Coherent allocation of risk capital

Michel Denault*
RiskLab
Swiss Federal Institute of Technology
Switzerland

26 October 1999

Abstract

The allocation problem stems from the diversification effect observed in risk measurements of financial portfolios: the sum of the risk measures of many portfolios is typically larger than the risk of all portfolios taken together. The allocation problem is to apportion this "diversification advantage" to the portfolios in a fair manner, to obtain new, firm-internal risk evaluations of the portfolios.

Our approach is axiomatic, in the sense that we first establish arguably necessary properties of an allocation scheme, and then study schemes that fulfill the properties. Important results from the area of game theory find a direct application, and are used here.

Keywords: allocation of risk; coherent risk measure; game theory; Shapley value; Aumann-Shapley prices; RORAC; risk-adjusted performance measure.

1 Introduction

The underlying theme of this paper is the sharing of costs within the different constituents of a firm. We call this sharing "allocation", as it is assumed that a higher authority exists within the firm, which has an interest in unilaterally dividing the costs between the constituents. The latter could be departments, business units, or, in the case of a financial firm, portfolios; the

*The author gratefully acknowledges the financial support of the RiskLab research institute in Zürich

important is that each such constituent yield some profit or loss, and that some uncertainty be involved in the level of this profit or loss for the future. As an insurance against this uncertainty, the firm could well, and may even be regulated to, hold an amount of cash aside, to better face unpleasant surprises the future may offer. We will call this reserve, the *risk capital* of the firm. The costs mentionned above are precisely this risk capital: indeed, from a financial perspective, holding a large amount of money dormant, i.e. in extremely low risk, low return money instruments, is a cost.

The problem of allocation of the risk capital to the constituents of the firm is interesting and non-trivial, because the sum of the costs (i.e. risk capitals) of each constituent, taken separately, is usually larger than the cost of the firm taken as a whole. This can be seen as large-scale economies, or better, as a diversification effect. That is, there is a decline in total costs to be expected by pooling the activities of the firm, and this *advantage* needs to be shared fairly between the constituents. In that sense, the allocated amounts are effectively *internal risk measures*, which account for the diversification effect.

The allocation exercise is basically performed for comparison purposes: knowing the profit *and* the risk taken by the components of the firm, allows for a much wiser comparison than knowing only of profits. This idea of a richer information set underlies the well-known concept of risk-adjusted performance measures (R.A.P.M.) and return on risk-adjusted capital (R.O.R.A.C.)

The article is divided in five main parts. We introduce the basic concepts in the next section. Section 3 describes the good qualities of an allocation principle. Section 4 considers one specific allocation principle, called the Shapley value. The following section, takes a slightly different perpective of what is a "good" allocation principle. Section 6 discusses the problem of negative allocation amounts.

We make, throughout this article, liberal use of the concepts and results of *game theory.* As we hope to convince the reader, game theory is an excellent structure on which to cast the allocation problem, and a eloquent language to discuss it.

2 Risk measure, risk capital, and allocation

While risk naturally evokes the same idea of danger, giving a clear and natural definition of it is far from trivial. Here, we shall define **risk** for a firm as the danger of having such a low net worth at some point in the future, that it must stop its activities. A risk measure ρ brings a quantification of the level of risk; more specifically, a risk measure is a mapping of a random variable X to the reals. The random variable represents the net worth of a portfolio or firm, at some point in the future; $\rho(X)$ is the amount of a specified numéraire (e.g. cash dollars) which, kept aside as "safety net", ensures that the firm will keep a high enough net worth, at the said point in the future. We call this safety net, **"risk capital"**.

In this paper, we will not be overly concerned with specific risk measures; instead, we will specify the properties that a generic risk measure ρ shall possess. In their paper [1], Artzner, Delbaen, Eber and Heath have suggested a set of properties that any risk measure should satisfy, thus defining the concept of *coherent measures of risk*:

Definition 1 *A risk measure ρ is* **coherent** *if it satisfies the properties:*

Subadditivity *For all random variables X and Y,*

$$\rho(X+Y) \leq \rho(X) + \rho(Y)$$

Monotonicity *For all random variables X and Y such that $X \leq Y$*[1]*,*

$$\rho(X) \geq \rho(Y)$$

Degree one homogeneity *For all $\lambda \geq 0$ and all random variable X,*

$$\rho(\lambda X) = \lambda \rho(X)$$

[1]The relation $X \leq Y$ between two random variables is taken to mean $X(\omega) \leq Y(\omega)\ \ \forall\, \omega \in \Omega$, in a probability space $(\Omega, \mathcal{F}, P)$.

Translation invariance *For all random variable X and all $\alpha \in \mathbb{R}$,*

$$\rho(X + \alpha r_f) = \rho(X) - \alpha$$

where r_f is the rate of return on a reference, riskless investment.

The axioms that define coherence of risk are to be understood as necessary conditions for a risk measure to be *reasonable.* Let us briefly justify the axioms (see [1] for a comprehensive view). Subadditivity reflects the diversification of portfolios, or that "a merger does not create extra risk" [1, p.209]. Monotonicity says that if a portfolio Y is *always* worth more than X, then Y cannot be riskier than X. Homogeneity is a sort of limit case of subadditivity, representing what happens when there is precisely *no* diversification effect. Translation invariance is a natural requirement, given the meaning of the risk measure as defined in [1] and above.

Throughout this paper, we will therefore assume the generic risk measure ρ to be coherent, although no specific risk measure (such as Expected Shortfall) will be discussed.

Suppose now that one computes the risk capital of a firm consisting of several portfolios, or departments, or business units (from now on, we will speak only of portfolios, but departments or business units can equaly well be understood). By the subadditivity of ρ, or equivalently because of diversification effects, the risk capital of the firm is less than the sum of the risk capitals of the individual portfolios. It is precisely this fact that makes the *allocation problem,* an interesting and nontrivial one.

