Additive Marginal Allocation of Risk Measures

Gary G Venter, John A Major, Rodney E Kreps

Abstract

Companies looking for risk adjusted profitability by business component often start by allocating capital to the components to provide the denominator for risk adjusted return. Typically the allocation is done by allocating a risk measure. This is consistent with the principle of basing decisions on marginal impacts only if the risk measure is allocated by a marginal method and the marginal impacts add up to the total risk measure. The method of co-measures always provides additive allocations, but these are not always marginal impacts. It will give marginal allocations when the risk measure is scalable and growth is homogeneous. There can be several co-measure allocations for the same risk measure, but only one has a marginal impact. For homogeneous growth, this one can be specified by a directional derivative. The first well known additive marginal allocations were the Myers-Read method and co-Tail Value at Risk. Now we see that with homogeneous growth there are many others. This allows the choice of risk measure to be based on economic meaning rather than the availability of an allocation method. When the growth method is not homogeneous, the situation is more complex. Risk measures based on transformed probabilities are shown to be marginal, additive and sometimes economically meaningful in this case.

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

Additive Marginal Allocation of Risk Measures

Insurers are finding it increasingly useful to have risk measures to quantify company risk and the contributions to it. Understanding total risk and its sources can help in performance measurement, strategic planning, pricing, and communication with regulators, rating agencies, and security analysts. We focus here on breaking down risk measures by the business components that drive them. One application of this is allocation of capital, which we discuss in some detail but do not particularly advocate. Other ways to compute component performance can also benefit from understanding the contributions to company risk.

Sometimes a preferred 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 be done in a marginal additive fashion. 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 capital is 3.8 time 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.

More abstractly, allocation 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. Allocating this risk measure to the components X_j requires finding a function r(X) so that $\rho(Y) = \Sigma r(X_j)$. For an insurer, Y could be total claims, or it could be change in capital, underwriting profit, etc. 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.

Section 1 describes the properties we focus on: additivity and marginal impact. Marginal allocation is easier if a company can make homogeneous changes to its business composition. This is discussed in section 2. Section 3 describes the allocation methods used: co-measures and the directional derivative, and discusses when these give additive and marginal allocations for the case of homogeneous changes. Section 4 provides a few examples. Section 5 goes into detail for risk measures that are related to quantification of insolvency. This may not be practical for many companies in that insolvency is difficult to quantify, and this section is not needed for the rest of the paper. Section 6 considers the non-homogeneous case, where growth requires adding discrete exposure units that may change the shape of the distribution. Additivity is not a problem here but marginal impact can be. This is different for different risk measures: some are far from marginal, some are close, and some are introduced that are additive and marginal even in that case. Section 7 considers performance measurement without allocating capital, and section 8 concludes. Two appendices go into calculation issues for directional derivatives.

1. Two Desirable Properties of Allocation Procedures

Two useful properties of allocation procedures are additivity and marginal impact.

An allocation procedure is additive if it produces allocations that sum up to the total risk measure without an adjustment factor for off-balance. (The off-balance is often called either "the benefit of pooling" or "fixed costs," depending on whether the adjustment factor is less than or greater than unity, but neither case is considered additive here.) For instance, the covariances of X_j with Y sum over the j's to the variance of Y, so allocation by covariance is an additive procedure. 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 an additive procedure absent independence, even though it can be forced to add up with an off-balance factor.

An allocation procedure produces a marginal impact if the change in the overall risk measure from a small change to the volume of a component is all allocated to that component. The covariance allocation of variance does not meet this criterion, as the change in overall variance from increasing a component by a factor can be shown to be double the change in the covariance of the component with the total.

Both properties are desirable. Additivity is necessary to decompose the risk measure entirely to the component pieces. Marginal impact is important if you want to make decisions about the effect of a change in a component to the overall measure. For instance if capital is allocated by a risk measure in order to calculate risk-adjusted return by line, then growing a line with higher than average return theoretically requires more capital. But since all that capital will be allocated to that line, which has higher return on capital, the average return on capital for the whole company will increase. Thus marginal allocation leads to appropriate decisions.

