COST RISK ALLOCATION THEORY AND PRACTICE

Christian Smart, Ph.D., CCEA
Operations Research Analyst
Missile Defense Agency
christian.smart@mda.mil
256-313-9566
July, 2010

COST RISK ALLOCATION THEORY AND PRACTICE

ABSTRACT

Risk allocation is the assignment of risk reserves from a total project or portfolio level to individual constituent elements. For example, cost risk at the total project level is allocated to individual work breakdown structure (WBS) elements. This is a non-trivial exercise in most instances, because of issues related to the aggregation of risks, such as the fact that percentiles do not add. For example, if a project is funded at a 70% confidence level then one cannot simply allocate that funding to WBS elements by assigning each its 70% confidence level estimate. This is because the resulting sum may (but not necessarily will) be larger than the total 70% confidence estimate for the entire project. One method for allocating risk that has commonly been used in practice and has been implemented in a cost estimating integration software package is to assign risk by assigning the element's standard deviation as a proportion of the sum of the standard deviations for all WBS elements (Sandberg 2007). Another method proposed as an improvement over this notes that risk is typically not symmetric, and looks at the relative contribution of the element's variation above the mean or other reference estimate. This was first presented by Book to a limited government audience in 2003 and presented to a wider audience three years later (Book 2006). This technique, based on the concept of "need," has been implemented in the NASA/Air Force Cost Model (Smart 2005). These contributions represent the current state-of-the-practice in cost analysis. The notion of considering positive semi-variance as an alternative to the needs method was brought forth by Book (Book 2007) and further propounded by Sandberg (Sandberg 2007). A new method proposed by Hermann (Hermann 2010) introduces the concept of optimality in risk allocation and proposes a one-sided moment method objective function for calculating the optimal allocation. Aside from Hermann's paper, cost risk allocation has typically not been associated with optimality, so neither the proportional standard deviation method nor the needs method guarantees the allocation scheme will be optimal or even necessarily desirable. Indeed, the twin topics of risk measurement and risk allocation have either been treated independently (Book 2006), or they have been treated as one and the same (Sandberg 2007). Regardless, the current situation is muddled, with no clear delineation between the two. In this paper, the author introduces to cost analysis the concept of gradient risk allocation, which has been recently used in the areas of finance and insurance (McNeil et al. 2005). Gradient allocation clearly illustrates that the notions of risk measure and risk allocation are distinct but intrinsically linked. This principle is shown to be an optimal method for allocation using three distinct arguments – axiomatic, game-theoretic, and economic (optimal is used in this context as desirable or good, not as the minimum or maximum of a specified objective function). It is also shown that the gradient risk allocation method is intrinsically tied to the method used to measure risk, a concept not heretofore considered in cost analysis. Gradient allocation is applied to five risk measures, resulting in five different allocation methods, each optimal for the risk measure from which they are derived. Considerations on when

the proportional standard deviation and needs method are optimal are discussed, and a link between Hermann's method and the proportional standard deviation method is demonstrated.

INTRODUCTION

"I can see that it works in practice, but does it work in theory?"
Garrett Fitzgerald, Prime Minister of Ireland 1981-1987

Risk is often measured at a variety of WBS levels. However, risk management is typically applied at the project level, with the focus on measuring and guarding against risk for the entire project, at least within funding allocations (such as design and development, production, and operations and sustainment). These "colors of money" have legal restrictions on how money can be moved from one portion of life-cycle funding to another. For a NASA satellite this may encompass both development and production since these are often unique missions with a single production unit.

Once risk is measured at the project level, it is a non-trivial exercise to determine how much of that total risk is attributable to each individual WBS element. This is because, even though total cost is the sum of the costs for each WBS element, most risk measures do not add. For example if percentile funding is used for risk measurement, it is often (but not always) the case that the sum of the percentiles will be greater than the percentile of the sums of the individual WBS elements. For example consider two independent and normally distributed random variables, X_1 and X_2 with $X_1 \sim N(100,20)$ and $X_2 \sim N(300,80)$. To combine these two distributions the means and the variances are (separately) aggregated, so that the total mean is $100+300 = 400$, and the total variance is $20^2+80^2 = 6800$. The standard deviation is the square root of this latter value, which is approximately ***82.5***. The combined random variable, X_1+X_2, is also normally distributed with mean equal to ***400*** and standard deviation equal to ***82.5***, i.e., $X_1+X_2 \sim N(400,82.5)$. The 80th percentile of X_1 can be calculated as

$$p_{.80} = \mu + z_{.80}\sigma \approx 100 + 0.8416 \cdot 20 \approx 116.8$$

where $z_{.80}$ is equal to the inverse of the standard normal distribution at the 80th percentile.

Similarly, the 80th percentile of X_2 is approximately 367.3, and the 80th percentile of X_1+X_2 is approximately 469.4. Thus the sum of the 80th percentiles for X_1 and X_2 is *116.8+367.3=484.1*, which is larger than the 80th percentile of the sum X_1+X_2. In other words, percentiles do not add. One cannot add the 80th percentiles and expect that this sum will be equal to the 80th percentile of the sum of the random variables. To see why this is the case in general for the sum of two independent normally distributed variables note that the percentiles are determined by the mean and standard deviation. The sum of the means of random variables is equal to the mean of the sum of the random variables, regardless of the distribution type, so the difference lies with the variance. For combining independent normal random variables, the variances are added rather than the standard deviations, which is key, since the sum of the variances, where the standard deviations of the individual normal random variables are denoted by $\boldsymbol{a}$ and $\boldsymbol{b}$, is equal to $\boldsymbol{a^2+b^2}$. The standard deviation of the sum is the square root of this quantity, i.e., $\sqrt{\boldsymbol{a^2+b^2}}$**.** Note that since $\boldsymbol{a^2+b^2 \le a^2+2ab+b^2}$**,** it follows that $\sqrt{\boldsymbol{a^2+b^2}} \le \boldsymbol{a+b}$, with strict inequality unless at least one of $\boldsymbol{a}$ or $\boldsymbol{b}$ is equal to zero. Note that $\sqrt{\boldsymbol{a^2+b^2}}$ represents the risk of the combined distributions. The quantity $\boldsymbol{a+b}$ represents the sum of the individual risks. Thus in this case, combining the independent elements is a diversification of risk. The total portfolio is not as risky on a relative basis as each individual project.

Since percentiles do not add, when funding at specific percentiles, risk allocation becomes a non-trivial exercise. More generally, risk measures are not typically additive, so whatever risk measure is being used at the total project level, some care is required to effectively and fully allocate risk. The goal of risk allocation is to apportion the total estimate to individual WBS elements so that each is funded in a manner so that the sum of the individual WBS allocations equals the total risk measurement.

RISK MEASUREMENT

Risk allocation begins with risk measurement. Without measuring risk, it cannot be properly allocated. Since there has been confusion about the two, with some authors treating the issues of risk measurement and allocation as independent problems, and others treating them as one and the same, a review the topic of risk measurement is needed. There are several popular ways to measure risk. Variance, and its square root, standard deviation, are popular measures for risk measurement. The notion of measuring risk via the standard deviation dates back to (at least) the work of H. Markowitz (Markowitz 1959).

Coherent Risk Measures

A risk measure is a single number that is used to represent cost risk for a project or program. The variance of the distribution is a risk measure since it quantifies the spread in the cost risk distribution. Value at risk is another risk measure and there are many others.

What properties should a risk measure have? This issue has been studied in insurance specifically and in risk measurement in general. A groundbreaking paper (Artzner et al., 1999) introduced the notion of coherent risk measures. One property important for a risk measure is that when two random variables are combined the risk measure of the portfolio should be no riskier than the sum of the individual random variables' risk measures. That is for any risk measure $\boldsymbol{\rho}$ it should be the case that

$$\rho(X+Y) \le \rho(X)+\rho(Y).$$

That is there should be some diversification benefit from combining risks, which is called *subadditivity*. A better-known term for subadditivity is the "portfolio effect" (Anderson 2004) which has been relied upon by policymakers in setting funding levels to individual projects to relatively low levels, such as the 70th percentile and below. Subadditivity embodies the principle of "l'union fait la force," that is, unity is strength.

For example if the cost of structures hardware is higher in every circumstance than thermal control hardware, then the 70th percentile of the cost risk distribution should be higher for the structures subsystem than for the 70th percentile of the cost risk distribution for the thermal control subsystem. This is the property of *monotonicity*, and can be stated in equation form as

$$X \leq Y \text{ for all possible outcomes } \Rightarrow \rho(X) \leq \rho(Y).$$

A third desirable property is that the risk measure should be invariant of the currency in which the risk is measured, or whether cost is accounted for in thousands or millions of dollars. Also it means that an increase or decrease in exposure to the risk requires an equivalent change in the amount of capital needed to guard against this risk. This is the property of *positive homogeneity*, and can be expressed as $\rho(cX) = c\rho(X)$ for a constant real number c. An example of this would be a joint U.S. and European project, whose cost risk could be measured in either dollars or euros. The risk should be the same regardless of the currency used, up to a currency conversion factor. Another example of this would be two components that are made at the same time, by the same manufacturer, and built to the exact same specifications and requirements and hence have the same cost risk. If X represents the cost random variable for one component, then it should be the case that two times the risk of one is the risk of both components considered together. It is also important in measuring risk that if we add some certain fixed, certain amount to a random variable, the risk does not change. This is the property of *translation invariance* and can be expressed as $\rho(X+c) = \rho(X)+c$. A coherent risk measure is defined as a risk measure $\rho(X)$ which has the four properties of *subadditivity*, *monoticity*, *positive homogeneity*, and *translation invariance*.

