Optimal Capital Allocation Principles*

Jan Dhaene[†‡] Andreas Tsanakas [§] Emiliano A. Valdez [¶] Steven Vanduffel [‖]

April 20, 2010

Abstract

This paper develops a unifying framework for allocating the aggregate capital of a financial firm to its business units. The approach relies on an optimisation argument, requiring that the weighted sum of measures for the deviations of the business unit's losses from their respective allocated capitals be minimised. The approach is fair insofar as it requires capital to be close to the risk that necessitates holding it. The approach is additionally very flexible in the sense that different forms of the objective function can reflect alternative definitions of corporate risk tolerance. Owing to this flexibility, the general framework reproduces several capital allocation methods that appear in the literature and allows for alternative interpretations and possible extensions.

Keywords: Capital allocation; risk measure; comonotonicity; Euler allocation; default option; optimisation.

1 Introduction

The level of the capital held by a bank or an insurance company is a key issue for its stakeholders. The regulator, primarily sharing the interests of depositors and policyholders, establishes rules to determine the required capital to be held by the company. The level of this capital is determined such that the company will be able to meet its financial obligations with a high probability as they fall due, even in adverse situations. Rating agencies rely on the level of available capital to assess the financial strength of a company. Shareholders and investors alike are concerned with the risk of their capital investment and the return that it will generate.

The determination of a sufficient amount of capital to hold is only part of a larger risk management and solvency policy. The practice of Enterprise Risk Management (ERM) enhances identifying, measuring, pricing, and controlling risks. An important component of an ERM framework is the exercise of *capital allocation*, a term referring to the subdivision of the aggregate capital held by the firm across its various constituents, e.g. business lines, types of exposure, territories or even individual products in a portfolio of insurance policies.

Most financial firms write several lines of business and may want their total capital allocated across these lines for a number of reasons. First, there is a need to redistribute the total (frictional or opportunity) cost associated with holding capital across various business lines so that this cost is equitably transferred back to the depositors or policyholders in the form of charges.

*An early draft of this paper was presented at the 9th International Congress on Insurance: Mathematics and Economics (IME2005) in Laval, Quebec, Canada.

†Corresponding author: jan.dhaene@econ.kuleuven.be (phone: +32 16 326750)

‡Faculty of Business and Economics, Katholieke Universiteit, Leuven, Belgium.

§Cass Business School, City University, London, United Kingdom.

¶Department of Mathematics, University of Connecticut, Storrs, Connecticut, USA.

‖Department of Economics and Political Science, Vrije Universiteit Brussel (VUB), Brussel, Belgium.

Secondly, the allocation of expenses across lines of business is a necessary activity for financial reporting purposes. Thirdly, capital allocation provides for a useful device of assessing and comparing the performance of the different lines of business by determining the return on allocated capital for each line. Comparing these returns allows one to distinguish the most profitable business lines and hence may assist in remunerating the business line managers. Finally, allocating capital may help to identify areas of risk consumption within a given organisation and support the decision making concerning business expansions, reductions or even eliminations.

There is a countless number of ways to allocate the aggregate capital of a company to its different business units. Mutual dependencies that may exist between the performances of the various business units make capital allocation a non-trivial exercise. Accordingly, there is an extensive amount of literature on this subject with a wide number of proposed capital allocation algorithms. Cummins (2000) provides an overview of several methods suggested for capital allocation in the insurance industry and relates capital allocation to management decision making tools such as RAROC (risk-adjusted return on capital) and EVA (economic value added). Myers and Read Jr. (2001) consider capital allocation principles based on the marginal contribution of each business unit to the company's default option. LeMaire (1984) and Denault (2001) discuss capital allocations based on game theoretic considerations, where a risk measure is used as a cost functional. In the case of coherent risk measures (see Artzner et al. (1999)), such capital allocations reduce to subdivisions according to marginal costs. Overbeck (2000) considers marginal contributions to the expected shortfall risk measure in a credit risk context. In closely related works, marginal ('Euler') capital allocations are proposed within a portfolio optimisation context by Tasche (2004) and an axiomatic allocation system is proposed by Kalkbrener (2005). A commentary on the various approaches to allocating capital has appeared in Venter (2004). A recent work by Kim and Hardy (2008) proposed a method based on a solvency exchange option and which explicitly accounts for the notion of limited liability.

Panjer (2001) considers the particular case of multivariate normally distributed risks and provides an explicit expression of marginal cost based allocations, when the risk measure used is Tail Value-at-Risk (TVaR). Landsman and Valdez (2003) extends these explicit capital allocation formulas to the case where risks belong to the class of multivariate elliptical distributions, for which the class of multivariate normal is a special case. Dhaene et al. (2008) re-derive the results of Landsman and Valdez (2003) in a more straightforward manner and apply these to sums that involve normal as well as lognormal risks. In Valdez and Chernih (2003), expressions for covariance-based allocations are derived for multivariate elliptical risks. Tsanakas (2004) studies allocations where the relevant risk measure belongs to the class of distortion risk measures, while Tsanakas (2008) extends these allocation principles to the more general class of convex risk measures including the exponential risk measures. By considering the link between solvency and a fair rate of return, Sherris (2006) developed allocation principles consistent with the economic value of a financial institution's balance sheet. Furman and Zitikis (2008b) introduce the class of weighted risk capital allocations "which stems from the weighted premium calculation principle".

The multitude of allocation methods proposed in the literature can be bewildering, with the justifications of allocation approaches varying between e.g. economic (Tasche, 2004), game-theoretic (Denault, 2001), and axiomatic (Kalkbrener, 2005) criteria, while some authors doubt the purpose itself of allocating capital (Gründl and Schmeiser, 2007; Phillips et al., 1998; Venter, 2004).

This paper constructs a unifying framework designed to address specific decision criteria in a comprehensive yet still highly stylised setting. We consider capital allocation as the outcome of a particular optimisation problem, in which the weighted sum of measures for the deviations of the business unit's losses from their respective allocated capitals is minimised. The proposed approach is justified as follows:

Electronic copy available at: http://ssrn.com/abstract=1332264

- The idea of capital being 'close' to the risk it is being allocated to is intuitive because allocated capital should be a reflection of the associated risk. Moreover, such closeness models a notion of fairness within an organisation: risky portfolios are penalised, less risky ones are rewarded.
- The objective function of our optimisation approach relies crucially on the function used to weigh, that is, define the significance, of different scenarios (states of the world). Consequently, capital is allocated such that it matches closely the risk under particular scenarios considered adverse by management. These may be scenarios affecting the whole portfolio or isolated business units, extreme or less extreme, depending solely on the company's aggregate risk or on broader market conditions. Thus the proposed approach allows the flexibility of aligning capital allocation with management's different notions of risk tolerance.

It is then shown that many capital allocation approaches appearing in the literature can be seen as special cases of our more general framework. Thus different allocation approaches, studied through a common framework, are made comparable and are offered an alternative interpretation. A quadratic deviation criterion gives rise to allocations that generally have the form of expectations. Allocations based on well-known risk measures such as the Conditional Tail Expectation (CTE) (e.g. as in Overbeck (2000)) as well as allocations taking into account the insurer's default option (e.g. as in Sherris (2006)) are derived in this setting. An absolute deviation criterion gives rise to quantile-based allocations, whereby diversification is reflected by lowering the confidence level of VaR measures applied at sub-portfolio level. In both cases, the allocations may or may not reflect the dependence structure of the portfolio, via dependence between the individual and the aggregate risks.

The purpose of our approach is not to choose a "best possible" capital allocation method, but instead, by considering very different capital allocation formulas as part of the same framework in order to make these more comparable and mutually illuminating.

We note that mathematical results related to the material in this paper have been presented by Zaks et al. (2006). These authors work in a premium allocation context, restricting themselves to quadratic deviation criteria and do not use the device of weighting different scenarios; the range of allocation approaches derived in that paper is therefore narrower in scope. It is to be noted that the idea of deriving capital allocation via optimisation arguments was also discussed by Dhaene et al. (2003) and Laeven and Goovaerts (2004). However, Dhaene et al. (2003) is within the scope of our general framework while Laeven and Goovaerts (2004) generalizes Dhaene et al. (2003) but in a different direction than ours.

The structure of the rest of the paper is as follows. In Section 2, risk measures and the capital allocation problem are discussed and an overview of some popular allocation methods is given. In Section 3, which forms the main contribution of the paper, a unifying optimal capital allocation approach is presented. From this general approach, a multitude of special cases is derived. Finally, brief conclusions are given in Section 4.

2 Capital allocation

2.1 Risk Measures

A risk measure is a mapping ρ from a set Γ of real-valued random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to the real line $\mathbb{R}$:

$$\rho : \Gamma \to \mathbb{R} : X \in \Gamma \to \rho\left[X\right]. \tag{1}$$

The random variable X refers to the loss associated with conducting a business. In actuarial science, risk measures have traditionally been used for determining insurance premiums (Goovaerts et al. (1984)). More recently, however, they have been applied in a risk management context, with $\rho[X]$ representing the amount of capital to be set aside in order to make the loss X an acceptable risk; see Artzner et al. (1999).

