MARGINAL DECOMPOSITION OF RISK MEASURES

BY

GARY G. VENTER, JOHN A. MAJOR AND RODNEY E. KREPS

ABSTRACT

The marginal approach to risk and return analysis compares the marginal return from a business decision to the marginal risk imposed. Allocation distributes the total company risk to business units and compares the profit/risk ratio of the units. These approaches coincide when the allocation actually assigns the marginal risk to each business unit, i.e., when the marginal impacts add up to the total risk measure. This is possible for one class of risk measures (scalable measures) under the assumption of homogeneous growth and by a subclass (transformed probability measures) otherwise. For homogeneous growth, the allocation of scalable measures can be accomplished by the directional derivative. The first well known additive marginal allocations were the Myers-Read method from Myers and Read (2001) and co-Tail Value at Risk, discussed in Tasche (2000). Now we see that there are many others, which allows the choice of risk measure to be based on economic meaning rather than the availability of an allocation method. We prefer the term "decomposition" to "allocation" here because of the use of the method of co-measures, which quantifies the component composition of a risk measure rather than allocating it proportionally to something.

Risk adjusted profitability calculations that do not rely on capital allocation still may involve decomposition of risk measures. Such a case is discussed. Calculation issues for directional derivatives are also explored.

MARGINAL DECOMPOSITION OF RISK MEASURES

Insurers are finding it increasingly useful to have risk measures to quantify company risk and its components. Understanding risk sources can help in performance measurement, strategic planning, pricing, and communication with regulators, rating agencies, and securities analysts. We focus here on decomposing risk measures by the business components that generate the risk. One possible application is allocation of capital, which we discuss but do not particularly advocate. Other ways to compute business component performance without allocating capital can also benefit from being able to quantify the contributions to company risk.

Sometimes a selected risk measure gets the label "economic capital." For instance, if a company has ruin probability of 0.0189%, it might decide that Value

at Risk for the negative of earnings at probability level 0.02% is a relevant risk measure, and call that economic capital. Then allocation of economic capital is purely an issue of allocation of that risk measure. The methods below show how this can often be done in a marginal additive fashion – that is the marginal impacts of the components on the company risk measure add up to the risk measure itself. Alternatively, allocating the Value at Risk at probability 0.0189% would allocate actual capital. The term "economic capital" is a bit misleading, as the fact that the discussion is about a risk measure can get overshadowed. Other risk measures could also be used in this context. For instance, if actual capital is 3.8 times TVaR of negative earnings at 10%, 3.5 times TVaR could be called economic capital. Again this is a risk measure and can be allocated using the methods below. However the same approach would work for actual capital, which may be more meaningful. We tend to use the term "decomposition" rather than "allocation" for the methods we develop below, as they look at how the risk is generated within the business units rather than assigning risk to units proportionally to some measure.

Decomposition of risk measures starts with a risk pool (sum of random variables) $Y = \Sigma X_j$ and a risk measure $\rho(Y)$ that expresses some aspect of the risk inherent in Y. The composition of this by the components X_j is a function $r(X)$ with $\rho(Y) = \Sigma r(X_j)$. For an insurer, Y could be total claims, or it could be some other account, such as change in capital, underwriting profit, etc. Taking Y as the negative of earnings is often useful, as then large values are bad and typical right-tail analysis is applicable. Some risk measures make more sense with a specific choice of Y, but others are fairly general. The components could be lines of business, or they could be marketing units, like all lines sold to dairy farmers, or geographic, like all business in urban areas, etc. Y is assumed here to be continuous, which for a pool of many individual losses is usually a realistic approximation.

Section 1 describes the properties we focus on: decomposition and marginal impact. Marginal decomposition is easier if a company can make homogeneous changes to its business. This is discussed in section 2. Section 3 describes the decomposition methods used: co-measures and the directional derivative, and their application in the case of homogeneous changes. Section 4 provides a few examples to illustrate the approach. Section 5 has some additional examples that all use risk measures related to insolvency. These may not be practical for many companies in that insolvency is difficult to quantify, so this section may be of more academic interest. Section 6 considers the non-homogeneous case, where growth is accomplished by adding discrete exposure units that change the shape of the distribution. The co-measure decomposition is not always marginal in this case, and marginal impacts are less likely to add up to the entire risk measure. This is different for different risk measures: for some the decomposition is far from marginal, for some it is close, and some have a marginal decomposition even in this case. Section 7 considers performance measurement without allocating capital, and Section 8 concludes. The appendices address calculation and derivation issues.

1. MARGINAL DECOMPOSITION OF A RISK MEASURE

If a risk measure for a company can be expressed as a sum over the business units of related business unit risk measures, those are considered a decomposition of the risk measure to the business units. For instance, the covariances of X_j with Y sum over the j's to the variance of Y, so the covariances are a decomposition of the variance. The variances of the X_j's do not sum to the variance of Y unless the X_j's are all independent, so allocation of variance by the variances of the components is not a decomposition absent independence, even though it can be forced to add up with an off-balance factor.

Marginal impact can be incremental or total, that is the impact could be that of a small change or of the entire business component. Here we adopt the incremental approach. That is, the marginal impact of a business component on a companywide risk measure is the proportional change in the overall risk measure due to an incremental change in the volume of the business unit.

A decomposition of a risk measure is marginal if each business component is assigned its marginal impact. The covariance decomposition of variance is not marginal, as it can be shown that an incremental change in a component increases the variance proportionally by double the covariance of the component with the total.

If capital is allocated by a risk measure in order to calculate risk-adjusted return by line (ratio of profit to allocated capital, which is equivalent to ratio of profit to allocated risk), you would like to be able to conclude that growing a line with a higher-than-average ratio will increase the ratio for the whole company. Growing any line is likely to increase the overall risk. But under marginal decomposition, the increment in company risk is the growth increment for the line times the risk allocated to the line. If that line has a higher than average return on risk, the average return on risk for the whole company will increase, as desired. This does not mean that the incremental risk from growing a line is all assigned to that line under marginal decomposition. In general, it is not. Growing a line, even incrementally, reconfigures the decomposition. Appendix 4 addresses this.

2. HOMOGENEOUS GROWTH

Growth of a business component is classified as homogeneous if the unit grows by a scalar factor $1 + a$. That is, $G(x)$, the probability distribution of X after the growth, is related to $F(x)$, the probability distribution before the growth, by $G(x + ax) = F(x)$. Then the probability of an overall loss to the component of x or less before the growth is the same as a loss of $x + ax$ or less after the growth. Here a can be positive or negative, and a reduction is called negative growth.

Homogeneous growth can occur by simple inflation to the component. Or if the component has quota share reinsurance, reducing the share ceded can produce positive homogeneous growth. Buying a new quota share can produce negative homogeneous growth.

However if a company aims to have positive growth it can often do so only by writing new business, which does not always produce a scalar change. The strongest results below are for homogenous growth, which is assumed for now. Some of the results are expressed with derivatives, which for convenience are defined by adding εX. They could equally well be expressed by subtracting εX, which can always be achieved by increasing a quota share. So marginal decomposition of the existing risk by derivatives is well defined for any insurer, but applying it to a positive growth strategy may require consideration of the non-homogeneous case as well.

3. DEFINITION OF METHODS

Two methods for determining marginal decompositions are co-measures and the directional derivative. Both of these are used to specify components' contributions to risk measures. If a selected risk measure quantifies something meaningful to the whole company, then management may like to know how much each business component contributes to the risk measure. For this exercise the risk measure's value for each component on a standalone basis, or for the company less the component, is not necessarily relevant. The focus is on the impact the component has on the total company risk measure, not the component's standalone risk. Properties of risk measures such as coherence are not needed for decomposition by co-measures. For example, *VaR*, which is not a coherent measure, is shown in section 4 to have a marginal decomposition.

3.1. Co-measures

Co-measures define $r(X)$ when $\rho(Y)$ is expressed as a conditional expected value:

$$\rho(Y) = E[\Sigma_i h_i(Y) L_i(Y) \,|\, \text{condition on } Y],$$

where the h_i's are additive functions, i.e., $h(V + W) = h(V) + h(W)$, and the L_i's are any functions for which this conditional expected value exists. Usually only one h and one L are needed. It turns out that many risk measures can be expressed in this form. The co-measure for component j for such a risk measure is the same formula but with Y replaced by X_j in the argument of the h functions, which comprise the linear part of the formula.

That is, the co-measure r is defined by:

$$r(X_j) = E[\Sigma_i h_i(X_j) L_i(Y) \,|\, \text{condition on } Y],$$

By the additivity of the h's this satisfies $\rho(Y) = \Sigma r(X_j)$.

As an example of a co-measure, excess tail value at risk ($XTVaR$) excess of level b, can be defined as:

$$\rho(Y) = E[(Y - EY) \mid Y > b]$$

Now $h(X)$ is $X - EX$, $L(Y) = 1$, the condition is $Y > b$, and $r(X_j) = E[(X_j - EX_j) \mid Y > b]$.

There may be more than one set of conditions and h and L functions that define the same risk measure, each one leading to a different co-measure. In fact, given a risk measure defined by L and h functions L_1, h_1, and another additive function h_2, setting $L_2 = L_1 h_1 / h_2$ gives $L_2 h_2 = L_1 h_1$, which defines the same risk measure. Also note that making the definition a conditional expectation is for convenience only. It could be simply an expectation, with the conditioning being done with indicator functions put into $L(Y)$.

3.2. Directional Derivative

The marginal impact of the jth business component can be formalized by the jth directional derivative defined as:

$r(X_j) = \lim_{\varepsilon \to 0} \dfrac{\rho(Y + \varepsilon X_j) - \rho(Y)}{\varepsilon}$. As mentioned above, ε might be negative.

Euler showed (see Appendix 2 for details) that this derivative produces a decomposition of the risk measure if ρ is scalable[1], i.e., $\rho[aY] = a\rho[Y]$. That is, for a scalable function ρ of several variables, the sum of its derivatives *wrt* each of the variables is ρ itself. Thus for scalable risk measures in the case of homogeneous growth, additive marginal decomposition is produced by the directional derivative.

Homogeneous growth is required here for εX_j to be defined in the limit. If growth is by adding in exposure units, and ΔX_j is the smallest unit you can add to component j, then the incremental proportional impact is $[\rho(Y + \Delta X_j) - \rho(y)]/\Delta$.

The strategic decision of growing a unit that has higher-than-average profit / risk can be more formally treated in this framework. Say P is the overall expected profit and P_j is that of the jth unit, and suppose $P/\rho(Y) < P_j/r(X_j)$, so $Pr(X_j) < P_j\rho(Y)$. Multiplying this by ε and adding $P\rho(Y)$ gives $P\varepsilon r(X_j) + P\rho(Y) < \varepsilon P_j\rho(Y) + P\rho(Y)$. By the definition of r, the left hand side is $P\rho(Y + \varepsilon X_j)$ in the limit, so $P/\rho(Y) < [P + \varepsilon P_j]/\rho(Y + \varepsilon X_j)$. Thus comparative to risk, the total company is more profitable with the incremental business.

Our approach is to find marginal decompositions of scalable risk measures by the directional derivative. (Only one co-measure is marginal because derivatives

---

[1] This is often called homogeneous of degree 1, but to avoid confusion with homogeneous growth the term "scalable" is used here.

are unique.) This co-measure is still a decomposition under non-homogeneous growth, and is often still close to marginal, as discussed further in Section 6.

It may help to use L'Hopital's rule to calculate the limit defining $r(X)$. Taking the derivative of the numerator and denominator *wrt* ε gives $r(X_i) = \rho'(Y + \varepsilon X_j)|_0$, where here the prime denotes the derivative *wrt* ε. Although this simplifies the calculation of the derivative, it can still get complicated. The appendices discuss methodology for taking the derivatives. For the most part derivatives in the text will be stated but not derived. The math might seem clearer if Y is considered a weighted sum of the X_j's, with weights w_j that reflect the amount of risk X_j taken. These could be 100% for every j if there is not a quota share in place. Then the derivatives can be considered to be *wrt* the weights, which are real numbers and in theory can change up or down by changing the quota share percentages, even above 100%. This approach is taken in Appendix 2 in order to apply Euler's Theorem.

3.3. *TVaR* Example

For *XTVaR*, $\rho(Y) = E[(Y - EY)|Y > b]$ is not scalable if b is a fixed constant amount. But if b is a fixed percentile of Y, say the αth percentile, then it is. Multiplying Y by a constant increases EY and every percentile of Y by the same factor. You can write *XTVaR* as

$\rho(Y) = E[(Y - EY)|F(Y) > 1 - \alpha]$. Then
$\rho(Y + \varepsilon X_i) = E[(Y + \varepsilon X_j - EY - \varepsilon EX_j)|F(Y + \varepsilon X_j) > 1 - \alpha]$ and
$\rho'(Y + \varepsilon X_j)|_0 = E[X_j - EX_j|F(Y) > 1 - \alpha]$, which is the co-measure.

The derivative is not obvious. Details are in Appendix 1.

