companies will look even more professionally towards their needed risk based capital, because there is a price to be paid for it. As seen in the recent aviation disasters it is clear that it is an advantage to have a large firm broken into subcompanies, meaning that the corresponding risk measures should reflect this. Even bankruptcy in company A immediately gives that no loss is left for the conglomerate which otherwise would have been carried by the global company. This perhaps is a question of corporate governance.

Example 6: Consider a uniform risk X in the interval $(9, 10)$ and compare it with a risk Y that is 20 with certainty. Clearly $\Pr[X < Y] = 1$ but in $X - E[X]$ there might be a risk of insolvency, while $Y - E[Y]$ represents no risk at all. Hence, a risk measure should incorporate a component reflecting the mean of the risk, or any other central tendency characteristic. Suppose a company splits its risk X as $X_I + X_R$ where X_I is the retained risk while X_R is the reinsured part. If one has subadditivity, then $\rho(X_I + X_R) \leq \rho(X_I) + \rho(X_R)$. Because the safety loading for a tail result like $X_R = (X - r)_+$ tends to be relatively high, for the reinsurer's risk measure we often will have $\rho_R(X_R) > \rho(X_R)$, and it follows that no reinsurance will be bought, since it will be considered too expensive. In case $\rho(X_I + X_R) \geq \rho(X_I) + \rho(X_R)$ it is possible that $\rho(X_I + X_R) \geq \rho(X_I) + \rho_R(X_R)$, and these are the reinsurance treaties that exist. There should be a relation between the expected gain and the remaining risk. The expected gain has to be seen in relationship with the change in riskiness.

Example 7: The condition of subadditivity, $\rho(X+Y) \leq \rho(X)+\rho(Y)$, for a translation invariant risk measure can be rewritten as $\rho(X + Y - \rho(Y)) \leq \rho(X)$. This can be interpreted in the following way. Suppose the risk measure is derived from an insurance premium principle. Then adding more risks to an existing portfolio is always advantageous, no matter what the dependency structure between the risks is. Now consider $0 \leq \rho(X) \leq 1$ for a Bernoulli(q) risk. Consider the sum of n such risks which are assumed to be comonotonic. Then the new surplus equals $u + n\,\rho(X)$ with probability $1-q$ and $u + n\,\rho(X) - n$ with probability q. From this it follows that for n large enough, the probability of getting ruined by adding risks equals q, which might be high, and from a solvency point of view, adding risks does not necessarily reduce the risk. Note that translation invariance implies that $\rho(X - \rho(X)) = 0$. Hence, a very insolvent situation might occur, unless $\rho(X) = Max[X]$.

Example 8: It is immediately clear, see e.g. Goovaerts *et al.* (1984), that the risk measures $E[X]$ and $Max[X]$ both satisfy the properties of coherence, see Artzner (1999), hence $\rho(X) \leq \rho(Y)$ if $\Pr[X \leq Y] = 1$, $\rho(aX + b) = a\rho(X) + b$ for all $a \geq 0$ and all b, as well as $\rho(X + Y) \leq \rho(X) + \rho(Y)$ for any X and Y. As premium principles, both are useless in practice, since the latter violates the no rip-off condition and breaks down for unbounded risks like exponential or Pareto risks, while the former does not involve a risk loading, hence leads to ruin with certainty in a ruin process. Also, it fails to makes a distinction between a risk and the increased risk arising when its distribution has been subjected to a ‘mean preserving spread’ in the sense of Rothschild and Stiglitz (1970, 1971).

Example 9: An interpretation of the solvency margin based on an economic reasoning could be obtained as follows. The price of reducing the total risk is the sum of the risk measure of the remaining risk added to the cost of the available capital to be paid to the shareholders. For instance in case one transfers the risk by a stop-loss insurance with a loaded premium, the cost is $(1+\alpha)E[(X-K)_+] + i_D\ K$. In this particular case, the optimal capital, minimizing the total cost, can be shown to be given by $K = F_X^{-1}(1 - i_D/(1+\alpha))$. In this example we get as a risk measure a particular quantile, where the probability is not arbitrary but can be determined from economic parameters.