Commonly Used Risk Measures and Coherence

Standard Deviation Principle

A simple and popular risk measure is the defined as the mean plus a fixed number of standard deviations, i.e., $\mu + k\sigma$ for some real number k which is called the *standard deviation principle*. Note that this risk measure is subadditive, since

$$\mu_{X+Y} + k\sigma_{X+Y} = \mu_X + \mu_Y + k\sigma_{X+Y} \leq \mu_X + \mu_Y + k\sigma_X + k\sigma_Y$$

$$= \mu_X + k\sigma_X + \mu_Y + k\sigma_Y.$$

Also, the standard deviation principle is positive homogeneous, since

$$\mu_{c(X+Y)} + k\sigma_{c(X+Y)} = c\mu_{X+Y} + ck\sigma_{X+Y} = c(\mu_{X+Y} + k\sigma_{X+Y}).$$

And since standard deviation is not affected by a translation of the random variable, but the mean is shifted by exactly the translation, the standard deviation principle is translation invariant.

However, the standard deviation principle is not monotonic. To see this, consider a bivariate random variable defined as

$$\boldsymbol{p(X,Y)} = \begin{cases} \mathbf{0.25 \text{ for } \boldsymbol{X} = 0, \boldsymbol{Y} = 4} \\ \mathbf{0.75 \text{ for } \boldsymbol{X} = 4, \boldsymbol{Y} = 4} \end{cases}$$

In this case, $\boldsymbol{\mu_X = 3, \mu_Y = 4, \sigma_X = \sqrt{3}, \sigma_Y = 0}$. Note that even though $\boldsymbol{X \leq Y}$ it is the case that

$$\boldsymbol{\mu_X + \sigma_X} = \mathbf{3} + \sqrt{\mathbf{3}} > 4 = 4 + 0 = \mu_{\boldsymbol{Y}} + \boldsymbol{\sigma_Y}.$$

Thus an important consequence of a risk measure not being monotonic is that an element's risk can be greater than the maximum value possible for the variable, which is illogical.

Value at Risk

Note that the standard deviation principle is not the same as Value at Risk, unless we restrict our attention to normally distributed random variables. In this case *VaR* is a special case of the standard deviation principle with ***k*** set to satisfy whichever percentile is selected. In the case of the 70th percentile, $\boldsymbol{k \approx 0.5244}$. In this case, *VaR* clearly satisfies the conditions of translation invariance, monotonicity, and positive homogeneity. It is also subadditive by the same rationale used for the standard deviation principle. Thus in the special case of normally distributed random variables, *VaR* is a coherent risk measure.

However, in general, *VaR*, as a percentile of a cost distribution, is translation invariant, monotonic, and has positive homogeneity. However, *VaR* is not guaranteed to be subadditive for non-normal random variables. A recent paper (Smart 2010) provides examples of two projects which when combined are actually superadditive when *VaR* is the risk measure, leading to a reverse portfolio effect!

VaR, or percentile funding, is commonly used for measuring risk for NASA and Department of Defense projects. The 50th, 70th, and 80th percentiles are commonly used to set budgets for these agencies' projects. One of the motivating factors for funding at such a low percentile is the prospect of a portfolio effect, or diversification. With diversification, funding individual projects at the 70th or 80th percentile will result in much higher confidence level when the entire agency is considered as a portfolio. At least, that is the hope. However, in reality, the portfolio effect is minimal at best (Smart 2009). Not only is such an effect guaranteed, but the prospect of superadditivity, as shown in a recent paper (Smart 2010), means that there can be a negative portfolio effect. So funding at the 70th or 80th percentile is no guarantee that the confidence level of the total budget is any higher than the 70th or the 80th percentile (respectively), and may in fact be lower. Funding at lower levels, such as the 50th percentile is even more problematic. In the case of skewed risk distributions, the 50th percentile is below the mean, and this can result in extremely low confidence levels for the entire agency.

Percentile funding is also problematic for other reasons. Percentile funding only indicates when there is a problem, and does not set aside funds for bad times. Thus percentile funding is not a true risk management policy. Also percentile funding ignores the right tail of the distribution. For more on these issues, see Smart (Smart 2010).

Expected Shortfall

Expected shortfall (ES) is similar to *VaR*, but it looks at the expected overrun past a fixed percentile. Thus is provides not only an indication that bad times have occurred (when the percentile is exceeded), but also a reserve set aside to deal with adverse conditions when they occur.

Expected shortfall is defined as

$$\boldsymbol{ES_\alpha = \frac{1}{1-F(Q_\alpha)}\int_{Q_\alpha}^{1} xf(x)dx = \frac{1}{1-\alpha}\int_{\alpha}^{1} VaR_u(X)du}$$

In the case of continuous cost risk distributions, this risk measure is referred to as Condtional Tail Expectation (*CTE*). For example, $\boldsymbol{Q_{0.95}}$ is the 95th percentile (McNeil et al., 2005). It is the "Tail Value at Risk" since in the case of continuous cost distributions it may be viewed as

$$\boldsymbol{CTE_\alpha = E[X \mid X > Q_\alpha]}$$

In the case of normally distributed cost risk,

$$\boldsymbol{ES_\alpha(X) = CTE_\alpha(X) = \mu + \sigma\frac{\phi\left(\Phi^{-1}(\alpha)\right)}{1-\alpha}}$$

And for lognormally distributed cost risk,

$$\boldsymbol{ES_\alpha(X) = CTE_\alpha(X) = \frac{\mathrm{E}[\mathrm{X}]\left[1-\Phi\left(\frac{\ln \mathrm{VaR}_\alpha - \mu - \sigma^2}{\sigma}\right)\right]}{1-\alpha}}$$

where $\boldsymbol{\phi}$ represent the standard normal density function, $\boldsymbol{\Phi}$ is the cumulative normal distribution function and $\boldsymbol{\Phi}^{-1}$ represents the inverse of the cumulative normal distribution. See Smart (Smart 2010) for derivations. Also note that expected shortfall is a coherent risk measure. For more on the merits of expected shortfall as a risk measure, see Smart (Smart 2010).

One-Sided Moments

One-sided moments make sense from a risk perspective, since they only look at risk above the mean, rather than uncertainty both above and below the mean. Thus they are an improvement over standard deviation as a risk measure, as advocated separately by Book and Sandberg for use in risk allocation (Book 2006 and Sandberg 2007). The use of one-sided moments dates back to at least the early 1950s, when they were advocated by the Nobel Prize-winning economist Markowitz (Markowitz 1959). However, their use did not become popular until much later.

Formally, the p^{th} one-sided (positive) moment about the mean is defined as

$$\boldsymbol{E}\left((\boldsymbol{X}-\boldsymbol{\mu})_{+}^{\boldsymbol{p}}\right)$$

or for a continuous cost risk distribution,

$$\int_{-\infty}^{\infty}(\boldsymbol{x}-\boldsymbol{\mu})_{+}^{\boldsymbol{p}}\,\boldsymbol{f}(\boldsymbol{x})\boldsymbol{dx}$$

where $(\boldsymbol{X}-\boldsymbol{\mu})_{+}=\boldsymbol{max}\,(\boldsymbol{0},\boldsymbol{X}-\boldsymbol{\mu})$.

The first one-sided (positive) moment about the mean is thus

$$\int_{-\infty}^{\infty}(\boldsymbol{x}-\boldsymbol{\mu})_{+}\,\boldsymbol{f}(\boldsymbol{x})\boldsymbol{dx}=\int_{\boldsymbol{\mu}}^{\infty}(\boldsymbol{x}-\boldsymbol{\mu})\,\boldsymbol{f}(\boldsymbol{x})\boldsymbol{dx}$$

Note the similarity to expected shortfall, instead of a percentile, the mean is used. Consider the risk measure defined by

$$\boldsymbol{\mu}+\ \boldsymbol{E}((\boldsymbol{X}-\boldsymbol{\mu})_{+})$$

which has been advocated (Hermann 2010) for use in risk allocation.

The second moment about the mean is

$$\boldsymbol{\sigma}_{+}^{2}=\int_{-\infty}^{\infty}(\boldsymbol{x}-\boldsymbol{\mu})_{+}^{2}\,\boldsymbol{f}(\boldsymbol{x})\boldsymbol{dx}=\int_{\boldsymbol{\mu}}^{\infty}(\boldsymbol{x}-\boldsymbol{\mu})^{2}\,\boldsymbol{f}(\boldsymbol{x})\boldsymbol{dx}$$

which is referred to as positive semi-variance. Consider the risk measure defined by

$$\boldsymbol{\mu}+\boldsymbol{\sigma}_{+}$$

where

$$\boldsymbol{\sigma}_{+}=\sqrt{\int_{\boldsymbol{\mu}}^{\infty}(\boldsymbol{x}-\boldsymbol{\mu})^{2}\,\boldsymbol{f}(\boldsymbol{x})\boldsymbol{dx}}$$

is the positive semi-deviation. This is similar to the standard deviation principle. Both of these risk measures are coherent.