One possibility for establishing a cutoff probability for tail risk measures would be to use the probability of having any loss of capital at all. Then *XTVaR* would be the average loss of capital when there is a loss of capital. Another possible choice is the probability that capital is exhausted. The former is arguably more relevant to capital allocation, in that it charges for any use of capital rather than focusing on the shortfalls upon its depletion. To apply it, suppose that total capital is 11 times the average loss of capital when there is any capital loss at all. Then allocation by loss of capital would preserve this ratio for all components. That is, each unit would get 11 times its average draw on capital in those cases where total capital is reduced.

On the other hand, policyholders tend to be sensitive to impairment or default. Studies[2] suggest that they demand premium reductions one or two orders of magnitude greater than the expected value of the default cost in order to accept less than certain recovery. This is in part due to undiversified

---

[2] E.g., Phillips, Cummins and Allen (1998).

purchases of insurance. Thus the value of default has meaningful pricing effects, and policyholder concerns become quite relevant to shareholders as well.

4. EXAMPLES OF MARGINAL DECOMPOSITION OF RISK MEASURES

This section illustrates the approach with marginal decompositions for *VaR*, standard deviation, an exponential moment measure, and spectral measures.

4.1. Variance and VaR

For some risk measures there is a natural definition as a conditional expected value. (Unconditional is a special case of conditional with a condition that always holds.) For these the co-measure is clear. For instance variance and value-at-risk work quite easily:

$$\rho(Y) = Variance(Y) = E[(Y - EY)^2], \text{ so } L(Y) = h(Y) = Y - EY$$
$$r(X_j) = Cov(X_j, Y) = E[(X - EX_j)(Y - EY)]$$

and

$$\rho(Y) = VaR_\alpha(Y) = E[Y | F(Y) = 1 - \alpha]$$
$$r(X_j) = Co\text{-}VaR_\alpha(X_j, Y) = E[X_j | F(Y) = 1 - \alpha]$$

In both cases $r(X)$ can be plausibly interpreted as the contribution of X to $\rho(Y)$. Variance does not meet the scalability criterion for decomposition by directional derivative, as $Variance(aY) = a^2 Variance(Y)$, but value at risk does and gives the same decomposition as above.

4.2. Standard Deviation

When there are alternative intuitively reasonable definitions of a risk measure as a conditional expected value, then the derivative could help determine the preferred decomposition. For instance, there are different ways to use co-measures to express standard deviation. If you take $h(X) = X$ and $L(Y) = Std(Y)/EY$, with the condition $Y = Y$, you get:

$$\rho(Y) = E[YStd(Y)/EY] = Std(Y)$$
Then $r(X_j) = E[X_j Std(Y)/EY] = Std(Y)EX_j/EY$

This just spreads the standard deviation in proportional to the mean of the components. Alternatively, you could take $h(X) = X - EX$ and $L(Y) = (Y - EY)/Std(Y)$. Then:

$$\rho(Y) = E[(Y - EY)^2/Std(Y)] = Std(Y) \text{ and}$$
$$r(X_j) = Cov(X_j, Y)/Std(Y)$$

This decomposes the standard deviation in proportion to the covariance of the component with the total.

The standard deviation is scalable so there should be a marginal decomposition. Taking the derivative of $Std(Y + \varepsilon X_j) = [Var\,Y + 2\varepsilon\,Cov(X_j, Y) + \varepsilon^2 Var(X_j)]^{\frac{1}{2}}$ gives at $\varepsilon = 0$

$$r(X_j) = Cov(X_j, Y)/Std(Y)$$

This agrees with the second form of the co-measure, so it provides h and L functions that lead to marginal decomposition. Thus the total change in $Std(Y)$ brought about by a small change in X_j can be attributed to j by this procedure.

4.3. An Exponential Moment Risk Measure

Quadratic risk measures like standard deviation or semi-variance do not appear very good at capturing market aversion to extreme loss risk. Tail risk measures can do this but require a somewhat arbitrary choice of cutoff. Even then, weight functions that add more weight to the tail probabilities seem to better capture risk preferences than do the usual tail measures (see Venter, Barnett and Owen (2004) for instance). Another useful alternative is the exponential moment, when it exists.

For example, let $\rho(Y) = E(Ye^{cY/EY})$. This is scalable as $\rho(aY) = a\rho(Y)$. Thus it should have a marginal decomposition. The simplest co-measure is $r_1(X_j) = E(X_j e^{cY/EY})$. Although these add up to $\rho(Y)$, this is not a marginal decomposition. Taking the directional derivative (straightforward if messy) yields the marginal decomposition:

$$r(X_j) = r_1(X_j) + c(EX_j/EY)\,E[Ye^{cY/EY}(X_j/EX_j - Y/EY)]$$

Without the excess ratio factor $(X_j/EX_j - Y/EY)$ the second term is an allocation of $c\rho(Y)$ by the ratio of means EX_j/EY. The $c\rho(Y)$ term is dominated by the large values of Y. When Y is large, the components of the company that are contributing most to the large losses would have $X_j/EX_j > Y/EY$, so the excess ratio factor gives them an increase in allocation. The other components would have a decrease in allocation.

To express the marginal decomposition as a co-measure, set $h_1(Y) = Y$, $L_1(Y) = e^{cY/EY} + cYe^{cY/EY}/EY$, $h_2(Y) = -EY$, and $L_2(Y) = cY^2 e^{cY/EY}/(EY)^2$. Evaluating the h's at Y for the risk measure and at X_j for the co-measure then reproduces the decomposition:

$$\rho(Y) = E(Ye^{cY/EY}) + E[cY^2 e^{cY/EY}/EY] - E[cY^2 e^{cY/EY}/EY] \text{ and}$$
$$r(X_j) = E(X_j e^{cY/EY}) + cE(X_j Ye^{cY/EY})/EY - cEX_j E[Y^2 e^{cY/EY}]/(EY)^2$$
$$= E(X_j e^{cY/EY}) + c(EX_j/EY)\,E[Ye^{cY/EY}(X_j/EX_j - Y/EY)]$$

Marginal co-measures of risk measures that are functions of Y/EY are discussed in greater generality in the final section of Appendix 1.

This risk measure emphasizes large loss risk without requiring the selection of a tail cutoff point. Calibration of c can be done by setting the risk measure equal to capital or a fraction of capital, or by setting the random variable of interest minus its mean to a reasonable fraction of capital. This can be done with either claims or negative profit for Y. With Y = claims, r_1 is always positive. With Y = negative profit, a line that has a profit when the company as a whole has large losses could be assigned a negative contribution. That is reasonable but may not meet all needs for allocated capital. There is some evidence (e.g., Venter, Barnett, Owen (2004)) that the market value of risk transfer requires exponential transforms at least in the far right tail, so this measure may be related to the market value of the risk it is measuring. There is theoretical support for such transforms as well, e.g., in Ballotta (2004) and Møller (2003). If the risk measure is going to be used to express risk-adjusted profit, then having it proportional to the market value of the risk would be useful.

4.4. Spectral Measures

A generalization of $TVaR$ is obtained if one considers $\rho = E[Y \cdot \eta(F(Y))]$ for nonnegative scalar functions η. These are known as spectral measures. If η is given by

$$\eta(p) = \begin{cases} 0, & p \leq 1 - q \\ 1/q, & 1 - q < p \end{cases}$$

then you get $TVaR$ with tail probability q. If η is an arbitrary step function, then you get linear combinations of $TVaR$. For example,

$$\eta(p) = \begin{cases} 0, & p \leq 1 - q \\ 1/(q - r), & 1 - q < p \leq 1 - r \\ 0, & 1 - r < p \end{cases}$$

results in a measure $\rho = (q/(q-r))TVaR(q) - (r/(q-r))TVaR(r)$. This might be referred to as the risk of a "managed layer." This measure is not coherent because it ignores the risk in the upper r tail, and therefore fails subadditivity. For example, (following Artzner et. al. (1999)), consider two independent random variables on the interval $[-1, 2]$ with uniform density 0.9 from -1 to 0 and density 0.05 from 0 to 2. Evaluated with $q = 0.1$ and $r = 0.08$, they each have associated $\rho = 0.2$, but their sum random variable has $\rho = 0.638$, more than the sum of their individual ρ values.

Another spectral measure is given by:

$$\eta(p) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2}\left(\frac{p - (1 - q)}{\sigma}\right)^2\right).$$

This can be interpreted as defining a "blurred" *VaR*. It, like *VaR*, is not coherent, failing subadditivity. In the limit, as $\sigma \to 0$, the η function becomes a Dirac delta-function at $p = 1 - q$ and the risk measure ρ becomes the *VaR*.

In general, the co-measure is $r_j = E\lfloor X_j \cdot \eta(F(Y)) \rfloor$. If η is a step function, following logic similar to that used to derive the result for *XTVaR* with fixed percentage excess, this co-measure is seen to be the derivative. For smooth η, it can also be shown that this is the derivative. If η is nondecreasing, then this risk measure family consists of all coherent, law-invariant,[3] comonotonic additive[4] risk measures (Tasche 2002, Acerbi 2002, Kusuoka 2001). Examples above of η functions that are not nondecreasing show that marginal decompositions still exist for spectral measures that are not coherent.

5. INSOLVENCY RELATED RISK MEASURES

This section gives some examples of marginal decomposition for risk measures related to insolvency. Risk measures associated with insolvency are of theoretical interest in that insolvency is an economically meaningful event. However they may not be practical in that the probability of insolvency is difficult to measure in practice. A few such measures (formulations of expected policyholder deficit, the default put option, and an analogue of the Myers-Read approach) are reviewed here for their historical and theoretical interest. Efforts at quantification appear worth continuing in order to bridge theory and practice.

5.1. Expected Policyholder Deficit and the Insolvency Put

Expected policyholder deficit, or EPD, can be formulated in different ways. Most straightforwardly, if *B* is capital (or book value) and *Y* is the negative of profit[5], expected policyholder deficit is:

$$\rho(Y) = Pr(Y > B)\,E[Y - B\,|\,Y > B]$$

This is not scalable because *B* does not scale. Other formulations below will scale *B* as a function of losses to make *EPD* scalable.

The value of the default put option is *EPD* under a risk-adjusted probability distribution that reflects the market value of the default:

$$\rho^*(Y) = Pr^*(Y > B)\,E^*[Y - B\,|\,Y > B]$$

---

[3]   "Law-invariant" means that two random variables with the same probability distribution end up with the same risk measure.

[4]   "Comonotonic additive" means that if *Y* and *X* are two random variables with a nondecreasing deterministic relationship between them, $Y = f(X)$, then $\rho(X + Y) = \rho(X) + \rho(Y)$.

[5]   The negative of profit is taken so that large is bad, as with losses.

There are also alternatives for how to express *EPD* or the default put, even with fixed *B*, as a co-measure. Basically for any a, you can set $h(X) = X - aEX$ and $L(Y) = Pr(Y > B)(Y - B)/(Y - aEY)$. Then:

$$r(X_j) = E[(X_j - aEX_j)L(Y)|Y > B)$$
$$= Pr(Y > B)E[(Y - B)(X_j - aEX_j)/(Y - aEY)|Y > B)$$

As examples consider $a = 0$ or $a = 1$. In the case $a = 1$, the allocation of the default amount $Y - B$ is the ratio of j's excess loss over its mean to the company's excess over its mean, conditional on default. When the allocation is negative, j has more than its mean profit on average when the company depletes capital, so j could be considered a capital contributor, and a negative allocation would make sense.

For *EPD* more intuitive is the case $a = 0$. Then:

$$r(X_j) = Pr(Y > B)E[(Y - B)X_j/Y|Y > B]$$
$$= Pr(Y > B)E[X_j - BX_j/Y|Y > B].$$

Here j gets its own expected shortfall in company failure, which is allocation from the customer viewpoint. This is an allocation of the expected shortfall the customers are subject to, so makes particular sense as a basis for reduction in the premium for the risk of insolvency. Sherris (2004) uses such an allocation with transformed probabilities to allocate the default put option.

Taking $a = B/EY$ simplifies things. Then $L(Y) = Pr(Y > B)$:

$$r(X_j) = Pr(Y > B)E[(X_j - BEX_j/EY)|Y > B]$$
$$= Pr(Y > B)\{E[X_j|Y > B] - BEX_j/EY\}$$

This amounts to expressing the *EPD* risk measure as:

$$\rho(Y) = Pr(Y > aEY)E[Y - aEY|Y > aEY]$$

For instance if book value is 15 times expected earnings, with $EY < 0$, then $a = -15$. This form is scalable so would work for marginal decomposition, which would give the same r as the co-measure. However capital would have to be able to be adjusted for small changes in EY in order to keep B/EY constant. This could be done by an agreement with capital providers, for example.

This decomposition can be viewed as first spreading capital by expected value, so $B_j = BEX_j/EY$. Then X_j's share of *EPD* is $Pr(Y > B)E[(X_j - B_j)|Y > B]$. This could be negative if component j does not use up such capital in the average of cases where the company goes under. This is not intuitively appealing, however, as j could still have used some of its capital, so it is not necessarily a capital contributor.