We consider two allocation approaches: co-measures and directional derivatives. Co-measures provide additive allocations of most risk measures, and those allocations are in some way an expression of the contribution of the component to the risk measure for the whole company. However they are not unique – a risk measure will have different allocations by different co-measures. Also these may or may not produce marginal impacts.

Directional derivatives, defined more precisely below, produce marginal impacts in many cases, but they may or may not be additive. For instance the directional derivative of the variance is twice the covariance, which is the marginal impact, but these sum to twice the variance.

2. Homogeneous Growth

Grow 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, so 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. Buying a negative quota share can produce positive growth.

A negative quota share is like selling short. A reinsurer who wants to bet against a company could pay it a proportion of its annual premiums and receive an equal share of its losses. An insurer or reinsurer that writes only shares of syndicated programs can achieve homogeneous growth in a program by increasing its shares.

However often if a company aims to have positive growth it can 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 adding a quota share. So marginal allocation 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 used for determining additive marginal allocations are co-measures and the directional derivative. Both of these are used to specify components' contributions to risk measures. If the risk measure quantifies something meaningful to the whole company, then management would like to know how much each component contributes to the risk measure. For this the risk measure applied to the component separately, or to the company less the component, is not usually relevant. The focus is on the impact the component is having on the total risk measure. Properties of risk measures such as coherence are also irrelevant in this process.

Co-measures provide a way to define $r(X)$ when $\rho(Y)$ can be 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.

In this set-up, 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 $\rho(Y) = \Sigma r(X_j)$.

Co-measures are not necessarily unique, as there may be different conditions and h and L functions that end up defining the same risk measure. In fact, given a risk measure defined by single L and h functions L_1, h_1, and another additive function h_2, setting $L_2 = L_1h_1/h_2$ gives $L_2h_2= L_1h_1$, 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).

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

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

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

The j^{th} directional derivative is defined as:

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

Euler essentially showed that this derivative produces an additive allocation of the risk measure if ρ is scalable[1], i.e., ρ[aY] = aρ[Y]. The overall approach here is to find additive marginal allocations of scalable risk measures using the directional derivative, then find co-measures which produce the same allocations. Co-measures are still additive for non-homogeneous growth, and often they are also still quite close to marginal, as discussed 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 technical details of the math work out better 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%. L'Hopital's rule also has some continuity assumptions.

For XTVaR, ρ(Y) = E[(Y – EY)|Y>b] is not scalable if b is a fixed constant amount. But if b is a percentile of Y, then it is. Multiplying Y by a constant increases EY and

---

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

every percentile of Y by the same factor. You can write XTVaR as

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

$\rho(Y+\varepsilon X_i) \quad = E[(Y + \varepsilon X_j - EY - \varepsilon EX_j) | F(Y+\varepsilon X_j) > 1 - \alpha]$ 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.

4. Examples of Additive Marginal Allocation of Risk Measures

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:

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

$r(X_j) = \text{Cov}(X_j, Y) = E[(X - EX_j)(Y - EY)]$

and

$\rho(Y) = \text{VaR}_\alpha(Y) = E[Y | F(Y) = 1 - \alpha]$

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

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

When there are alternative intuitively reasonable definitions of a risk measure as a conditional expected value, then the derivative could help determine the preferred allocation. For instance, there are different ways to use co-measures to allocate 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 $$r(X_j) = E[X_jStd(Y)/EY] = Std(Y)EX_j/EY$$

This just allocates 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 allocates 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 allocation. Taking the derivative of $Std(Y+\varepsilon X_j) = [VarY+2\varepsilon Cov(X_j,Y)+\varepsilon^2 Var(X_j)]^{1/2}$ gives at $\varepsilon = 0$

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

This agrees with the second form of the co-measure, so this is the marginal allocation. Thus the total change in Std(Y) brought about by a small change in X_j would be allocated to j by this procedure.