3 Coherence of the allocation principle

In a similar way to what the authors of [1] did in the case of risk measures, we provide in this section a set of axioms, that we suggest are necessary properties of a "reasonable" allocation principle. We will call *coherent* the allocation principles that satisfy the set of axioms.

We will use the following notation:

- X is a random variable representing a firm's total net worth at some point in the future T.
- X_i, $i \in \{1, 2, \dots, n\} = N$, is the net worth at the time T of the i^{th} portfolio of the firm (alternatively, one can speak of the i^{th} business unit of the firm). We assume that the n^{th} portfolio is a riskless intrument with net worth at time T equal to $X_n = \alpha r_f$, r_f being the riskless interest rate over that period. We assume that the relation $X = \sum_{i=1}^{n} X_i$ holds.
- $K = \rho(X)$ is the risk capital as measured by ρ, for the complete firm.
- K_i is the amount of risk capital allocated to portfolio $i \in N$, by using some allocation principle.

We suggest the following definition of a coherent allocation:

Definition 2 *An allocation K_i, $i \in N$, is* **coherent** *if it satisfies the four properties:*

1) Full allocation

$$\sum_{i \in N} K_i \; = \; \rho\left(\sum_{i \in N} X_i\right)$$

2) No undercut[2]

$$\forall\, M \subseteq N, \; \sum_{i \in M} K_i \; \leq \; \rho\left(\sum_{i \in M} X_i\right)$$

3) Symmetry *If by joining any subset of players $M \subseteq N \backslash \{i, j\}$, i and j both make the same contribution in risk, then $K_i = K_j$.*

[2]The justification of this name comes from the following equivalent formulation:

$$\nexists\, M \subseteq N \text{ and } H \in \mathbb{R}^{|M|} \text{ such that } \sum_{i \in M} H_i = \rho(\sum_{i \in M} X_i) \text{ and } H_i < K_i \;\forall\, i \in M.$$

That is, there is no subset M of the set of portfolios, such that an allocation of the subset's risk capital exists, which is cheaper for every single portfolio in M.

5) Riskless allocation

$$K_n = \rho(\alpha r_f) = -\alpha$$

Recall that by definition $X_n = \alpha r_f$

Furthermore, we call **nonnegative coherent allocation** a coherent allocation which satisfies $K_i \geq 0 \, \forall i \in N$.

The above axioms can be justified as follows. The full allocation property is necessary so that the risk capital of the firm be completely allocated; a cost allocation exercise is futile if costs disappear into thin air. The "no undercut" property ensures that no portfolio manager, or coalition of portfolio managers, can argue that it would be better off on its own than with the firm, and as a consequence request a lower allocation of risk capital. The symmetry property ensures that a portfolio's allocation depends only on its contribution to risk within the firm, and nothing else. According to the riskless allocation axiom, a riskless portfolio should be allocated exactly its risk measure, which indidentally will be negative, as long as the riskless interest rate is positive. It also means that, all other things being equal, a portfolio that increases its cash position, should see its allocated capital decrease by the same amount, in the sense $K_i + K_n = K_i - \alpha, \; i \neq n$.

More generally, while the "no undercut" property ensures the *stability* of the solution, symmetry and translation invariance concern the *fairness* of the solution. Full allocation is needed for the problem to be non trivial.

The nonnegativity property's implications are less straightforward, and will be discussed in section 6.

3.1 Game theory and the allocation problem

Game theory is the study of situations where participants, called *players* adopt various behaviours, or strategies, to best attain their individual goals; what the goals are, what behaviours are possible, is determined by a set of rules, which constitutes the *game*.

We suggest that game theory is a very useful framework for the modeling and study of risk capital allocation problems; in fact, *cost allocation* games are a recurrent theme in game theory.

More specifically, we will consider here *cooperative* games, i.e. games where the players can do best by cooperating with each other. If we associate the concept of player with the portfolios of a firm, then the goal of each player is to minimize the capital it is allocated. Given the subadditivity of risks, the players have on one hand an incentive to operate as a single firm, since cooperation brings a net improvement of the total risk charge. On the other hand, the players will bring arguments to each other to keep their allocation as low as possible. This last idea is modelled through the consideration of *coalitions:* subsets of players who can argue that they are treated unfairly by the allocation principle. The cost allocation game considered here will then also be called coalitional.

Some background on game theory is needed at this point.

Definition 3 *A* game in characteristic function form*, or coalitional form, consists of*

- *a finite set N of players*
- *a function ρ that associates to every subset S of N (a* **coalition***) a real number $\rho(S)$.*

The game is then represented as (N, ρ).

In our case, the players represent the portfolios, and ρ, the the risk measure. We shall abuse our notation, and define $\rho(S) \triangleq \rho(\sum_{i \in S} X_i)$ where S is a subset of N; one can infer from the context whether the argument is a set or a random variable. $\rho(S)$ is then the amount of risk incurred by the coalition S, that is, the total risk capital charged to the portfolios in S if they form as one firm.

It is assumed that the risk, or cost function ρ, is **transferable**, in the sense that one unit of risk has the same meaning, or disutility, for all the players. Other coalitional games are said to have *nontransferable costs.*

In accordance with both the definition of coherent risk measures and with coalitional game theory, we make the assumption that characteristic functions are *subadditive*, as per definition 1 above: $\rho(X+Y) \leq \rho(X)+\rho(Y)$. If the arguments of ρ are subsets, then subadditivity is written $\rho(S \cup T) \leq \rho(S) + \rho(T)$ for all subsets S and T of N with empty intersection.

It is important to point out that the characteristic functions of coalitional games are usually, in the game theory literature, taken to be *superadditive* ($\rho(X + Y) \geq \rho(X) + \rho(Y)$), rather than subadditive; the interpretation of such characteristic functions is that they represent payments (payoffs) to the players, as opposed to costs charged to them. Of course, these are two sides of the same coin, and considering one or the other makes no difference, except perhaps when the two viewpoints are discussed in the same text, which can bring some confusion. For the sake of clarity, we assume throughout this paper the viewpoint of a cost allocation, and thus of subadditive characteristic functions.