Various allocations can give negative capital to components. It seems logical to avoid methods that give negative allocations to components that do not

make profits on average when the company as a whole has a loss. When Y is negative profit, the risk measure $\rho(Y) = Pr(Y > 0)E[Y | Y > 0]$ is the expected capital drawdown, with the co-measure $r(X_j) = Pr(Y > 0)E[X_j | Y > 0]$. This is j's expected draw on capital when capital is reduced, and if negative it can be interpreted as an expected capital contribution. This risk measure is scalable and the co-measure is the marginal impact.

Another scalable expression of EPD could be obtained by having a mechanism for adjusting capital to keep the probability of survival $1 - \alpha = F_Y(B)$ constant with small changes in Y. That is, $B = F_Y^{-1}(1 - \alpha)$. Then

$$\rho(Y) = \alpha E[Y - B | F(Y) > 1 - \alpha]$$

In this case the directional derivative gives (see Appendix 1):

$$r(X_j) = \alpha[E(X_j | Y > B) - E(X_j | Y = B)]$$

That sums up to $\alpha[E(Y | Y > B) - E(Y | Y = B)] = \alpha[E(Y | Y > B) - B] = \alpha E(Y - B | Y > B) = \rho(Y)$. This r can be formulated as a co-measure by taking $h(X) = X - E(X | Y = B)$, $L(Y) = Pr(Y > B)$ and the condition $Y > B$. This appears to be a sensible way to express EPD as a risk measure for the insurer. The derivative is the change in EPD due to a small change in X_j, given that capital is adjusted to keep the probability of ruin the same before and after the change. The assignment to j is j's expected contribution to losses in default over what its losses would be if capital were exactly used up but not defaulted. That is the amount by which j is pushing the company into default. When the risk measure is the expected amount of default for the whole company, this makes sense as j's contribution. Also the expression in brackets is $co\text{-}TVaR - co\text{-}VaR$.

5.2. Myers-Read Approach

Adjusting capital to keep the probability of ruin constant is similar to Myers-Read, but they would adjust capital enough to keep the value of the insolvency put as a portion of expected claims the same before and after the change, and charge each component by the change in capital needed to do this. Thus their risk measure is capital itself, contingent on the insolvency put maintaining a specific percent of expected claims, and the decomposition is the derivative of the risk measure.

Both claims and negative profit are needed to express this measure. So keep $Y = \Sigma X_j$ as negative profit and let $C = \Sigma D_j$ be the claims. Then denote the Myers-Read capital as $\rho(Y)$, which is defined implicitly by[6]:

$$\int_{\rho(Y)}^{\infty} [y - \rho(Y)] f_Y(y)\, dy = a \int_0^{\infty} c L_C(c)\, dc$$

---

[6] Myers and Read use this general criterion but their formulation also includes investment income, which links the variables Y and C, and an explicit probability transform.

where a is the desired ratio to expected claims, $L_C(c)$ is the density of C and $f_Y(y)$ is the modified density of negative profit used to price the option. The criterion for capital can be expressed as $Pr^*[Y > \rho(Y)]E^*[Y - \rho(Y)|Y > \rho(Y)] = aEC$, which can be solved for $\rho(Y)$ to yield the expression:

$$\text{capital} = \rho(Y) = E^*[Y|Y > \rho(Y)] - aEC/Pr^*[Y > \rho(Y)].$$

Here aEC is the value of the default put, so $aEC/Pr^*[Y > \rho(Y)]$ is the value of the conditional expected shortfall given that there is a default. The most direct co-measure decomposition of this is $aED_j/Pr^*[Y > \rho(Y)]$, which is j's portion of the conditional expected shortfall, proportional to expected losses. Doing this for $E^*[Y|Y > \rho(Y)]$ gives $E^*[X_j|Y > \rho(Y)]$. The difference then distributes capital by:

$$r(X_j) = E^*[X_j|Y > \rho(Y)] - aED_j/Pr^*[Y > \rho(Y)]$$

This may seem a bit convoluted, but it is also the decomposition from the directional derivative, so it is the marginal impact of a line of business on capital. Although this is a decomposition of capital based on the insolvency put, it is not a decomposition of the put value itself. If there is a linear relationship between claims and loss, so $D_j = h_2(X_j)$, this decomposition can be expressed as a co-measure, with $h_1(X) = X$, $g_1(Y) = 1$, and $g_2(Y) = -a/Pr^*[Y > \rho(Y)]$.

6. Non-homogeneous Growth

For allocation purposes and taking directional derivatives, any company can be assumed to have homogeneous growth opportunities, as that can be accomplished through quota share, at least for negative growth. However, strategic planning often looks to grow business components by adding new exposure units. Even reducing business is usually done by reducing exposure units rather than by increasing quota-share reinsurance, unless the reduction is designed to be temporary. For large profit centers, adding units is often quite similar to homogeneous growth, but in other cases the differences can be large. Adding a dollar of payroll or of sales to a book of business, when these are the exposure units, is usually close to homogeneous growth, but adding a satellite launch or super-tanker can be different. The following example, based on one from Glenn Meyers, illustrates this point.

A company has two independent components (Lines 1 and 2), each with negative binomial claim count and constant severity. Each has mean count of $v = 100$. Line j has severity j and claim count variance $v + v^2/(100j)$, and loss variance $j^2[v + v^2/(100j)] = j^2v + jv^2/100$. These are 200 and 600 respectively. The standard deviation of total company losses is $800^{\frac{1}{2}} = 28.284$. Since the lines are independent, each one's covariance with total losses is its own loss variance. Allocating standard deviation by the co-standard deviation $Cov(X,Y)/Std(Y)$ gives 7.071 and 21.213 respectively.

Now suppose that the smallest possible increase for Line 1 is by 1%, so its new mean is 101 and its variance is 203.01. The total standard deviation then increases to 28.337 which is an increase of 0.05316. The marginal change proportional to the change in mean (or exposure) of 1% is thus 5.316. For Line 2 this is 14.142. These are approximately $\frac{3}{4}$ and $\frac{2}{3}$, respectively, of the allocated values. Since the risk here is basically frequency, and frequency variance tends to increase by exposure, frequency standard deviation tends to reduce proportionally to exposure when exposure increases. Thus the marginal change when adding exposure is less than the marginal increase from a homogeneous change. Allocation does not work as well as pure marginal analysis in this case. Meyers points out that this is a small company problem. As the volume increases in this example, the co-standard deviation allocation becomes very close to marginal. For instance, with $v = 10,000$, adding one unit to v produces proportional marginal impacts of 575 and 1157 for Lines 1 and 2 compared to the co-standard deviation allocations of 578 and 1168.

This is an interesting example in that the lines are independent of each other but the exposure units within a line are not. If they were, adding exposure units would increase the claim count mean and variance by the same factor. Such a case can arise in practice when the exposure units are conditionally independent given some common factor that affects them all, like the weather.

The allocation of $TVaR$ is fairly close to marginal even for this small book. For instance $TVaR$ at 80% for the company is 340.24 by simulation, and this increases to 341.33 with an increase of 1 expected claim in Line 1. This overall increase of 1.09 is 109 proportionally, compared to the co-$TVaR$ of approximately 115.

The exponential moment co-measure is also reasonably close to marginal even with this small book of business. Taking $c = 0.094431$ makes $E[Ye^{cY/EY}] = 330$, which is 10% over EY. The co-measure allocation can be found numerically to be about 110 and 220. Increasing Line 1 by 1 expected claim increases the overall moment to 331.20, which gives a proportional marginal impact of 120.

Thus just being scalable is not enough to guarantee that a risk measure has a marginal decomposition with non-homogeneous growth, but some risk measures are close to this even for a small book. Other risk measures have perfectly marginal decomposition for non-homogeneous growth even for a small book like this.

The mean of a transformed probability distribution is one such risk measure. In the case of non-homogeneous growth you can separately transform frequency and severity. Suppose for instance that for the negative binomial model above the transform maps $a \rightarrow 1.11a$ for each line. It is a good idea in transforming probabilities to keep the same set of possible events, which in this case would mean no change in the severity distribution. The risk measure is the transformed mean. The co-measure is the transformed mean of the component. So the co-measure for Line 1 is 111 and for Line 2 is 222, with a total risk measure of 333. When Line 1 increases by 1 the total risk becomes 334.11, which gives a proportional marginal impact of 111, agreeing with the co-measure. Thus this is a marginal decomposition even for non-homogeneous growth. If the

risk measure were the mean, the marginal impacts would add up. The fact that the transformed mean uses different probabilities does not change this.

Risk pricing can be done through probability transforms as well, so it should be possible to find a transform that represents the market value of the risk. This would be a risk measure with a direct economic interpretation. Venter (1991) showed that covariance loadings can be expressed as probability transforms[7], so $CAPM$ etc. are special cases of this approach.

A well supported transform for compound Poisson distributions is the frequency-severity combined Esscher transform[8], e.g., see Ballotta (2004) and Venter, Barnett and Owen (2004). This uses a constant c and transforms the severity density $f_X(x)$ to $f_X^*(x) = f_X(x)e^{x/c}/E(e^{X/c})$ and the frequency λ to $\lambda E(e^{X/c})$.

Each component or even peril within component could have its own c parameter and still preserve marginal decomposition. However it might also be appropriate to transform the joint dependencies among the frequencies and severities of the components, perhaps by doing copula transforms. There could be such transforms that would not give marginal decomposition. For instance if one line is dependent on the actual number of claims from another line, adding business to one line could affect the expected losses of the other line, so the transformed means would not be marginal. But if this is avoided there could be quite a few possible dependencies included. For instance, parameter change could be correlated, like the frequency per exposure unit for different lines changing according to correlated stochastic processes, frequencies per exposure unit and severities could be dependent, etc. There are a lot of possibilities for joint transforms and working out the most appropriate is an open problem.

7. Allocating Firm Value

A typical application of allocation of a risk measure assumed throughout the above is to express the capital of an insurer as a risk measure of the losses, and then allocate the capital in proportion to the allocation of the risk measure. Then risk-adjusted returns could be calculated using the return on allocated capital.

An alternative way of evaluating the value contribution of each component would be to distribute firm value to component. For instance, if firm value can be expressed as the risk-adjusted present value of future earnings, then it could be written as an expected value of earnings under transformed probabilities:

$V(Y) = E_Q(Y)$, where now Y is not losses but earnings of the firm. Then:

$V(X_j) = E_Q(X_j)$ is j's contribution to firm value.

It is the co-value of the firm under the Q measure.

---

[7] That paper shows that a probability transform of X that produces its covariance with Y can be achieved by taking $f^*(x) = f(x)[1 + cE(Y|x) - cE(Y)]$ for small enough c.

[8] This can be derived as an application of the minimal entropy martingale measure from information theory to incomplete market pricing, and has been worked out for the compound Poisson process. It assumes that jump risk is inherently non-diversifiable and so is always priced.

This is related to measuring financial performance of the business units using economic value added in a capital consumption framework. Say for instance that the cost to the firm of maintaining a business component is the economic value of the component's right to access firm capital if it needs to. This was suggested by Merton-Perold (1993). The firm is essentially providing the component a stop-loss cover attaching at the point its cash flows become negative. This is a contingent claim. But the firm has a contingent claim on the profits of the component, attaching if they are positive. The firm gets all the profits or losses of the component so these two claims are complementary and not contingent combined. The value of their difference is the economic value added of the component.

The overall value of the firm is a risk measure of the firm. This can be estimated in different ways – like expected future cash flows discounted at a rate that reflects their risk, or expected cash flows under transformed probabilities, discounted at the risk-free rate. The capital cost (value of the right to use capital) is greater for a component that would use the capital at the same time as other units. The value of profits is greater for a component that has profits when the rest of the company has losses. In any case, the value of the components should sum to the firm value and consider the correlations. The co-measures of the value calculation do just that. For instance in a simulation, scenarios with a large drop in value contribute large negative value to lines with losses and large positive value to lines with profits.

8. Conclusion

Decomposition of risk by co-measures is a quite general method, but has a weakness of not having a unique solution, and it is not always marginal. When decomposition by the directional derivative is possible (scalable measures), doing this shows which formulation of the co-measure is marginal in the case of homogeneous growth. For non-homogeneous growth the same co-measure is often close to marginal. Transformed probability risk measures have a marginal decomposition even in the non-homogeneous growth case and can be formulated to approximate the market value of the risk being measured. One advantage of using marginal decomposition is that growing the higher return-to-risk units increases the return to risk of the entire company.

Appendix 1. Taking Derivatives

This appendix calculates the derivatives of some common risk measures. The marginal co-measures for *XTVaR*, *EPD* (expected policyholder deficit), and risk measures that are functions of Y/EY are derived, and an abstraction of the Myers-Read approach is also developed. An alternative method of taking derivatives is presented in Appendix 3.

A1.1. Mathematical Preliminaries

It is convenient to use indicator functions. The indicator for $x = 0$ is denoted $\delta(x)$. It can be thought of as the limit of a normal distribution density with standard deviation ε as $\varepsilon \to 0$. Thus the δ function integrates to one and is zero for non-zero argument. The indicator for $x > 0$ is denoted $\Theta(x)$. It can be thought of as the distribution function with the δ density, so its derivative is $\delta(x)$.