Some well known properties that risk measures may or may not satisfy are law invariance, monotonicity, positive homogeneity, translation invariance (or equivariance) and subadditivity. They are formally defined as:

- *Law invariance*: For any $X_1, X_2 \in \Gamma$ with $\mathbb{P}[X_1 \leq x] = \mathbb{P}[X_2 \leq x]$ for all $x \in \mathbb{R}$, $\rho[X_1] = \rho[X_2]$.
- *Monotonicity*: For any $X_1, X_2 \in \Gamma$, $X_1 \leq X_2$ implies $\rho[X_1] \leq \rho[X_2]$.
- *Positive homogeneity*: For any $X \in \Gamma$ and $a > 0$, $\rho[aX] = a\rho[X]$.
- *Translation invariance*: For any $X \in \Gamma$ and $b \in \mathbb{R}$, $\rho[X + b] = \rho[X] + b$.
- *Subadditivity*: For any $X_1, X_2 \in \Gamma$, $\rho[X_1 + X_2] \leq \rho[X_1] + \rho[X_2]$.

Artzner et al. (1999) call any risk measure that satisfies the last four properties a *coherent* risk measure. Föllmer and Schied (2002) provide weaker sets of properties and discuss the desirability or otherwise of the properties of coherent risk measures.

2.2 The allocation problem

Consider a portfolio of n individual losses $X_1, X_2, ..., X_n$ materialising at a fixed future date T. Assume that $(X_1, X_2, ..., X_n)$ is a random vector on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Throughout the paper, we will always assume that any loss X_i has a finite mean. The distribution function $\mathbb{P}[X_i \leq x]$ of X_i will be denoted by $F_{X_i}(x)$.

The aggregate loss is defined by the sum

$$S = \sum_{i=1}^{n} X_i, \tag{2}$$

where this aggregate loss S can be interpreted as:

- the total loss of a corporate, e.g. an insurance company, with the individual losses corresponding to the losses of the respective business units;
- the loss from an insurance portfolio, with the individual losses being those arising from the different policies; or
- the loss suffered by a financial conglomerate, while the different individual losses correspond to the losses suffered by its subsidiaries.

It is the first of these interpretations we will use throughout this article. Hence S is the aggregate loss faced by an insurance company and X_i the loss of business unit i. We assume that the company has already determined the aggregate level of capital and denote this total risk capital by K. (This may or may not include technical provisions, but is not reflective of market premiums). The company now wishes to allocate this exogenously given total risk capital K

across its various business units, that is, to determine non-negative real numbers $K_1, \ldots, K_n$ satisfying the full allocation requirement:

$$\sum_{i=1}^{n} K_i = K. \tag{3}$$

This allocation is in some sense a notional exercise; it does not mean that capital is physically shifted across the various units, as the company's assets and liabilities continue to be pooled. The allocation exercise could be made in order to rank the business units according to levels of profitability. This task can be performed, for example, by determining the returns on the allocated capital for the respective business units.

Given that a capital allocation can be carried out in a countless number of ways, additional criteria must be set up in order to determine the most suitable. A reasonable start is to require the allocated capital amounts K_i to be 'close' to their corresponding losses X_i in some appropriately defined sense. This underlies the approach proposed in the present paper. Prior to introducing the idea of 'closeness' between individual loss and allocated capital, we revisit some well-known capital allocation methods.

2.3 Some known allocation formulas

For a given probability level $p \in (0,1)$, we denote the Value-at-Risk (VaR) or quantile of the loss random variable X by $F_X^{-1}(p)$. As usual, it is defined by

$$F_X^{-1}(p) = \inf\left\{x \in \mathbb{R} \mid F_X(x) \geq p\right\}, \qquad p \in [0,1]. \tag{4}$$

with $\inf\{\emptyset\} = +\infty$ by convention. Below we will also need so-called α–mixed inverse distribution functions; see Dhaene et al. (2002). Therefore, we first define the inverse distribution function $F_X^{-1+}(p)$ of the random variable X by

$$F_X^{-1+}(p) = \sup\left\{x \in \mathbb{R} \mid F_X(x) \leq p\right\}, \qquad p \in [0,1], \tag{5}$$

with $\sup\{\emptyset\} = -\infty$. The α–mixed inverse distribution function $F_X^{-1(\alpha)}$ of X is then defined as follows:

$$F_X^{-1(\alpha)}(p) = \alpha F_X^{-1}(p) + (1-\alpha) F_X^{-1+}(p), \qquad p \in (0,1)\ \alpha \in [0,1]. \tag{6}$$

From this definition, one immediately finds that for any random variable X and for all x with $0 < F_X(x) < 1$, there exists an $\alpha_x \in [0,1]$ such that$F_X^{-1(\alpha_x)}(F_X(x)) = x$.

2.3.1 The haircut allocation principle

It is a common industry practice, driven by banking and insurance regulations, to measure stand-alone losses by a VaR for a given probability level p. In line with such practice, a straightforward allocation method consists of allocating the capital $K_i = \gamma F_{X_i}^{-1}(p)$, $i = 1, \ldots, n$, to business unit i, where the factor γ is chosen such that the full allocation requirement (3) is satisfied. This gives rise to the *haircut allocation principle:*

$$K_i = \frac{K}{\sum_{j=1}^{n} F_{X_j}^{-1}(p)} F_{X_i}^{-1}(p), \qquad i = 1, \ldots, n. \tag{7}$$

For an exogenously given value of K, this principle leads to an allocation that is not influenced by the dependence structure between the losses X_i of the different business units. In this sense, one can say that the allocation method is independent of the portfolio context within which the individual losses X_i are embedded.

It is well-known that the quantile risk measure is not always subadditive. Consequently, using the p-quantile as stand-alone risk measure will not necessarily imply that the subportfolios will benefit from a pooling effect. This means that it may happen that the allocated capitals K_i exceed the respective stand-alone capitals $F_{X_i}^{-1}(p)$.

2.3.2 The quantile allocation principle

The haircut allocation rule (7) allocates to each business unit i a proportion γ of its p-quantile, with γ chosen such that the full allocation condition is fulfilled. This means that a constant proportional reduction (or increase) is applied on each of the quantiles $F_{X_i}^{-1}(p)$. Instead of applying a proportional cut on the monetary amounts $F_{X_i}^{-1}(p)$, one could adopt the probability level p equally among all business units and determine an α–mixed inverse with $\alpha \in [0,1]$, such that the full allocation requirement is again satisfied. This approach gives rise to the *quantile allocation principle* with allocated capital amounts K_i given by

$$K_i = F_{X_i}^{-1(\alpha)}(\beta p), \text{ with } \alpha \text{ and } \beta \text{ such that} \sum_{i=1}^{n} K_i = K. \tag{8}$$

Similar to the haircut allocation principle, for a given aggregate capital K, the allocated capitals K_i are not influenced by the dependence structure between the different losses X_i, $i = 1, \ldots, n$.

The quantile allocation rule is in compliance with the principle of using equal quantiles to measure the risk associated with the different business units. If it is considered 'consistent' to measure each stand-alone loss X_i by the corresponding quantile $F_{X_i}^{-1}(p)$, then it makes sense to measure each 'pooled' loss by $F_{X_i}^{-1(\alpha)}(\beta p)$ where α and β are chosen such that the full allocation requirement is satisfied. This means that all losses X_i continue to be evaluated at the same probability level '$\beta \times p$' and the benefits from pooling are in some sense 'subdivided neutrally' across the different business units.

The haircut allocation principle (7) will in general not lead to a quantile-based allocation with the same probability level for all business units. Companies and regulators when debating that all risks should be evaluated using the same p-quantile measure may prefer the quantile allocation principle rather than the haircut principle.

The appropriate levels of α and β are to be determined as the solutions to

$$K = \sum_{i=1}^{n} F_{X_i}^{-1(\alpha)}(\beta p), \tag{9}$$

In order to solve this problem, we need to introduce the concept of a *comonotonic sum* S^c defined by

$$S^c = \sum_{i=1}^{n} F_{X_i}^{-1}(U), \tag{10}$$

where U is a uniform random variable on $(0,1)$. It then holds that

$$K = F_{S^c}^{-1(\alpha)}(\beta p), \tag{11}$$

which leads us to

$$\beta p = F_{S^c}(K). \tag{12}$$

Furthermore, we have that

$$K = F_{S^c}^{-1(\alpha)}(F_{S^c}(K)). \tag{13}$$

Further additional details can be found in Dhaene et al. (2002). The quantile allocation rule in (8) can then be re-expressed as

$$K_i = F_{X_i}^{-1(\alpha)}\left(F_{S^c}(K)\right), \qquad i = 1, \ldots, n, \tag{14}$$

with α determined from (13).