Definition 4 *Given a coalition S of N, we call a vector $K \in \mathbb{R}^n$ an* **S-feasible allocation** *if $\sum_{i \in S} K_i = \rho(S)$; an N-feasible allocation is simply called a* **feasible allocation.**

One crucial definition is that of the core of a game:

Definition 5 *The* core of a coalitional game *(set of players N, characteristic function ρ) is the set of feasible allocations $K \in \mathbb{R}^n$ for which $\sum_{i \in S} K_i \leq \rho(S)$ for all coalitions S.*

Clearly, the nonemptiness of the core is a necessary condition for the existence of a coherent allocation, since an element of the core by definition fulfils the axioms "full allocation" and "no undercut". Since the core is defined as a system of linear inequalities, a condition for its nonemptiness can be derived using classical linear algebra results on separating hyperplanes, as will be done in the next section. However, the special structure of games and cores allows a more specific condition called the *Bondareva-Shapley* *theorem*. Let $\mathcal{C}$ be the set of all coalitions of N, let us denote by $1_S \in \mathbb{R}^n$ the characteristic vector of the coalition S:

$$(1_S)_i = \begin{cases} 1 & \text{if } \ i \in S \\ 0 & \text{otherwise} \end{cases}$$

A **balanced collection of weights** is a collection of $|\mathcal{C}|$ numbers λ_S in $[0, 1]$ such that for any player i, the sum of the λ_S over all coalitions S that contain i, is 1, i.e. $\sum_{S \in \mathcal{C}} \lambda_S 1_S = 1_N$. In words, player i can distribute her time between different coalitions, but cannot participate in two coalitions at the same time. A game is **balanced** if $\sum_{S \in \mathcal{C}} \lambda_S \rho(S) \geq \rho(N)$ for all balanced collections of weights. We then can state:

Theorem 1 (Bondareva-Shapley, [6], [14]) *A coalitional game with transferable costs has a nonempty core if and only if it is balanced.*

Proof: see e.g. [9]. Note that this theorem remains an application, if specific, of the separating hyperplane theorem.

A direct consequence of this theorem, we have the following:

Theorem 2 *If the allocation problem is modelled as a coalitional game with transferable utility whose characteristic function is a* coherent *risk measure, then the core of this game is nonempty.*

Proof:

$$\begin{aligned} \sum_{S \in \mathcal{C}} \lambda_S \, \rho(S) &= \sum_{S \in \mathcal{C}} \rho \left(\sum_{i \in S} \lambda_S X_i \right) \\ &\geq \rho \left(\sum_{S \in \mathcal{C}} \left(\sum_{i \in S} \lambda_S X_i \right) \right) \\ &= \rho \left(\sum_{i \in N} \left(\sum_{S \in \mathcal{C},\, S \ni i} \lambda_S X_i \right) \right) \\ &= \rho(N) \end{aligned}$$

□

As a result, an allocation problem based on a coherent risk measure has a solution satisfying the first two properties of a coherent allocation.

Another condition can be given for the core of a game to be nonempty, this time pertaining to the *strong subadditivity* of the game:

Definition 6 *A coalitional game with transferable costs is* strongly subadditive *if it is based on a strongly subadditive*[3] *characteristic function:*

$$\rho(S) + \rho(T) \geq \rho(S \cup T) + \rho(S \cap T)$$

for all coalitions S and T.

Theorem 3 *A strongly subadditive game has a nonempty core.*

Proof: see [9, p. 260].

While the non-emptiness of the core is a necessary condition for the existence of coherent allocation, it is not sufficient, nor does the core, as a rule, yield a unique solution. We address both points in the following section.

4 The Shapley value as allocation principle

We discuss in this section a specific *allocation principle*, i.e. a way of choosing one allocation among others, namely the Shapley value. We thereby also address the topic of sufficient conditions for the coherence of the allocation. Finally, we make a fews on other eventual allocation principles.

4.1 The Shapley value

The Shapley value was introduced by Lloyd Shapley [16] and has ever since received a considerable amount of interest (see for example [12]). Let us first define what a value is.

Definition 7 *A* **value** *is a function that maps each game (N, ρ) into a unique, feasible allocation, i.e.*

$$\Phi : (N, \rho) \longmapsto \begin{bmatrix} \Phi_1(N, \rho) \\ \Phi_2(N, \rho) \\ \vdots \\ \Phi_n(N, \rho) \end{bmatrix} = \begin{bmatrix} K_1 \\ K_2 \\ \vdots \\ K_n \end{bmatrix} \text{ where } \sum_{i \in N} K_i = \rho(N)$$

[3] By definition, a strongly subadditive set function is subadditive. We follow Shapley [17] in our terminology; note that he calls convex, a function satisfying the reverse relation of strong subadditivity

(the $K-$ notation is used when the arguments are clear from the context)

Clearly, a value is simply an allocation principle in the game-theoretic parlance. The Shapley value is a value as its name indicates and can therefore be used as solution concept yielding a unique solution. We give two characterizations of the Shapley value, one axiomatic and one algebraic.

We use the abbreviation $\Delta_i(S) = \rho(S \cup i) - \rho(S)$ for any set $S \subset N, i \notin S$. Two players i and j are *interchangeable in* (N, ρ) if either one makes the same contribution to any coalition S it may join, that contains neither i nor j: $\Delta_i(S) = \Delta_j(S)$. A player is a *dummy* if it brings the contribution $\rho(i)$ to any coalition S that does not contain it already: $\Delta_i(S) = \rho(i)$. We now define three properties that a value may exhibit:

Symmetry If players i and j are interchangeable, then $\Phi(N, \rho)_i = \Phi(N, \rho)_j$

Dummy player For a dummy player, $\Phi(N, \rho)_i = \rho(i)$

Additivity over games For two games (N, ρ_1) and (N, ρ_2), $\Phi(N, \rho_1 + \rho_2) = \Phi(N, \rho_1) + \Phi(N, \rho_2)$, where the game $(N, \rho_1 + \rho_2)$ is defined by $(\rho_1 + \rho_2)(S) = \rho_1(S) + \rho_2(S)\ \forall S \subset N$.