We use capitals for the random variables, F for distribution functions, and $S = 1 - F$. We are positing a joint distribution function $F(x_1, x_2, \ldots, x_N) \equiv F(\mathbf{x})$ and a corresponding joint density function $f(x_1, x_2, \ldots, x_N) \equiv f(\mathbf{x})$. We implicitly assume the density is everywhere finite (no point masses). This assumption could easily be relaxed in what follows.

The density for the total $Y \equiv \sum_{n=1}^{N} X_n$ is $f_Y(y) = \iint \delta\left(y - \sum_{n=1}^{N} x_n\right) f(\mathbf{x}) d\mathbf{x}$. This is just a way of writing the density of Y using the multivariate distribution of the X's. The mean of Y is $\mu = \iint \sum_{n=1}^{N} x_n f(\mathbf{x}) d\mathbf{x} = \int y f_Y(y) dy$ and the mean of X_k is $\mu_k = \iint x_k f(\mathbf{x}) d\mathbf{x}$, so of course $\mu = \sum_{n=1}^{N} \mu_n$ as it should.

Another (bivariate) distribution which will turn out to be of interest is the joint distribution of Y and X_k: $f_k(y, x) = \iint \delta\left(y - \sum_{n=1}^{N} x_n\right) \delta(x - x_k) f(\mathbf{x}) d\mathbf{x}$. We have $f_Y(y) = \int f_k(y, x) dx$ for any k. We recognize that the distribution for X_k conditional on Y being fixed at b is $f_k(x \mid y = b) = \dfrac{f_k(b, x)}{f_Y(b)}$. An expression we will encounter often is $\int x f_k(b, x) dx = E[X_k \mid Y = b] f_Y(b)$. This is the mean value of X_k given that $Y = b$ times the density function for the total evaluated at b.

A1.2. *XTVaR* with Fixed Excess Point

The generalized *XTVaR* risk function excess of a fixed level b for any constant c is

$$\rho = E[Y - c\mu \mid Y > b]$$
$$= \int (y - c\mu)\, \Theta(y - b)\, f_Y(y)\, dy \Big/ \int \Theta(y - b)\, f_Y(y)\, dy$$
$$= \iint \left(\sum_{n=1}^{N} x_n - c\mu\right) \Theta\left(\sum_{n=1}^{N} x_n - b\right) f(\mathbf{x})\, d\mathbf{x} / S(b).$$

The co-measures are

$$\rho_k = \iint (x_k - c\mu_k)\, \Theta\left(\sum_{n=1}^{N} x_n - b\right) f(\mathbf{x})\, d\mathbf{x} \Big/ S(b) = E[X_k - c\mu_k \mid Y > b].$$

The co-measures sum to the total measure: $\rho = \sum_{n=1}^{N} \rho_n$.

The question is, if we think of ρ as $\rho(Y)$, what is meant by $\rho(Y+Z)$? Basically it is the same functional form with another variable added. In the present case,

$$\rho(Y+Z) =$$

$$\iint \left(z + \sum_{n=1}^{N} x_n - c\mu - cE(Z) \right) \Theta \left(z + \sum_{n=1}^{N} x_n - b \right) f(\mathbf{x}, z)\, d\mathbf{x}\, dz \Big/ S_z(b)$$

where the joint density function now depends on the additional variable and the survivor function is $S_z(b) = \iint \Theta \left(z + \sum_{n=1}^{N} x_n - b \right) f(\mathbf{x}, z)\, d\mathbf{x}\, dz$.

We are interested in $\rho(\varepsilon) = \rho(Y + \varepsilon X_k)$ and in particular want to evaluate the derivative $\frac{d}{d\varepsilon}\rho(\varepsilon)$ at $\varepsilon = 0$. As discussed in the main text, this is the derivative in the direction X_k and represents the marginal contribution of that variable to the whole. Note that we have not indicated explicitly the dependence on k of $\rho(\varepsilon)$ but it is of course always present. Thus the case of interest is where $Z = \varepsilon X_k$. One way of looking at this is that it is the same as the original with a transformed variable $Z_k = (1 + \varepsilon) X_k$ replacing X_k. The joint distribution function is

$$F_\varepsilon(x_1, x_2, ..., x_{k-1}, z_k, x_{k+1}, ..., x_N) = F\left(x_1, x_2, ..., x_{k-1}, \frac{z_k}{1+\varepsilon}, x_{k+1}, ..., x_N \right)$$

so the density function is

$$f_\varepsilon(x_1, x_2, ..., x_{k-1}, z_k, x_{k+1}, ..., x_N) = f\left(x_1, x_2, ..., x_{k-1}, \frac{z_k}{1+\varepsilon}, x_{k+1}, ..., x_N \right) \frac{1}{1+\varepsilon}$$

Putting this into the equation for the risk measure and changing variables from z_k to $x_k = \frac{z_k}{1+\varepsilon}$ we end up with

$$\rho(\varepsilon) = \iint \left(\varepsilon x_k + \sum_{n=1}^{N} x_n - c\mu - \varepsilon c\mu_k \right) \Theta \left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b \right) f(\mathbf{x})\, d\mathbf{x} \Big/ S(\varepsilon, b)$$

and $S(\varepsilon, b) \equiv \iint \Theta \left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b \right) f(\mathbf{x})\, d\mathbf{x}$. Our convention will be $S(b) \equiv S(0, b)$ in accord with previous notation.

A more formal way to get the same result with much less effort is to say that the joint density function is $f(\mathbf{x}, z) = f(\mathbf{x})\, \delta(z - \varepsilon x_k)$. Doing the integration over z directly leads to the above results.

We will now directly evaluate the terms in the derivative and then take the limit. We need to write the terms so that the limiting process can be easily

seen. The first term below separates out the elements of $\rho(\varepsilon)$ that are multiples of ε.

$$\frac{\rho(\varepsilon) - \rho}{\varepsilon} = \iint (x_k - c\mu_k)\, \Theta\!\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b\right) f(\mathbf{x})\, d\mathbf{x}\, / S(\varepsilon, b)$$

$$+ \iint \left(\sum_{n=1}^{N} x_n - c\mu\right) \frac{1}{\varepsilon} \left[\frac{\Theta\!\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b\right)}{S(\varepsilon, b)} - \frac{\Theta\!\left(\sum_{n=1}^{N} x_n - b\right)}{S(b)} \right] f(\mathbf{x})\, d\mathbf{x}$$

$$= \iint (x_k - c\mu_k)\, \Theta\!\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b\right) f(\mathbf{x})\, d\mathbf{x}\, / S(\varepsilon, b)$$

$$+ \frac{1}{S(b)} \iint \left(\sum_{n=1}^{N} x_n - c\mu\right) \left[\frac{\Theta\!\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b\right) - \Theta\!\left(\sum_{n=1}^{N} x_n - b\right)}{\varepsilon} \right] f(\mathbf{x})\, d\mathbf{x}$$

$$+ \frac{S(b) - S(\varepsilon, b)}{\varepsilon S(b)\, S(\varepsilon, b)} \iint \left(\sum_{n=1}^{N} x_n - c\mu\right) \Theta\!\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b\right) f(\mathbf{x})\, d\mathbf{x}$$

The last two terms of the second expression sum to the last term of the first. All three terms are now in a form where we may go to the limit $\varepsilon \to 0$. The first term goes to $\iint (x - c\mu_k)\, \Theta(y - b)\, f_k(y, x)\, dx / S(b) = \rho_k$. We may write the second term as

$$\frac{1}{S(b)} \iint (y - c\mu) \frac{\Theta(\varepsilon x + y - b) - \Theta(y - b)}{\varepsilon} f_k(y, x)\, dy\, dx$$

$$= \frac{1}{S(b)} \int dx \left\{ \frac{1}{\varepsilon} \int_{b-\varepsilon x}^{b} (y - c\mu)\, f_k(y, x)\, dy \right\}$$

$$\to \frac{b - c\mu}{S(b)} \int x f_k(b, x)\, dx.$$

For the limit, you can think of the integrand as a constant at $y = b$ over a rectangle of width εx, so the integral is εx times that integrand.

For the third term, similarly,

$$\frac{S(\varepsilon, b) - S(b)}{\varepsilon} = \int dx \left\{ \frac{1}{\varepsilon} \int_{b-\varepsilon x}^{b} f_k(y, x)\, dy \right\} \to \int x f_k(b, x)\, dx$$

and the limit of the integral is ρ, so in the limit the third term is $-\dfrac{\int x f_k(b, x)\, dx}{S(b)}\, \rho$.

Thus in the limit as $\varepsilon \to 0$ we obtain

$$\frac{d}{d\varepsilon}\rho(\varepsilon)\bigg|_{\varepsilon=0} = \rho_k + \frac{b - c\mu - \rho}{S(b)}\int xf_k(b,x)\,dx$$

$$= E[X_k - \mu_k \mid Y > b] + \frac{f_Y(b)}{S(b)}E[X_k \mid Y = b]\{b - c\mu - E[Y - \mu \mid Y > b]\}$$

The first term is the co-measure, and the second term is proportional to the expected value of X_k given $Y = b$. Given appropriate values for the constants, the curly bracket could vanish. However in general the fixed threshold b keeps the risk measure from being homogeneous, so the derivative is not the co-measure, and the derivatives do not add to the risk measure. Now let us try the same derivation in a simpler fashion, using the formal relation $\frac{d}{dx}\Theta(x) = \delta(x)$. We repeat the original form

$$\rho(\varepsilon) = \iint\left(\varepsilon x_k + \sum_{n=1}^{N}x_n - \mu - \varepsilon\mu_k\right)\Theta\left(\varepsilon x_k + \sum_{n=1}^{N}x_n - b\right)f(\mathbf{x})\,d\mathbf{x}\,/\,S(\varepsilon)$$

where $S(\varepsilon) = \iint\Theta\left(\varepsilon x_k + \sum_{n=1}^{N}x_n - b\right)f(\mathbf{x})\,d\mathbf{x}$

Again, we note that $S(0)$ is what we earlier called $S(b)$, the survivor function on the total at b. We will use a prime to indicate differentiation.

$$\rho'(0) = \iint(x_k - c\mu_k)\,\Theta\left(\sum_{n=1}^{N}x_n - b\right)f(\mathbf{x})\,d\mathbf{x}\,/\,S(0)$$

$$+ \iint\left(\sum_{n=1}^{N}x_n - c\mu\right)x_k\delta\left(\sum_{n=1}^{N}x_n - b\right)f(\mathbf{x})\,d\mathbf{x}\,/\,S(0)$$

$$- \frac{S'(0)}{S(0)^2}\iint\left(\sum_{n=1}^{N}x_n - c\mu\right)\Theta\left(\sum_{n=1}^{N}x_n - b\right)f(\mathbf{x})\,d\mathbf{x}$$

and

$$S'(0) = \iint x_k\delta\left(\sum_{n=1}^{N}x_n - b\right)f(\mathbf{x})\,d\mathbf{x} = \int xf_k(b,x)\,dx$$

Thus the terms directly reduce to

$$\rho'(0) = \rho_k + \frac{b - c\mu}{S(b)}\int xf_k(b,x)\,dx - \frac{\rho}{S(b)}\int xf_k(b,x)\,dx$$

$$= \rho_k + \frac{b - c\mu - \rho}{S(b)}\int xf_k(b,x)\,dx$$

which is the result above, of course.

A1.3. *XTVaR* with Fixed Percentage Excess

Let us try a somewhat different measure: We will use *XTVaR* excess of a fixed probability, rather than a fixed amount. We define the statistic by the probability of the total being excess of it is α. Formally, $b = F_Y^{-1}(1 - \alpha)$ or more explicitly in terms of the densities

$$\alpha = S(b) = \iint \Theta\left(\sum_{n=1}^{N} x_n - b\right) f(\mathbf{x})\, d\mathbf{x} = \int \Theta(y - b) f_Y(y)\, dy$$

$$\rho = E\left[Y - c\mu \,\middle|\, Y > b = F_Y^{-1}(1 - \alpha)\right]$$

$$= \int (y - c\mu)\, \Theta(y - b) f_Y(y)\, dy \,/\, \alpha$$

$$= \iint \left(\sum_{n=1}^{N} x_n - c\mu\right) \Theta\left(\sum_{n=1}^{N} x_n - b\right) f(\mathbf{x})\, d\mathbf{x} \,/\, \alpha$$

When we introduce the additional variable to do the directional derivative, we have

$$\alpha = \iint \Theta\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b(\varepsilon)\right) f(\mathbf{x})\, d\mathbf{x}$$

and so taking the derivative of this equation with respect to ε we get

$$0 = \iint \delta\left(\sum_{n=1}^{N} x_n - b\right)(x_k - b'(0)) f(\mathbf{x})\, d\mathbf{x} \Rightarrow b'(0)$$

$$= \frac{\int x f_k(b, x)\, dx}{f_Y(b)} = E[X_k | Y = b]$$

and

$$\rho'(0) = \iint (x_k - c\mu_k)\, \Theta\left(\sum_{n=1}^{N} x_n - b\right) f(\mathbf{x})\, d\mathbf{x} \,/\, \alpha$$

$$+ \iint \left(\sum_{n=1}^{N} x_n - c\mu\right) \delta\left(\sum_{n=1}^{N} x_n - b\right)(x_k - b'(0)) f(\mathbf{x})\, d\mathbf{x} \,/\, \alpha$$

$$= \rho_k$$

since the second term is zero. In this case the directional derivative is exactly the co-measure.