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. Another useful alternative is exponential moments, when these exist.

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

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

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

$$\rho(Y) = E(Ye^{cY/EY}) + E[cY^2e^{cY/EY}/EY] - E[cY^2e^{cY/EY}/EY] \text{ and}$$

$$r(X_j) = E(X_je^{cY/EY}) + cE(X_jYe^{cY/EY})/EY - cEX_jE[Y^2e^{cY/EY}]/(EY)^2$$

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

Marginal co-measures of risk measures that are functions of Y/EY are discussed in greater generality at the end 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 risk less the mean as a fraction of capital. This can use 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 get a negative allocation. That is reasonable but may not meet all needs for allocated capital. This measure is related to the Esscher transform and minimum entropy pricing, so is connected to estimates of the market value of the risk taken.

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

5. Insolvency Related Risk Measures

Risk measures associated with insolvency are of theoretical interest in that insolvency is an economically meaningful event. However they are not very practical because the probability of insolvency is difficult to measure in practice. A few such measures are reviewed here for their historical and theoretical interest. Efforts at quantification appear worth continuing in order to bridge theory and practice.

If B is capital (or book value) and Y is the negative of profit[2], expected policyholder deficit, or EPD is:

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

This is not scalable. The value of the default put option is EPD under a risk-adjusted probability distribution that reflects the market value of the default:

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

EPD or the default put can be written as a co-measure in several ways. Basically for any a, you can set $h(X) = X - aEX$ and $L(Y) = \Pr(Y>B)(Y - B)/(Y - aEY)$. Then:

$$r(X_j) = E[(X_j - aEX_j)L(Y) \mid Y>B)$$

$$= \Pr(Y>B)E[(Y - B)(X_j - aEX_j)/(Y - aEY) \mid Y>B]$$

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 the 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:

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

Here j gets its own expected shortfall in company failure, which is a look at 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 chance of insolvency. Sherris (2004) uses such an allocation with transformed probabilities to allocate the default put option to customer.

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

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

This amounts to expressing the EPD risk measure as:

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

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 allocation, 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 allocation can be viewed as first allocating capital by expected value, so $B_j = BEX_j/EY$. Then X_j's share of EPD is $Pr(Y>B)E[(X_j - B_j) \mid Y>B]$. This could be negative if component j does not use up its allocated capital in the average of cases where

---

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

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 exclude 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 allocation 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 allocate 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 allocation 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.

This is similar to Meyers-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 allocation 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 ρ(Y), which is defined implicitly by:

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

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>ρ(Y)]E*[Y– ρ(Y)|Y>ρ(Y)] = aEC, which can be solved for ρ(Y) to yield the expression:

capital = ρ(Y) = E*[Y|Y>ρ(Y)] – aEC/Pr*[Y>ρ(Y)].

Here aEC is the value of the default put, so aEC/Pr*[Y>ρ(Y)] is the value of the conditional expected shortfall given that there is a default. The most direct co-measure allocation of this is aED_j/Pr*[Y>ρ(Y)], which is j's portion of the conditional expected shortfall, proportional to expected losses. The similar allocation of E*[Y|Y>ρ(Y)] is E*[X_j|Y>ρ(Y)]. The difference then allocates 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 allocation from the directional derivative, so it is the marginal impact of a line of business on capital. Although this is an allocation of capital based on the insolvency put, it is not an allocation of the put value itself. If there is a linear relationship between claims and loss, so $D_j = h_2(X_j)$, this allocation can be expressed as a co-measure, with $h_1(X) = X$, $g_1(Y) = 1$, and $g_2(Y) =– a/Pr^*[Y>\rho(Y)]$.