The axiomatic characterization of the Shapley value is then:

Definition 8 ([16]) *The* **Shapley value** *is the only value that satisfies the properties of symmetry, dummy player, and additivity over games.*

The reader will have recognized that the Shapley value can be an important piece of the allocation puzzle: using the Shapley value to allocate risk capital, automatically **yields a coherent allocation, but for the "no undercut" axiom.** Full allocation and symmetry are satisfied by definition. The riskless allocation axiom of Definition 2 is equivalent to the dummy player axiom: from our definitions of section 3, the reference, riskless instrument (cash and equivalents) is a dummy player.

A note on additivity over games: no such property is required of coherent allocations, as it conflicts with the coherence of the risk measures, see section 5.4. The uniqueness question, for coherent allocation principles, thus remains open.

When, then, does the Shapley value satisfy the "no undercut" property, yielding a coherent allocation ? Equivalently, when is the Shapley value in the core of the game ? The only pertaining result to our knowledge is that of Shapley (1971), and involves the strong subadditivity of the game:

Theorem 4 ([17]) *If a game (N, ρ) is strongly subadditive, its core contains the Shapley value.*

This is perhaps disappointing, as strong subadditivity is more stringent than the subadditivity we like to require of ρ, and we thus fall short of a convincing proof of the existence of coherent allocations. Of course in practice, one could check if the strong subadditivity property of ρ is not indeed satisfied. Alternatively, one can turn to section 5 where a different view of the allocation problem is given, with stronger existence arguments.

Before closing this section, let us turn to the algebraic definition of the Shapley value, which provides both an interpretation, and an explicit computational approach.

Definition 9 *The Shapley value K^{Sh} for the game (N, ρ) is defined as:*

$$K_i^{Sh} = \sum_{S \in \mathcal{C}_i} \frac{(s-1)!(n-s)!}{n!} \left(\rho\left(\sum_{j \in S} X_j \right) - \rho\left(\sum_{j \in S/\{i\}} X_j \right) \right), \quad i \in N$$

where $\mathcal{C}_i$ represent all coalitions of N that contain i and $s = |S|$.

The Shapley value can then be given the following interpretation: let the players agree to meet in a room at a certain time. Assume that they will arrive at the meeting at slightly different, random times, and that all orders of arrival are equally likely. The Shapley value is the *expected* contribution of a player to the risk measure of the group's in the room, as she arrives.

From a computational point of view, a risk measure evaluation is required for each of the 2^n coalitions (i.e. all possible subsets), a task which quickly become impossible, even for moderate n. For example, should a risk evaluation (i.e. computing $\rho(S)$ for some $S \subset N$) last one hour, independently of the size of S then a week of (serial) computation would be necessary to obtain the Shapley value of a game with seven portfolios.

Let us however add that 1) in many applications, n could indeed be rather small (say, in the 5 to 10 range) 2) computation time is not a critical issue here. This being said, the interested reader will pursue until section 5, where a similar type of coherent allocation is described, with better computational properties.

4.2 Remarks on other allocation principles

Other allocation principles can be considered. A simple approach could be to allocate the risk capital proportionately to the risk measure of each portfolio:

$$K_i = \frac{\rho(X_i)}{\sum_{j \in N} \rho(X_j)} K.$$

This naive principle can quite easily be shown not to satisfy coherence.

Two more principles based on the statistical properties of the random variables X_i are the *covariance principle:*

$$K_i = \frac{\mathrm{cov}(X_i, X)}{\mathrm{var}(X)} K$$

and the *conditional expected shortfall*

$$K_i = E\,[\, X_i \mid X \le q_\alpha \,]$$

where q_α is the α-quantile of the distribution of X. Such principles cannot be analysed in abstraction of the specific risk measure ρ used, as they are not functions of the (finite) number of *evaluations* of ρ for the coalitions. With the exception of some remarks at the end of next section, these approaches will not be discussed further in this paper.

5 Allocation of costs to scalable players

In the previous sections, the components of a firm were portrayed as players of game, each of them *indivisible*. Given that our players are portfolios, this indivisibility assumption need not necessarily hold, as one could consider coalitions involving *fractions* of players. The purpose of this section is to examine a variant of the "allocation game" allowing divisible players.

We also stress that the results given in this section have a clear computational advantage over the Shapley principle, in that the computation of the risk measures of every coalition is avoided.

5.1 Games with scalable players

As mentioned above, the allocation game can be viewed from a different point of view from the one taken so far. The theory of coalitional, cooperative games has been extended to "continuous players" who need be neither in nor out of a coalition, but who have a "scalable presence". This point of view seems much less uncongruous if the players in question are, for example, portfolios; a coalition can then consist of sixty percent of portfolio A, and fifty percent of portfolio B. Of course, this means "x percent of *each* instrument in the portfolio".

The seminal work for the development of the game concepts discussed in this section, was Aumann and Shapley's book "Values of Non-Atomic Games" ([2]). There, the interval $[0, 1]$ represents the set of all players, and coalitions are measurable subintervals (in fact, elements of a σ-algebra). Any subinterval contains one of smaller measure, so that there are no *atoms*, i.e. smallest entities that could be called players; hence the name "nonatomic games". We prefer the more intuitive and practical approach taken later by various authors, see [3], [4], [7], who let a vector $\lambda \in \mathbb{R}^n_+$ represent the "level of presence" of the players, each component being associated to a player. The goal of the cost allocation game is then to find a price vector, each component of which represents the *per-unit* cost allocated to the corresponding player.