In the special case that all distribution functions F_{X_i} are strictly increasing and continuous, this rule reduces to

$$K_i = F_{X_i}^{-1}\left(F_{S^c}(K)\right), \qquad i = 1, \ldots, n. \tag{15}$$

This allocation principle was proposed in Dhaene et al. (2003), where it was derived as the solution of an appropriate optimisation problem; See also Section 3.3. Notice that for strictly increasing and continuous distribution functions F_{X_i}, the quantile allocation principle can be considered as a special case of the haircut allocation principle (7) by choosing $p = F_{S^c}(K)$.

2.3.3 The covariance allocation principle

The *covariance allocation principle* proposed by e.g. Overbeck (2000) is given by

$$K_i = \frac{K}{\mathrm{Var}[S]}\mathrm{Cov}[X_i, S], \qquad i = 1, \ldots, n, \tag{16}$$

where Cov$[X_i, S]$ is the covariance between the individual loss X_i and the aggregate loss S and Var$[S]$ is the variance of the aggregate loss S. Because clearly the sum of these individual covariances is equal to the variance of the aggregate loss, the full allocation requirement is automatically satisfied in this case.

The covariance allocation rule, unlike the haircut and the quantile allocation principles, explicitly takes into account the dependence structure of the random losses $(X_1, X_2, ..., X_n)$. Business units with a loss that is more correlated with the aggregate portfolio loss S are penalised by requiring them to hold a larger amount of capital than those which are less correlated.

2.3.4 The CTE allocation principle

For a given probability level $p \in (0, 1)$, the Conditional Tail Expectation (CTE) of the aggregate loss S is defined as

$$\mathrm{CTE}_p[S] = \mathbb{E}\left[S \mid S > F_S^{-1}(p)\right]. \tag{17}$$

At a fixed level p, it gives the average of the top $(1-p)$% losses. In general, the CTE as a risk measure does not necessarily satisfy the subadditivity property. However, it is known to be a coherent risk measure in case we restrict to random variables with continuous distribution function. See e.g. Acerbi and Tasche (2002) and Remark 4.2.3. in Dhaene et al. (2006).

The *CTE allocation principle*, for some fixed probability level $p \in (0, 1)$, has the form

$$K_i = \frac{K}{\mathrm{CTE}_p[S]}\mathbb{E}\left[X_i \left| S > F_S^{-1}(p)\right.\right], \qquad i = 1, \ldots, n \tag{18}$$

In the particular case that $K = \mathrm{CTE}_p[S]$, formula (18) essentially reduces to the "contributions to expected shortfall" allocation suggested by Overbeck (2000) and, as a special case, by Denault (2001). In fact, the CTE allocation principle is a special case of marginal or Euler allocations discussed in detail by Tasche (2004).

The CTE allocation rule explicitly takes into account the dependence structure of the random losses $(X_1, X_2, ..., X_n)$. Imterpreting the event '$S > F_S^{-1}(p)$' as 'the aggregate portfolio loss S is large', we see from (18) that business units with larger conditional expected loss, given that the aggregate loss S is large, will be penalised with a larger amount of capital required than those with lesser conditional expected loss.

2.3.5 Proportional allocations

The capital allocation methods we have discussed so far can also be viewed as special cases of a more general class. Each member of this class is obtained by first choosing a risk measure ρ and then attributing the capital $K_i = \alpha\rho\,[X_i]$ to each business unit i, $i = 1, \ldots, n$. The factor α is chosen such that the full allocation requirement (3) is satisfied. This gives rise to the *proportional allocation principle*:

$$K_i = \frac{K}{\sum_{j=1}^{n} \rho[X_j]} \rho[X_i], \qquad i = 1, \ldots, n. \tag{19}$$

The allocation principles discussed in the previous subsections follow from (19) by choosing the appropriate risk measure ρ:

$$\begin{aligned}
\text{Haircut allocation:} \qquad & \rho[X_i] = F_{X_i}^{-1}(p), && (20)\\
\text{Quantile allocation:} \qquad & \rho[X_i] = F_{X_i}^{-1}\left(F_{S^c}\left(K\right)\right) \text{ with } \alpha \text{ from (13)}, && (21)\\
\text{Covariance allocation:} \qquad & \rho[X_i] = \text{Cov}\left[X_i, S\right], \text{ and} && (22)\\
\text{CTE allocation:} \qquad & \rho[X_i] = \mathbb{E}\left[X_i \,\middle|\, S > F_S^{-1}(p)\right]. && (23)
\end{aligned}$$

We note that in the last two allocations, the risk measure $\rho(X)$ does not depend only on the distribution of X, that is, ρ is not law invariant. If ρ is law invariant (first two allocations), the proportional allocation derived from ρ is not influenced by the dependence structure between the losses X_i of the different business units.

Let us assume that stand-alone losses are measured by a risk measure ρ. This means that $K = \rho[S]$ and also that the risk of business unit i, considered as a stand-alone unit, is measured by $\rho[X_i]$. From (19) one finds that in case of a proportional allocation, each business unit benefits from a pooling effect in the sense that $K_i \leq \rho[X_i]$ if and only if

$$K = \rho[S] \leq \sum_{j=1}^{n} \rho[X_j]. \tag{24}$$

This condition is fulfilled for subadditive risk measures ρ. As we have observed before, the haircut allocation method, which chooses a VaR as a stand-alone risk measure, may lead to a positive or a negative pooling effect. On the other hand, choosing Tail Value-at-Risk (TVaR) as stand-alone risk measure such as in the CTE allocation method, will lead to $K_i \leq \rho[X_i]$.

In Section 3, we develop a unifying optimal capital allocation approach and show that the quantile allocation (21), the covariance allocation (22) and the CTE allocation (23) fall as special cases of this approach. In contrast, the haircut allocation (20) does not seem to be reconcilable with our general framework. Note however that for strictly increasing and continuous distribution functions F_{X_i}, the quantile allocation principle can be considered as a special case of the haircut allocation principle by choosing $p = F_{S^c}\left(K\right)$.

2.3.6 Location-scale families of distributions

This section investigates the relationship that exists between the different allocation rules presented above, in case the losses X_i belong to the same location-scale family of distributions. Herewith, we assume that there exists a random variable Z with a zero mean and constants $a_i > 0$ and b_i such that

$$X_i \stackrel{d}{=} a_i Z + b_i, \qquad i = 1, \ldots, n, \tag{25}$$

where $\stackrel{d}{=}$ stands for 'equality in distribution'. For simplicity, we further assume that F_Z is strictly increasing and continuous on $\mathbb{R}$.

Let us first consider the general proportional allocation principle (19) where ρ is assumed to be law invariant, translation invariant and positive homogeneous. In this case, all $\rho[X_i]$ can be expressed as

$$\rho[X_i] = F_{X_i}^{-1}(p) = a_i + z_p b_i, \qquad i = 1, \ldots, n, \tag{26}$$

where z_p is the p-th quantile of Z for some fixed probability level $p \in (0,1)$; see e.g. Dhaene et al. (2009). This means that under the stated conditions, the proportional allocation principle (19) reduces to the haircut allocation principle (7).

Next we consider the CTE allocation principle (18) with $K = \text{CTE}_p[S]$ and where the vector of business losses $(X_1, X_2, ..., X_n)$ is multivariate elliptically distributed with $\mathbb{E}\left[X_i\right] = 0,\ i = 1, \ldots, n$. The assumption that all X_i have a zero mean may be relevant for practical situations where a provision equal to the expected value of the aggregate loss is set aside, and in addition capital is used as a buffer to protect against the 'uncertainty of the aggregate loss around its mean'. In this case, each X_i has to be interpreted as the loss of business unit i minus its expectation.

From Landsman and Valdez (2003), we find that in this case the allocated capitals K_i are given by

$$K_i = \mathbb{E}\left[X_i \,\middle|\, S > F_S^{-1}(p)\right] = \frac{\text{CTE}_p[S]}{\text{Var}[S]} \text{Cov}\left[X_i, S\right], \qquad i = 1, \ldots, n. \tag{27}$$

Hence, we can conclude that when $(X_1, X_2, ..., X_n)$ is multivariate elliptically distributed with zero means and in addition $K = \text{CTE}_p[S]$, the CTE allocation principle (18) coincides with the covariance allocation principle (16).

2.3.7 Allocation and the default option

A somewhat different class of approaches, based on the arguments of Myers and Read Jr. (2001), produces a capital allocation procedure by considering the value of the insurer's default option. Since the shareholders of the company have limited liability, in the event of default, i.e. when $S > K$, they are not, in principle, obligated to pay the excess loss $S - K$. Therefore, the protection that the collective of policyholders purchases is $\min(S, K)$, which can be written as

$$S - (S - K)_+. \tag{28}$$

The quantity $(S-K)_+$ is called the *policyholder deficit* or alternatively the *insurer's default option*.