To summarize, there are different allocations of the insolvency put and EPD, each based on a different formulation of the effect on capital of a change in volume of business. The put can be expressed as $\rho^*(Y) = Pr^*(Y>B)E^*[Y – B | Y>B]$, where Y is the random variable for the negative of profit.

Setting capital as a multiple of expected profit, i.e., $B = aEY$, where hopefully a is negative, gives the allocation:

$$r(X_j) = Pr^*(Y>B)E^*[X_j – BEX_j/EY | Y>B]$$

This is a marginal allocation based on j's excess of actual losses, in the event of ruin, over its proportion of capital based on expected profits. This formulation is like a price-earnings ratio, but instead of market value, the price is book value.

It seems more satisfactory to set capital as a probability level, i.e., $B = F_Y^{-1}(1 – \alpha)$. This gives the marginal allocation:

$$r(X_j) = Pr^*(Y>B)[E^*(X_j | Y>B) – E(X_j | Y=B)]$$

This can be expressed as a co-measure with $h(X) = X – E(X | Y=B)$ and $L(Y) =$

Pr*(Y>B), and can be interpreted as the market value of j's contribution to default.

These allocations use capital set to maintain either a multiple of expected profit or a probability of ruin. Myers-Read is in a similar vein, and expresses capital as the amount needed to maintain the value of the default put at a target ratio of expected claims. This does not actually provide an allocation of the default put value, but does produce a marginal allocation of capital:

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

This is the value of j's conditional expected loss given default less j's proportion, based on expected claims, of the value of the amount of default.

Finally, from the customers' viewpoint, the allocations would be based on their expected recovery shortfalls:

$$r(X_j) = Pr^*(Y>B)E^*[(Y - B)X_j/Y \mid Y>B] = Pr^*(Y>B)E^*[X_j - BX_j/Y \mid Y>B].$$

This looks at j's losses in default less its part $E^*[X_j/Y]$ of capital under default. However it is not a marginal allocation in that B is a constant.

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 $800^{½} = 28.284$. 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 increase Line 1 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 allocated 7.164, 21.173. The overall standard deviation has increased by 0.053, which changes the two allocated standard deviations by 0.093 and -0.040. This is quite non-marginal, as the increase in overall standard deviation affects both lines' allocations by a substantial percentage of the overall change. 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.

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 marginal increases in standard deviation from an increase of 1% in each line, divided by 1%, are 5.316 and 14.142. These do not add to the total standard deviation. Thus there is a conflict between marginal and additive principles. For strategic planning purposes, economic theory says price by marginal cost. Adding an expected claim in Line 1 increases the overall standard deviation by 0.053. If capital is held at the level of 6 sigma, the capital need would thus increase by 0.319. Comparing the cost of capital to the profit from the extra business would determine whether the growth would increase or decrease the overall profit.

The allocation of TVaR is close to marginal even for this small book. For instance TVaR at 80% for the company is about 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 composed of an increase in Line 1 of 1.21 and a decrease in Line 2 of 0.12.

The exponential moment co-measure is also already almost marginal even with this small book of business. Taking c=0.094431 makes $E[Ye^{cY/EY}] = 330$, which is 10% over EY. The allocation can be found numerically to be 109.95 and 220.05. Increasing Line 1 by 1 expected claim increases the overall moment to 331.20, which gets allocated 111.08 and 220.12. So of the total 1.20 increase, Line 2 gets 0.07 instead of the marginal value of 0. The 1% increase in Line 1 increases its allocated risk by 1.03% and Line 2's by 0.03%.

Thus just being scalable is not enough to guarantee that a risk measure has an additive marginal allocation with non-homogeneous growth, but some risk measures are close to this even for a small book. Other risk measures have perfectly additive marginal allocations 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 its co-measure increases to 112.11 and the total risk becomes 334.11. The entire increase of 1.11 is allocated to Line 1, and Line 2 gets no change in risk measure. Thus this is an additive marginal allocation even for non-homogeneous growth.