Note that the articles [4] and [7] in fact did not make use of game theoretic concepts, and couched their results in economic terms only. We keep here the language of game theory, mainly because of its clarity.

We thus define a second type of games, as follows.

Definition 10 *A* **coalitional game with scalable players** *(N, Λ, r) consists of*

- *a finite set N of players, with $|N| = n$;*
- *a positive vector $\Lambda \in \mathbb{R}^n_+$, an amount representing for each of n players his full involvement.*
- *a real-valued characteristic function $r : \mathbb{R}^n \to \mathbb{R}$, $r : \lambda \mapsto r(\lambda)$ defined on $0 \leq \lambda \leq \Lambda$*

As before, the players are identified with portfolios, or business units within a firm. The vector Λ represents, for each portfolio, the "size" of the portfolio or its "activity level", where the reference unit for each portfolio can be chosen at will. This Λ can, for example, represent the business volume of the business units, in a reference currency. Keeping X_i with the same meaning as earlier (a random variable representing the net worth of portfolio i at a future time T), we introduce for future use the variables Y_i defined as

$$Y_i := \frac{X_i}{\Lambda_i},\ i \in \{1, 2, \dots, n\} = N$$

(X_n keeps its earlier "special case" definition, $X_n = \alpha r_f$) The characteristic function r is again a risk measure, identified with the risk measure ρ of the previous section, through

$$r(\lambda) := \rho\left(\sum_{i \in N} \lambda_i Y_i\right)$$

so that $r(\Lambda) = \rho(N)$.

The definition of coherent risk measure given as Definition 1 is adapted as follows:

Definition 11 *A risk measure r is* **coherent** *if it satisfies the four properties:*

Subadditivity *For all* λ_1 *and* λ_2 *in* $\mathbb{R}^n$ *such that* $0 \leq \lambda_1 \leq \Lambda$, $0 \leq \lambda_2 \leq \Lambda$, *and* $0 \leq \lambda_1 + \lambda_2 \leq \Lambda$,

$$r(\lambda_1 + \lambda_2) \leq r(\lambda_1) + r(\lambda_2)$$

Monotonicity *For all* λ_1 *and* λ_2 *in* $\mathbb{R}^n$ *such that* $0 \leq \lambda_1 \leq \Lambda$, $0 \leq \lambda_2 \leq \Lambda$,

$$\lambda_1^t X \leq \lambda_2^t X \Rightarrow r(\lambda_1) \geq r(\lambda_2)$$

where the left-hand side inequality is again understood as in footnote 1.

Degree one homogeneity *For all* $\lambda \in \mathbb{R}^n$, $0 \leq \lambda \leq \Lambda$, *and for all* $\gamma \in \mathbb{R}_+$ *such that* $0 \leq \gamma\lambda \leq \Lambda$,

$$r(\gamma\lambda) = \gamma r(\lambda)$$

Translation invariance *For all* λ *in* $0 \leq \lambda \leq \Lambda$,

$$r(\lambda) \quad = \quad r\left(\begin{bmatrix} \lambda^{(1)} \\ \lambda^{(2)} \\ \vdots \\ \lambda^{(n-1)} \\ 0 \end{bmatrix}\right) - \frac{\lambda^{(n)}}{\Lambda^{(n)}}\alpha$$

Clearly, r is coherent if ρ is, and vice-versa also within the domain of variables $\sum_i \lambda_i X_i$, $0 \leq \lambda_i \leq \Lambda_i$. As in the previous section, a coherent risk measure will be the basis of the allocation principle.

5.2 Cost allocation to scalable players

We introduced above future net worth on a *per unit* basis; the allocation shall similarly be discussed “per unit”. To this end, we introduce a vector $k \in \mathbb{R}^n$, each component of which represent the *per unit* allocation of risk capital to each player (or portfolio); we also call this vector “price vector”, to emphasize its *per-unit* nature. The capital allocated to each player is obtained by a simple Hadamard (or component-wise) product

$$\Lambda .* \; k \; = \; K \tag{1}$$

Let us also define, in a manner equivalent to the concepts of section 3.1:

Definition 12 *A vector $k \in \mathbb{R}^n$ is a* **feasible** *per-unit allocation vector of the game $(N,\, \Lambda,\, r)$ if*

$$\Lambda^t k = r(\Lambda) \tag{2}$$

A **value** *is a function assigning to each coalitional game with scalable players $(N,\, \Lambda,\, r)$ a unique feasible per-unit allocation vector:*

$$\phi : (N, \Lambda, r) \longmapsto \begin{bmatrix} \phi_1(N,\Lambda,r) \\ \phi_2(N,\Lambda,r) \\ \vdots \\ \phi_n(N,\Lambda,r) \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}$$

We now give the definitions of the properties of per-unit allocations that are used later.

Definition 13 *The following properties are defined for a game with scalable players $(N,\, \Lambda,\, r)$ and a per-unit allocation vector $\phi(N,\Lambda,r) = k$:*

- **No undercut** *For all λ in $[0,\Lambda]$,*

$$\lambda^t k \leq r(\lambda) \tag{3}$$

- **Aggregation invariance** *Suppose the risk measure r satisfies $r(\lambda) = r'(\Gamma\lambda)$ for all λ in $[0,\Lambda]$ and $m \times n$ constant matrix Γ, then*

$$\phi_i(N,\Lambda,r) = \phi_i(N,\Gamma\Lambda,r')\Gamma \tag{4}$$

- **Dummy player** *If i is a dummy player, in the sense that*

$$r(\lambda) - r(\lambda') = (\lambda_i - \lambda'_i)\frac{\rho(X_i)}{\Lambda_i}$$

whenever $0 \leq \lambda \leq \Lambda$ and $\lambda' = \lambda$ except in the i^{th} component, then

$$k_i = \frac{\rho(X_i)}{\Lambda_i} \tag{5}$$

- **Monotonicity** *If* $r(\lambda) \le r(\lambda')$ *for all* λ *and* λ' *such that* $0 \le \lambda \le \lambda' \le \Lambda$*, then*

$$\phi(N, \Lambda, r) \ge 0 \tag{6}$$

We are now in a position of defining "coherence" in the setting of allocation to scalable players:

Definition 14 *A* **coherent per-unit allocation vector** *for the game (N, Λ, r) is a vector* $k \in \mathbb{R}^n$ *which is* feasible *(2),* aggregation invariant *(4),* monotone *(6), satisfies the* dummy player property *(5) and allows* no undercut *(3) for that game.*

To put these definitions in context, we would like to find a *value*, that is a method of assigning to each allocation problem a *feasible* price vector; we furthermore want that price vector be *coherent*.