Myers and Read Jr. (2001) assume that markets are complete and that a proportional increase in the exposure to a particular line of business produces a proportional increase in its allocated capital. They subsequently allocate the value of the default option via the marginal contributions of each line of business to that value.

It can be shown that the Myers-Read allocation is given by the general formula (which does not appear in their paper in this form),

$$\mathbb{E}[(S - K)_+] = \sum_{j=1}^{n} \mathbb{E}[(X_j - K_j)\ \mathbb{I}(S > K)], \tag{29}$$

where $\mathbb{I}(A)$ is the indicator function of event A and expectations may be taken under a risk neutral measure.

It is worth noting that (29) is not a capital allocation formula in itself, as the $K_1 \ldots, K_n$ are considered as given and only the default option value is allocated. Capital allocation methods that consider the value of the default options have been derived by e.g. Sherris (2006).

3 Optimal capital allocations

3.1 General setting

The allocation of the exogenously given aggregate capital K to n parts $K_1, \ldots, K_n$, corresponding to the different subportfolios or business units, can be carried out in an infinite number of ways, some of which were illustrated in the previous section. At first glance, there seems to be a lack of a clear motivation for preferring to choose one method over another, although it appears obvious that different capital allocations must in some sense correspond to different questions that can be asked within the context of risk management. Hereafter we systematise capital allocation methods by viewing them as solutions to a particular decision problem. For that we need to formulate a decision criterion, such as:

Capital should be allocated such that for each business unit the allocated capital and the loss are sufficiently close to each other.

In order to cast this statement in a more formal setting, consider the aggregate portfolio loss $S = X_1 + \cdots + X_n$ with aggregate capital K. Once the aggregate capital is allocated, the difference between aggregate loss and aggregate capital can be expressed as

$$S - K = \sum_{j=1}^{n} (X_j - K_j), \tag{30}$$

where the quantity $(X_j - K_j)$ expresses the loss minus the allocated capital for subportfolio j. Important to notice is that in this setting, the subportfolios are cross-subsidising each other, in the sense that the occurence of the event '$X_k > K_k$' does not necessarily lead to 'ruin'; such unfavorable performance of subportfolio k may be compensated by a favorable outcome for one or more values $(X_l - K_l)$ of the other subportfolios.

We propose to determine the appropriate allocation by the following optimisation problem:

Optimal capital allocation problem: *Given the aggregate capital $K > 0$, determine the allocated capitals $K_i,\ i = 1, \ldots, n$, from the following optimisation problem:*

$$\min_{K_1,\ldots,K_n} \sum_{j=1}^{n} v_j \mathbb{E}\left[\zeta_j\ D\left(\frac{X_j - K_j}{v_j}\right)\right],\ \textit{such that} \sum_{j=1}^{n} K_j = K, \tag{31}$$

where the v_j are non-negative real numbers such that $\sum_{j=1}^{n} v_j = 1$, the ζ_j are non-negative random variables such that $\mathbb{E}[\zeta_j] = 1$ and D is a non-negative function.

Before solving the general optimal capital allocation problem (31), we first elaborate on its various elements.

- v_j : The non-negative real number v_j is a measure of exposure or business volume of the j-th unit, such as revenue, insurance premium, etc. These scalar quantities are chosen such that they sum to 1. Their inclusion in the expression $D\left(\frac{X_j - K_j}{v_j}\right)$ normalises the deviations of loss from allocated capital across business units to make them relatively more comparable. At the same time, the v_j's are used as weights attached to the different values of $\mathbb{E}\left[\zeta_j\ D\left(\frac{X_j-K_j}{v_j}\right)\right]$ in the minimisation problem (31), in order to reflect the relative importance of the different business units.

$D\left(\frac{X_j-K_j}{v_j}\right)$: For simplicity, we first assume that $v_j = 1$ and also that $\zeta_j \equiv 1$. The terms $D(X_j - K_j)$ quantify the deviations of the outcomes of the losses X_j from their allocated capital K_j. Minimising the sum of the expectations of these quantities essentially reflects the requirement that the allocated capitals should be 'as close as possible' to the losses they are allocated to. Examples of distance measures are "squared or quadratic deviations" and "absolute deviations".

ζ_j : The deviations of the losses X_j from their respective allocated capital levels K_j are measured by the terms $\mathbb{E}\left[\zeta_j\, D(X_j - K_j)\right]$. These expectations involve non-negative random variables ζ_j with $\mathbb{E}[\zeta_j] = 1$ that are used as weight factors to the different possible outcomes of $D(X_j - K_j)$.
One possible choice for the ζ_j could be $\zeta_j = h(X_j)$ for some non-negative and non-decreasing function h. In this case, the heaviest weights are attached to deviations that correspond to states-of-the-world leading to the largest outcomes of X_j. We will call allocations based on such a choice for the ζ_j *business unit driven allocations*.
Another choice is to let $\zeta_j = h(S)$ for some non-negative and non-decreasing function h, such that the outcomes of the deviations are weighted with respect to the aggregate portfolio performance. In this case, heavier weights are attached to deviations that correspond to states-of-the-world leading to larger outcomes of S. Allocations based on such a choice for the random variables ζ_j will be called *aggregate portfolio driven allocations*.
A yet different approach is to let $\zeta_j = \zeta_M$ for all j, where ζ_M can be interpreted as the loss on a reference (or market) portfolio. In this case, the weighting is market driven and the corresponding allocation is said to be a *market driven allocation*.

In summary, we propose in (31) to fully allocate the aggregate capital K to the different business units such that the exposure-weighted sum of the expectations of the weighted and normalised deviations of the losses X_j from their respective allocated capitals K_j, is minimised.

3.2 The quadratic optimisation criterion

3.2.1 General solution of the quadratic allocation problem

In this subsection we discuss optimal allocation under a quadratic criterion, that is, by letting

$$D(x) = x^2. \tag{32}$$

In this case, the optimal allocation problem (31) reduces to

$$\min_{K_1,\ldots,K_n} \sum_{j=1}^{n} \mathbb{E}\left[\zeta_j \frac{(X_j - K_j)^2}{v_j}\right], \text{ such that } \sum_{j=1}^{n} K_j = K. \tag{33}$$

The solution to this minimisation problem is given in the following theorem.

Theorem 1 *The optimal allocation problem (33) has the following unique solution:*

$$K_i = \mathbb{E}[\zeta_i X_i] + v_i \left(K - \sum_{j=1}^{n} \mathbb{E}[\zeta_j X_j]\right), \qquad i = 1, \ldots, n. \tag{34}$$

Proof. Problem (33) can be solved via a Lagrange optimisation. However, we will give a simple geometric proof, based on an argument given in Zaks et al. (2006) who considered the minimisation problem (33) in the special case that all $\zeta_j \equiv 1$.

Taking into account the relations

$$\mathbb{E}\left[\zeta_j\left(X_j-K_j\right)^2\right]=\left(\mathbb{E}\left[\zeta_j X_j\right]-K_j\right)^2+\mathbb{E}\left[\zeta_j X_j^2\right]-\left(\mathbb{E}\left[\zeta_j X_j\right]\right)^2, \qquad j=1,\ldots,n,$$

we have that the solution of the minimisation problem (33) is identical to the solution of the following minimisation problem:

$$\min_{K_1,\ldots,K_n}\sum_{j=1}^{n}\frac{\left(\mathbb{E}\left[\zeta_j X_j\right]-K_j\right)^2}{v_j}, \text{ such that } \sum_{j=1}^{n}K_j=K. \tag{35}$$

Clearly, eliminating the term $\sum_{j=1}^{n}\left(\mathbb{E}\left[\zeta_j X_j^2\right]-\left(\mathbb{E}\left[\zeta_j X_j\right]\right)^2\right)$ does not change the optimal allocation.