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[3], so CAPM etc. are special cases of this approach.

---

[3] Adding covariance with Y can be achieved by $f^*(x) = f(x)[1 + cE(Y|x) - cE(Y)]$ for small enough c.

A well supported transform for compound Poisson distributions is the frequency-severity combined Esscher transform[4], 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 $\lambda E(e^{X/c})$. This factor can be applied for other frequency means as well, which should suffice for calculating the transformed mean, but to fully specify the transform in this case, further details would have to be supplied on how exactly the frequency probabilities change.

Each component or even peril within component could have its own c parameter and still preserve marginal additivity. 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 allocations. 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.

---

[4] 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

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 allocate 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.

This is related to measuring financial performance of the business units by economic value added, or as it is sometimes called, capital consumption. 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) for instance. The firm is essentially providing the component a stop-loss cover attaching at the point its cash flows become nega-

---

for the compound Poisson process. It assumes that jump risk is inherently non-diversifiable and so is always priced.

tive. 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 allocate large negative value to lines with losses and large positive value to lines with profits.

8. Conclusion

Allocation of risk by co-measures is a quite general method for additive allocation, but has a weakness of not having a unique solution, and it is not always marginal. When allocation 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 additive and is often very close to marginal. Transformed probability risk measures are additive and marginal even in the non-homogeneous growth case and can be formulated to approximate the market value of the risk being measured.

Appendix 1 – Taking Derivatives

The question is, what are the derivatives on some common co-measures? We will answer by doing a sample detailed derivation first, and then using more streamlined methods to get further results.

Mathematical preliminaries:

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

The density function for the total $Y \equiv \sum_{n=1}^{N} X_n$ is $f_Y(y) = \iint \delta\left(y - \sum_{n=1}^{N} x_n\right) f(\vec{x})\, d\vec{x}$, where the δ function integrates to one and is zero for non-zero argument. The δ function provides formal convenience in derivations and expressions. The mean of Y is $\mu = \iint \sum_{n=1}^{N} x_n f(\vec{x})\, d\vec{x} = \int y f_Y(y)\, dy$ and the mean of X_k is $\mu_k = \iint x_k f(\vec{x})\, d\vec{x}$, so of course $\mu = \sum_{n=1}^{N} \mu_n$ as it should.

Another (bivariate) distribution which will turn out to be of great interest is the joint distribution of Y and X_k: $f_k(y, x) = \iint \delta\left(y - \sum_{n=1}^{N} x_n\right) \delta(x - x_k) f(\vec{x})\, d\vec{x}$. 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 $f_k(x \mid y=b) = \frac{f_k(b,x)}{f_Y(b)}$. An expression we will encounter often is $\int x f_k(b,x)\,dx = E\left[X_k \mid Y=b\right] f_Y(b)$. I.e., it is the mean value of X_k given that $Y=b$ times the density function for the total evaluated at b.

XTVAR with fixed excess point:

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

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

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

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 $\rho(Y+Z)$? The direct answer is that it is the same functional form with another variable added. In the present case,

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

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

We are interested in $\rho(\varepsilon) \equiv \rho(Y + \varepsilon X_k)$ and in particular want to evaluate the derivative $\frac{d}{d\varepsilon}\rho(\varepsilon)$ at $\varepsilon = 0$. 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

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

so the density function is

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

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

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

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

$$f\left(\vec{x}, z\right) = f\left(\vec{x}\right)\delta\left(z - \varepsilon x_k\right)$$

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.