A1.4. Expected Policy-Holder Deficit

Here

$$\rho = E[Y - b \,|\, Y > b]\, S(b) = \int (y - b)\, \Theta(y - b)\, f_Y(y)\, dy$$

$$= \iint \left(\sum_{n=1}^{N} x_n - b \right) \Theta \left(\sum_{n=1}^{N} x_n - b \right) f(\mathbf{x})\, d\mathbf{x}$$

For the moment we leave open the question of whether b is fixed, defined by a percentage, or has some other dependence on ε. We have

$$\rho(\varepsilon) = \iint \left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b(\varepsilon) \right) \Theta \left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b(\varepsilon) \right) f(\mathbf{x})\, d\mathbf{x}$$

and

$$\rho'(0) = \iint (x_k - b'(0))\, \Theta \left(\sum_{n=1}^{N} x_n - b \right) f(\mathbf{x})\, d\mathbf{x}$$

$$+ \iint \left(\sum_{n=1}^{N} x_n - b \right) x_k\, \delta \left(\sum_{n=1}^{N} x_n - b \right) f(\mathbf{x})\, d\mathbf{x}$$

The second term is immediately zero, so we may write

$$\rho'(0) = \{ E[X_k \,|\, Y > b] - b'(0) \}\, S(b)$$

In the case where $b = F_Y^{-1}(1 - \alpha)$ we know that $b'(0) = E[X_k \,|\, Y = b]$ and so the directional derivative becomes $\rho'(0) = \{ E[X_k \,|\, Y > b] - E[X_k \,|\, Y = b] \}\, \alpha$. The sum over all N variables is

$$\alpha \sum_{n=1}^{N} \{ E[X_n \,|\, Y > b] - E[X_n \,|\, Y = b] \} = \alpha \{ E[Y \,|\, Y > b] - E[Y \,|\, Y = b] \}$$

$$= \alpha \{ E[Y \,|\, Y > b] - b \} = \alpha E[Y - b \,|\, Y > b] = \rho$$

A1.5. Myers-Read

This is an analogue to the Myers-Read approach, but does not consider every element they address, such as investment income. Here we have a risk-adjusted distribution in addition to the usual probability distribution. We use a superscript * to indicated the risk-adjusted distribution and any quantities derived from it, for example the density $f^*(x_1, x_2, \ldots, x_N) \equiv f^*(\mathbf{x})$. The risk measure ρ is the capital. It satisfies the relation that capital is set to the level needed so that the value of the default put is α times mean losses for a target small number α. That is, $\{ E^*[Y \,|\, Y > \rho] - \rho \}\, S^*(\rho) = \alpha\mu$ which is also

$$\int_{\rho}^{\infty} (y - \rho)\, f_Y^*(y)\, dy = \alpha \int y f_Y(y)\, dy$$

In terms of the actual underlying variables, this defining relation is

$$\iint \left(\sum_{n=1}^{N} x_n - \rho \right) \Theta\left(\sum_{n=1}^{N} x_n - \rho \right) f^*(\mathbf{x})\, d\mathbf{x} = \alpha \iint \left(\sum_{n=1}^{N} x_n \right) f(\mathbf{x})\, d\mathbf{x}$$

For the directional derivative we need

$$\iint \left(\varepsilon x_k + \sum_{n=1}^{N} x_n - \rho(\varepsilon) \right) \Theta\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - \rho(\varepsilon) \right) f^*(\mathbf{x})\, d\mathbf{x}$$
$$= \alpha \iint \left(\varepsilon x_k + \sum_{n=1}^{N} x_n \right) f(\mathbf{x})\, d\mathbf{x}$$

We take the derivative of this relation with respect to ε and evaluate at $\varepsilon = 0$. Putting the right-hand side first, we have

$$\alpha \iint x_k f(\mathbf{x})\, d\mathbf{x} = \iint (x_k - \rho'(0))\, \Theta\left(\sum_{n=1}^{N} x_n - \rho \right) f^*(\mathbf{x})\, d\mathbf{x}$$
$$+ \iint \left(\sum_{n=1}^{N} x_n - \rho \right) (x_k - \rho'(0))\, \delta\left(\sum_{n=1}^{N} x_n - \rho \right) f^*(\mathbf{x})\, d\mathbf{x}$$

Hence $\alpha \iint x f_k(y, x)\, dy dx = \iint (x - \rho'(0))\, \Theta(y - \rho)\, f_k^*(y, x)\, dy dx$

since the term containing the delta function contributes zero. Solving for the derivative, which is the proposed allocation to the variable X_k,

$$\rho'(0) = \frac{\iint x\, \Theta(y - \rho)\, f_k^*(y, x)\, dy dx - \alpha \mu_k}{\iint \Theta(y - \rho)\, f_k^*(y, x)\, dy dx}$$
$$= E^*[X_k \mid Y > b] - \frac{\alpha \mu_k}{S^*(\rho)}$$

Note that the mean of the kth variable μ_k is the true mean, not the risk-adjusted mean. Summing these up, we get $E^*[Y \mid Y > \rho] - \frac{\alpha \mu}{S^*(\rho)}$. Going back to the defining relation for ρ and solving for ρ, we get this same quantity and can see that it is equal to the sum of the allocations.

A1.6. Functions of Y/EY

Take $\rho = E[h(Y) L(Y)] = \iint h\left(\sum_{n=1}^{N} x_n \right) L\left(\sum_{n=1}^{N} x_n \right) f(\mathbf{x})\, d\mathbf{x} = \int h(y) L(y) f_Y(y)\, dy$

We want h to be a linear functional: $h(aX + bY) = ah(X) + bh(Y)$ for all constants a and b and all random variables X and Y. One example of this would be

$h(Y) = aY + bE[Y] + cCov(Y, M)$ with M an external variable such as a market. The function L is any integrable function. In the preceding examples L often gave the conditional expectations, for example $L(Y) = \Theta\big(Y - F_Y^{-1}(1-\alpha)\big)/\alpha$. It essentially describes management's attitude toward risk.

Let us see when we can get a marginal co-measure by differentiating with respect to an increase of ε in X_k. We will again have

$$\rho(\varepsilon) = \iint h\left(\varepsilon x_k + \sum_{n=1}^{N} x_n\right) L\left(\varepsilon x_k + \sum_{n=1}^{N} x_n\right) f(\mathbf{x})\, d\mathbf{x}$$

Since h is a linear functional, we have $h\left(\varepsilon x_k + \sum_{n=1}^{N} x_n\right) = \varepsilon h(x_k) + \sum_{n=1}^{N} h(x_n)$.

Let us look at a specific example, where L is any integrable function of the ratio $\zeta = \dfrac{Y}{E[Y]}$. Then the marginal co-measure is

$$\rho_k = E\big[h(X_k) L(\zeta)\big] + \frac{1}{E[Y]^2} E\big[h(Y) L'(\zeta)\{E[Y] X_k - Y E[X_k]\}\big]$$

It is clear that when summed, the term in the curly brackets is zero. It is essentially a correction to the simple co-measure to make it a marginal co-measure.

As an example, take the exponential $L(\zeta) = \exp(c\zeta)$ for c some constant. This becomes

$$\rho_k = E\big[h(X_k)\exp(c\zeta)\big] + \frac{c}{E[Y]^2} E\big[h(Y)\exp(c\zeta)\{E[Y] X_k - Y E[X_k]\}\big]$$

$$= E\big[h(X_k)\exp(c\zeta)\big] + \frac{c}{E[Y]^2} E\big[X_k\{h(Y)\exp(c\zeta) E[Y] - E[Y h(Y)\exp(c\zeta)]\}\big]$$

APPENDIX 2. EULER RELATION

Here, we discuss the application of Euler's Relation on Homogeneous Functions to risk measures. In the body of the paper, we use the term "scalable" to mean homogeneous degree 1.

A2.1. Euler's Relation

Definition: A vector function $f : \mathfrak{R}^n \to \mathfrak{R}$ is said to be homogeneous of degree d if, for any scalar t, $f(t \cdot \mathbf{X}) = t^d \cdot f(\mathbf{X})$.

Theorem (Euler): if f is a smooth function, homogeneous of degree d, then $\nabla f(\mathbf{X}) \cdot \mathbf{X} = d \cdot f(\mathbf{X})$. Here, the ∇ operator is the vector of partial derivatives;

by convention it is a row vector whereas $\mathbf{X}$ is a column vector. Therefore, written out in terms of components, this equation is $\sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(x_1, x_2, \ldots, x_n) \cdot x_i = d \cdot f(x_1, x_2, \ldots, x_n)$.

Proof: Take the derivative of both sides of $f(t \cdot \mathbf{X}) = t^d \cdot f(\mathbf{X})$ with respect to t. On the left, use the chain rule to obtain $\frac{d}{dt} f(t \cdot \mathbf{X}) = \nabla f \cdot \frac{d}{dt} (t \cdot \mathbf{X}) = \nabla f \cdot \mathbf{X}$. On the right, $d \cdot t^{d-1} \cdot f(\mathbf{X})$. Set $t = 1$ to complete the proof.

Comments: in particular, if f is homogeneous degree 1, $f(\mathbf{X}) = \nabla f(\mathbf{X}) \cdot \mathbf{X}$. The right hand side is sometimes called the "Euler Operator" applied to f and, in that context, homogeneous functions are seen to be eigenfunctions of the Euler Operator with eigenvalue d.

A2.2. Application to Risk Measures

A scalable risk measure ρ satisfies $\rho\left(t \cdot \sum_i X_i\right) = t \cdot \rho\left(\sum_i X_i\right)$ for any random variables X. This is almost, but not quite, enough to be considered homogeneous degree 1 in the sense of Euler. The problem here is that ρ is not a function of a real variable but rather a functional of a random variable.

Under conditions of homogeneous growth, we may regard the vector r.v. X as being "in the background" or fixed in the sense that its distribution (not its value) is fixed, and use variable weights w to express changes in the argument to ρ. Thus, we can write $f(w) = \rho\left(\sum_i X_i \cdot w_i\right)$ and the scalable risk measure is now expressed as a function of the real-valued vector variable w. As it happens, the initial value of interest is $w = \mathbf{1} = \langle 1\, 1\, \ldots\, 1 \rangle'$, but that will change once increments are considered.

Thus, we have $f(t \cdot w) = \rho\left(t \cdot \sum_i X_i \cdot w_i\right) = t \cdot \rho\left(\sum_i X_i \cdot w_i\right) = t \cdot f(w)$ satisfies the hypotheses of the theorem as long as ρ is sufficiently smooth.

Applying the theorem, we get

$$\rho\left(\sum X\right) = f(\mathbf{1}) = \nabla f(\mathbf{1}) \cdot \mathbf{1} = \sum_i \frac{\partial}{\partial w_i} f = \sum_i \frac{\partial}{\partial w_i} \rho\left(\sum_j X_j \cdot w_j\right)\Bigg|_{w=\mathbf{1}} = \sum_i r_i\left(\sum X\right)$$

where $r_i\left(\sum X\right) = \lim\limits_{\varepsilon \to 0} \dfrac{\rho\left(\sum X + \varepsilon \cdot X_i\right) - \rho\left(\sum X\right)}{\varepsilon}$.

APPENDIX 3. INTEGRAL OVER THE SURFACE FORMULA

A3.1. The "IOS" formula

An alternative approach to calculating derivatives of risk measures is by reference to the "Integral over the surface formula" (Uryasev, [1995a, b; 1999]).

Let the domain of integration be defined by $I(\psi) = \{x \in R^N \,|\, \gamma(x, \psi) \leq 0\}$ and the boundary of this set by ∂I. Consider the volume integral

$$H(\psi) = \int_{I(\psi)} \phi(x, \psi) \, dV$$

where dV denotes volume integration over all the x-elements.

There are two important parts here. First, the integrand ϕ usually includes a factor $f(x)$, a probability density function. It may include other factors, however, if the integral is to represent a moment. Second, the volume I over which the integration takes place is equally important. In the case of risk measures with conditions, such as *XTVaR* or *EPD*, this domain will not be the entire R^N space. Either component, the integrand or the domain, may or may not be a function of the parameter ψ with respect to which we want to differentiate.

If the constraint function γ (which defines the domain I) is differentiable and the following integrals exist, then the gradient of H with respect to ψ is given by:

$$\nabla_\psi H(\psi) = \int_{I(\psi)} \nabla_\psi \phi(x, \psi) \, dV - \int_{\partial I(\psi)} \frac{\nabla_\psi \gamma(x, \psi)}{\|\nabla_x \gamma(x, \psi)\|} \cdot \phi(x, \psi) \, dS$$

where dS denotes (hyper)surface element and $\|v\|$ denotes vector norm.

This result is stated in terms of vectors and gradients (vector of partial derivatives) so as to allow ψ to be a vector parameter. If it is merely a scalar, then ∇ just means the scalar derivative, and vector norm just means absolute value.