Example 10: Expanding on the previous example, consider a risk business facing a net loss $X = Y - E[Y]$. The economic capital K is made available to the portfolio by the shareholders at a price of i per unit. To minimize the cost of capital iK, we should take K as small as possible, but to minimize the insolvency risk $E[(X-K)_+]$, the capital K should be large. Just like in the previous example, we have to minimize the total of these two cost components, hence

$$E[(X-K)_+] + iK. \tag{2}$$

Note that a stop-loss premium can be expressed as follows

$$E[(X-K)_+] = \int_K^\infty [1 - F_X(x)]dx. \tag{3}$$

To account for the riskiness of the tail, we use Yaari’s dual theory, introducing a distortion function g with $g(0) = 0$, $g(1) = 1$, $g(x)$ increasing and $g(x) \geq x$. See, e.g., Wang (1996). Then we can compute the ‘cost of avoiding insolvency’ by

$$\int_K^\infty g(1 - F_X(x))dx. \tag{4}$$

Therefore, instead of simply the total cost as above, we minimize the expression

$$\int_K^\infty g(1 - F_X(x))dx + iK. \tag{5}$$

The optimal solution is given by

$$K = F_X^{-1}(1 - g^{-1}(i)). \tag{6}$$

Assuming $i = 0.1$, without distortion, hence with $g(x) \equiv x$, the optimal working capital K equals the 90% percentile of X, but assuming $g(0.01) = 0.1$, it is the 99% percentile. Hence, the optimal threshold depends on i and on the way that we blow up the tail. It should be noted that the percentile is not the risk measure, but in fact it is the value of K corresponding to the (minimized) total cost. Also note that one might incorporate the time-aspect in this model, assuming capital to grow in a year by the risk-free rate r. Note that i must be interpreted as the difference between the cost of capital (seen as a risk premium for a risk capital per unit of capital) and a risk free rate of interest. For the portfolio under consideration the situation at the end of the year reads $\pi(Y) - Y - iK = G$, equal to the gain or loss of the year result of the portfolio. The situation before taxation of the shareholders, only taking into account part of the risk based capital covered, then equals $(i+r)K+G$ in case $Y \leq E[Y]$, $(i+r)K-(Y-E[Y])+G$ in case $E[Y] < Y \leq E[Y]+K$, and $(i+r)K - K + G$ in case $K + E[Y] \leq Y$.

3 A risk measure based on exponential premiums

Consider the discrete time ruin model for insuring a certain portfolio of risks. We have a surplus U_t at time t which increases because of collected premiums c and decreases in the event of claims, the total of which is S_t in year t, leading to the model studied by Bühlmann (1985), see also Kaas *et al.* (2001, Ch. 5):

$$U_t = U_{t-1} + c - S_t \qquad t = 1, 2, \ldots \tag{7}$$

Ruin occurs if $U_t < 0$ for some t. We assume that the annual total claims S_t, $t = 1, 2, \ldots$ are independent and identically distributed, say $S_t \sim S$, with $c - S$ having positive mean but, to make things interesting, also positive probability of being negative. We suppose that S has exponentially bounded tails, hence that the moment generating function exists to the right of the origin. The case of heavy tails could be handled by the same method, using another bound for the ruin probability. The initial question that arises is how large the initial capital $U_0 = u$ and the premium $c = \pi(S)$, that shall be used as a risk measure $\rho(S)$, should be for ruin not to occur with a sufficiently high probability. In e.g. Kaas *et al.* (2001, Ch. 4) it is shown that the probability of ruin is bounded from above by e^{-Ru} where R denotes the so-called adjustment coefficient, which is the positive root of the equation $e^{Rc} = E[e^{RS}]$ or equivalently of $c = \frac{1}{R} \log E[e^{RS}]$. Hence, we get a ruin probability of at most ε when $R = \frac{1}{u} |\log \varepsilon|$. The corresponding premium to be asked is

$$\pi(S) = \rho(S) = c = \frac{1}{R} \log E[e^{RS}]. \tag{8}$$

In the framework of insurance, the premium (8) clearly is an adequate risk measure. Since for the exponential utility functions $-e^{-\alpha w}$ with risk aversion $\alpha > 0$, the utility preserving premium can be shown to be $\frac{1}{\alpha} \log E[e^{\alpha X}]$, the adjustment coefficient can be interpreted as the degree of risk aversion that leads to the actual premium c being in fact the correct exponential premium. Such exponential premiums have been thoroughly studied by actuaries.
This ruin consistent risk measure (8) satisfies the following properties, as can easily be verified.