$$
\begin{aligned}
\frac{\rho(\varepsilon)-\rho}{\varepsilon} &= \iint \left(x_k - c\mu_k\right)\Theta\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b\right) f\left(\vec{x}\right) d\vec{x} \Big/ S(\varepsilon, b) \\
&\quad + \iint \left(\sum_{n=1}^{N} x_n - c\mu\right)\frac{1}{\varepsilon}\left[\frac{\Theta\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b\right)}{S(\varepsilon,b)} - \frac{\Theta\left(\sum_{n=1}^{N} x_n - b\right)}{S(b)}\right] f\left(\vec{x}\right) d\vec{x} \\
&= \iint \left(x_k - c\mu_k\right)\Theta\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b\right) f\left(\vec{x}\right) d\vec{x} \Big/ S(\varepsilon, b) \\
&\quad + \frac{1}{S(b)}\iint \left(\sum_{n=1}^{N} x_n - c\mu\right)\left[\frac{\Theta\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b\right) - \Theta\left(\sum_{n=1}^{N} x_n - b\right)}{\varepsilon}\right] f\left(\vec{x}\right) d\vec{x} \\
&\quad + \frac{S(b) - S(\varepsilon . b)}{\varepsilon S(b) S(\varepsilon . b)}\iint \left(\sum_{n=1}^{N} x_n - c\mu\right)\Theta\left(\varepsilon x_k + \sum_{n=1}^{N} x_n - b\right) f\left(\vec{x}\right) d\vec{x}
\end{aligned}
$$

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

$$
\begin{aligned}
&\frac{1}{S(b)}\iint (y - c\mu)\frac{\Theta(\varepsilon x + y - b) - \Theta(y-b)}{\varepsilon} f_k(y,x)\,dydx \\
&= \frac{1}{S(b)}\int dx \left\{\frac{1}{\varepsilon}\int_{b-\varepsilon x}^{b} (y - c\mu) f_k(y,x)\,dy\right\} \\
&\to \frac{b - c\mu}{S(b)}\int x f_k(b,x)\,dx
\end{aligned}
$$

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 in the limit the third term is $-\frac{\int x f_k(b,x)\,dx}{S(b)}\rho$. Thus in the limit as $\varepsilon \to 0$ we obtain

$$\left.\frac{d}{d\varepsilon}\rho(\varepsilon)\right|_{\varepsilon=0} = \rho_k + \frac{b - c\mu - \rho}{S(b)}\int x f_k(b,x)\,dx$$
$$= E\left[X_k - \mu_k \mid Y > b\right] + \frac{f_Y(b)}{S(b)} E\left[X_k \mid Y = b\right]\left\{b - c\mu - E\left[Y - \mu \mid Y > b\right]\right\}$$

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.

Now let us try this in a simpler fashion, using the formal relation $\frac{d}{dx}\Theta(x) = \delta(x)$.

We repeat the original form

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

where

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

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.

$$\rho'(0) = \iint \left(x_k - c\mu_k\right)\Theta\left(\sum_{n=1}^{N} x_n - b\right) f\left(\vec{x}\right) d\vec{x} \Big/ S(0)$$
$$+ \iint \left(\sum_{n=1}^{N} x_n - c\mu\right) x_k \delta\left(\sum_{n=1}^{N} x_n - b\right) f\left(\vec{x}\right) d\vec{x} \Big/ S(0)$$
$$- \frac{S'(0)}{S(0)^2} \iint \left(\sum_{n=1}^{N} x_n - c\mu\right)\Theta\left(\sum_{n=1}^{N} x_n - b\right) f\left(\vec{x}\right) d\vec{x}$$

and

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

Thus the terms directly reduce to

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

which is the result above, of course.