To see this note that the first one-sided moment is clearly positively homogenous, due to properties of moments, as well as translation invariant. And since

$$\left(\boldsymbol{X}+\boldsymbol{Y}-(\boldsymbol{\mu_x}+\boldsymbol{\mu_y})\right)_+ \leq (\boldsymbol{X}-\boldsymbol{\mu_x})_+ + (\boldsymbol{Y}-\boldsymbol{\mu_y})_+$$

this measure is also subadditive. For monotonicity, suppose that $\boldsymbol{Y} \leq \boldsymbol{X}$. Then

$$\boldsymbol{Y}-\boldsymbol{\mu_y}-(\boldsymbol{X}-\boldsymbol{\mu_x}) \geq \boldsymbol{Y}-\boldsymbol{\mu_y}-(\boldsymbol{Y}-\boldsymbol{\mu_x}) = \boldsymbol{\mu_x}-\boldsymbol{\mu_y}$$

for all $\boldsymbol{X},\boldsymbol{Y}$, so this inequality holds when $\boldsymbol{X} \geq \boldsymbol{\mu_x}$ and $\boldsymbol{Y} \geq \boldsymbol{\mu_y}$. Thus

$$(\boldsymbol{Y}-\boldsymbol{\mu_y})_+ - (\boldsymbol{X}-\boldsymbol{\mu_x})_+ \geq \boldsymbol{\mu_x}-\boldsymbol{\mu_y}$$

So

$$\boldsymbol{\mu_x} + (\boldsymbol{X}-\boldsymbol{\mu_x})_+ \leq \boldsymbol{\mu_y} + (\boldsymbol{Y}-\boldsymbol{\mu_y})_+$$

Applying expected values to both sides yields the desired result.

For the semi-standard deviation principle, positive homogeneity and translation invariance hold for the same reasons as with the standard deviation principle. For subadditivity, note that

$$\left(\boldsymbol{X}+\boldsymbol{Y}-(\boldsymbol{\mu_x}+\boldsymbol{\mu_y})\right)_+ \leq (\boldsymbol{X}-\boldsymbol{\mu_x})_+ + (\boldsymbol{Y}-\boldsymbol{\mu_y})_+$$

which implies that

$$\sqrt{\boldsymbol{E}(\boldsymbol{X}+\boldsymbol{Y}-\left(\boldsymbol{\mu_x}+\boldsymbol{\mu_y}\right))_+^2} \leq \sqrt{\boldsymbol{E}((\boldsymbol{X}-\boldsymbol{\mu_x})_+ + \left(\boldsymbol{Y}-\boldsymbol{\mu_y}\right)_+)^2}$$

and

$$\sqrt{\boldsymbol{E}(\boldsymbol{X}+\boldsymbol{Y}-\left(\boldsymbol{\mu_x}+\boldsymbol{\mu_y}\right))_+^2} \leq \sqrt{\boldsymbol{E}(\boldsymbol{X}-\boldsymbol{\mu_x})_+^2} + \sqrt{\boldsymbol{E}(\boldsymbol{Y}-\boldsymbol{\mu_y})_+^2}$$

by Minkowski's inequality, so

$$\sqrt{\boldsymbol{E}(\boldsymbol{X}+\boldsymbol{Y}-\left(\boldsymbol{\mu_x}+\boldsymbol{\mu_y}\right))_+^2} \leq \sqrt{\boldsymbol{E}(\boldsymbol{X}-\boldsymbol{\mu_x})_+^2} + \sqrt{\boldsymbol{E}(\boldsymbol{Y}-\boldsymbol{\mu_y})_+^2}$$

Since the means are additive, subadditivity holds.

To show monotonicity, let $\boldsymbol{Y} \leq \boldsymbol{X}$. Now in this case,

$$\boldsymbol{X}-\boldsymbol{\mu_x} \leq \boldsymbol{\mu_y}-\boldsymbol{\mu_x}$$

which for $\boldsymbol{X-\mu_x \geq 0}$ it holds that

$$\boldsymbol{(X-\mu_x)_+^2 \leq (\mu_y-\mu_x)^2}$$

which implies that

$$\boldsymbol{E(X-\mu_x)_+^2 \leq (\mu_y-\mu_x)^2}$$

so that

$$\boldsymbol{\sqrt{E(X-\mu_x)_+^2} \leq (\mu_y-\mu_x)}$$

and thus

$$\boldsymbol{\mu_x+\sqrt{E(X-\mu_x)_+^2} \leq \mu_y \leq \mu_y+\sqrt{E\left(Y-\mu_y\right)_+^2}}$$

One-sided moments are likely to provide lower amounts than those calculated for expected shortfall and thus may prove to be more palatable to management in setting policy. They also take into account the entire tail of the distribution, although this measure is not as sensitive to large or what is referred to as “fat” right tails as is expected shortfall.

A Note on the Mean

Note that the mean is a coherent risk measure. It takes into account the entire right tail, and it is strictly additive, rather than subadditive, so no diversification benefits are seen from funding at the mean. On the other hand, additivity is appealing since it simplifies the risk allocation process greatly, is easy to explain to management, and is easy to communicate with budget analysts and accountants. It is also the smallest of the coherent risk measures considered, so it may be palatable to project management for that very reason. Indeed the mean is used as a risk measure for some government agencies, and this is a better, more sensible funding policy than percentile funding.

Comparison of Risk Measures – Example

Consider the 10 individual projects shown in Table 1. Each is assumed to be lognormally distributed, with common correlation equal to 20% among all projects.

Project	Mean	Standard Deviation
Project 1	1501	556
Project 2	804	219
Project 3	907	302
Project 4	875	400
Project 5	1450	420
Project 6	1271	419
Project 7	874	541
Project 8	1001	229
Project 9	1139	392
Project 10	981	485
Total	10803	

Table 1. Notional Example of 10 Projects.

Now consider the total risk for all 10 projects combined, as a portfolio. Note that this is purely a notional example, and any resemblance to the cost of actual projects, either historical or currently in development, is purely coincidental.

The total risk of these 10 projects was aggregated using a 50,000 trial Latin hypercube simulation. The six risk measures discussed in this paper applied to the total risk aggregation are shown in Table 2.

Risk Measure	Value	Coherent?
Mean	$10,803	Yes
1st One-Sided Moment	$11,629	Yes
Value at Risk (70th Percentile)	$11,695	No
Semi-Standard Deviation Principle	$12,413	Yes
Standard Deviation Principle	$12,909	No
Expected Shortfall ($VaR_{0.70}$)	$13,331	Yes

Table 2. Comparison of Six Risk Measures.

Policymakers – take note! Using a coherent risk measure does not necessarily translate into more risk dollars. The lowest risk measure in the example is the mean, followed by the first one-sided moment. Since the one-sided moment risk measures only consider risk, and ignore opportunity (uncertainty below the mean), they always translate into fewer risk dollars than their analogous two-sided risk measure.

RISK ALLOCATION

Standard Deviation-Based Methods for Allocating Risk

The current established state-of-the-practice methods for allocating risk are based on standard deviation as the measure of risk. The first method is conceptually simple. It involves apportioning the risk by setting the amount allocated for a WBS element equal to its ratio to the sum of the total standard deviations. For example for a project with two elements with standard deviations equal to 100 and 200, the sum of the standard deviations is 300, and the ratio of the first element to the sum is 100/300 = 1/3, so it is allocated one-third of the total risk reserve, while the second is allocated the remaining two-thirds. This method, which is referred to in this paper as the proportional standard deviation method, is easy to understand and easy to implement in a spreadsheet. In the past it has typically been used to allocate risk when risk is measured as a percentile.

Allocating percentile funding via proportional standard deviation begins with calculating the specified percentile, such as the 70th or 80th. When a normal or lognormal probability distribution is used to represent cost risk the mean and standard deviation describe the distribution and are typically used as the parameters to define it. Note that in the case of *n* independent WBS elements the total standard deviation can be calculated as

$$\sigma_{Total} = \sqrt{\sum_{i=1}^{n} \sigma_i^2}$$

For normal and lognormal distributions, once the mean and standard deviation have been determined, a percentile, such as the 80th percentile, may be calculated. For a normal distribution the calculated 80th percentile is

$$\mu_{Total} + z_{0.80}\sigma_{Total}$$

If the mean is used as the point estimate (i.e., the non risk-adjusted cost estimate), since the sum of the individual WBS elements' means is the total mean, the risk dollars to be allocated back is simply

$$\mu_{Total} + z_{0.80}\sigma_{Total} - \mu_{Total} = z_{0.80}\sigma_{Total}$$

These risk dollars are what is allocated back to each individual WBS element, and amounts to the risk reserve above the mean.

Now that the risk dollars have been determined, we calculate the WBS element's portion of this total, i.e.,

$$\boldsymbol{p_i} = \frac{\boldsymbol{\sigma_i}}{\sum_{j=1}^{N} \boldsymbol{\sigma_j}}$$

These risk dollars are then allocated to each specific WBS element. In the case of a normal or distribution the amount of risk dollars assigned to a specific WBS element is

$$\boldsymbol{\mu_i + p_i(z_{0.80}\sigma_{Total})}$$

Note that these individual amounts add to the total 80th percentile since

$$\sum_{i=1}^{n} \mu_i + p_i(z_{0.80}\sigma_{Total}) = \sum_{i=1}^{n} \mu_i + \sum_{i=1}^{n} p_i(z_{0.80}\sigma_{Total}) = \mu + (z_{0.80}\sigma_{Total}) \sum_{i=1}^{n} p_i$$

$$= \mu + z_{0.80}\sigma_{Total}$$

and the allocation weights $\boldsymbol{p_i}$ sum to *1*.

Allocating via proportional standard deviation is not an optimal way to allocate risk when percentile funding is used as the risk measure, as we shall see when we look at gradient allocation later in this paper.

The proportional standard deviation method has some drawbacks. The most obvious is that summing the standard deviations is a heuristic without basis in theory. The sum of standard deviations is not the total standard deviation for example, or really any useful statistic at the total project level (However in a later section it will be shown that under some specific circumstances this is an optimal way to allocate risk). Another glaring drawback is that it ignores correlation, and so may allocate risk to individual elements in a non-optimal manner. Also, the proportional standard deviation method equates risk with standard deviation. However, the two are not the same. When risk is high, uncertainty is also necessarily high, but the converse is not always true. It may be the case that uncertainty is high but risk is low. Consider the example of two triangular distributions displayed in Figure 1 due to Book (Book 2006).