1. In case S and T are *independent*, one gets $\rho(S + T) = \rho(S) + \rho(T)$;
2. If $S \leq_{cx} T$, then $\rho(S) \leq \rho(T)$;
3. ρ is invariant for a proportional change in monetary units;
4. $\rho(S + T) \leq \rho(S^\bullet + T^\bullet)$ whenever $(S^\bullet, T^\bullet)$ is 'more related' than (S, T), with equality only if (S, T) and $(S^\bullet, T^\bullet)$ have the same joint cdf.

A pair of random variables is defined to be 'more related' than another with the same marginals if the probability of simultaneously obtaining small values, hence the joint cdf, is uniformly larger, see e.g. the textbook Kaas *et al.* (2001, Section 10.6). Comonotonicity is the extreme case when this joint cdf equals an upper bound for it; for a review of its properties and applications, see Dhaene *et al.* (2002a, 2002b). It can be shown that $S + T \leq_{cx} S^\bullet + T^\bullet$ is valid. Translated to the discrete time ruin model above, if the yearly results $S_t^\bullet$ and $T_t^\bullet$ would be PQD (positive quadrant dependent, i.e., more related than in case of independence), the risk, measured as that exponential premium such that the ruin probability with initial capital u has the same bound ε, is larger than when the results are independent. From the first and last properties it follows that $\rho(S^\bullet) + \rho(T^\bullet) \leq \rho(S^\bullet + T^\bullet)$ if $(S^\bullet, T^\bullet)$ is PQD; note that $S^\bullet$ and $T^\bullet$ cannot be degenerate in the discrete time ruin model.

Remark: We have argued that considering the pure variance as a risk measure seems in principle wrong. Taking the initial capital into account is vital. One could argue that by summing $\rho(S)+\rho(T)$ and asking for subadditivity an economic principle is respected, but this is not the case in general, because accepting both risks, seen as marginal risks, might in reality be accepting S^c+T^c and for addition to be meaningful, $\rho(S^c+T^c) = \rho(S)+\rho(T)$ must hold, hence we would have to impose additivity as a desirable property. This problem has been considered by Wang (1996). From this it is certain that, a priori, addition of risk in principle does not make sense. Only after having constructed a risk measure one might verify whether it has nice properties. From the example above one sees that additivity should perhaps be required for comonotonic risks, but then of course it does not hold for risks that are not comonotonic. Any risk measure which is additive for arbitrary pairs of risks attaches the same risk to $S + T$ in all three situations that S and T form a complete hedge (correlation -1), are independent (correlation 0) or are completely dependent (correlation $+1$), which is, to put it mildly, counter-intuitive. Any model that implies such additivity has requirements that are too strict.

3.1 Optimal asset allocation in case of marginal information

The situation that marginal information is available generally occurs when we consider the point of view of the regulating authority. We now have the following problem of allocation of economic capital. Assume that a conglomerate (or insurance regulator) is faced with a total risk $S = S_1+S_2+\ldots+S_n$, having an economic capital $u = u_1+u_2+\ldots+u_n$ to be distributed among the daughter companies. Then the question arises if, even with an exponential principle that is not subadditive, we can determine a subdivision $u_1, \ldots, u_n$ such that the risk measure of the conglomerate is smaller than the sum of risk measures of each of the daughter companies. This has nothing to do with the additivity property of the risk measures themselves since independence is required no longer. From an economic point of view the splitting of the conglomerate in different subcompanies should increase the total amount of risk, because problems arise as soon as one subcompany is ruined, while within the integrated conglomerate compensations between the companies still may avoid ruin. On the other hand the dangerous dependence should be measured too. If a risk measure gives a lower sum of risks for the different daughter companies than for the complete conglomerate, this risk measure should be unacceptable, but this has nothing to do with the mathematical property of subaddidivity of risk measures.
The exponential premium is not subadditive because if $(X^\bullet, Y^\bullet)$ is PQD, the premium for $X^\bullet + Y^\bullet$ is larger than the sum of the individual premiums. Also, it is not superadditive, as can

be seen by looking at any pair $(X^{\bullet\bullet}, Y^{\bullet\bullet})$ which is negative quadrant dependent. Consider the following theorem, a proof of which can be found in Gerber (1979), as well as in Section 5.6 of the textbook Kaas *et al.* (2001).