XTVAR with fixed percentage excess:

Let us try a somewhat different measure: We will use XTVAR with a fixed percentage, rather than a fixed quantile. We define the statistic $b(Y)$ 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

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

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

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

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

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

$$0 = \iint \delta\left(\sum_{n=1}^{N} x_n - b \right)\left(x_k - b^{'}(0) \right) f\left(\vec{x}\right) d\vec{x} \Rightarrow b^{'}(0) = \frac{\int x f_k(b,x)\,dx}{f_Y(b)} = E\left[X_k \mid Y = b \right]$$

and

$$\begin{aligned} \rho^{'}(0) &= \iint \left(x_k - c\mu_k \right) \Theta\left(\sum_{n=1}^{N} x_n - b \right) f\left(\vec{x}\right) d\vec{x} \Big/ \alpha \\ &+ \iint \left(\sum_{n=1}^{N} x_n - c\mu \right) \delta\left(\sum_{n=1}^{N} x_n - b \right)\left(x_k - b^{'}(0) \right) f\left(\vec{x}\right) d\vec{x} \Big/ \alpha \\ &= \rho_k \end{aligned}$$

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

Expected Policy-Holder Deficit:

Here

$$\begin{aligned} \rho &= E\left[Y - b \mid Y > b \right] S(b) = \int (y-b)\Theta(y-b) f_Y(y)\,dy \\ &= \iint \left(\sum_{n=1}^{N} x_n - b \right) \Theta\left(\sum_{n=1}^{N} x_n - b \right) f\left(\vec{x}\right) d\vec{x} \end{aligned}$$

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

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

and

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

The second term is immediately zero, so we may write

$$\rho'(0) = \left\{E\left[X_k \mid Y > b\right] - b'(0)\right\} S(b)$$

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

$$\alpha \sum_{n=1}^{N} \left\{E\left[X_n \mid Y > b\right] - E\left[X_n \mid Y = b\right]\right\} = \alpha\left\{E\left[Y \mid Y > b\right] - E\left[Y \mid Y = b\right]\right\}$$
$$= \alpha\left\{E\left[Y \mid Y > b\right] - b\right\} = \alpha E\left[Y - b \mid Y > b\right] = \rho$$

Myers-Read:

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 $f^*\left(x_1, x_2, \ldots, x_N\right) \equiv f^*\left(\vec{x}\right)$.

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, $\left\{E^*\left[Y \mid Y > \rho\right] - \rho\right\} S^*(\rho) = \alpha\mu$ 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

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

For the directional derivative we need

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

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

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

Hence $\alpha\iint xf_k(y,x)\,dydx=\iint\left(x-\rho'(0)\right)\Theta(y-\rho)f_k^*(y,x)\,dydx$

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

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

Note that the mean of the k<u>th</u> variable μ_k is the true mean, not the risk-adjusted mean. Summing these up, we get $E^*\left[Y\mid Y>\rho\right]-\frac{\alpha\mu}{S^*(\rho)}$. 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.

Marginal Co-measures:

Take

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

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 $L(Y)=\Theta\left(Y-F_Y^{-1}(1-\alpha)\right)/\alpha$. 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

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

Since *h* is a linear functional, we have $h\left(\varepsilon x_k + \sum_{n=1}^{N} x_n\right) = \varepsilon h(x_k) + \sum_{n=1}^{N} h(x_n)$. The derivative with respect to ε at zero is the proposed co-measure for the variable X_k:

$$\rho'(0) = \iint h(x_k) L\left(\sum_{n=1}^{N} x_n\right) f\left(\vec{x}\right) d\vec{x} + \iint h\left(\sum_{n=1}^{N} x_n\right) \frac{d}{d\varepsilon} L\left(\varepsilon x_k + \sum_{n=1}^{N} x_n\right)\Bigg|_{\varepsilon=0} f\left(\vec{x}\right) d\vec{x}$$

If we want the sum of these to be the original risk measure, the first term directly gives it so the sum over the second term must be zero. A sufficient condition is that $L(\vec{x})$ is homogenous of order zero: $L(\lambda\vec{x}) = L(\vec{x})$. Then the sum of the second terms is zero, because the integrand is zero.