The properties required of a coherent price vector can be justified essentially in the same manner as their equivalent in Definition 2. The "no undercut" property ensures that every portfolio, group of portfolios, or part thereof, receives an allocation such that it cannot be better off than with the whole firm. Aggregation invariance is akin to the symmetry property: equivalent risks should receive equivalent allocations. The dummy player property is the equivalent of the riskless allocation of Definition 2, and is necessary to give "risk capital" the sense we gave it in section 2: an amount of riskless instrument necessary to make a portfolio acceptable, riskwise. The monotonicity condition, (sometimes also called "nonnegativity", see [19]), is a much restricted version of the "at-large" nonnegativity of section 3. Here too, we reserve the name nonnegativity to mean that allocations are unconditionally nonnegative. The interpretation of monotonicity is that if the risk $r(\lambda)$ incurred by each of the n portfolios increases monotonically over $0 \le \lambda \le \Lambda$, then the per-unit allocation of risk capital should be nonnegative.

Much less is known about this allocation problem than is known about the similar problem described in section 3.1. On the other hand, one solution concept has been well investigated, which we discuss now: the Aumann-Shapley pricing principle.

5.3 Aumann-Shapley pricing

In their original work [2], Aumann and Shapley extended the concept of Shapley value to the nonatomic setting: that is, they established that if some specific but reasonable conditions were required of a cost (or gain) allocation, then this allocation was unique[4] and well-defined. The result was the Aumann-Shapley prices, nonatomic equivalent of the Shapley value for atomic games.

Their main result, as re-written later by Billera and Heath, leads directly to the following:

Theorem 5 *Consider the game with scalable players (N, r, Λ) with r having continuous first partial derivatives, $r(0) = 0$, and $\Lambda > 0$. Then there is a per-unit allocation vector $k \in \mathbb{R}^n$ that satisfies feasibility, aggregation invariance, and monotonicity (properties (2), (4), (6)). It is given by*

$$\phi_i^{AS}(N, \Lambda, r) = k_i^{AS} = \int_0^1 \frac{\partial r}{\partial \lambda_i}(\gamma\Lambda)\, d\gamma \tag{7}$$

The per-unit cost k_i^{AS} is an average of the marginal costs of the i^{th} portfolio, as the level of activity or volume increases uniformly for all portfolios from 0 to Λ.

Properties "no undercut" (3) and "dummy player" (5) are absent from the above definition. The "dummy player" property will be discussed in lemma 2, while the "no undercut" will be considered in lemma 3.

But first, let us state from standard calculus:

Lemma 1 *If f is a k–homogeneous function, i.e. $f(\gamma x) = \gamma^k f(x)$, then $\frac{\partial f(x)}{\partial x_i}$ is $(k-1)$-homogeneous.*

[4]See section 5.4 below

As a result, since r is 1–homogeneous,

$$\phi_i^{AS}(N,\Lambda,r) = k_i^{AS} = \frac{\partial r(\Lambda)}{\partial \lambda_i} \tag{8}$$

in the theorem above, and the per-unit allocation vector $k^{AS} = \phi(N,\Lambda,r)$ is the gradient of the mapping r, evaluated at the "full presence" level Λ. We call this gradient "Aumann-Shapley per-unit allocation vector", or simply **Aumann-Shapley prices**. The amount of risk capital allocated to each portfolio is then given by the components of the vector

$$K^{AS} = k^{AS} .* \Lambda \tag{9}$$

Given (8), the very meaning of a dummy player in Definition 13 implies:

Lemma 2 *When the allocation process is based on a coherent risk measure r, the Aumann-Shapley prices (8) satisfy the dummy player property.*

It was already noted in [4] that under a decreasing marginal costs condition over $[0,\Lambda]$, the Aumann-Shapley per-unit allocation (7) fulfils the "no undercut" property:

Lemma 3 *Suppose that r has nonincreasing marginal costs, in the sense that for any γ and γ' such that $0 \leq \gamma \leq \gamma' \leq \Lambda$,*

$$\frac{\partial r(\gamma)}{\partial \lambda_i} \;\geq\; \frac{\partial r(\gamma')}{\partial \lambda_i} \qquad \forall\, i \in N.$$

Then

$$\lambda^t\, \phi^{AS}(N,\Lambda,r) \;\leq\; r(\lambda) \qquad \forall\, \lambda \in [0,\Lambda]$$

Proof: This is a direct consequence of (8) and a well-known differential calculus result, **Euler's theorem**, which states that if F is a real, n–variables, homogeneous function of degree k, then

$$x_1\frac{\partial F(x)}{\partial x_1} + x_2\frac{\partial F(x)}{\partial x_2} + \ldots + x_n\frac{\partial F(x)}{\partial x_n} \;=\; kF(x)$$

Indeed,

$$\begin{aligned}\lambda^t\,\phi(N,\Lambda,r) &= \sum_{i\in N}\lambda_i\frac{\partial r(\Lambda)}{\partial\lambda_i}\\ &\leq \sum_{i\in N}\lambda_i\frac{\partial r(\lambda)}{\partial\lambda_i}\\ &= r(\lambda)\end{aligned}\qquad\square$$

Third, in a very recent report, Tasche [18] gives conditions under which some quantile-based and shortfall-based risk measures are differentiable.