Theorem: If $\frac{1}{\alpha} = \frac{1}{\alpha_1} + \frac{1}{\alpha_2} + \ldots + \frac{1}{\alpha_n}$, then $\frac{1}{\alpha} \log E[\exp(\alpha \sum_{i=1}^n S_i)] \leq \sum_{i=1}^n \frac{1}{\alpha_i} \log E[e^{\alpha_i S_i}]$.

In the references cited, this theorem serves to prove that given the risk aversions α_i of the daughter companies, the way to redistribute the total risk $S = \sum S_i$ over these daughters in such a way that the lowest total of exponential premiums is needed is to let daughter i carry $\frac{\alpha}{\alpha_i} S$. The optimal economic capital allocation $u_1, ..., u_n$ with $\sum u_i = u$ is the one obtained by minimizing the total premium $\sum_i \frac{1}{\alpha_i} \log E[e^{\alpha_i S_i}]$ needed, where $\alpha_i = \frac{|\log \varepsilon|}{u_i}$ leads to an exponential premium to be asked that ensures ultimate survival of the daughter company with probability $1 - \varepsilon$, see the beginning of this section. From the above theorem, it follows that the total premium quoted is also more than enough to keep the mother company out of ruin with probability $1 - \varepsilon$. The capital allocation problem can be formulated as:

$$\text{Minimize} \sum_{i=1}^n \frac{u_i}{|\log \varepsilon|} \log E\left[\exp\left(\frac{|\log \varepsilon|}{u_i} X_i\right)\right] \text{ over all } u_i \text{ with} \sum u_i = u. \tag{9}$$

The solution can be obtained by means of the Lagrange method. It can easily be shown that, with the notation

$$\rho^i_{exp}(X_i) = \frac{u_i}{|\log \varepsilon|} \log E[e^{\frac{|\log \varepsilon|}{u_i} X_i}]; \quad \rho^i_{Ess}(X_i) = \frac{E[X_i e^{\frac{|\log \varepsilon|}{u_i} X_i}]}{E[e^{\frac{|\log \varepsilon|}{u_i} X_i}]} \tag{10}$$

which are the exponential and Esscher premiums for X_i with parameter $\frac{|\log \varepsilon|}{u_i}$ respectively, the optimal solution satisfies the following system of equations

$$\frac{1}{u_j}\left(\rho^j_{exp}(X_j) - \rho^j_{Ess}(X_j)\right) = \frac{1}{u} \sum_i \left(\rho^i_{exp}(X_i) - \rho^i_{Ess}(X_i)\right). \tag{11}$$

Hence choosing the u_i in this way, the total risk measure of the conglomerate is under control using only the marginal information of the different daughter companies.
Recall that the Esscher premium for S with parameter $h > 0$ equals $E[Se^{hS}]/E[e^{hS}] = \kappa'(h)$, where $\kappa(t) = \log m_S(t)$ is the cumulant generating function. The exponential premium in case of risk aversion h is just $\frac{1}{h}\kappa(h)$. Since $\kappa(t) = E[S]t + Var[S]\frac{t^2}{2!} + O(t^3)$ for small t, for large values of the capitals u_j, hence small values of the parameters $\frac{|\log \varepsilon|}{u_i}$ for the Esscher and the exponential premiums, solution (11) can be written in the following form:

$$\frac{u_j}{u} \approx \frac{Var[X_j]/(2u_j)}{\sum_i Var[X_i]/(2u_i)}. \tag{12}$$

Hence, the optimal capital to allocate to daughter j is approximately proportional to the safety loading contained in the exponential (and Esscher) premium to be asked in case the capital is optimally allocated.

The following pair of inequalities is illuminating: if (X, Y) is comonotonic and $u = u_1 + u_2$, while $\pi[S; u] = \frac{u}{|\log \varepsilon|} \log E\left[\exp\left(\frac{|\log \varepsilon|}{u} S\right)\right]$, we have

$$\pi[X; u] + \pi[Y; u] \leq \pi[X + Y; u] \leq \pi[X; u_1] + \pi[Y; u_2]. \tag{13}$$

The latter comparison makes sense economically; the capital u cannot serve both for covering X and Y. The first inequality is an equality when (X, Y) are independent, and is reversed when (X, Y) are countermonotonic.