There are several important components to note on the right hand side. First, notice that there are two terms, say $\nabla H = A - B$. The first term, A, is a volume integral. Intuitively, it makes perfect sense: the integrand ϕ defined within H is being differentiated with respect to ψ, and the result is being integrated over the same domain as H. The second term, B, is more complicated. It accounts for the fact that the domain of integration I itself might be a function of ψ. B is the integral of the same integrand ϕ defined by H, but now it is a *surface* integral over the *boundary* of the domain I. Moreover, the integrand in B includes a factor involving ratios of derivatives to account for how fast the boundary changes as ψ is varied.

This ratio $\dfrac{\nabla_\psi \gamma(x, \psi)}{\|\nabla_x \gamma(x, \psi)\|}$ is probably the most difficult term to comprehend. The numerator is the derivative of the boundary defining condition γ with respect to ψ, so it is the rate of change of γ with respect to ψ. The denominator is the norm of a vector of partial derivatives (or the absolute value of a scalar derivative) of γ with respect to the *coordinates* x of the random variable X. This is the rate of change of γ with respect to motion in the x space. By taking the ratio of these two, we obtain the ratio that equates the effect of changes in ψ with movements in the x space. In other words, the ratio is a measure of how fast ψ moves the boundary of I (which is defined by γ).

Application of this theorem to risk measures will be shown by several examples.

A3.2. Exceedance probabilities

Let $S(\varepsilon, b) = \Pr\{\varepsilon \cdot X_k + \sum_i X_i \geq b\}$. First, we shall compute the derivative of S with respect to b at $\varepsilon = 0$. Putting this problem into the IOS format, we have $\psi = b$, $H(\psi) = S(0, b)$, $\phi(x, \psi) = f(\mathbf{x})$ (i.e. the joint p.d.f. of the X_i), and $\gamma(x, \psi) = b - \sum_i X_i$. There are two terms to the solution for the gradient (in this case just a one-dimensional derivative). Since ϕ is not a function of ψ, the first term vanishes, and we are left with

$$\frac{\partial S(\varepsilon, b)}{\partial b} = -\int\limits_{\sum_i x_i = b} \frac{1}{\left\| \nabla_x \sum_{i=1}^{N} x_i \right\|} \cdot f(x)\, dS = -\frac{1}{\sqrt{N}} \cdot \int\limits_{Y = b} f(x)\, dS.$$

This result can be shown to equal the probability density of $Y = \sum_i X_i$ at $Y = b$. (The square root of N emerges from the Jacobian of a variable transformation in getting the representation to include Y.)

Now consider the derivative of S with respect to ε at $\varepsilon = 0$. Rearranging what is considered the parameter, the IOS formulation for this becomes: $\psi = \varepsilon$, $H(\psi) = S(\varepsilon, b)$, $\phi(x, \psi) = f(\mathbf{x})$ (as before), and $\gamma(x, \psi) = b - \varepsilon \cdot x_k - \sum_i X_i$. Again, the first term in the derivative expression is zero, and we have

$$\frac{\partial S(\varepsilon, b)}{\partial \varepsilon} = -\int\limits_{\sum_i x_i = b} \frac{-x_k}{\left\| \nabla_x \sum_{i=1}^{N} x_i \right\|} \cdot f(x)\, dS = \frac{1}{\sqrt{N}} \cdot \int\limits_{Y = b} x_k \cdot f(x)\, dS.$$

A3.3. Value at Risk

The Value at Risk $(VaR)\rho(Y)$ at probability level q can be defined implicitly by $S(0, \rho(Y)) = q$. To take the directional derivative of ρ with respect to X_k, we can apply implicit differentiation to the equation $S(0, \rho(Y)) - q = 0$. This gives us

$$\frac{\partial S}{\partial \varepsilon} + \frac{\partial S}{\partial b} \cdot \frac{\partial \rho}{\partial \varepsilon} = 0 \quad \Rightarrow \quad \frac{\partial \rho}{\partial \varepsilon} = -\frac{\partial S}{\partial \varepsilon} \bigg/ \frac{\partial S}{\partial b}.$$

Substituting the previously-obtained derivatives for S, we get

$$\frac{\partial \rho}{\partial \varepsilon} = \frac{\partial S(\varepsilon, b)/\partial \varepsilon}{\partial S(\varepsilon, b)/\partial b} = \frac{-\dfrac{1}{\sqrt{N}} \cdot \displaystyle\int\limits_{Y = b} x_k \cdot f(x)\, dS}{-\dfrac{1}{\sqrt{N}} \cdot \displaystyle\int\limits_{Y = b} f(x)\, dS} = \frac{\displaystyle\int\limits_{Y = b} x_k \cdot f(x)\, dS}{\displaystyle\int\limits_{Y = b} f(x)\, dS} = E\big[x_k \,|\, Y = b\big]$$

This method is treated in more generality, and with more examples, in Major (2004). For $TVaR$ with quantile threshold, the first ("A") term is nonzero and the second ("B") term is ultimately seen to vanish. In $TVaR$ with fixed threshold, both terms are important.

Appendix 4. Reallocation of Co-measures
after Marginal Change

Even when the marginal change in a component is assigned to that component as its risk contribution, actually making a marginal change in one component can change the assignments to other components. This is investigated in this appendix.

We can look on derivative co-measures as functionals of the random variables. When we introduce a homogeneous perturbation in a particular variable, we can ask for the change not only in the risk measure itself, but in all the co-measures. In general these changes are not zero for co-measures for variables other than the variable of perturbation. In important special cases, the co-measure of the variable of perturbation gets more than the change of the entire measure.

A4.1. Formulation

We work with the techniques and conventions of Appendix 1. Introduce the notation $\rho(\varepsilon, k)$ for the risk measure with a homogeneous addition ε to X_k. Then $\rho_k = \lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} \rho(\varepsilon, k)$. We can now introduce a change in the first variable and look at the subsequent allocation to the kth variable $\rho_k(\varepsilon, 1)$. Because these are co-measures, we know that $\sum_{k=1}^{N} \rho_k(\varepsilon, 1) = \rho(\varepsilon, 1)$ and so

$$\sum_{k=1}^{N} \lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} \rho_k(\varepsilon, 1) = \lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} \rho(\varepsilon, 1) = \rho_1$$

We want to evaluate the marginal effect on the co-measures, namely $\lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} \rho_k(\varepsilon, 1)$. The naïve hope is that for $k \neq 1$ it is zero, and that for $k = 1$ we get ρ_1. In general this is not true. We will do some specific examples, and repeat some earlier work for clarity and for the sake of being self-contained here.

To clarify a bit, the naïve hope is fulfilled when the risk measure is $\rho = E[Y]$, and so it also holds when the risk measure is the expected value under transformed probabilities. For the mean, in the notation of this appendix, $\rho(\varepsilon, k) = E[Y + \varepsilon X_k]$ and taking the derivative wrt ε gives $\rho_k = E[X_k]$. The derivative of $\rho(\varepsilon, k)$ in the direction of any X_j, $j \neq k$, is just $E[X_j]$ and the derivative in the k direction is $E[X_k(1 + \varepsilon)]$. Thus $\rho_k(\varepsilon, 1) = E[X_k(1 + \varepsilon \delta_{1k})]$, using indicator $\delta_{1k} = (1$ for $k = 1$ and zero otherwise), and finally

$$\lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} \rho_k(\varepsilon, 1) = \delta_{1k} E[X_1]$$

This will also be true in the more general case for which $\rho = E[h(Y)]$ with $h(Y)$ a linear functional. In what follows, the expression for $\lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} \rho_k(\varepsilon, 1)$ always begins with $\delta_{1k} \rho_1$ but has other terms.

A4.2. Standard Deviation

The measure is

$$\rho = \left\{ E\left[(Y - \mu)^2 \right] \right\}^{1/2} = \left\{ E\left[\iint \left(\sum_{n=1}^{N} x_n - \mu \right)^2 f(\mathbf{x}) \, d\mathbf{x} \right] \right\}^{1/2}$$

When we introduce a change in the k direction, we get

$$\rho(\varepsilon, k) = \left\{ \iint \left(\varepsilon x_k + \sum_{n=1}^{N} x_n - \mu - \varepsilon \mu_k \right)^2 f(\mathbf{x}) \, d\mathbf{x} \right\}^{1/2}$$

The derivative co-measures are, as before,

$$\rho_k = \lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} \rho(\varepsilon, k) = \frac{\iint (x_k - \mu_k) \left(\sum_{n=1}^{N} x_n - \mu \right) f(\mathbf{x}) \, d\mathbf{x}}{\rho}$$

$$= \mathrm{Cov}(X_k, Y) / \mathrm{Std}(Y)$$

Now, let us consider an increment in the first direction. Its allocation to the kth component is the co-measure just above but with the first variable increased by a factor of ε.

$$\rho_k(\varepsilon, 1) = \frac{\iint (x_k - \mu_k)(1 + \varepsilon \delta_{1k}) \left(\varepsilon x_1 + \sum_{n=1}^{N} x_n - \mu - \varepsilon \mu_1 \right) f(\mathbf{x}) \, d\mathbf{x}}{\rho(\varepsilon, 1)}$$

We still have explicitly that the co-measures add to the measure.

$$\sum_{k=1}^{N} \rho_k(\varepsilon, 1) = \rho(\varepsilon, 1)$$

We want to look at the change in the co-measures.

$$\lim_{\varepsilon \to 0} \frac{d\rho_k(\varepsilon, 1)}{d\varepsilon} = \frac{\delta_{1k} \iint (x_1 - \mu_1) \left(\sum_{n=1}^{N} x_n - \mu \right) f(\mathbf{x}) \, d\mathbf{x}}{\rho}$$

$$+ \frac{\iint (x_k - \mu_k)(x_1 - \mu_1) f(\mathbf{x}) \, d\mathbf{x}}{\rho} - \frac{\iint (x_k - \mu_k) \left(\sum_{n=1}^{N} x_n - \mu \right) f(\mathbf{x}) \, d\mathbf{x}}{\rho^2} \rho_1$$

$$= \left(\delta_{1k} - \frac{\rho_k}{\rho} \right) \rho_1 + \frac{\iint (x_k - \mu_k)(x_1 - \mu_1) f(\mathbf{x}) \, d\mathbf{x}}{\rho}$$

$$= \delta_{1k} \rho_1 + \frac{\mathrm{Var}(Y) \, \mathrm{Cov}(X_k, X_1) - \mathrm{Cov}(X_k, Y) \, \mathrm{Cov}(X_1, Y)}{\mathrm{Std}(Y)^3}$$

If the first term were all that existed, the naïve expectation would be fulfilled. We may also note that when $k = 1$ the second term is intrinsically positive, unless X_1 correlates perfectly with Y in which case it is zero.

Thus, the decomposition in the perturbation variable receives more than the change in the overall risk measure. That is, its new allocation is greater than its marginal impact. Since the sum over the second terms is zero, some of the other decompositions must be negative.

A4.3. XTVaR

The risk measure is

$$\rho = E\left[Y - c\mu \,\middle|\, Y > y(\alpha)\right]$$

$$= \frac{1}{\alpha} \iint \left(\sum_{n=1}^{N} x_n - c\mu \right) \Theta\left(\sum_{n=1}^{N} x_n - y(\alpha) \right) f(\mathbf{x})\, d\mathbf{x}$$

with c and α fixed and $y(\alpha)$ implicitly defined by

$$\alpha = \iint \Theta\left(\sum_{n=1}^{N} x_n - y(\alpha) \right) f(\mathbf{x})\, d\mathbf{x}$$

When we introduce an amount in the k variable, we get

$$\rho(\varepsilon, k) = \frac{1}{\alpha} \iint \left(\varepsilon x_k + \sum_{n=1}^{N} x_n - c\mu - c\varepsilon\mu_k \right) \Theta\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - y(\alpha, \varepsilon, k) \right) f(\mathbf{x})\, d\mathbf{x}$$

and

$$\alpha = \iint \Theta\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - y(\alpha, \varepsilon, k) \right) f(\mathbf{x})\, d\mathbf{x}$$

As in Appendix 1, we can get

$$\frac{\partial}{\partial \varepsilon} y(\alpha, 0, k) = \frac{\iint \delta\left(\sum_{n=1}^{N} x_n - y(\alpha) \right) x_k f(\mathbf{x})\, d\mathbf{x}}{\iint \delta\left(\sum_{n=1}^{N} x_n - y(\alpha) \right) f(\mathbf{x})\, d\mathbf{x}} = E\left[X_k \,\middle|\, Y = y(\alpha)\right] \equiv X_k^\alpha$$

This is the mean value of X_k on the hyperplane of fixed Y, which we will use later and have denoted by X_k^α. The derivative co-measures are

$$\rho_k = \lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} \rho(\varepsilon, k)$$

$$= \frac{1}{\alpha} \iint (x_k - c\mu_k) \Theta\left(\sum_{n=1}^{N} x_n - y(\alpha) \right) f(\mathbf{x})\, d\mathbf{x}$$

$$+ \frac{(y(\alpha) - c\mu)}{\alpha} \iint \delta\left(\sum_{n=1}^{N} x_n - y(\alpha)\right)\left(x_k - \frac{\partial}{\partial \varepsilon} y(\alpha, 0, k)\right) f(\mathbf{x})\, d\mathbf{x}$$

$$= \frac{1}{\alpha} \iint (x_k - c\mu_k)\, \Theta\left(\sum_{n=1}^{N} x_n - y(\alpha)\right) f(\mathbf{x})\, d\mathbf{x}$$

$$= E[X_k - c\mu_k | Y = y(\alpha)]$$

The second term is proportional to the derivative of α which is of course zero. These are the results we had earlier.