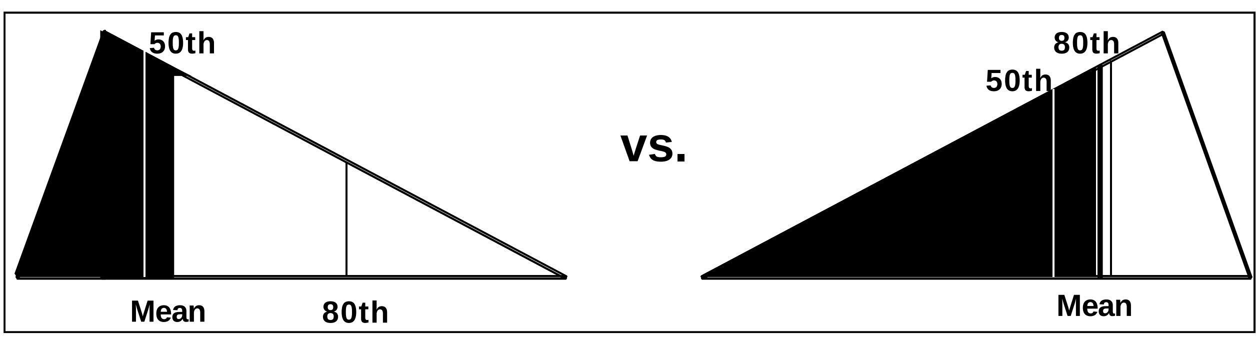

Figure 1. Two triangular distributions with the same level of uncertainty but different amounts of risk.

The two triangular distributions displayed in the figure are symmetric and hence have the same amount of uncertainty and the same standard deviation. However, if the point estimate is represented by the mean and risk is measured as the 80th percentile, then the triangular distribution on the left has much more risk than the triangular distribution on the right. As is evident from the graph the triangular distribution on the right requires few risk dollars above the mean to achieve the 80th percentile.

The correlation issue can be overcome; the method could be changed to one based on covariance contributions. The covariance principle is based on the notion that in the general case when we consider correlation among WBS elements, the total variance is equal to

$$\boldsymbol{Total\,Variance = \sigma^T P \sigma}$$

where $\boldsymbol{P}$ is the $\boldsymbol{N \times N}$ correlation matrix, and $\boldsymbol{\sigma}$ is the $\boldsymbol{N \times 1}$ vector of standard deviations for the individual WBS elements. The amount the i^{th} WBS element contributes to the total variance is equal to

$$\boldsymbol{\Gamma_i = \sigma_i \sum_{j=1}^{N} \rho_{ij} \sigma_j}$$

where $\boldsymbol{\rho_{ij}}$ is the correlation between the i^{th} and j^{th} WBS elements. The covariance principle then allocates risk as

$$\boldsymbol{p_i = \frac{\Gamma_i}{Total\,Variance}}$$

Note that since

$$\boldsymbol{\sum_{i=1}^{N} p_i = \sum_{i=1}^{N} \frac{\Gamma_i}{Total\,Variance} = 1}$$

so the risk allocation fully distributes the risk to the WBS elements. Note that this is the only specific requirement of these risk allocation schemes, which is that they are complete – they distribute the total risk dollar among the elements, no more and no less. There are others that will be considered when optimal allocation is covered. Another potential issue, pointed out to the author (Hermann 2010), is the possibility of a negative allocation if negative correlations are present. While not a conceptual problem, it may be hard to communicate with management.

Even though the covariance principle overcomes the specific issue of correlation that is not considered by the proportional standard deviation method, it too does not distinguish between downside opportunities for cost savings and upside risk of cost growth. In order to distinguish between upside risk and downside opportunity, the notion of need was introduced (Book 2006).

This idea is based on the concept of semi-variance. Semi-variance only looks at the second moment above the mean. For a continuous random variable $\boldsymbol{X}$ this is defined as

$$\int_{-\infty}^{\infty}(\boldsymbol{x}-\boldsymbol{\mu})_{+}^{2}\,\boldsymbol{f}(\boldsymbol{x})\boldsymbol{dx}=\int_{\boldsymbol{\mu}}^{\infty}(\boldsymbol{x}-\boldsymbol{\mu})^{2}\,\boldsymbol{f}(\boldsymbol{x})\boldsymbol{dx}$$

where $\boldsymbol{Y}_{+}=\boldsymbol{max}(\boldsymbol{Y},\boldsymbol{0}).$

The notion of need considers the difference between a selected percentile, such as the 80th, and a point estimate, such as the mean. The difference between these two values at the total project level is the amount of risk dollars. Similar to covariance, the total need base is calculated as

$$\boldsymbol{Need\ Base}=\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\rho}_{ij}\boldsymbol{Need}_{i}\,\boldsymbol{Need}_{j}$$

For the i^{th} element, if the project's percentile risk measure (denoted by $\boldsymbol{\pi_i}$) is lower than the reference point estimate (denoted by $\boldsymbol{c_i}$), e.g., the mean, then the need for the i^{th} element is $\boldsymbol{0}$. Otherwise the need for the i^{th} element is positive, and the need contribution of the i^{th} element to the total need is calculated as

$$\sum_{j=1}^{N}\boldsymbol{\rho}_{ji}\boldsymbol{Need}_{j}\boldsymbol{Need}_{i}$$

where

$$\boldsymbol{Need}_{i}=\mathbf{max}\,(\mathbf{0},\boldsymbol{\pi}_{i}-\boldsymbol{c}_{i})$$

The percentage of risk dollars allocated to i^{th} element is then calculated as

$$\boldsymbol{p}_{i}=\begin{cases}\dfrac{\sum_{j=1}^{N}\boldsymbol{\rho}_{ji}\boldsymbol{Need}_{j}\boldsymbol{Need}_{i}}{\boldsymbol{Need\ Base}} & \boldsymbol{if\ Need}_{i}>0\\ \mathbf{0} & \boldsymbol{if\ Need}_{i}=\mathbf{0}\end{cases}$$

The need concept also has its drawbacks. Negative correlation can lead to negative need allocations, just as with the covariance principle. More importantly, the concept of need ignores the tail of the distribution (Sandberg 2007). This is an important consideration is risk measurement, as discussed in a recent ISPA/SCEA presentation (Smart 2010). First brought to the attention of the cost community by Book (Book 2007), Sandberg (Sandberg 2007) leverages this idea by proposing using semi-variance at the element level to allocate risk and replaces the need at the element level with the positive semi-variance. The idea of using semi-variance to measure risk is a long-standing one. One of the first proponents of semi-variance in finance was the Nobel Prize-winning economist Markowitz (Markowitz 1959). While taking into account the right tail of the distribution, Sandberg does not consider the relationship between the sum of the semi-variance contributions, and the total semi-variance, so again risk allocation and risk measurement are considered as two separate, independent problems. Sandberg (Sandberg 2007) considers the issue of optimization but only in the narrow context that the method he proposes by definition minimizes a specific quantity.

Optimal Allocation

In a recent technical note risk allocation is posed as an explicit optimization problem (Hermann 2010). This begins with the risk measure

$$\boldsymbol{r = \int_{\mu}^{\infty} (x-\mu) f(x) dx}$$

and proceeds to consider allocating this risk to individual WBS elements ***1,…,n*** by minimizing the sum of the individual expected shortfalls across WBS elements, i.e.,

$$\boldsymbol{Minimize \sum_{i=1}^{n} r_i^* = Minimize \sum_{i=1}^{n} \int_{\mu_i + r_i}^{\infty} (x_i - \mu_i - r_i) f(x) dx}$$

subject to the restriction that $\boldsymbol{\sum_{i=1}^{n} r_i = r_T}$ and $\boldsymbol{r_i \geq 0}$ for all ***i = 1,…n***. Note that in this paper risk measure functions are denoted by the letter ***r*** and are functions of a single variable. Risk allocation, when the context is clear to which set risk is being allocated from a larger superset, is denoted by $\boldsymbol{r_i}$, such as the risk attributed to the i^{th} element from the total project level. When greater clarity is needed, the notation ***r(X,Y)*** is used, which means the allocation of risk from set ***Y*** to a subset ***X***, using risk measure ***r***.

This novel method is notable for considering the issue of allocation as an optimization problem, and for taking into consideration the entire right tail of the cost risk distribution in the allocation process. The motivating factor for minimizing the sum of the expected shortfalls could be that risk dollars are not fungible across WBS elements. That is, once committed, risk dollars cannot be re-allocated if one element needs more than expected or another needs less. But this is not typically what is seen in practice. Risk is measured and allocated within a specific funding category. Financial managers then have the ability to juggle and re-juggle allocations as needed.