3.2 Optimal asset allocation in case of multinormal distributions

We now study the situation of capital allocation as considered by Panjer (2001), where he assumed that the joint distribution of $S_1, S_2, \ldots, S_n$ is multivariate normal with given mean and variance-covariance structure. He bases his reasoning on the 'tail var' for calculating the allocation of capital:

$$u_j = E[S_j \mid S > x_q], \tag{14}$$

where x_q is the q-percentile and $S = \sum S_i$. In many practical situations x_q will not arise as a percentile, but is just a given capital u, from which q follows. This way of distributing the economic capital can be justified by making use of the following conditional risk measure. Let

$$\rho(S) = E[(S - u)_+^2] = qE[(S - u)^2 \mid S > u] \tag{15}$$

Because we consider u as a q-quantile, we have $1 - F_S(u) = q$ and hence by conditioning, a quadratic loss function results. The conditioning is introduced because a quadratic loss function gives a TRM in case the expectation is calculated. While essentially the capital allocation problem should be solved by means of a ORM, the conditioning is essential. The advantage in Panjer's approach consists in the easy way in which (15) can be calculated analytically for the multinormal situation. Using the same conditioning for the daughter companies, the sum of the risk measures then equals (after dividing by q)

$$\sum_i E[(S_i - u_i)^2 \mid S > u] \tag{16}$$

where we take $\sum u_i = u$. Unfortunately we are not able to cope with the danger that arises due to the dependencies in case we add together $S_1 + S_2 + \ldots + S_n$ basing ourselves on the above risk measures, since $\rho(S, u) \leq \sum \rho(S_i, u_i)$ is not necessarily true. In addition, $\rho(S)$ cannot be expressed in monetary units. We are only in the position to optimize (16) with respect to the u_i, which of course makes sense, too. By the method of Lagrange multipliers, optimal $u_1, u_2, \ldots, u_n$ are obtained as the solution of the set of equations:

$$E[S_i - u_i \mid S > u] = \lambda \tag{17}$$

Adding up all these equations we get

$$n\lambda = E[S - u \mid S > u] \tag{18}$$

such that

$$u_j = E[S_j \mid S > u] - \frac{1}{n} E[(S - u) \mid S > u]. \tag{19}$$

This result can also be written as

$$u_j = \frac{u}{n} + \left\{ E[S_j \mid S > u] - \frac{1}{n} E[S \mid S > u] \right\}. \tag{20}$$

We might replace the condition $S > u$ by $S > t$ for $t = F_S^{-1}(q)$ but still requiring $\sum u_i = u$. We get the same solution, only with the conditioning event $S > u$ replaced by $S > t$. Then in the special case that $u = E[S \mid S > t]$, we get Panjer's (2001) result:

$$u_i = E[S_i \mid S > t]. \tag{21}$$

Remarks:

1. In principle the results are based on a quadratic loss function, which is questionable. There are other loss functions that provide a ORM, necessary because economic capital allocation is a one-tail problem. After redistributing the economic capital, the remaining risk is related to the right end tail.

2. The disadvantage of the present situation is that although one minimizes the sum of risk measures based on the marginal risks, the final situation is not a guarantee for having a risk measure for the global conglomerate which is an upper bound of the risk. This has nothing to do with the additivity (or not) property of the risk measure based on loss functions.

3. The advantage of the approach by Panjer, using multinormal $S_1, ..., S_n$, lies in the fact that in the multivariate normal case the dependencies can be taken into account.

4 Risk measures based on convex order

We already noted that when comparing random variables X and Y having the same mean $E[X] = E[Y]$, stochastic ordering is not relevant. As a next step, we consider convex order, where $X \leq_{cx} Y$ if $E[(X-t)_+] \leq E[(Y-t)_+]$ holds for all real t, as well as $E[X] = E[Y]$. As being convex larger denotes increased risk, a desirable property is clearly that $X \leq_{cx} Y \Rightarrow \rho(X) \leq \rho(Y)$. A risk measure is called comonotonicity consistent if

$$\rho(X_1 + ... + X_n) \leq \rho(X_1^c + ... + X_n^c) \tag{22}$$

For example, risk measures that can be written as $E[\phi(X)]$ for some convex function ϕ are comonotonicity consistent, as is any risk measure that respects convex order, hence has the property $X \leq_{cx} Y \Longrightarrow \rho(X) \leq \rho(Y)$.