Now, let us again consider an increment in the first variable.

$$\rho_k(\varepsilon, 1) = \frac{1}{\alpha} \iint (x_k - c\mu_k)(1 + \varepsilon\delta_{1k})\, \Theta\left(\varepsilon x_1 + \sum_{n=1}^{N} x_n - y(\alpha, \varepsilon, 1)\right) f(\mathbf{x})\, d\mathbf{x}$$

$$\alpha = \iint \Theta\left(\varepsilon x_1 + \sum_{n=1}^{N} x_n - y(\alpha, \varepsilon, 1)\right) f(\mathbf{x})\, d\mathbf{x}$$

We still have explicitly $\sum_{k=1}^{N} \rho_k(\varepsilon, 1) = \rho(\varepsilon, 1)$.

Again, we go to the limit of small ε. For $k \neq 1$

$$\lim_{\varepsilon \to 0} \frac{d\rho_k(\varepsilon, 1)}{d\varepsilon} = \frac{1}{\alpha} \iint (x_k - c\mu_k)\, \delta\left(\sum_{n=1}^{N} x_n - y(\alpha)\right)\left(x_1 - \lim_{\varepsilon \to 0} \frac{\partial y(\alpha, \varepsilon, 1)}{\partial \varepsilon}\right) f(\mathbf{x})\, d\mathbf{x}$$

$$= \frac{1}{\alpha} \iint (x_k - c\mu_k)\left(x_1 - X_1^\alpha\right) \delta\left(\sum_{n=1}^{N} x_n - y(\alpha)\right) f(\mathbf{x})\, d\mathbf{x}$$

$$= \frac{f_Y(y(\alpha))}{\alpha} E\left[(X_k - c\mu_k)\left(x_1 - X_1^\alpha\right) \Big| Y = y(\alpha)\right]$$

$$= \frac{f_Y(y(\alpha))}{\alpha} E\left[X_k\left(X_1 - X_1^\alpha\right) \Big| Y = y(\alpha)\right]$$

where $f_Y(y) \equiv \iint \delta\left(\sum_{n=1}^{N} x_n - y\right) f(\mathbf{x})\, d\mathbf{x}$ is the density function on the total and we have used $E\left[X_1 - X_1^\alpha | Y = y(\alpha)\right] = 0$. These are not apparently zero.

$$\lim_{\varepsilon \to 0} \frac{d\rho_1(\varepsilon, 1)}{d\varepsilon} = \frac{1}{\alpha} \iint (x_1 - c\mu_1)\, \Theta\left(\sum_{n=1}^{N} x_n - y(\alpha)\right) f(\mathbf{x})\, d\mathbf{x}$$

$$+ \frac{1}{\alpha} \iint (x_1 - c\mu_1)\, \delta\left(\sum_{n=1}^{N} x_n - y(\alpha)\right)\left(x_1 - \lim_{\varepsilon \to 0} \frac{\partial y(\alpha, \varepsilon, 1)}{\partial \varepsilon}\right) f(\mathbf{x})\, d\mathbf{x}$$

$$= \rho_1 + \frac{1}{\alpha} \iint (x_1 - c\mu_1)\left(x_1 - X_1^\alpha\right) \delta\left(\sum_{n=1}^{N} x_n - y(\alpha)\right) f(\mathbf{x})\, d\mathbf{x}$$

$$= \rho_1 + \frac{f_Y(y(\alpha))}{\alpha} E\left[(X_1 - c\mu_1)(X_1 - X_1^\alpha) \,\middle|\, Y = y(\alpha)\right]$$

$$= \rho_1 + \frac{f_Y(y(\alpha))}{\alpha} E\left[(X_1 - X_1^\alpha)^2 \,\middle|\, Y = y(\alpha)\right]$$

We may write both forms as

$$\lim_{\varepsilon \to 0} \frac{d\rho_k(\varepsilon, 1)}{d\varepsilon} = \rho_1 \delta_{1k} + \frac{f_Y(y(\alpha))}{\alpha} E\left[(X_k - X_k^\alpha)(X_1 - X_1^\alpha) \,\middle|\, Y = y(\alpha)\right]$$

We may interpret the additional terms as proportional to a covariance matrix limited to the hyperplane. It is not obvious how large these terms may be.

The notable point again is that the addition to ρ_1 beyond the prior allocation is intrinsically non-negative. This means that also under an *XTVaR* allocation a component gets more than its marginal contribution to the previous risk.

A4.4. Function of $Y/E[Y]$

Take L is a smooth integrable function of the ratio $\zeta \equiv \frac{Y}{E[Y]}$. The risk measure is

$$\rho = E[h(Y)L(\zeta)]$$

with $h(Y)$ linear. Then explicitly

$$\rho(\varepsilon, k) = \iint h\left(\varepsilon x_k + \sum_{n=1}^N x_n\right) L\left(\frac{\varepsilon x_k + \sum_{n=1}^N x_n}{\iint \left(\varepsilon z_k + \sum_{n=1}^N z_n\right) f(\vec{z}) \, d\vec{z}}\right) f(\mathbf{x}) \, d\mathbf{x}$$

We know

$$\lim_{\varepsilon \to 0} \frac{d\zeta(\varepsilon, k)}{d\varepsilon} = \frac{E[Y]X_k - YE[X_k]}{E[Y]^2} = \frac{\mu X_k - Y\mu_k}{\mu^2} = \frac{X_k - \zeta\mu_k}{\mu}$$

where $\mu \equiv \iint \left(\sum_{n=1}^N x_n\right) f(\mathbf{x}) \, d\mathbf{x}$ and $\mu_k \equiv \iint x_k f(\mathbf{x}) \, d\mathbf{x}$.

The derivative co-measure is

$$\rho_k = \lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} \rho(\varepsilon, k) = E[h(X_k)L(\zeta)]$$

$$+ \frac{1}{E[Y]^2} E\left[h(Y)L'(\zeta)\{E[Y]X_k - YE[X_k]\}\right]$$

$$= E[h(X_k)L(\zeta)] + E\left[h(Y)L'(\zeta)\frac{X_k - \zeta\mu_k}{\mu}\right]$$

$$= \iint h(x_k)L\left(\frac{\sum_{n=1}^{N} x_n}{\mu}\right)f(\mathbf{x})\,d\mathbf{x}$$

$$+ \iint \frac{h\left(\sum_{n=1}^{N} x_n\right)}{\mu^2}L'\left(\frac{\sum_{n=1}^{N} x_n}{\mu}\right)\left\{\mu x_k - \mu_k \sum_{n=1}^{N} x_n\right\}f(\mathbf{x})\,d\mathbf{x}$$

Now again add an increment in the first variable. The co-measures for finite ε become

$$\rho_k(\varepsilon, 1) = \iint h\left((1 + \varepsilon\delta_{k1})x_k\right)L\left(\frac{\varepsilon x_1 + \sum_{n=1}^{N} x_n}{\mu(\varepsilon, 1)}\right)f(\mathbf{x})\,d\mathbf{x}$$

$$+ \iint \frac{h\left(\varepsilon x_1 + \sum_{n=1}^{N} x_n\right)}{\mu(\varepsilon, 1)^2}L'\left(\frac{\varepsilon x_1 + \sum_{n=1}^{N} x_n}{\mu(\varepsilon, 1)}\right)\left\{\begin{aligned}&\mu(\varepsilon, 1)(1 + \varepsilon\delta_{k1})x_k\\&-\mu_k(\varepsilon, 1)\left[\varepsilon x_1 + \sum_{n=1}^{N} x_n\right]\end{aligned}\right\}f(\mathbf{x})\,d\mathbf{x}$$

We also have

$$\mu(\varepsilon, 1) \equiv \iint\left(\varepsilon x_1 + \sum_{n=1}^{N} x_n\right)f(\mathbf{x})\,d\mathbf{x} = \varepsilon\mu_1 + \mu$$

and

$$\mu_k(\varepsilon, 1) \equiv \iint(1 + \varepsilon\delta_{k1})x_k\,f(\mathbf{x})\,d\mathbf{x} = (1 + \varepsilon\delta_{k1})\mu_k$$

Putting these last two forms in, we have

$$\rho_k(\varepsilon, 1) = \iint h\left((1 + \varepsilon\delta_{k1})x_k\right)L\left(\frac{\varepsilon x_1 + \sum_{n=1}^{N} x_n}{\mu + \varepsilon\mu_1}\right)f(\mathbf{x})\,d\mathbf{x}$$

$$+ \iint \frac{h\left(\varepsilon x_1 + \sum_{n=1}^{N} x_n\right)}{(\mu + \varepsilon\mu_1)^2}L'\left(\frac{\varepsilon x_1 + \sum_{n=1}^{N} x_n}{\mu + \varepsilon\mu_1}\right)(1 + \varepsilon\delta_{k1})\left\{\begin{aligned}&(\mu + \varepsilon\mu_1)x_k\\&-\mu_k\left[\varepsilon x_1 + \sum_{n=1}^{N} x_n\right]\end{aligned}\right\}f(\mathbf{x})\,d\mathbf{x}$$

If we sum these over k, we are not surprised to get an exact and complete cancellation from the second line and

$$\sum_{k=1}^{N} \rho_k(\varepsilon, 1) = \iint h\left(\varepsilon x_1 + \sum_{n=1}^{N} x_n\right) L\left(\frac{\varepsilon x_1 + \sum_{n=1}^{N} x_n}{(\mu + \varepsilon \mu_1)}\right) f(\mathbf{x}) \, d\mathbf{x} = \rho(\varepsilon, 1)$$

as it should.

We may write the co-measure as an expectation

$$\rho_k(\varepsilon, 1) = E\left[h\left((1 + \varepsilon \delta_{k1}) X_k\right) L(\zeta(\varepsilon, 1))\right]$$

$$+ (1 + \varepsilon \delta_{k1}) E\left[\frac{h(\varepsilon X_1 + Y)}{(\mu + \varepsilon \mu_1)^2} L'(\zeta(\varepsilon, 1)) \left\{(\mu + \varepsilon \mu_1) X_k - \mu_k [\varepsilon X_1 + Y]\right\}\right]$$

Then

$$\lim_{\varepsilon \to 0} \frac{d\rho_k(\varepsilon, 1)}{d\varepsilon} = \delta_{k1} E\left[h(X_1) L(\zeta)\right] + E\left[h(X_k) L'(\zeta) \frac{X_1 - \zeta \mu_1}{\mu}\right]$$

$$+ \delta_{k1} E\left[h(Y) L'(\zeta) \frac{X_1 - \zeta \mu_1}{\mu}\right]$$

$$+ \lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} E\left[h(\varepsilon X_1 + Y) L'(\zeta(\varepsilon, 1)) \frac{(\mu + \varepsilon \mu_1) X_k - \mu_k [\varepsilon X_1 + Y]}{(\mu + \varepsilon \mu_1)^2}\right]$$

$$= \delta_{k1} \rho_k + E\left[h(X_k) L'(\zeta) \frac{X_1 - \zeta \mu_1}{\mu}\right] + E\left[h(X_1) L'(\zeta) \frac{X_k - \zeta \mu_k}{\mu}\right]$$

$$+ E\left[h(Y) L''(\zeta) \frac{X_1 - \zeta \mu_1}{\mu} \frac{X_k - \zeta \mu_k}{\mu}\right]$$

$$+ E\left[h(Y) L'(\zeta) \left\{\frac{X_k \mu_1 - X_1 \mu_k}{\mu^2} - 2 \frac{X_k - \zeta \mu_k}{\mu^2} \mu_1\right\}\right]$$

$$= \delta_{k1} \rho_k + E\left[\frac{L'(\zeta)}{\mu} \begin{cases} h(X_k)(X_1 - \zeta \mu_1) + h(X_1)(X_k - \zeta \mu_k) \\ + h(\zeta)(2 \zeta \mu_k \mu_1 - X_k \mu_1 - X_1 \mu_k) \end{cases}\right]$$

$$+ E\left[h(\zeta) \frac{L''(\zeta)}{\mu} (X_1 - \zeta \mu_1)(X_k - \zeta \mu_k)\right]$$

Looking at the last form, it is at the very least not obvious that these derivatives for $k \neq 1$ are zero.

For $k = 1$

$$\lim_{\varepsilon \to 0} \frac{d\rho_1(\varepsilon, 1)}{d\varepsilon} = \rho_1 + E\left[2\frac{h(X_1) - \mu_1 h(\zeta)}{\mu} L'(\zeta)(X_1 - \zeta\mu_1)\right]$$
$$+ E\left[\frac{h(\zeta)}{\mu} L''(\zeta)(X_1 - \zeta\mu_1)^2\right]$$

If the whole change in ρ were going to be in ρ_1 then we would want the last terms to be zero. We cannot say more without specifying the forms of h and L, but generally speaking it would be surprising if the addition were not positive.