By introducing the notation

$$x_j=\frac{K_j-\mathbb{E}\left[\zeta_j X_j\right]}{\sqrt{v_j}} \qquad j=1,\ldots,n, \tag{36}$$

we can transform (35) into

$$\min_{x_1,\ldots,x_n}\sum_{j=1}^{n}x_j^2, \text{ such that } \sum_{j=1}^{n}\sqrt{v_j}x_j=\left(K-\sum_{j=1}^{n}\mathbb{E}[\zeta_j X_j]\right). \tag{37}$$

Let us now interpret the set

$$\left\{(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n \;\middle|\; \sum_{j=1}^{n}\sqrt{v_j}x_j=\left(K-\sum_{j=1}^{n}\mathbb{E}[\zeta_j X_j]\right)\right\}$$

as a hyperplane in $\mathbb{R}^n$. The solution of (37) can then be interpreted as the point $(x_1,x_2,\ldots,x_n)$ on the hyperplane $\sum_{j=1}^{n}\sqrt{v_j}x_j=\left(K-\sum_{j=1}^{n}\mathbb{E}[\zeta_j X_j]\right)$ that is closest to the origin $(0,0,\ldots,0)$. Hence,

$$x_i=\sqrt{v_i}\left(K-\sum_{j=1}^{n}\mathbb{E}[\zeta_j X_j]\right), \qquad i=1,\ldots,n.$$

Translating this result in terms of the K_i via (36) immediately leads to (34). ■

The capital K_i given by (34) equals the weighted expected loss of X_i, in addition to a term proportional to the volume of the unit. This second term is a redistribution of the difference between the amount of aggregate capital K held and $\sum_{j=1}^{n}\mathbb{E}[\zeta_j X]$. This redistribution is assigned using weights v_i based on the 'volume' or on some other measure of the 'riskiness' of the corresponding business units.

In the particular case that the volume weights are given by

$$v_i=\frac{\mathbb{E}[\zeta_i X_i]}{\sum_{j=1}^{n}\mathbb{E}[\zeta_j X_j]}, \tag{38}$$

it immediately follows that (34) reduces to

$$K_i=\frac{K}{\sum_{j=1}^{n}\mathbb{E}[\zeta_j X_j]}\mathbb{E}[\zeta_i X_i], \qquad i=1,\ldots,n. \tag{39}$$

This allocation rule can be seen as a special case of the proportional allocation rule (19), by choosing

$$\rho[X_i] = \mathbb{E}[\zeta_i X_i], \qquad i = 1, \ldots, n. \tag{40}$$

Notice that (39) can be rearranged as

$$\frac{K_i - \mathbb{E}[\zeta_i X_i]}{\mathbb{E}[\zeta_i X_i]} = \frac{K - \sum_{j=1}^{n} \mathbb{E}[\zeta_j X_j]}{\sum_{j=1}^{n} \mathbb{E}[\zeta_j X_j]}, \qquad i = 1, \ldots, n. \tag{41}$$

In the special case that the aggregate capital K is given by $K = \sum_{j=1}^{n} \mathbb{E}[\zeta_j X_j]$, the allocation rule (34) reduces to

$$K_i = \mathbb{E}[\zeta_i X_i], \qquad i = 1, \ldots, n. \tag{42}$$

This class of allocations is investigated in Furman and Zitikis (2008b) who call the members of this class *weighted risk capital allocations*.

3.2.2 Business unit driven allocations

In this subsection we consider the case where the weighting random variables ζ_i in the quadratic allocation problem (33) are given by

$$\zeta_i = h_i(X_i), \tag{43}$$

with h_i being a non-negative and non-decreasing function such that $\mathbb{E}[h_i(X_i)] = 1$, for $i = 1, \ldots, n$. Hence, for each business unit i, the states-of-the-world to which we want to assign the heaviest weights are those under which the business unit performs the worst. As earlier pointed out, we call allocations based on (43) *business unit driven allocations*. In this case, the allocation rule (34) can be rewritten as

$$K_i = \mathbb{E}[X_i h_i(X_i)] + v_i \left(K - \sum_{j=1}^{n} \mathbb{E}[X_j h_i(X_j)] \right), \qquad i = 1, \ldots, n. \tag{44}$$

For an exogeneously given value of K, the allocations K_i are not influenced by the mutual dependence structure between the losses X_i of the different business units. In this sense, one can say that the allocation principle (44) is independent of the portfolio context within which the X_i's are embedded, and hence, is indeed business unit driven. Such allocations might be a useful instrument for determining the performance bonuses of the business unit managers, in case one assumes that each manager should be rewarded for the performance of his own business unit, but not extra rewarded (or penalised) for the interrelationship that exists between the performance of his business unit and that of the other units of the company. One should however note that disregarding in this way diversification between business units, the allocation may give incentives to managers that are at odds with overall portfolio optimization criteria.

The law invariant risk measure $\mathrm{E}[X_i h_i(X_i)]$ assigns to any loss X_i the expected value of the weighted outcomes of this loss, where higher weights correspond to larger outcomes of the loss, that is, to more adverse scenarios. Risk measures and premium principles of this general type have been proposed and investigated in Heilmann (1989), Tsanakas (2007) and Furman and Zitikis (2008a).

A particular choice of the random variables $h_i(X_i)$ considered in (44) is given by

$$h_i(X_i) = \frac{\mathbb{I}\left(X_i > F_{X_i}^{-1}(p)\right)}{1 - F_{X_i}\left(F_{X_i}^{-1}(p)\right)}, \qquad i = 1, \ldots, n, \tag{45}$$

for some $p \in (0,1)$. In this case, we find that $\mathbb{E}[X_i h_i(X_i)]$ transforms into

$$\mathbb{E}[X_i h_i(X_i)] = \text{CTE}_p\left[X_i\right], \qquad i = 1, \ldots, n. \tag{46}$$

More generally, consider the random variables $h_i(X_i)$ defined by

$$h_i(X_i) = g'\left(\overline{F}_{X_i}(X_i)\right), \qquad i = 1, \ldots, n, \tag{47}$$

with $g : [0,1] \to [0,1]$ an increasing and concave function with derivative g' if it exists, and $\overline{F}_X$ the decumulative function of X. We then find that

$$\mathbb{E}[X_i h_i(X_i)] = \mathbb{E}[X_i\, g'\left(\overline{F}_{X_i}(X_i)\right)], \qquad i = 1, \ldots, n, \tag{48}$$

and $\mathbb{E}[X_i h_i(X_i)]$ is a concave distortion risk measure, also called spectral risk measure. See Wang (1996), Acerbi (2002), or Dhaene et al. (2006).

Other examples of risk measures of the form $\mathbb{E}[X_i h_i(X_i)]$ are the standard deviation principle, the Esscher principle and the exponential principle. These are summarised in Table 1.

Defining the volumes v_i by

$$v_i = \frac{\mathbb{E}[X_i h_i(X_i)]}{\sum_{j=1}^{n} \mathbb{E}[X_j h_j(X_j)]}, \qquad i = 1, \ldots, n, \tag{49}$$

we find that the allocation principle (44) reduces to

$$K_i = \frac{K}{\sum_{j=1}^{n} \mathbb{E}[X_j h_j(X_j)]} \mathbb{E}[X_i h_i(X_i)], \qquad i = 1, \ldots, n, \tag{50}$$

which is a special case of the proportional allocation principles discussed in Section 2.3.5.

Table 1: *Business unit driven capital allocation*

Reference	$h_i(X_i)$	$\mathbb{E}[X_i h_i(X_i)]$
Standard deviation principle Bühlmann (1970)	$1 + a\dfrac{X_i - \mathbb{E}[X_i]}{\sigma_{X_i}},\ a \geq 0$	$\mathbb{E}[X_i] + a\sigma_{X_i}$
Conditional tail expectation Overbeck (2000)	$\dfrac{1}{1-p}\mathbb{I}\left(X_i > F_{X_i}^{-1}(p)\right),\ p \in (0,1)$	$\text{CTE}_p\left[X_i\right]$
Distortion risk measure Wang (1996), Acerbi (2002)	$g'\left(\overline{F}_{X_i}(X_i)\right), g : [0,1] \mapsto [0,1],$ $g' > 0, g'' < 0$	$\mathbb{E}\left[X_i g'\left(\overline{F}_{X_i}(X_i)\right)\right]$
Exponential principle Gerber (1974)	$\int_0^1 \dfrac{e^{\gamma a X_i}}{\mathbb{E}[e^{\gamma a X_i}]} d\gamma,\ a > 0$	$\dfrac{1}{a}\ln \mathbb{E}\left[e^{aX_i}\right]$
Esscher principle Gerber (1981)	$\dfrac{e^{aX_i}}{\mathbb{E}[e^{aX_i}]},\ a > 0$	$\dfrac{\mathbb{E}[X_i e^{aX_i}]}{\mathbb{E}[e^{aX_i}]}$

It is to be noted that, with the exception of the CTE allocation, none of the allocation rules discussed in Section 2 can be readily seen as examples of *business unit driven allocation rules* that we discussed here.