4.1 Optimal allocation in case of marginal information for a comonotonicity consistent risk measure

Assume that the total solvency risk of a conglomerate $X_1 + X_2 + \ldots + X_n$ with n subcompanies is measured by $E[(X_1 + X_2 + \ldots + X_n - d)_+]$ where in principle all dependencies between the random variables $X_1, X_2, \ldots, X_n$ are possible. This risk measure is used in examining the subcompanies from a global point of view. On the other hand for company j we consider as a risk measure

also $E[(X_j - d)_+]$. In this way, we order risks within the subcompanies also by means of a risk measure respecting comonotone ordering. It is clear that addition of risk measures, hence comparing $E[(\sum_j X_j - d)_+]$ and $\sum_j E[(X_j - d)_+]$, makes no sense here, indeed for a suitable choice of d, $E[(X_j - d)_+]$ could be close to 0 for all j, while $E[(\sum_j X_j - d)_+]$ might be quite large.
For any convex function $\phi(.)$, the risk measures $E[\phi(X)]$, $E[\phi(X - d)]$ and $E[\phi((X - d)_+)]$ are comonotonicity consistent. But the question remains how we can introduce realistic addition of risk measures of this special type. This is done by interpreting d as the economic capital. This, and that is the only importance of talking about subadditivity of risk measures, can be achieved by the capital to be allocated. Indeed the problem that arises is the distribution of u among the companies with amounts u_i such that $u = u_1 + u_2 + \ldots + u_n$. It is clear that

$$E\left(X_1 + X_2 + \ldots + X_n - u\right)_+ \leq \sum_{i=1}^{n} E(X_i - u_i)_+, \tag{23}$$

since the inequality between the random variables on both sides in fact holds with probability one. This indicates that a risk measure, a ORM in particular, has to be related to other economic variables than just the pure risk variable. Indeed the available capital is an important parameter. Now the last inequality has to be approached from two sides. Indeed we can consider:

Problem A:

$$\text{Minimize } \sum_{i=1}^{n} E[(X_i - u_i)_+] \text{ over the } u_i \text{ with } u = \sum u_i \tag{24}$$

Problem B:

$$\text{Maximize } E[(X_1 + X_2 + \ldots + X_n - u)_+] \text{ over the dependency structure} \tag{25}$$

In case of marginal information the maximum has to be taken over all possible dependencies with these marginal distributions (Fréchet space, see e.g. Dhaene *et al.,* 2002a, 2002b), because in case there is no co-operation the only statistics available to the decision maker (the one who allocates the economic capital) are the marginal data. Once again it is clear that a subadditivity property is not a necessary requirement for constructing risk measures. One should construct risk measures that deal with the risk (such as ruin probability, cost of having insufficient economic capital). From the results on comonotonic risks, see Kaas *et al.* (2001, Section 10.6), it immediately follows that the solutions of the problem of maximizing the conglomerate risk and minimizing the sum of the risks of the daughter companies, both give the same value of the problems

$$E[(X_1^c + X_2^c + \ldots + X_n^c - u)_+] = \sum_j E[(X_j - F_{X_j}^{-1}(F_W(u))_+], \tag{26}$$

where $F_W^{-1}(u) = \sum_j F_{X_j}^{-1}(u)$. While the expectation of any convex function gives a risk measure which historically has shown its value, this shows that addition of utilities as an axiom is not necessarily a realistic manipulation. In the class of risk measures based on the expectations of convex functions $E[\phi(X)]$, it can be shown that the only choice of ϕ that leads to an equality of the optimal values of both problems A and B is of stop-loss type. This result can be found in Goovaerts *et al.* (2001, Theorem 5).

4.2 Optimal asset allocation for multivariate normal distributions

Suppose now we consider again the situation of Section 3.2 this time using as a risk measure $\rho(S) = E[(S-u)_+]$, where $S = S_1 + S_2 + \ldots + S_n$. Because the distribution of S is known in the allocation of capital for companies who cooperate as well as provide statistics describing dependencies, $\rho(S)$ is cast into the form $\rho(S,u) = E[(S-u)I_{S>u}]$, where $I_b = 1$ if the boolean expression b is true, 0 otherwise. For each of the daughter companies we have

$$\rho(S_j,u_j) = E[(S_j-u_j)_+] = E[(S_j-u_j)_+ I_{S_j>u_j}] \geq E[(S_j-u_j)_+ I_{S>u}] =: \rho_{inf}(S_j,u_j) \quad (27)$$