As a result of looking at the sum of expected shortfalls, this method does not incorporate the impact of correlation, and thus is similar to the proportional standard deviation method. And this is not the only connection between these two methods. Note that in the case of normally distributed random variables,

$$\boldsymbol{r = \int_{\mu}^{\infty} \frac{x-\mu}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \frac{\sigma}{\sqrt{2\pi}}}$$

Note that the funding level for this risk measure is

$$\boldsymbol{\mu + r = \mu + \frac{\sigma}{\sqrt{2\pi}} \approx 65.5th\ percentile}$$

Given funding to $\boldsymbol{\mu + r}$ the "remaining expected risk exposure" (as defined by Hermann (Hermann 2010) is, for a normally distributed random variable, equal to

$$\boldsymbol{r^* = \int_{\mu+r}^{\infty} \frac{x-\mu-r}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx}$$

$$\boldsymbol{= \int_{\mu+r}^{\infty} \frac{x-\mu}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx - \int_{\mu+r}^{\infty} \frac{r}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx}$$

$$\boldsymbol{= \frac{\sigma}{\sqrt{2\pi}} exp\ (-\frac{r^2}{2\sigma^2}) - \int_{\mu+r}^{\infty} \frac{r}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx}$$

Employing a change of variable, letting $\boldsymbol{u = \frac{x-\mu}{\sigma}}$ yields

$$\boldsymbol{\frac{\sigma}{\sqrt{2\pi}} exp\ (-\frac{r^2}{2\sigma^2}) - r\int_{\frac{r}{\sigma}}^{\infty} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{u^2}{2}\right) du}$$

$$\boldsymbol{= \frac{\sigma}{\sqrt{2\pi}} \exp\ (-\frac{r^2}{2\sigma^2}) - r\int_{\frac{r}{\sigma}}^{\infty} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{u^2}{2}\right) du}$$

$$\boldsymbol{= \frac{\sigma}{\sqrt{2\pi}} exp\ (-\frac{r^2}{2\sigma^2}) - r(1 - \Phi\left(\frac{r}{\sigma}\right))}$$

For Hermann's method, using Lagrangian multipliers, the objective function with embedded constraint can be written as

$$\boldsymbol{\Lambda(r_1, \dots, r_n, \lambda) = \sum_{i=1}^{n} \int_{\mu_i+r_i}^{\infty} (x_i - \mu_i - r_i) f(x) dx - \lambda(\sum_{i=1}^{n} r_i - r_T)}$$

In the case of normally distributed random variables,

$$\boldsymbol{\frac{\partial\Lambda}{\partial r_i} = -\frac{r_i}{\sqrt{2\pi\sigma_i^2}} exp\left(-\frac{r_i^2}{2\sigma_i^2}\right) - \left(1 - \Phi\left(\frac{r_i}{\sigma_i}\right)\right) + r_i\left(\phi\left(\frac{r_i}{\sigma_i}\right)\frac{1}{\sigma_i}\right) - \lambda = 0}$$

which simplifies to

$$\boldsymbol{\Phi\left(\frac{r_i}{\sigma_i}\right) - 1 - \lambda = 0}$$

which means that

$$\frac{\mathbf{r_i}}{\boldsymbol{\sigma}_\mathbf{i}} = \boldsymbol{\Phi}^{-1}(\mathbf{1}+\boldsymbol{\lambda})$$

Note that the right side of this equation is constant for all $i = 1, \ldots, n$, which implies that

$$\frac{r_i}{\sigma_i} = \frac{r_j}{\sigma_j}$$

or

$$r_i = \sigma_i \frac{r_j}{\sigma_j}$$

for all $i,j = 1,\ldots,n$.

Note that the constraint

$$\sum_{i=1}^{n} r_i = r_T$$

can be written as

$$r_1 + \frac{\sigma_2}{\sigma_1} r_1 + \cdots + \frac{\sigma_n}{\sigma_1} r_1 = r_T$$

and thus

$$r_1 = \frac{\sigma_1}{\sum_{i=1}^{n} \sigma_i} r_T$$

and in general

$$r_j = \frac{\sigma_j}{\sum_{i=1}^{n} \sigma_i} r_T \; for\ all\ j = 1, \ldots, n$$

Thus the percentage allocation is the proportional contribution of the j^{th} random variable to the sum of the standard deviation values. This will not often occur in practice since even in the case of a relatively large WBS, for which the central limit theorem begins to take effect, and thus for which the total risk distribution is approximately normal, the individual risks are typically skewed and not normally distributed.

Allocating Along the Gradient

Returning to the specific question of risk allocation, what are some reasonable criteria for allocating risk? Heretofore analysts have largely restricted their attention to something that seems reasonable (heuristic), subject only to the strict criteria that the allocation be complete, that is, that for a risk measure $\boldsymbol{r_T}$ and $\boldsymbol{n}$ WBS elements with respective allocation $\boldsymbol{r_1,\ldots r_n}$, that

$$\sum_{i=1}^{n} \boldsymbol{r_i} = \boldsymbol{r_T}$$

The proportional standard deviation method is illustrative of this type of approach, and allocation according to needs is an improvement on some of its inherent shortcomings such as treating uncertainty and risk as equivalent. Hermann (Hermann 2010) is among the first in the cost analysis community to suggest optimizing a specific quantity in the allocation. However, what has still been lacking is the ability to clearly discern between risk measurement, and risk allocation, and how the two relate. Also the need for other criteria in establishing a good, reasonable, or perhaps optimal risk allocation has gone largely unnoticed. In the case that risk dollars cannot be reallocated among WBS elements, Hermann's approach offers a good optimization criterion, and he constrains his allocation to be a complete one. However, this is not typically seen in risk management practice, so its practical utility in most cases is limited.

Gradient allocation is a commonly used way to allocate risk in finance and insurance and has been found by different authors in different fields using different approaches to meet the single allocation method that meets simple criteria for allocating risk. This method is also consistent with coherent risk measures. And as long as the risk measure meets a simple criterion, it is guaranteed to be total.

Gradient allocation involves allocating to each WBS element an amount equal to the gradient of the risk measure. To define this, consider an $\boldsymbol{n}$-element WBS with cost random variables denoted by $\boldsymbol{X_1,\ldots X_n}$ and portfolio weights denoted by $\boldsymbol{\lambda_1,\ldots \lambda_n}$. The total cost for the project is found by summing the individual WBS elements, accounting for the weights, i.e.

$$\boldsymbol{x(\lambda)} = \sum_{i=1}^{n} \boldsymbol{\lambda_i x_i}$$

If the total risk measure is denoted by r and the risk measure for each individual WBS is denoted by $\boldsymbol{r_i}$ then the gradient of $\boldsymbol{r}$ is defined as

$$\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\lambda_i}}$$

which reflects the rate of change in the total risk relative to the rate of change in the portfolio weight for individual WBS elements. As long as the risk measure is positive homogeneous (a property shared by all the risk measures discussed in this paper), such a risk allocation is guaranteed to be a complete allocation. This is due to Euler's homogeneous function theorem, which (for the purposes of this paper) states that a continuously differentiable real-valued function $\boldsymbol{r}: \mathbb{R}^{\boldsymbol{n}} \longrightarrow \mathbb{R}$ is positively homogeneous if and only if

$$\sum_{i=1}^{n} \lambda_i \frac{\partial r}{\partial \lambda_i} = r(\lambda)$$

Note that if $\boldsymbol{r}$ is positive homogeneous, then $\boldsymbol{r}(\boldsymbol{\alpha\lambda}) = \boldsymbol{\alpha r}(\boldsymbol{\lambda})$ for $\boldsymbol{\alpha} \in \mathbb{R}_+$.

Then clearly

$$\frac{d\alpha r(\lambda)}{d\alpha} = r(\lambda)$$

Then taking the derivative of $\boldsymbol{r}(\boldsymbol{\alpha\lambda})$ with respect to $\boldsymbol{\alpha}$

$$\sum_{i=1}^{n} \frac{\partial r(\alpha\lambda)}{\partial(\alpha\lambda_i)} \frac{\partial(\alpha\lambda_i)}{\partial\alpha} = \sum_{i=1}^{n} \frac{\partial r(\alpha\lambda)}{\partial(\alpha\lambda)} \lambda_i$$

Since this is true for all $\boldsymbol{\alpha}$ it is also true for $\boldsymbol{\alpha=1}$, so

$$r(\lambda) = \sum_{i=1}^{n} \frac{\partial r(\lambda)}{\partial\lambda} \lambda_i$$

Thus, if a risk measure is positive homogenous, then the risk can be allocated to each constituent element by its gradient. This allocation is complete, so no additional constraint is needed to ensure this property holds. Also, it provides a natural connection between risk measure and risk allocation. Given a risk measure, the allocation method is derived directly and is specific to the method used to measure risk. It is popular in finance and insurance, where it is also referred to as the Euler principle, due to its connection to Euler's theorem. In the last decade several authors have written papers urging its use, and have derived it from relatively simple criteria (Tasche 1999, Denault 2001, Kalkbrenner 2005). Arguments for its use have been based on economic principles, simple axioms related to continuity and diversification, and even game theory.

In terms of economic principles, it has been argued that risk should be viewed as relative to its performance (Tasche 1999). In terms of cost analysis, this would be the cost relative to the risk, as measured by the ratio

$$\frac{E(X)}{r}$$

The economic performance criteria is then defined (Tasche 1999) as

$$\frac{\partial}{\partial \lambda_i}\left(\frac{E(X(\lambda)}{r(\lambda)}\right)\begin{cases} > 0 \ if \ \dfrac{E(X_i)}{r_i} > \dfrac{E(X(\lambda)}{r(\lambda)} \\ < 0 \ if \ \dfrac{E(X_i)}{r_i} < \dfrac{E(X(\lambda)}{r(\lambda)} \end{cases}$$

Tasche (1999) demonstrated that gradient allocation is the only allocation method that meets this condition under the condition that the gradient is continuous. To see this, first assume that gradient allocation is the method used. Then applying the product rule yields

$$\frac{\partial}{\partial \lambda_i}\left(\frac{E\big(X(\lambda)\big)}{r(\lambda)}\right) = \frac{1}{r(\lambda)}\frac{\partial E\big(X(\lambda)\big)}{\partial \lambda_i} - \frac{E(X(\lambda))}{r(\lambda)^2}\frac{\partial r}{\partial \lambda_i}$$

Since

$$\frac{\partial E\big(X(\lambda)\big)}{\partial \lambda_i} = \frac{\partial r}{\partial \lambda_i}\big(\lambda_1 E(X_1) + \cdots + \lambda_n E(X_n)\big) = E(X_i)$$

and

$$\frac{\partial r}{\partial \lambda_i} = r_i$$

for $i = 1,\ldots,n$ by assumption, the economic performance condition is seen to hold.