3.2.3 Aggregate portfolio driven allocations

Let us now consider the case where

$$\zeta_i = h(S), \qquad i = 1, \dots, n, \tag{51}$$

with h being a non-negative and non-decreasing function such that $\mathbb{E}[h(S)] = 1$. In this case, the states-of-the-world to which we assign the heaviest weights are those under which the aggregate portfolio performs worst. Therefore, we call such allocations *aggregate portfolio driven allocations*. The allocation rule (34) can now be rewritten as

$$K_i = \mathbb{E}[X_i h(S)] + v_i \left(K - \mathbb{E}[S h(S)]\right), \qquad i = 1, \dots, n. \tag{52}$$

Hence, the capital K_i allocated to unit i is determined using a weighted expectation of the loss X_i, with higher weights attached to states-of-the-world that involve a large aggregate loss S. Notice that the allocation principle (52) can be reformulated as

$$K_i = \mathbb{E}[X_i] + \text{Cov}[X_i, h(S)] + v_i \left(K - \mathbb{E}[S h(S)]\right), \qquad i = 1, \dots, n. \tag{53}$$

This means that the capital allocated to the i-th business unit is given by the sum of the expected loss $\mathbb{E}[X_i]$, a loading which depends on the covariance between the individual and aggregate losses X_i and $h(S)$, plus a term proportional to the volume of the business unit. A strong positive correlation between X_i and $h(S)$, which reflects that X_i could be a substantial driver of the aggregate loss S, produces a higher allocated capital K_i. Allocation principles of the form (52) are closely related to the 'Euler' allocations proposed in Tasche (2004).

Using aggregate portfolio driven allocations might be appropriate when one wants to investigate each individual portfolio's contribution to the aggregate loss of the entire company. In other words, the company wishes to evaluate the subportfolio performances, e.g. the returns on the allocated capitals, in the presence of the other subportfolios. This can provide relevant information to the company within which it can further be used to evaluate either business expansions or reductions.

A particular choice of the random variable $h(S)$ considered in (52) is given by

$$h(S) = \frac{\mathbb{I}\left(S > F_S^{-1}(p)\right)}{1 - F_S\left(F_S^{-1}(p)\right)}, \qquad i = 1, \dots, n, \tag{54}$$

for some $p \in (0, 1)$. In this case, we find that $\mathbb{E}[X_i h(S)]$ and $\mathbb{E}[S h(S)]$ transform into

$$\mathbb{E}[X_i h(S)] = \mathbb{E}\left[X_i \,\middle|\, S > F_S^{-1}(p)\right], \qquad i = 1, \dots, n \tag{55}$$

and

$$\mathbb{E}[S h(S)] = \text{CTE}_p\left[S\right], \tag{56}$$

respectively. Furthermore, by taking

$$h(S) = S - \mathbb{E}[S], \tag{57}$$

we find

$$\mathbb{E}[X_i h(S)] = \text{Cov}\left[X_i, S\right], \qquad i = 1, \dots, n \tag{58}$$

and

$$\mathbb{E}[S h(S)] = \text{Var}[S]. \tag{59}$$

Other choices for the random variable $h(S)$ and the related expressions for $\text{E}[X_i h(S)]$ can be found in Table 2. They correspond to capital allocation principles that have been considered in Overbeck (2000), Tsanakas (2004), and Tsanakas (2008).

Defining the exposures v_i by

$$v_i = \frac{\mathbb{E}[X_i h(S)]}{\mathbb{E}[Sh(S)]}, \qquad i = 1, \ldots, n, \tag{60}$$

we find that the allocation principle (52) reduces to the proportional allocation rule

$$K_i = \frac{K}{\mathbb{E}[Sh(S)]} \mathbb{E}[X_i h(S)], \qquad i = 1, \ldots, n. \tag{61}$$

Now, the CTE allocation principle (23) that we discussed in Section 2.3.4 follows as a special case of the allocation principle (61) by choosing $h(S)$ as in (54). Furthermore, taking $h(S)$ as in (57) means that the allocation principle (61) effectively reduces to the covariance allocation principle (22) discussed in Section 2.3.3.

Table 2: *Aggregate portfolio driven allocations*

Reference	$h(S)$	$\mathbb{E}[X_i h(S)]$
Overbeck (2000)	$1 + a\dfrac{S - \mathbb{E}[S]}{\sigma_S}, \ a \geq 0$	$\mathbb{E}[X_i] + a\dfrac{\text{Cov}[X_i, S]}{\sigma_S}$
Overbeck (2000)	$\dfrac{1}{1-p}\mathbb{I}\left(S > F_S^{-1}(p)\right), \ p \in (0,1)$	$\mathbb{E}[X_i \vert S > F_S^{-1}(p)]$
Tsanakas (2004)	$g'(\overline{F}_S(S)), g : [0,1] \mapsto [0,1], g' > 0,$ $g'' < 0$	$\mathbb{E}\left[X_i g'(\overline{F}_S(S))\right]$
Tsanakas (2008)	$\displaystyle\int_0^1 \frac{e^{\gamma a S}}{\mathbb{E}[e^{\gamma a S}]} d\gamma$, $a > 0$	$\mathbb{E}\left[X_i \displaystyle\int_0^1 \frac{e^{\gamma a S}}{E[e^{\gamma a S}]} d\gamma\right]$
Wang (2007)	$\dfrac{e^{aS}}{\mathbb{E}[e^{aS}]}, \ a > 0$	$\dfrac{\mathbb{E}[X_i e^{aS}]}{\mathbb{E}[e^{aS}]}$

3.2.4 Market driven allocations

Let ζ_M be a random variable such that market-consistent values of the aggregate portfolio loss S and the business unit losses X_i are given by

$$\pi[S] = \mathbb{E}[\zeta_M S] \tag{62}$$

and

$$\pi[X_i] = \mathbb{E}[\zeta_M X_i], \qquad i = 1, \ldots, n, \tag{63}$$

respectively. Further suppose that at the aggregate portfolio level, a provision $\pi[S]$ is set aside to cover future liabilities S. Apart from the aggregate provision $\pi[S]$, the aggregate portfolio has an available solvency capital equal to $(K - \pi[S])$. The *solvency ratio* of the aggregate portolio is then given by

$$\frac{K - \pi[S]}{\pi[S]}. \tag{64}$$

In order to determine an optimal capital allocation over the different business units, we let in (31) $\zeta_i = \zeta_M, \ i = 1, \ldots, n$, thus allowing the market to determine which states-of-the-world are to be regarded adverse. This yields:

$$K_i = \pi[X_i] + v_i \left(K - \pi[S]\right). \tag{65}$$

If we now use the market-consistent prices as volume measures, after substituting

$$v_i = \frac{\pi[X_i]}{\pi[S]}, \qquad i = 1, \ldots, n, \tag{66}$$

in (65), we find

$$K_i = \frac{K}{\pi[S]}\pi[X_i], \qquad i = 1, \ldots, n. \tag{67}$$

Rearranging these expressions leads to

$$\frac{K_i - \pi[X_i]}{\pi[X_i]} = \frac{K - \pi[S]}{\pi[S]}, \qquad i = 1, \ldots, n. \tag{68}$$

The quantities $\pi[X_i]$ and $(K_i - \pi[X_i])$ can be interpreted as the market-consistent provision and the solvency capital attached to business unit i, while $\frac{K_i - \pi[X_i]}{\pi[X_i]}$ is its corresponding solvency ratio. From (68), we can conclude that the optimisation criterion (33) with $\zeta_i = \zeta_M,\ i = 1, \ldots, n$, and volume measures given by (66), leads to a capital allocation whereby the solvency ratio for each business unit is the same and equal to that of the aggregate portfolio. A similar allocation principle has been proposed by Sherris (2006) within the context of allocating the company's total equity to the different business units "to determine an expected return on equity by line of business".

3.2.5 Allocation with respect to the default option

An alternative choice for the weighting random variable ζ_i is given by

$$\zeta_i = h(S) = \frac{\mathbb{I}(S > K)}{\mathbb{P}[S > K]}, \qquad i = 1, \ldots, n, \tag{69}$$

such that only those states-of-the-world that correspond to insolvency are considered when determining the expectations $\mathbb{E}[X_i h(S)]$. The allocation rule (34) then becomes

$$K_i = \mathbb{E}\left[X_i \mid S > K\right] + v_i \left(K - \mathbb{E}\left[S \mid S > K\right]\right). \tag{70}$$

Notice the similarity between (69) and the choice of $h(S)$ made in (54) which led to the CTE allocation rule.

Expression (70) can be rearranged as follows:

$$\mathbb{E}[(X_i - K_i)\, \mathbb{I}(S > K)] = v_i \mathbb{E}[(S - K)_+], \qquad i = 1, \ldots, n. \tag{71}$$

Summing the left and right hand sides of this expression over $i = 1, \ldots, n$, leads to expression (29). The quantity $\mathbb{E}[(S - K)_+]$ represents the *expected policyholder deficit* or alternatively, the expected value of the *default option* that shareholders of an insurance company hold, given their limited liability. The allocation principle in (71) is such that the marginal contribution of each business unit to the expected value of the policyholder deficit is the same per unit of business volume, and hence is consistent with the arguments of Myers and Read Jr. (2001).