Let us look at a specific example, where *L* is any integrable function of the ratio

$\zeta \equiv \frac{Y}{E[Y]}$. Then the marginal co-measure is

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

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

$$\begin{aligned}\rho_k &= E\left[h(X_k)\exp(c\zeta)\right] + \frac{c}{E[Y]^2}E\left[h(Y)\exp(c\zeta)\left\{E[Y]X_k - YE\left[X_k\right]\right\}\right] \\ &= E\left[h(X_k)\exp(c\zeta)\right] + \frac{c}{E[Y]^2}E\left[X_k\left\{h(Y)\exp(c\zeta)E[Y] - E\left[Yh(Y)\exp(c\zeta)\right]\right\}\right]\end{aligned}$$

Appendix 2 – Integral Over the Surface Formula

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 $I(\boldsymbol{\psi}) = \left\{ \mathbf{x} \in R^N \,|\, \gamma(\mathbf{x}, \boldsymbol{\psi}) \le 0 \right\}$ and the boundary of this set by ∂I. Consider the volume integral

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

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:

$$\nabla_{\boldsymbol{\psi}} H(\boldsymbol{\psi}) = \int_{\rho(\boldsymbol{\psi})} \nabla_{\boldsymbol{\psi}} \phi(\mathbf{x}, \boldsymbol{\psi}) dV - \int_{\partial\rho(\boldsymbol{\psi})} \frac{\nabla_{\boldsymbol{\psi}} \gamma(\mathbf{x}, \boldsymbol{\psi})}{\left\| \nabla_{\mathbf{x}} \gamma(\mathbf{x}, \boldsymbol{\psi}) \right\|} \cdot \phi(\mathbf{x}, \boldsymbol{\psi}) dS$$

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 $\frac{\nabla_{\boldsymbol{\psi}}\gamma(\mathbf{x},\boldsymbol{\psi})}{\|\nabla_{\mathbf{x}}\gamma(\mathbf{x},\boldsymbol{\psi})\|}$ 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).

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

Exceedance probabilities. Let $S(\varepsilon, b) = \Pr\left\{\varepsilon \cdot X_k + \sum_i X_i \geq b\right\}$. First, we shall compute the derivative of S with respect to b at ε=0. Putting this problem into the IOS format, we have ψ=*b*, *H*(ψ)=*S*(0,*b*), $\phi(\mathbf{x}, \boldsymbol{\psi}) = f(\vec{\mathbf{x}})$ (i.e. the joint p.d.f. of the X_i), and $\gamma(\mathbf{x}, \boldsymbol{\psi}) = b - \sum_i x_i$. The 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

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

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 Jacobean of a variable transformation in getting the representation to include *Y*.)

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

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

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

$$\frac{\partial \rho}{\partial \varepsilon} = -\frac{\partial S(\varepsilon, b)/\partial \varepsilon}{\partial S(\varepsilon, b)/\partial b} = \frac{-\frac{1}{\sqrt{N}} \cdot \int\limits_{Y=b} x_k \cdot f(\mathbf{x}) dS}{-\frac{1}{\sqrt{N}} \cdot \int\limits_{Y=b} f(\mathbf{x}) dS} = \frac{\int\limits_{Y=b} x_k \cdot f(\mathbf{x}) dS}{\int\limits_{Y=b} f(\mathbf{x}) dS} = E\left[x_k \middle| Y = b\right].$$

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.

References

Artzner, P., Delbaen, F., Eber, J.-M., 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", *PCAS*

Mango, D. F. (2003) "Capital Consumption: An Alternative Methodology for Pricing Reinsurance," Casualty Actuarial Society *Forum*, Winter pp. 351-378. Online at http://www.casact.org/pubs/forum/03wforum/03wf351.pdf .

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" *NAAJ* 8:2

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. Online at www.casact.org/coneduc/specsem/sp2003/papers/

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.

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

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

____ (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

____ (2004) "Capital Allocation Survey With Commentary", NAAJ, 8:2.

Venter, G., Barnett, J. and Owen, M. (2004) Market Value of Risk Transfer: Catastrophe Reinsurance Case, AFIR Colloquium