Now assume that the performance condition holds. Recalling the product rule expansion of $\frac{\partial}{\partial \lambda_i}\left(\frac{E(X(\lambda))}{r(\lambda)}\right)$, then if

$$\frac{E(X_i)}{r_i} = \frac{E(X(\lambda)}{r(\lambda)}$$

and

$$\frac{1}{r(\lambda)}\frac{\partial E\big(X(\lambda)\big)}{\partial \lambda_i} - \frac{E(X(\lambda))}{r(\lambda)^2}\frac{\partial r}{\partial \lambda_i} = 0$$

It follows that

$$\frac{\partial r}{\partial \lambda_i} = \frac{E(X_i)}{E(X(\lambda)}r(\lambda) = r_i$$

But since $\frac{\partial}{\partial \lambda_i}$ is a continuous function, a limiting series of values such that

$$\lim_{n\to\infty}\left(\frac{\boldsymbol{E(X_i)}}{\boldsymbol{r_i}}\right)_n = \frac{\boldsymbol{E(X(\lambda)}}{\boldsymbol{r(\lambda)}}$$

means that

$$\lim_{n\to\infty}\left(\frac{\boldsymbol{\partial}}{\boldsymbol{\partial\lambda_i}}\left(\frac{\boldsymbol{E\big(X(\lambda)\big)}}{\boldsymbol{r(\lambda)}}\right)\right)_n = 0$$

which implies that

$$\frac{\boldsymbol{\partial r}}{\boldsymbol{\partial \lambda_i}} = \boldsymbol{r_i}$$

Using a simple set of criteria, another derivation of gradient allocation was obtained (Kalkbrenner 2005). Kalkbrenner defines three conditions. First, the allocation must be a linear function. Second, the allocation must be diversifying in the sense that $\boldsymbol{r_i(X_i)} \le \boldsymbol{r(X_i)}$ for all $\boldsymbol{i=1,\dots,n}$. That is the risk allocated to the i^{th} element should be no larger than the risk measure for that particular element. Third, the risk allocation function must be continuous. Given these conditions, the gradient allocation method is the only one that meets all three criteria. Before proceeding with the proof, note that for $\boldsymbol{\epsilon} \in \mathbb{R}$

$$\boldsymbol{r(\lambda_1,\dots,\lambda_i+\epsilon,\dots,\lambda_n)} = \boldsymbol{r}\Big(\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}\boldsymbol{\lambda_i X_i} + \boldsymbol{\epsilon X_i}\Big) = \boldsymbol{r(X+\epsilon X_i)}$$

Also note that $\boldsymbol{r_i}$ denotes the allocation of $\boldsymbol{X}$ to the i^{th} element, X_i. This could also be written as $\boldsymbol{r_i(X_i,X)}$ or simply as $\boldsymbol{r(X_i,X)}$ when both sets are specified.

Since the allocation is diversifying, for $\boldsymbol{\epsilon},\boldsymbol{\epsilon^*} \in \mathbb{R}$

$$\boldsymbol{r(X+\epsilon^* X_i)} \ge \boldsymbol{r(X+\epsilon^* X_i, X+\epsilon X_i)}$$

And since it is linear the latter term is equivalent to

$$\boldsymbol{r(X+\epsilon X_i+(\epsilon^*-\epsilon)X_i, X+\epsilon X_i)} = \boldsymbol{r(X+\epsilon X_i)} + \boldsymbol{(\epsilon^*-\epsilon)r(X_i, X+\epsilon X_i)}$$

Without loss of generality assume $\boldsymbol{\epsilon} < \boldsymbol{\epsilon^*}$. Then

$$\boldsymbol{r(X_i, X+\epsilon X_i)} \le \frac{\boldsymbol{r(X+\epsilon^* X_i)} - \boldsymbol{r(X+\epsilon X_i)}}{\boldsymbol{\epsilon^*-\epsilon}}$$

and swapping the $\boldsymbol{\epsilon' s}$ and the $\boldsymbol{\epsilon^{*\prime} s}$ yields

$$\frac{\boldsymbol{r(X+\epsilon^* X_i)} - \boldsymbol{r(X+\epsilon X_i)}}{\boldsymbol{\epsilon^*-\epsilon}} \le \boldsymbol{r(X_i, X+\epsilon^* X_i)}$$

Putting the two inequalities together and taking the limit as $\boldsymbol{\epsilon^*} \to \boldsymbol{0}$ it is found that

$$\boldsymbol{r(X_i, X) \leq \frac{r(X) - r(X + \epsilon X_i)}{\epsilon} \leq r(X_i, X + \epsilon X_i)}$$

Taking the limit as $\boldsymbol{\epsilon \to 0}$ and noting that the allocation function is continous, the result is that the allocation is equal to $\frac{\partial r}{\partial \lambda_i}$ for all $\boldsymbol{i = 1,\ldots,n}$.

The gradient allocation principle has also been derived from game-theoretic arguments. Rather than the non-cooperative game theory that most people are familiar with, such as popularized in the prisoner's dilemma and in the film *A Beautiful Mind*, risk allocation can be viewed as a cooperative game, where the coalitions or elements work in accordance to allocate total risk. A game that allows fractional allocations, such as with cost risk allocation, is a "fuzzy" game. It has been found with some simple criteria that the only allocation principle consistent with them is gradient allocation (Denault 2001). These criteria are the diversifying allocation principle (also used by Kalkbrenner (Kalkbrenner 2005)); the property of symmetry, which means that if by adding any set to the portfolio, any two subportfolios that contribute the same amount of risk will also receive the same allocation; and a riskless item will receive only its cost in the allocation scheme, no more, no less. It is interesting to note that in game theory, gradient allocation is referred to as the Aumann-Shapley value. For more information see Denault (2001) and Aubin (2007).

In terms of criteria for cost risk allocation, the notion of economic performance may be a good one for activities involving profit and loss, but is not as big a motivating factor for the cost of government projects where the activities are determined according to scientific pursuits, technological objectives, or the needs of national defense. On the other hand, diversification for an allocation makes sense regardless of the application. The amount allocated to a specific WBS element should be less than or equal to the contribution of that element to the overall risk. Gradient allocation thus meets logical, sound criteria, is linked with and thus consistent with the risk measure used, and is naturally a complete allocation without requiring an explicit constraint.

Hermann demonstrates that his risk allocation method is similar to a gradient allocation (Hermann 2010). However, his method only looks at the gradient in the direction of the individual WBS element without considering dependencies (correlation) between the elements.

Application of Gradient Allocation

Note that gradient allocation potentially indicates different risk allocation algorithms for different methods of risk measurement, since the method of allocation explicitly depends upon the risk measurement. In this section, the gradient allocations for the risk measurement methods discussed in a previous section are calculated. In what follows, an $\boldsymbol{n}$-element WBS with cost random variables denoted by $\boldsymbol{X_1,\ldots X_n}$ with portfolio weights $\boldsymbol{\lambda=(\lambda_1,\ldots,\lambda_n)}$ is assumed.

Standard Deviation Principle

Denote the covariance matrix for the WBS by $\boldsymbol{\Sigma}$. Note that

$$\boldsymbol{r}(\boldsymbol{\lambda}) = \boldsymbol{\mu} + \boldsymbol{k\sigma} = \boldsymbol{\mu} + \boldsymbol{k}\sqrt{\boldsymbol{\lambda}'\boldsymbol{\Sigma}\boldsymbol{\lambda}}$$

Then applying gradient allocation, it is found that

$$\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\lambda}_i} = \boldsymbol{\mu}_i + \boldsymbol{k}\frac{\boldsymbol{\Sigma}\boldsymbol{\lambda}_i}{\sqrt{\boldsymbol{\lambda}'\boldsymbol{\Sigma}\boldsymbol{\lambda}}} = \boldsymbol{\mu}_i + \boldsymbol{k}\frac{\sum_{j=1}^{n} \boldsymbol{Cov}(\boldsymbol{X}_i\boldsymbol{X}_j)\boldsymbol{\lambda}_j}{\sqrt{\boldsymbol{\lambda}'\boldsymbol{\Sigma}\boldsymbol{\lambda}}}$$

Setting $\boldsymbol{\lambda_i=1}$ for all $\boldsymbol{i=1,\ldots,n}$ it is easy to see that

$$\sum_{i=1}^{n} \boldsymbol{\mu}_i + \boldsymbol{k}\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} \boldsymbol{Cov}(\boldsymbol{X}_i\boldsymbol{X}_j)}{\boldsymbol{Std.Dev.}(\boldsymbol{X})} = \boldsymbol{\mu} + \boldsymbol{k\sigma}$$

since

$$\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} \boldsymbol{Cov}(\boldsymbol{X}_i\boldsymbol{X}_j)}{\boldsymbol{Std.Dev.}(\boldsymbol{X})} = \frac{\boldsymbol{\sigma}^2}{\boldsymbol{\sigma}} = \boldsymbol{\sigma}$$

This allocation method is also referred to as the covariance principle (McNeil et al., 2005).