There is therefore hope that the conditions put on the risk measure by this second definition of allocation coherence, are not overly stringent. The topic of such an “appropriate” risk measure, i.e. one that is coherent, continuously differentiable, with nonincreasing marginal costs etc., is being studied and will be discussed elsewhere.

To recapitulate, given the existence of an “appropriate” risk measure, **there is a coherent allocation principle, given by**

$$k_i^{AS}(r,\Lambda)=\frac{\partial r(\Lambda)}{\partial\lambda_i};$$

of course, the *coherence* meant here is that of the scalable players context. On uniqueness, see section 5.4.

Two comments are in order, concerning the allocation principle (8). One is that the feasibility (2) of the allocation vector follows directly from Euler’s theorem (see the proof above), and that out of consideration for this, some authors have called the allocation principle (8) the *Euler principle*. See for example the attachment to the report of Patrik, Bernegger, and Rüegg [11], which provides some properties of this principle.

Second comment, the report of Tasche [18] comes fundamentally to the same result obtained in this section, namely that given some differentiability conditions on the risk measure ρ, a correct way of allocating risk capital is (8). Tasche’s justification of this contention is however completely different; he defines as “suitable”, capital allocations such that if the risk-adjusted return of a portfolio is “above average”, then, at least locally, increasing the share of this portfolio improves the overall return of the firm (and vice-versa for below average returns). Finally, note that the work of Schmock and Straumann [15] points again to the same conclusion. In the approach of [18] and [15], the Aumann-Shapley prices are in fact the unique satisfactory allocation principle.

We shall end this section by drawing the attention to the importance of the coherence of the risk *measure* for the allocation.

The subadditivity of the risk measure: is a necessary condition for the existence of an allocation with no undercut, in both the scalable and pure coalitional settings.

The homogeneity of the risk measure: ensures the simple form 8 of the Aumann-Shapley prices.

Both subadditivity and homogeneity: are used to prove that the core in nonempty (theorem 2), i.e. that an allocation without undercut exists. They are also used in the nonnegativity proof of addendum B.

The riskless property: is central to the definition of the riskless allocation (dummy player) property.

5.4 The uniqueness of allocation principles

Just as was the case with the Shapley value, the original results of the Aumann and Shapley, and Billera and Heath papers, involve a supplementary property, called *additivity*: If $r(\lambda) = r_1(\lambda) + r_2(\lambda)\ \forall \lambda \in [0, \Lambda]$, then

$$\phi_i(N, \Lambda, r) = \phi_i(N, \Lambda, r_1) + \phi_i(N, \Lambda, r_2)\ i = 1, \dots, n \tag{10}$$

If additivity is also to be fulfilled, then the Aumann-Shapley prices are the *only* possible allocation principle. In our case however, additivity has to be ruled out if we want to consider only coherent risk measures. Indeed, because of the riskless condition, a coherent risk measure cannot be the sum of two other coherent risk measures, as it leads to the contradiction:

$$\begin{aligned}\rho(X)-\alpha &= \rho(X+\alpha\, r_f) = \rho_1(X+\alpha\, r_f)+\rho_2(X+\alpha\, r_f)\\ &= \rho_1(X)+\rho_2(X)-2\alpha\\ &= \rho(X)-2\alpha\end{aligned}$$

Since the antecedent of additivity cannot hold with all three measures coherent, we leave the property aside. Whether or not the Aumann-Shapley prices remain unique, or under which situation, remains to be proved.

Note: the above contradiction concerning coherent risk measures has the important consequence that, for example, a coherent *market* risk measure and a coherent *credit* risk measuse will *never* sum up to a coherent measure.

6 The nonnegativity of the allocation

Given the concept of risk measure defined in [1] and in section 3, the risk of a portfolio may well be a negative value, with the interpretation that the portfolio is then *safer* than deemed necessary.

Similarly, there is no justification *per se* to enforce that the risk allocated to a portfolio be nonnegative; that is, the allocation of a negative risk amount does not pose a conceptual problem. Unfortunately, in the *application* we would like to make of the allocated capital, nonnegativity is a problem. If the amount is to be used in a RAPM quotient of the type $\frac{\text{return}}{\text{allocated capital}}$, negativity has a rather nasty drawback, as a portfolio with an allocated capital slightly below zero ends up with a negative risk-adjusted measure of large magnitude, whose interpretation is less than obvious. A negative allocation is therefore not so much a concern with the allocation itself, than with the use we would like to make of it.

A "crossed-fingers", and perhaps most pragmatic approach, is to hope that the coherent allocation is inherently nonnegative, in which case there is no further problem. In fact, one could reasonably expect nonnegative allocations to be the norm, in real-life situations. For example, provided each portfolio of the firm increases the risk measure when added to any subset of portfolios of the firm, i.e.

$$\rho(M \cup \{i\}) \geq \rho(M) \quad \forall\, M \subseteq N, \forall\, i \in N$$

then the Shapley value is necessarily nonnegative.

Should the coherent allocation *not* be nonnegative, two avenues can be considered:

1. Force the allocation principle to yield nonnegative solutions, eventually changing the definition of coherence to ensure existence of a solution.
2. Map allocated capitals to the positivie axis, so that the RAPM quotients behave in a more reasonable manner.