3.3 The absolute deviation optimisation criterion

In this section we discuss the optimal allocation problem under an absolute deviation criterion, that is, by letting

$$D(x) = |x|. \tag{72}$$

In this case the optimisation problem (31) reduces to

$$\min_{K_1,\ldots,K_n} \sum_{j=1}^{n} \mathbb{E}\left[\zeta_j \left|X_j - K_j\right|\right], \text{such that} \sum_{j=1}^{n} K_j = K. \tag{73}$$

From the relation

$$|x| = 2\,(x)_+ - x, \tag{74}$$

we immediately find that the optimal solution of (73) is identical to the solution of the following problem:

$$\min_{K_1,\ldots,K_n} \sum_{j=1}^{n} \mathbb{E}\left[\zeta_j \left(X_j - K_j\right)_+\right], \text{ such that} \sum_{j=1}^{n} K_j = K. \tag{75}$$

This means that the absolute deviation optimisation problem (73) only takes into account the outcomes of the business unit losses X_j that lead to technical insolvency $X_j > K_j$ in that unit.

In order to solve the optimisation problem (75), we first consider the special case where all ζ_j's are identical to 1.

Theorem 2 *Assuming that $F_{S^c}^{-1+}(0) < K < F_{S^c}^{-1}(1)$, the optimal allocation problem*

$$\min_{K_1,\ldots,K_n} \sum_{j=1}^{n} \mathbb{E}\left[\left(X_j - K_j\right)_+\right] \text{ such that} \sum_{j=1}^{n} K_j = K \tag{76}$$

has the following solution:

$$K_i = F_{X_i}^{-1(\alpha)}(F_{S^c}(K)), \qquad i = 1, \ldots, n, \tag{77}$$

where S^c is defined in (10) and $\alpha \in [0,1]$ follows from

$$F_{S^c}^{-1(\alpha)}(F_{S^c}(K)) = K. \tag{78}$$

Proof. Let α be determined from (78). Then we immediately find from Dhaene et al. (2002) that

$$\begin{aligned} \mathbb{E}\left[\left(S^c - K\right)_+\right] &= \sum_{j=1}^{n} \mathbb{E}\left[\left(X_j - F_{X_i}^{-1(\alpha)}(F_{S^c}(K))\right)_+\right] \\ &\leq \sum_{j=1}^{n} \mathbb{E}\left[\left(X_j - K_j\right)_+\right], \qquad K \in \left(F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1)\right), \end{aligned}$$

holds for all $(K_1, K_2, ..., K_n)$ such that $\sum_{j=1}^{n} K_j = K$. This proves the stated result. ■

We can conclude that the quantile allocation principle (14) considered in Section 2.3.2 is a solution of the minimisation problem (76). This optimisation problem and its solution were previously considered in Dhaene et al. (2003) in the particular case that the F_{X_i}'s are all strictly increasing. A proof of Theorem 2 using Lagrange techniques can be found in Laeven and Goovaerts (2004).

The solution to the general optimisation problem (73) will be expressed in terms of functions $F_{X_i}^{(\zeta_i)}$ defined as follows:

$$F_{X_i}^{(\zeta_i)}(x) = \mathbb{E}[\zeta_i \, \mathbb{I}(X_i \leq x)] = \mathbb{E}[\zeta_i \mid X_i \leq x] \; F_{X_i}(x), \qquad i = 1, \ldots, n. \tag{79}$$

One can prove that each function $F_{X_i}^{(\zeta_i)}$ defines a proper distribution function, which we will call the ζ_i–weighted distribution of X_i; see Rao (1997), Furman and Zitikis (2008a) and the references therein. The decumulative distribution function $\overline{F}_{X_i}^{(\zeta_i)}(x) = 1 - F_{X_i}^{(\zeta_i)}(x)$ is given by

$$\overline{F}_{X_i}^{(\zeta_i)}(x) = \mathbb{E}\left[\zeta_i\, \mathbb{I}(X_i > x)\right] = \mathbb{E}[\zeta_i \mid X_i > x]\, \overline{F}_{X_i}(x), \qquad i = 1, \dots, n. \tag{80}$$

A sufficient condition for $F_{X_i}^{(\zeta_i)}$ to be continuous is that F_{X_i} be continuous. A sufficient condition for $F_{X_i}^{(\zeta_i)}$ to be strictly increasing is that F_{X_i} be strictly increasing and that $\mathbb{P}[\zeta_i > 0] = 1$. For any $p \in (0,1)$ and any $\alpha \in [0,1]$, we denote the α–mixed inverse of $F_{X_i}^{(\zeta_i)}$ at level p by $\left(F_{X_i}^{(\zeta_i)}\right)^{-1(\alpha)}(p)$.

In the following lemma, we prove that the deviation measure $\mathbb{E}\left[\zeta_i \left(X_i - K_i\right)_+\right]$ can be transformed to a stop-loss premium of X_i with retention K_i, where the expectation is taken with respect to the ζ_i–weighted distribution of X_i.

Lemma 3 *Let U be a uniform random variable on the unit interval $(0,1)$. Then it holds that*

$$\mathbb{E}\left[\zeta_i \left(X_i - K_i\right)_+\right] = \mathbb{E}\left[\left(\left(F_{X_i}^{(\zeta_i)}\right)^{-1}(U) - K_i\right)_+\right], \qquad i = 1, \dots, n. \tag{81}$$

Proof. From the tower property of the expectation operator, we find

$$\mathbb{E}\left[\zeta_i \left(X_i - K_i\right)_+\right] = \mathbb{E}\left[\zeta_i \mathbb{E}\left[\left(X_i - K_i\right)_+ \mid \zeta_i\right]\right].$$

Substituting $\mathbb{E}\left[\left(X_i - K_i\right)_+ \mid \zeta_i\right]$ by $\int_{K_i}^{\infty} \mathbb{P}[X_i > x \mid \zeta_i]\, dx$ and changing the order of the integrations, we find

$$\begin{aligned}
\mathbb{E}\left[\zeta_i \left(X_i - K_i\right)_+\right] &= \int_{K_i}^{\infty} \mathbb{E}\left[\zeta_i\, \mathbb{P}[X_i > x \mid \zeta_i]\right]\, dx \\
&= \int_{K_i}^{\infty} \mathbb{E}\left[\zeta_i\, \mathbb{E}[\mathbb{I}(X_i > x) \mid \zeta_i]\right]\, dx.
\end{aligned}$$

Taking into account the tower property once more leads to

$$\mathbb{E}\left[\zeta_i \left(X_i - K_i\right)_+\right] = \int_{K_i}^{\infty} \mathbb{E}\left[\zeta_i\, \mathbb{I}(X_i > x)\right]\, dx = \int_{K_i}^{\infty} \overline{F}_{X_i}^{(\zeta_i)}(x)\, dx.$$

The stated result follows then from observing that the distribution function of $\left(F_{X_i}^{(\zeta_i)}\right)^{-1}(U)$ is given by $F_{X_i}^{(\zeta_i)}$ so that $\int_{K_i}^{\infty} \overline{F}_{X_i}^{(\zeta_i)}(x)\, dx$ is an expression for the stop-loss premium of $\left(F_{X_i}^{(\zeta_i)}\right)^{-1}(U)$ with retention K_i. ■

Now we are able to prove our main result concerning the absolute deviation optimisation problem.

Theorem 4 *Let $\overline{S}^c$ be the comonotonic sum defined by*

$$\overline{S}^c = \sum_{i=1}^{n} \left(F_{X_i}^{(\zeta_i)}\right)^{-1}(U), \tag{82}$$

where the random variable U is uniformly distributed on the unit interval $(0,1)$. In case $F_{\overline{S}^c}^{-1+}(0) < K < F_{\overline{S}^c}^{-1}(1)$, the optimal allocation problem (73) has the following solution:

$$K_i = \left(F_{X_i}^{(\zeta_i)}\right)^{-1(\alpha)}(F_{\overline{S}^c}K), \qquad i = 1, \ldots, n, \tag{83}$$

where $\alpha \in [0,1]$ follows from

$$F_{\overline{S}^c}^{-1(\alpha)}(F_{\overline{S}^c}(K)) = K. \tag{84}$$

Proof. From Lemma 3, we find that the optimisation problem (73) can be rewritten as

$$\min_{K_1,\ldots,K_n} \sum_{j=1}^{n} \mathbb{E}\left[\left(\left(F_{X_j}^{(\zeta_j)}\right)^{-1}(U) - K_j\right)_+\right], \text{ such that } \sum_{j=1}^{n} K_j = K.$$

The stated result follows then by applying Theorem 2. ■

From the theorem above, we can conclude that the mean absolute deviation optimality criterion (73) gives rise to a quantile-based allocation principle: Each allocated capital K_i is given by the α–mixed inverse of the ζ_i–weighted distribution function of X_i at a fixed probability level, which is chosen such that the full allocation requirement is satisfied. From (83), we find that the optimal allocations K_i satisfy the following conditions:

$$F_{X_i}^{(\zeta_i)}(K_i) = F_{\overline{S}^c}(K), \qquad i = 1, \ldots, n. \tag{85}$$