We may note that if $h(X) = aX$ then

$$\lim_{\varepsilon \to 0} \frac{d\rho_1(\varepsilon, 1)}{d\varepsilon} = \rho_1 + \frac{a}{\mu} E\left[(X_1 - \zeta\mu_1)^2 (2L'(\zeta) + \zeta L''(\zeta))\right]$$

We could take $L(\zeta) = A - B/\zeta$ so that $2L'(\zeta) + \zeta L''(\zeta) = 0$ and then the whole change in ρ is in ρ_1, but this is a fairly restrictive and probably useless choice. We can note that

$$2L'(\zeta) + \zeta L''(\zeta) = \frac{1}{\zeta}\frac{d}{d\zeta}\left(\zeta^2 L'(\zeta)\right)$$

so that if $\zeta^2 L'(\zeta)$ is an increasing function and $\zeta > 0$ this is always positive.

If we take $h(X) = X - c\mu_x$ then

$$\lim_{\varepsilon \to 0} \frac{d\rho_1(\varepsilon, 1)}{d\varepsilon} = \rho_1 + E\left[\frac{(X_1 - \zeta\mu_1)^2}{\mu}\{2L'(\zeta) + (\zeta - c)L''(\zeta)\}\right]$$

A4.5. Function of $F(Y)$

The risk measure is, for η a smooth function

$$\rho = E[Y\eta\{F(Y)\}]$$
$$= \iint \sum_{n=1}^{N} x_n \eta\left\{\iint \Theta\left(\sum_{n=1}^{N} x_n - \sum_{n=1}^{N} z_n\right) f(\mathbf{z})\, d\mathbf{z}\right\} f(\mathbf{x})\, d\mathbf{x}$$

When we introduce an amount in the k variable, we get

$$\rho(\varepsilon, k) = \iint\left(\varepsilon x_k + \sum_{n=1}^{N} x_n\right)\eta\left\{\iint \Theta\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - \sum_{n=1}^{N} z_n - \varepsilon z_k\right) f(\mathbf{z})\, d\mathbf{z}\right\} f(\mathbf{x})\, d\mathbf{x}$$

The presence of the term in z_k comes from going back to the basic notion of the change being an additional variable with a delta function connection to the

kth variable. The notation is rapidly becoming difficult, so we introduce auxiliary functions. Let us again define the mean value of a variable on the hyperplane given by the total being fixed. In section 2 we defined this for fixed survivor function value.

$$\mu_k(y) \equiv \iint x_k \, \delta\left(y - \sum_{n=1}^{N} x_n\right) f(\mathbf{x}) \, d\mathbf{x}$$

We will also again need the density function for the total, which is the mean value of 1 on the same hyperplane:

$$f_Y(y) \equiv \iint \delta\left(y - \sum_{n=1}^{N} x_n\right) f(\mathbf{x}) \, d\mathbf{x}$$

The cumulative distribution function is as used above

$$F(y) = \iint \Theta\left(y - \sum_{n=1}^{N} x_n\right) f(\mathbf{x}) \, d\mathbf{x}$$

We may write the risk measure as

$$\rho = \int y \eta\{F(y)\} f_Y(y) \, dy$$

The derivative co-measures are

$$\rho_k = \lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} \rho(\varepsilon, k)$$

$$= \iint x_k \, \eta\left\{F\left(\sum_{n=1}^{N} x_n\right)\right\} f(\mathbf{x}) \, d\mathbf{x}$$

$$+ \iint \left(\sum_{n=1}^{N} x_n\right) \eta'\left\{F\left(\sum_{n=1}^{N} x_n\right)\right\} \left\{\iint (x_k - z_k) \delta\left(\sum_{n=1}^{N} x_n - \sum_{n=1}^{N} z_n\right) f(\mathbf{z}) \, d\mathbf{z}\right\} f(\mathbf{x}) \, d\mathbf{x}$$

$$= \iint x_k \, \eta\left\{F\left(\sum_{n=1}^{N} x_n\right)\right\} f(\mathbf{x}) \, d\mathbf{x}$$

$$+ \iint \left(\sum_{n=1}^{N} x_n\right) \eta'\left\{F\left(\sum_{n=1}^{N} x_n\right)\right\} \left\{x_k \, f_Y\left(\sum_{n=1}^{N} x_n\right) - \mu_k\left(\sum_{n=1}^{N} x_n\right)\right\} f(\mathbf{x}) \, d\mathbf{x}$$

However,

$$\iint \left(\sum_{n=1}^{N} x_n\right) \eta'\left\{F\left(\sum_{n=1}^{N} x_n\right)\right\} x_k \, f_Y\left(\sum_{n=1}^{N} x_n\right) f(\mathbf{x}) \, d\mathbf{x}$$

$$= \int y \eta'\{F(y)\} f_Y(y) \left[\iint \delta\left(y - \sum_{n=1}^{N} x_n\right) x_k \, f(\mathbf{x}) \, d\mathbf{x}\right] dy$$

$$= \int y \eta'\{F(y)\} f_Y(y) \mu_k(y) \, dy$$

and

$$\iint \left(\sum_{n=1}^{N} x_n\right) \eta' \left\{F\left(\sum_{n=1}^{N} x_n\right)\right\} \mu_k \left(\sum_{n=1}^{N} x_n\right) f(\mathbf{x})\, d\mathbf{x}$$

$$= \int y \eta'\{F(y)\} \mu_k(y) \left[\iint \delta\left(y - \sum_{n=1}^{N} x_n\right) f(\mathbf{x})\, d\mathbf{x}\right] dy$$

$$= \int y \eta'\{F(y)\} \mu_k(y) f_Y(y)\, dy$$

so that the last two terms sum to zero and we have only

$$\rho_k = \iint x_k\, \eta \left\{F\left(\sum_{n=1}^{N} x_n\right)\right\} f(\mathbf{x})\, d\mathbf{x} = \int \eta\{F(y)\} \mu_k(y)\, dy$$

Since

$$\sum_{k=1}^{N} \mu_k(y) = \iint \sum_{k=1}^{N} x_k\, \delta\left(y - \sum_{n=1}^{N} x_n\right) f(\mathbf{x})\, d\mathbf{x} = y f_Y(y)$$

we clearly have $\sum_{k=1}^{N} \rho_k = \rho$.

Now, let us again consider an increment in the first variable.

$$\rho_k(\varepsilon, 1) = \iint x_k (1 + \varepsilon \delta_{1k}) \eta \left\{F(\varepsilon, 1)\left(\varepsilon x_1 + \sum_{n=1}^{N} x_n\right)\right\} f(\mathbf{x})\, d\mathbf{x}$$

$$F(\varepsilon, 1)(y) = \iint \Theta\left(y - \sum_{n=1}^{N} z_n - \varepsilon z_1\right) f(\mathbf{z})\, d\mathbf{z}$$

We still have explicitly $\sum_{k=1}^{N} \rho_k(\varepsilon, 1) = \rho(\varepsilon, 1)$. We will need the covariance-similar function

$$\mu_{km}(y) \equiv \iint x_k x_m\, \delta\left(y - \sum_{n=1}^{N} x_n\right) f(\mathbf{x})\, d\mathbf{x}$$

$$= E\left[X_k X_m \mid Y = y\right]$$

Again, we go to the limit of small ε.

$$\lim_{\varepsilon \to 0} \frac{d\rho_k(\varepsilon, 1)}{d\varepsilon}$$

$$= \delta_{1k} \rho_1 + \iint x_k\, \eta'\left\{F\left(\sum_{n=1}^{N} x_n\right)\right\}\left\{\iint (x_1 - z_1)\, \delta\left(\sum_{n=1}^{N} x_n - \sum_{n=1}^{N} z_n\right) f(\mathbf{z})\, d\mathbf{z}\right\} f(\mathbf{x})\, d\mathbf{x}$$

$$= \delta_{1k} \rho_1 + \iint x_k\, \eta'\left\{F\left(\sum_{n=1}^{N} x_n\right)\right\}\left\{x_1 f_Y\left(\sum_{n=1}^{N} x_n\right) - \mu_1\left(\sum_{n=1}^{N} x_n\right)\right\} f(\mathbf{x})\, d\mathbf{x}$$

$$= \delta_{1k}\rho_1 + \int f_Y(y)\,\eta'\{F(y)\} \iint x_k\, x_1 \delta\left(y - \sum_{n=1}^{N} x_n\right) f(\mathbf{x})\, d\mathbf{x}\, dy$$

$$- \int \mu_1(y)\,\eta'\{F(y)\}\, \mu_k(y)\, dy$$

$$= \delta_{1k}\rho_1 + \int f_Y(y)\,\eta'\{F(y)\}\{f_Y(y)\,\mu_{k1}(y) - \mu_k(y)\,\mu_1(y)\}\, dy$$

For $k = 1$ the curly bracket in the second term is

$$E[1\,|\,Y = y]\, E[(X_1)^2\,|\,Y = y] - E[X_1\,|\,Y = y]^2$$

and this is always positive. If η is monotonic increasing, then the contribution is positive.

REFERENCES

ACERBI, C. (2002) Spectral measures of risk: A coherent representation of subjective risk aversion, *Journal of Banking & Finance*, **26**.

ARTZNER, P., DELBAEN, F., EBER, J.-M. and HEATH, D. (1999) Coherent measures of risk, *Mathematical Finance*, **9(3)**, 203-228.

BALLOTTA, L. (2004) Alternative framework for the fair valuation of life insurance contracts: a Lévy process based model, *AFIR Colloquium*.

KALKBRENER, M. (2005) An Axiomatic Approach to Capital Allocation, *Mathematical Finance*, July.

KREPS, R. (2005) Riskiness Leverage Models, *Proceedings of the Casualty Actuarial Society*.

KUSUOKA, S. (2001) On law invariant coherent risk measures, in: *Advances in Mathematical Economics*, **3**, 83-95, Springer: Tokyo.

MANGO, D.F. (2003) Capital Consumption: An Alternative Methodology for Pricing Reinsurance, *Casualty Actuarial Society Forum*, Winter, 351-378.

MERTON, R. and PEROLD, A. (1993) Theory of Risk Capital in Financial Firms, *Journal of Applied Corporate Finance*, Fall, 16-32.

MAJOR, J.A. (2004) Gradients of Risk Measures: Theory and Application to Catastrophe Risk Management and Reinsurance Pricing, *Casualty Actuarial Society Forum*, Winter.

MØLLER, T. (2003) Stochastic orders in dynamic reinsurance markets, *ASTIN Colloquium Papers*.

MYERS, S. and READ, J. (2001) Capital Allocation for Insurance Companies, *Journal of Risk and Insurance*, **68(4)**.

MILDENHALL, S. (2004) A Note on the Myers and Read Capital Allocation Formula, *North American Actuarial Journal*, **8(2)**.

PHILLIPS, R., CUMMINS, J.D. and ALLEN, F. (1998) Financial Pricing of Insurance in the Multiple-Line Insurance Company, *The Journal of Risk and Insurance*, **65(4)**, 597-636.

RUHM, D. (2003) Distribution-Based Pricing Formulas Are Not Arbitrage-Free, *Proceedings of the Casualty Actuarial Society*.

RUHM, D. and MANGO, D. (2003) A Risk Charge Calculation Based on Conditional Probability, *Thomas P. Bowles Jr. Symposium*, Georgia State University.

SHERRIS, M. (2004) *Solvency, Capital Allocation and Fair Rate of Return in Insurance*, UNSW working paper.

TASCHE, D. (2000) Risk contributions and performance measurement, *Report of the Lehrstuhl fur mathematische Statistik*, T.U. Munchen.

TASCHE, D. (2002) Expected shortfall and beyond, *Journal of Banking and Finance*, **26**, 1519-1533.

URYASEV, S. (1995a) Derivatives of Probability Functions and Integrals over Sets Given by Inequalities, *Journal of Computational and Applied Mathematics*, **56**, 197-223.

URYASEV, S. (1995b) Derivatives of Probability Functions and Some Applications, *Annals of Operations Research*, **56**, 287-311.

URYASEV, S. (1999) Derivatives of Probability and Integral Functions: General Theory and Examples, in *Encyclopedia of Optimization*, C.A. Floudas and P.M. Pardalos, eds., Kluwer Academic Publishers.

VENTER, G. (1991) Premium Calculation Implications of Reinsurance without Arbitrage, *ASTIN Bulletin*, **21(2)**.

VENTER, G. (2004) Capital Allocation Survey With Commentary, *North American Actuarial Journal*, **8(2)**.

VENTER, G., BARNETT, J. and OWEN, M. (2004) Market Value of Risk Transfer: Catastrophe Reinsurance Case, *AFIR Colloquium*.

GARY G. VENTER
E-mail: gary.g.venter@guycarp.com

JOHN A. MAJOR
E-mail: john.a.major@guycarp.com

RODNEY E. KREPS
E-mail: mrkreps@gmail.com