It follows from this that, because we have information concerning the dependency structure between the daughter companies of the conglomerate, the risk measure to be considered is $\rho_{inf}(S_j,u_j)$ [to be interpreted as $\rho(\cdot,\cdot)$ with additional information] which has the same structure as the risk measure of the conglomerate, of course with an adapted economic capital. As before, we have with probability one

$$\Big(\sum S_j - u\Big)_+ \leq \sum (S_j-u_j)_+. \quad (28)$$

Multiplying on both sides by $I_{S>u}$ and taking expectations, we get

$$E[(\sum S_j - u)_+ I_{S>u}] \leq \sum E[(S_j-u_j)_+ I_{S>u}]. \quad (29)$$

To get the best upper bound, we have to solve the following problem in the framework of an optimal allocation solution:

$$\text{Minimize} \sum E[(S_j-u_j)_+ \mid S>u] \text{ such that } \sum u_j = u \quad (30)$$

By means of a Lagrange multiplier method one obtains that for each j,

$$\Pr[S_j > u_j \mid S>u] = \lambda. \quad (31)$$

The joint distribution of S_j and S can easily be obtained because, writing $\rho_j = \rho(S_j,S)$ and $\sigma_j^2 = Var[S_j]$,

$$\begin{aligned} S_j \quad | \quad S = s &\sim N(\mu_j + \rho_j \frac{\sigma_j}{\sigma_S}(s-\mu_S), \sigma_j\sqrt{1-\rho_j^2}) \\ S \quad &\sim \quad N(\mu_S, \sigma_S) \end{aligned} \quad (32)$$

Hence the values $u_1, \ldots, u_n$ have to be determined as roots of the following system of equations

$$\int_0^\infty (1 - F_{S_j|S}(u_j,x) f_S(x))dx = \lambda(1 - F_{S_j}(u)) \quad (33)$$

with λ chosen such that $u_1 + u_2 + \ldots + u_n = u$.

4.3 Conclusions

In this paper it is argued that the combination of so-called desirable properties for risk measures or insurance premium principles that have to hold for *all* situations, hence for all types of dependencies simultaneously, often violates what could be called best practice, and sometimes leads to inconsistencies. In addition it is shown that risk measures should be added only if this is sensible, which it is not for instance when premiums for the same risks are compared that are quoted by companies with different strategic goals, for instance because they admit very different ruin probabilities. As an example we consider a risk measure describing the economic risk, taking into account economic factors as well as the contingent claims, and not depending on the difference between the economic capital and the risk variable. This risk measure provides a nice tool for determining optimal capital allocation. The relevance of some other capital allocation proposals is investigated. The importance of considering the remaining risk after allocating the economic capital is stressed.

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. (1985). "Premium calculation from top-down", *ASTIN Bulletin*, 15, 89-101.

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

[4] Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. & Vyncke, D. (2002b). "The concept of comonotonicity in actuarial science and finance: Applications", *Insurance: Mathematics and Economics*, forthcoming.

[5] Gerber, H.U. (1979). "An introduction to mathematical risk theory", Huebner Foundation Monograph 8, distributed by Richard D. Irwin, Homewood Illinois.

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

[7] Goovaerts, M.J., Dhaene, J. & Kaas, R. (2001). "Risk Measures, Measures for Insolvency Risk and Economic Capital Allocation", *Tijdschrift voor Economie en Management* XLVI, 4, 545-559.

[8] Kaas, R.; Goovaerts, M.J.; Dhaene, J.; Denuit, M. (2001). "Modern Actuarial Risk Theory", Kluwer Academic Publishers, Boston, Dordrecht, London.

[9] Panjer, H. H. (2001). "Measurement of risk, solvency requirements and allocation of capital within financial conglomerates", *Research Report 01-15*, Institute of Insurance and Pension Research, University of Waterloo.

[10] Rothschild, M. & Stiglitz, J.E. (1970). "Increasing risk: I. A Definition", *Journal of Economic Theory* 2, 225-243.

[11] Rothschild, M. & Stiglitz, J.E. (1971). "Increasing risk II: Its Economic Consequences", *Journal of Economic Theory* 3, 66-84.

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