In the case where

$$\mathbb{P}[\zeta_i > 0] = 1, \qquad i = 1, \ldots, n, \tag{86}$$

and the distributions $F_{X_j}^{(\zeta_j)}$ are strictly increasing, then the optimal allocations in (83) reduce to

$$K_i = \left(F_{X_i}^{(\zeta_i)}\right)^{-1}(F_{\overline{S}^c}K), \qquad i = 1, \ldots, n. \tag{87}$$

We end this subsection with an example of an absolute deviation allocation principle. Consider the following choice for the weighting random variables:

$$\zeta_i = \frac{\mathbb{I}(S > K)}{\mathbb{P}[S > K]}, \qquad i = 1, \ldots, n. \tag{88}$$

This means that the optimisation procedure only considers those outcomes that lead to insolvency, i.e. the case where $S > K$, on the aggregate portfolio level. In this instance, the optimisation problem (75) reduces to

$$\min_{K_1,\ldots,K_n} \sum_{j=1}^{n} \mathbb{E}\left[(X_j - K_j)_+ \mid S > K\right], \text{ such that } \sum_{j=1}^{n} K_j = K. \tag{89}$$

From (79), we find that the ζ_i–weighted distribution function of X_i is given by

$$F_{X_i}^{(\zeta_i)}(x) = \mathbb{P}\left[X_i \le x | S > K\right], \qquad i = 1, \ldots, n. \tag{90}$$

In the particular case that $x = K_i$, we find from (85) that:

$$\overline{F}_{X_i}^{(\zeta_i)}(K_i) = \mathbb{P}\left[X_i > K_i | S > K\right] = \overline{F}_{\overline{S}^c}(K), \qquad i = 1, \ldots, n. \tag{91}$$

Hence, capital is allocated such that the conditional probability of a business unit's loss exceeding its allocated capital, given that the whole company defaults, is identical for all business units.

4 Conclusion

In this article, we developed a general and unifying optimisation framework that produces several of the capital allocation approaches that are encountered both in the literature and in practice. This general framework is based on the idea of minimising the sum of the divergences between the losses and the allocated capital of the different subportfolios.

Depending on how this divergence is defined, several alternative allocation methods arise. In particular, choice of the functions ζ_i determines which scenarios (states of the world), e.g. portfolio or unit-specific, carry most weight in the capital allocation. We believe that this approach allows for a closer alignment between capital allocation and the definition of management's risk tolerance.

Finally, the framework presented in this paper provides the flexibility to produce new capital allocation methods, by varying, for example, the choices of the weights ζ_i and v_i. This is not an avenue we pursued here at great lengths but remains a subject of importance for future work.

Acknowledgement: The authors thank Qihe Tang, Yaniv Zaks and Peter England for their helpful comments and suggestions on an earlier version of this paper. Jan Dhaene acknowledges the financial support of the Onderzoeksfonds K.U. Leuven (GOA/07: Risk Modeling and Valuation of Insurance and Financial Cash Flows, with Applications to Pricing, Provisioning and Solvency), and of Fortis through the K.U. Leuven Fortis Chair in Financial and Actuarial Risk Management.

References

C. Acerbi. Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking and Finance*, 26(7):1505–1518, 2002.

C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking and Finance*, 26(7):1487–1503, 2002.

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

H. Bühlmann. *Mathematical Methods in Risk Theory*. Springer-Verlag, Berlin, 1970.

J.D. Cummins. Allocation of capital in the insurance industry. *Risk Management and Insurance Review*, 3(1):7–27, 2000.

M. Denault. Coherent allocation of risk capital. *Journal of Risk*, 4(1):1–34, 2001.

J. Dhaene, M. Denuit, M.J. Goovaerts, R. Kaas, and D. Vyncke. The concept of comonotonicity in actuarial science and finance: theory. *Insurance: Mathematics and Economics*, 31:3–33, 2002.

J. Dhaene, M.J. Goovaerts, and R. Kaas. Economic capital allocation derived from risk measures. *North American Actuarial Journal*, 7(2):44–59, 2003.

J. Dhaene, S. Vanduffel, Q. Tang, M.J. Goovaerts, R. Kaas, and D. Vyncke. Risk measures and comonotonicity: a review. *Stochastic Models*, 22(4):573–606, 2006.

J. Dhaene, L. Henrard, Z. Landsman, A. Vandendorpe, and S. Vanduffel. Some results on the CTE based capital allocation rule. *Insurance: Mathematics and Economics*, 42(2):855–863, 2008.

J. Dhaene, M. Denuit, and S. Vanduffel. Correlation order, merging and diversification. Submitted, 2009. Available at `http://www.econ.kuleuven.be/insurance/research.htm`.

H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447, 2002.

E. Furman and R. Zitikis. Weighted premium calculation principles. *Insurance: Mathematics and Economics*, 42(1):459–465, 2008a.

E. Furman and R. Zitikis. Weighted risk capital allocations. *Insurance: Mathematics and Economics*, 43(2):263–269, 2008b.

H.U. Gerber. On additive premium calculation principles. *ASTIN Bulletin*, 7(3):215–222, 1974.

H.U. Gerber. The Esscher premium principle: a criticism. comment. *ASTIN Bulletin*, 12(2): 139–140, 1981.

M.J. Goovaerts, F. deVylder, and J. Haezendonck. *Insurance Premiums: Theory and Applications*. Amsterdam, North Holland, 1984.

H. Gründl and H. Schmeiser. Capital allocation for insurance companies: what good is it? *Journal of Risk and Insurance*, 74(2):301–317, 2007.

W.R. Heilmann. Decision theoretic foundations of credibility theory. *Insurance: Mathematics and Economics*, 8:77–95, 1989.

M. Kalkbrener. An axiomatic approach to capital allocation. *Mathematical Finance*, 15(3):425–437, 2005.

J. Kim and M. Hardy. A capital allocation based on a solvency exchange option. *Insurance: Mathematics and Economics*, 2008. In press.

R.J.A. Laeven and M.J. Goovaerts. An optimization approach to the dynamic allocation of economic capital. *Insurance: Mathematics and Economics*, 35(2):299–319, 2004.

Z. Landsman and E.A. Valdez. Tail conditional expectations for elliptical distributions. *North American Actuarial Journal*, 7(4):55–71, 2003.

J. LeMaire. An application of game theory: cost allocation. *ASTIN Bulletin*, 14(1):61–81, 1984.

S.C. Myers and J.A. Read Jr. Capital allocation for insurance companies. *Journal of Risk and Insurance*, 68(4):545–580, 2001.

L. Overbeck. Allocation of economic capital in loan portfolios. in J. Franke, W. Haerdle and G. Stahl (eds.), Measuring Risk in Complex Systems, Springer, 2000.

H.H. Panjer. Measurement of risk, solvency requirements, and allocation of capital within financial conglomerates. Research Report 01-14. Institute of Insurance and Pension Research, University of Waterloo, Canada, 2001.

R.D. Phillips, J.D. Cummins, and F. Allen. Financial pricing of insurance in the multiple-line insurance company. *Journal of Risk and Insurance*, 65(4):597–636, 1998.

C.R. Rao. *Statistics and Truth: Putting Chance to Work*. World Scientific Publishing, River Edge, NJ, 1997.

M. Sherris. Solvency, capital allocation and fair rate of return in insurance. *Journal of Risk and Insurance*, 73(1):71–96, 2006.

D. Tasche. Allocating portfolio economic capital to sub-portfolios. In A. Dev (ed.), Economic Capital: A Practitioner's Guide, Risk Books, pp. 275-302, 2004.

A. Tsanakas. Dynamic capital allocation with distortion risk measures. *Insurance: Mathematics and Economics*, 35(2):223–243, 2004.

A. Tsanakas. Capital allocation with risk measures. Proceedings of the 5th Actuarial and Financial Mathematics Day, 9 February, Brussels, pp. 3-17, 2007.

A. Tsanakas. To split or not to split: capital allocation with convex risk measures. *Insurance: Mathematics and Economics*, 2008. In press.

E.A. Valdez and A. Chernih. Wang's capital allocation formula for elliptically contoured distributions. *Insurance: Mathematics and Economics*, 33(3):517–532, 2003.

G.G. Venter. Capital allocation survey with commentary. *North American Actuarial Journal*, 8 (2):96–107, 2004.

S.S. Wang. Premium calculation by transforming the premium layer density. *ASTIN Bulletin*, 26 (1):71–92, 1996.

S.S. Wang. Normalized exponential tilting: pricing and measuring multivariate risks. *North American Actuarial Journal*, 11(3):89–99, 2007.

Y. Zaks, E. Frostig, and B. Levikson. Optimal pricing of a heterogeneous portfolio for a given risk level. *ASTIN Bulletin*, 36(1):161–185, 2006.