Value at Risk

Tasche (Tasche 2000) demonstrated that allocation along the gradient for *VaR* amounts to

$$\boldsymbol{E}(\boldsymbol{X}_i|\boldsymbol{X} = \boldsymbol{VaR}(\boldsymbol{X}))$$

To see this, assume that $\boldsymbol{n} \geq \boldsymbol{2}$ and that $\boldsymbol{(X_1,\ldots,X_n)}$ has a joint density. Let $\boldsymbol{(\lambda_1,\ldots,\lambda_n)}$ be a vector of portfolio weights with $\boldsymbol{\lambda_1} > 0$. Note that

$$\boldsymbol{P}(\boldsymbol{X}(\boldsymbol{\lambda}) \leq \boldsymbol{t}) = \boldsymbol{E}\big(\boldsymbol{P}(\boldsymbol{X}(\boldsymbol{\lambda}) \leq \boldsymbol{t}|\boldsymbol{X}_2, \ldots, \boldsymbol{X}_n)\big)$$

$$= \boldsymbol{E}(\boldsymbol{P}\left(\boldsymbol{X}_1 \leq \frac{\boldsymbol{t}-\sum_{j=2}^{n}\boldsymbol{\lambda}_j\boldsymbol{X}_j}{\boldsymbol{\lambda}_1}\middle|\boldsymbol{X}_2, \ldots, \boldsymbol{X}_n\right))$$

$$= \boldsymbol{E}(\int_{\boldsymbol{0}}^{\boldsymbol{\lambda}_1^{-1}(\boldsymbol{t}-\sum_{j=2}^{n}\boldsymbol{\lambda}_j\boldsymbol{X}_j)} \boldsymbol{f}_{\boldsymbol{X}_1|\boldsymbol{X}_2,\ldots\boldsymbol{X}_n}(\boldsymbol{u}, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)\boldsymbol{du})$$

where $\boldsymbol{f}_{\boldsymbol{X}_1|\boldsymbol{X}_2,\ldots\boldsymbol{X}_n}$ is the conditional density function of $\boldsymbol{X_1}$.

Taking the derivative under the expectation yields

$$\boldsymbol{f_{X(\lambda)}(t) = \frac{1}{\lambda_1} E\left(\left(f_{X_1|X_2,\dots X_n}\left(\frac{1}{\lambda_1}\left(t - \sum_{j=2}^{n} \lambda_j x_j\right), x_2, \dots, x_n\right)\right)\right)}$$

Also note that

$$\boldsymbol{E(X_i|X(\lambda) = t) = \lim_{\delta \to 0} \frac{\delta^{-1} E(X_i I_{\{t \le X(\lambda) \le t+\delta\}})}{\delta^{-1} P(t < X(\lambda) \le t + \delta)}}$$

$$= \frac{\frac{\partial}{\partial t} E(X_i I_{\{X(\lambda) \le t\}})}{f_{X(\lambda)}(t)}$$

$$= \frac{\frac{\partial}{\partial t} E(X_i \int_0^{\lambda_1^{-1}(t - \sum_{j=2}^{n} \lambda_j X_j)} f_{X_1|X_2,\dots X_n}(u, x_2, \dots, x_n) du)}{f_{X(\lambda)}(t)}$$

as long as $\boldsymbol{f_{X(\lambda)}(t) \neq 0}$ and $\boldsymbol{i \ge 2.}$ Taking the derivative under the expectation and simplifying yields

$$\boldsymbol{E(X_i|X(\lambda) = t) = \frac{E(X_i f_{X_1|X_2,\dots X_n}(\frac{1}{\lambda_1}(t - \sum_{j=2}^{n} \lambda_j x_j), x_2, \dots, x_n))}{E\left(\left(f_{X_1|X_2,\dots X_n}\left(\frac{1}{\lambda_1}(t - \sum_{j=2}^{n} \lambda_j x_j), x_2, \dots, x_n\right)\right)\right)}}$$

In the case of ***i=1*** it can be derived in similar fashion that

$$\boldsymbol{E(X_i|X(\lambda) = t) = \frac{E(\frac{t - \sum_{j=2}^{n} \lambda_j x_j}{\lambda_1} f_{X_1|X_2,\dots X_n}(\frac{1}{\lambda_1}(t - \sum_{j=2}^{n} \lambda_j x_j), x_2, \dots, x_n))}{E\left(\left(f_{X_1|X_2,\dots X_n}\left(\frac{1}{\lambda_1}(t - \sum_{j=2}^{n} \lambda_j x_j), x_2, \dots, x_n\right)\right)\right)}}$$

Note that

$$\boldsymbol{\alpha = P\big(L(\lambda) \le VaR_\alpha(\lambda)\big) = E(\int_0^{\frac{VaR_\alpha(\lambda) - \sum_{j=2}^{n} \lambda_j X_j}{\lambda_1}} f_{X_1|X_2,\dots X_n}(u, x_2, \dots, x_n) du)}$$

Taking the derivative with respect to $\boldsymbol{\lambda_i}$***, i = 2,...n***, yields

$$\boldsymbol{0 = \lambda_1^{-1} E((\frac{\partial VaR(\lambda)}{\partial \lambda_i} - X_i) f_{X_1|X_1,\dots X_n}\left(\frac{1}{\lambda_1}\left(VaR_\alpha(\lambda) - \sum_{j=2}^{n} \lambda_j x_j\right), x_2, \dots, x_n\right))}$$

Solving for $\frac{\partial VaR(\lambda)}{\partial \lambda_i}$ and substituting the result derived for $E\left(X_i \middle| X(\lambda) = VaR_\alpha(\lambda)\right)$**,** it is found that

$$\frac{\partial VaR(\lambda)}{\partial \lambda_i} = E\left(X_i \middle| X(\lambda) = VaR_\alpha(\lambda)\right)$$

again, for ***i = 2,…,n***.

Taking the derivative of

$$\alpha = P\left(L(\lambda) \le VaR_\alpha(\lambda)\right) = E\left(\int_0^{\frac{VaR_\alpha(\lambda) - \sum_{j=2}^{n} \lambda_j x_j}{\lambda_1}} f_{X_1|X_2,\dots X_n}(u, x_2, \dots, x_n)\,du\right)$$

with respect to λ_1 yields

$$0 = E\left(\Gamma \cdot f_{X_1|X_1,\dots X_n}\left(\frac{1}{\lambda_1}\left(VaR_\alpha(\lambda) - \sum_{j=2}^{n} \lambda_j x_j\right), x_2, \dots, x_n\right)\right)$$

where $\Gamma = \left(\frac{\frac{\partial VaR_\alpha(\lambda)}{\partial \lambda_i}}{\lambda_1} + \frac{\sum_{j=2}^{n} \lambda_j X_j - VaR_\alpha(\lambda)}{\lambda_1^2}\right)$. Solving for $\frac{\partial VaR_\alpha(\lambda)}{\partial \lambda_i}$ and substituting the result derived for ***i= 1*** gives

$$\frac{\partial VaR(\lambda)}{\partial \lambda_i} = E\left(X_i \middle| X(\lambda) = VaR_\alpha(\lambda)\right)$$

The result is simple, and even intuitive, even though the mechanics of the derivation are complicated. However, even though the formula appears simple, it is not easy to calculate in practice. This is not a simple, straightforward conditional expected value calculation, since for continuous distributions, the probability that $X(\lambda) = VaR_\alpha(\lambda)$ will be zero. In the case of continuous distributions, a simple linear approximation can be found by noting that in the subject of linear regression, $E\left(X_i \middle| X(\lambda)\right)$ represents the best estimate of X_i by X. Thus a simple linear approximation can be found by minimizing $E((X_i - a - bX)^2)$ which is well known as

$$b = \frac{Cov(X_i, X)}{Var(X)}$$

$$a = E(Y) - bE(X)$$

where *Cov(X,Y)* is the covariance between X and Y. Plugging these values into the estimate yields

$$E(\widehat{X_i|X}(\lambda)) = E(X_i) + \frac{Cov(X_i, X)}{Var(X)}(VaR_\alpha - E(X))$$

Note that "Var" in the formula denotes variance while "VaR" denotes the "value at risk" or percentile. Note that this approximation amounts to applying the covariance principle to the difference of the percentile at which the project is funded and the total expected value, or mean.

In the case of Monte Carlo simulations, kernel smoothing or some other smoothing technique will likely be needed to overcome the issue that it is possible that none of the sample values will likely have a value such that $X(\lambda) = VaR_\alpha(\lambda)$. That is a subject deserving of a paper of its own and thus it is not covered in more detail here.

Expected Shortfall

Suppose VaR is set at the α^{th} percentile. Then the expected shortfall risk is defined as

$$r(\lambda) = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(\lambda)du$$

Calculating the gradient with respect to λ yields

$$\frac{\partial r}{\partial \lambda_i} = \frac{1}{1-\alpha}\int_\alpha^1 \frac{\partial VaR_u}{\partial \lambda_i}du$$

Plugging in the formula for the partial derivative of *VaR* with respect to λ_i obtained in the preceding section it is found that

$$\frac{\partial r}{\partial \lambda_i} = \frac{1}{1-\alpha}\int_\alpha^1 E\big(X_i|X(\lambda) = VaR_u(\lambda)\big)du$$

Let $v = VaR_u\big(X(\lambda)\big) = F_{X(\lambda)}^{-1}(u)$. Then since $f_{X(\lambda)}(v)dv = du$

$$\frac{1}{1-\alpha}\int_\alpha^1 E\big(X_i|X(\lambda) = VaR_u(\lambda)\big)du = \frac{1}{1-\alpha}\int_{VaR_\alpha}^\infty E(X_i|X(\lambda) = v)f_{X(\lambda)}(v)dv$$

$$= \frac{1}{1-\alpha}E(X_i; X(\lambda) \geq VaR_\alpha)$$

$$= E(X_i|X(\lambda) \geq VaR_\alpha)$$

That is, the gradient capital allocation is

$$\frac{\partial r}{\partial \lambda_i} = E(X_i|X(\lambda) \geq VaR_\alpha(\lambda))$$

While similar in form to the capital allocation for *VaR* (the only difference is that the equality in the conditioned expectation is now an inequality), this is more intuitive and easier to calculate than the *VaR* allocation. For a Monte Carlo simulation, it is simply the contribution of the i^{th} element to the expected shortfall.

One-Sided Moments

For the risk measure associated with the $\boldsymbol{p}^{th}$ one-sided moment

$$\boldsymbol{r(\lambda) = \mu + \left(\int_{-\infty}^{\infty} (x(\lambda) - \mu)_+^p\right)^{\frac{1}{p}} f(x)dx}$$

The gradient allocation principle is straightforward to apply. In this case,

$$\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\lambda_i}} = \boldsymbol{E(X_i) + \frac{1}{p}\left(\int_{-\infty}^{\infty} (x(\lambda) - \mu)_+^p\right)^{\frac{1}{p}-1} f(x)dx \cdot \int_{-\infty}^{\infty} p(x(\lambda) - \mu)_+^{p-1} f(x)dx \cdot (X_i - E(X_i))}$$

$$= \boldsymbol{E(X_i) + \left(\sigma_{+,p}\right)^{1-p} E((X_i - E(X_i)) \cdot \left(X - E(X)\right)_+^{p-1})}$$

where the second expectation is over the values for which $\sum_{i=1}^{n} \boldsymbol{\lambda_i X_i} > E(X)$.

When ***p* = *1***, which is Hermann's suggested risk measure (Hermann 2010)

$$\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\lambda_i}} = \boldsymbol{E(X_i) + E(\left(X_i - E(X_i)\right) \cdot 1_{\{\sum \lambda_i X_i > E(X)\}})}$$

$$= \boldsymbol{E(X_i) + P(\sum_{i=1}^{n} \lambda_i X_i} > E(X)) \cdot E(\boldsymbol{X_i - E(X_i)}| \sum_{i=1}^{n} \boldsymbol{\lambda_i X_i} > E(X))$$

When ***p* = *2***, the semi-standard deviation principle, the allocation scheme simplifies to

$$\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\lambda_i}} = \boldsymbol{E(X_i) + \frac{E(\left(X_i - E(X_i)\right)\left(X - E(X)\right)_+)}{\sigma_{+,2}}}$$

which is a one-sided covariance principle.

Summary of Gradient Allocation Formulas

A summary of the gradient allocation formulas presented in this section are summarized in Table 3.

Risk Measure	Associated Gradient Allocation Formula
Standard Deviation Principle	$\boldsymbol{\mu_i + k \frac{\sum_{j=1}^{n} Cov(X_i X_j)}{Std.Dev.(X)}}$
Value at Risk	$\boldsymbol{E(X_i \mid X = VaR_\alpha)}$
Expected Shortfall	$\boldsymbol{E(X_i \mid X \geq VaR_\alpha)}$
1st One-Sided (Positive) Moment	$\boldsymbol{E(X_i) + P(\sum_{i=1}^{n} X_i > E(X)) \cdot E(X_i - E(X_i) \mid \sum_{i=1}^{n} X_i > E(X))}$
Semi-Standard Deviation Principle	$\boldsymbol{E(X_i) + \frac{E((X_i - E(X_i))(X - E(X))_+)}{\sigma_{+,2}}}$

Table 3. Summary of Gradient Allocation Formulas for Five Risk Measures.

For the 10-project example given in Table 1, the results of applying the risk allocation methods discussed in this paper are shown in Table 4.

Allocation Method	Project 1	Project 2	Project 3	Project 4	Project 5	Project 6	Project 7	Project 8	Project 9	Project 10
Proportional Standard Deviation	14.0%	5.5%	7.6%	10.1%	10.6%	10.6%	13.7%	5.8%	9.9%	12.2%
Needs Method	16.2%	5.8%	7.8%	8.8%	12.8%	11.9%	9.1%	6.5%	10.7%	10.5%
Hermann's Method	14.8%	7.1%	8.7%	8.6%	13.2%	12.2%	7.0%	8.1%	11.0%	9.4%
Gradient - Standard Deviation Principle	15.3%	4.7%	6.9%	9.9%	10.5%	10.5%	14.7%	5.0%	9.6%	12.7%
Gradient - Value at Risk*	15.3%	4.7%	6.9%	9.9%	10.5%	10.5%	14.7%	5.0%	9.6%	12.7%
Gradient - Expected Shortfall	9.3%	8.4%	10.4%	9.8%	14.2%	6.9%	8.1%	8.4%	12.9%	11.5%
Gradient - 1st One-Sided Moment	13.7%	5.1%	9.9%	12.5%	15.7%	4.8%	7.0%	9.8%	10.8%	10.8%
Gradient - Semi-Standard Deviation Principle	15.6%	4.7%	9.5%	13.1%	15.4%	4.5%	6.8%	10.0%	10.2%	10.3%
* linear approximation applied										

Table 4. Comparison of Eight Risk Allocation Methods.

Note the differences in the allocations, which are significant. Hermann's method differs significantly from the first one-sided (positive) moment gradient allocation, for which the risk measure is the same. This may be due to the numerical methods required to find the optimal allocation method.

CONCLUSION

Current risk allocation theory and practice and relatively new methods for risk allocation have been discussed. The proportional standard deviation method and the needs method are heuristics that do not necessarily have optimal properties. A new method that explicitly seeks to minimize the sum of the allocated expected shortfalls beyond the mean was discussed. Risk allocation methods have not sought to distinguish between measurement and allocation, so risk measurement was also summarized. It was pointed out that the problems of risk measurement and risk allocation are separate and distinct but related topics. The concept of coherence for risk measures was discussed, and relatively new coherent risk measurement methods, such as expected shortfall were treated in depth. A new risk allocation method that is becoming increasingly popular in finance and insurance was discussed, which is gradient allocation.

Gradient allocation links together risk measurement and risk management, and in given certain criteria for allocating risk, proves to be the best method for an associated risk measurement method.

It was found that current risk allocation methods fall within this theoretical framework. The proportional standard deviation method is a special case of Hermann's allocation method, under the condition that all risks are normally distributed and the risk allocations are not reversible – once allocated, no re-allocation is possible. These are not practical conditions – in practice risk allocations can be re-allocated, and even if the WBS is large enough to apply the central limit theorem at the total level, individual risks are typically skewed, and are better represented by distributions capable of modeling this skew, such as lognormal distributions. Hermann's suggested risk measure turns out to be coherent, and thus a better measure of risk than percentile funding. This risk measure, in the example shown, is the smallest risk measure above the mean. And in many cases, such as for the normal distribution, this amount will be less than the 70th percentile, making it a potentially attractive risk measure for policymakers, who are constrained by tight budgets.

The needs method falls within the gradient allocation framework – it is similar to the semi-standard deviation method, and when applied to the difference between a percentile and the mean, is similar to the best linear estimator for gradient allocation when value at risk is used for risk measurement.

It is likely that analysts and policymakers are using one or more of the risk measures in this presentation. For each, an optimal risk allocation method has been presented and summarized. The author advocates the use of coherent risk measures, and associated gradient allocations. Non-coherent risk measures, such as value at risk, do not present effective risk management solutions, especially since value at risk lacks a portfolio effect. And not using gradient allocation does not meet the criteria of diversification for risk allocations, which is not desirable. The current methods available are practical heuristics, but looking at the theory of risk allocation indicates that there are better methods, gradient allocation among them.

REFERENCES

1. Anderson, T. P. "The Trouble With Budgeting to the 80th Percentile", presented at the 72nd Military Operations Research Society Symposium, 2004.

2. Artzner, P. et al., "Coherent Measures of Risk," Mathematical Finance 9:203-228, 1999.

3. Aubin, P., *Mathematical Methods of Game and Economic Theory*, Dover, Amsterdam, 2007.

4.Book, S.A. "Allocating 'Risk Dollars' Back to WBS Elements" LLC SSCAG/EACE/SCAF Meeting 19-21 September 2006.

5. Book, S.A., “Allocating ‘Risk Dollars’ Back to WBS Elements,” presented to the SSCAG Risk Subgroup, 17-18 January 2007.

6. Denault, M., “Coherent Allocation of Risk Capital,” Journal of Risk 4(1):7-21.

7. Hermann, W., “Optimal Risk Allocation: Theory, Heuristics, and Best Approximations,” Ballistic Missile Defense System Analysis Team technical note, May, 2010.

8. Hermann, W., personal communication, April, 2010.

9. Kalkbrenner, M., “An Axiomatic Approach to Capital Allocation,” Mathematical Finance 15(3):425-437, 2005.

10. Markowitz, H., *Portfolio Selection: Efficient Diversification of Investments*, Wiley, Yale University Press, 1959.

11. McNeil,A.J., et al., *Quantitative Risk Management*, Princeton University Press, Princeton, 2005.

12 Sandberg, J., “Cost Risk Allocation: Objectives, Tendencies, and Limitations,” presented at the Joint ISPA/SCEA Conference, June, 2007.

13. Smart, C., “Risk Analysis in the NASA/Air Force Cost Model,” presented at the Joint ISPA/SCEA Conference, June 2005.

14. Smart, C., “Here There Be Dragons: Considering the Right Tail in Risk Management,” presented at the Joint ISPA/SCEA Conference, June, 2010.

15. Smart, C., “The Portfolio Effect and the Free Lunch,” presented at the Joint ISPA/SCEA Conference, June, 2009.

16. Tasche, D., “Risk contributions and performance measurement,” Preprint, Department of Mathematics, TU-Munchen, 1999.

17. Tasche, D., “Conditional Expectation as a Quantile Derivative,” Preprint, Department of Mathematics, TU-Munchen, 2000.

18. Tasche, D. “Allokation des ökonomischen Kapitals auf Teilportfolios und Transaktionen,” in *Handbuch Ökonomisches Kapital*, A. Becker, V. Gehrmann und H. Schulte-Mattler (Herausgeber), Fritz-Knapp Verlag, pp. 193-212, 2008.