With respect to the first possibility, note for exemple that the properties "full allocation", "no undercut" and "nonnegativity" in Definition 2 form a set of linear inequalities (and one linear equality) on the variables K_i, so that with respect to these properties, the existence question is equivalent to the existence of a solution to a linear system. Specifically, a hyperplane separation argument proves that an allocation satisfying the three above properties will exist if the following condition on ρ holds:

$$\forall\, \lambda \in \mathbb{R}^n_+, \quad \rho\left(\sum_{i \in N} X_i\right) \min_{i \in N} \{\lambda_i\} \;\; \leq \;\; \rho\left(\sum_{i \in N} \lambda_i X_i\right) \tag{11}$$

The proof is given in addendum. The condition could be interpreted as follows. Firts assume that $\rho\left(\sum_{i \in N} X_i\right) > 0$, which is reasonable, if we are indeed to allocate some risk capital. Then (11) says that there is no positive linear combination of (each and every) portfolios, that runs no risk. In other words, a perfectly hedged portfolio cannot be attained by simply re-weighting the portfolios, if all portfolios are to have a positive weight.

The second avenue suggested above consists in computing the coherent allocation, and then applying a mapping of $\mathbb{R}^n$ to $\mathbb{R}^n_+$, to enforce the nonnegativity of the allocation. One could for example take $K^*_i = a \exp(b\, K_i)$

as the new allocation, for some constants a and b. The problem of choosing a mapping and its parameters, has not been investigated.

In both cases, there is a legitimacy problem that still needs to be resolved: how can one justify such operations ? The mere but obvious conclusion concerning the issue of the nonnegativity, is that it remains unsatisfactorily resolved for the moment.

7 Conclusion

In this article, we have analysed the allocation problem, mainly from a game-theoretic point of view. We suggest two sets of properties that define the *coherence* of risk capital allocation; the two are very similar, the difference being mainly in the view we hold of the firm and its portfolios.

The two definition of coherence are restrictive enough to limit the possibilities to a unique allocation principle for each definition. Further study is needed, to understand clearly how the two principles relate to each other.

Further work is also required to experiment the computational aspects of the problem, and to analyse the interaction between specific risk measures and the allocation principles.

Acknowledgments: The author would like to thank the members of RiskLab (Switzerland) for their input and encouragement; among others, the discussions with S. Bernegger, F. Delbaen, H.-J. Lüthi, and D. Straumann have been very useful. Prof. F. Delbaen is to be thanked in particular, as he graciously provided the fundamental idea that game theory and the Shapley value were promising avenues for the allocation of risk. He also suggested to us the condition (11), on the nonnegativity of allocations.

References

[1] Ph. Artzner, F. Delbaen, J.-M. Eber, D. Heath, "Coherent mea-

sures of risk", *Mathematical Finance*, vol. 9, no. 3 (1999), 203–228.

[2] R. Aumann, L. Shapley, "Values of Non-Atomic Games" (1974), Princeton University Press, Princeton, New Jersey.

[3] L. J. Billera, D. C. Heath, J. Raanan, "Internal Telephone Billing Rates—A Novel Application of Non-Atomic Game Theory", *Operations Research* vol. 26, no. 6 (1978), 956–965.

[4] L. J. Billera, D. C. Heath, "Allocation of shared costs: a set of axioms yielding a unique procedure", *Mathematics of Operations Research* vol. 7, no. 1 (1982) 32–39.

[5] L. J. Billera, J. Raanan, "Cores of nonatomic linear production games", *Mathematics of Operations Research* vol. 6, no. 3 (1981) 420–423.

[6] O.N. Bondareva, "Some applications of the methods of linear programming to the theory of cooperative games" (in Russian), *Problemy Kibernetiki*, no. 10 (1963), 119–139.

[7] L. J. Mirman, Y. Tauman, "Demand compatible equitable cost sharing prices", *Mathematics of Operations Research* vol. 7, no. 1 (1982) 40–56.

[8] National Association of Securities Dealers, "Reprint of the Manual", July 1996, C.C.R., Chicago.

[9] M. Osborne, A. Rubinstein, "A course in game theory" (1994), MIT Press, Massachusetts.

[10] G. Owen, "On the core of linear production games", *Mathematics of Operations Research* vol. 9, (1975) 358–370.

[11] G.S. Patrik, S. Bernegger, M. B. Rüegg, "The Use of Risk Adjusted Capital to Support Business Decision-Making", in the "Casualty Actuarial Society Forum, Spring 1999 Edition", pp. 243-334.

[12] A. Roth, editor, "The Shapley value. Essays in honor of Lloyd S. Shapley" (1988), Cambridge University Press, Cambridge.

[13] A. Rudd, M. Schroeder, "The calculation of minimum margin", *Management Science*, 28–12, 1982.

[14] H. Scarf, "The Core of an N-person game", *Econometrica* 35 (1967), 50–69.

[15] U. Schmock, D. Straumann, private communication.

[16] L. Shapley, "A value for n-person games", *Contributions to the theory of games*, volume II (Annals of mathematics studies, 28, Tucker and Luce, eds.) (1953), Princeton University Press, Princeton, New Jersey

[17] L. Shapley, "Cores of convex games", *International Journal of Game Theory* vol. 1, issue 1 (1971-1972), pp. 11–26.

[18] D. Tasche, "Risk contributions and performance measuremen t", Report of the Lehrstuhl für mathematische Statistik, T.U. München, 1999.

[19] H.P. Young, "Cost allocation", chapter 34 of "Handbook of Game Theory, Volume 2", R.J. Aumann and S. Hart, editors, Elsevier Science, 1994.

Appendices

A Allocation with an S.E.C.-like risk measure

In this section, we provide some examples of the coherent allocation principles discussed above, i.e. the Shapley value in the usual game-theoretic case, and the Aumann-Shapley value in the scalable players framework.

The risk measure we choose to use is derived from the Securities and Exchange Commission (SEC) rules for margin requirements, as described in the National Association of Securities Dealers (NASD) document [8]. These rules are used by stock exchanges to establish the margins required of their members, as guarantee against the risk that the members' portfolios involve (the Chicago Board of Options Exchange is one such exchange). The rules themselves are not constructive, in that they do not specify how the margin should be computed; in fact, this computation is left to each member of the exchange, who must then find the smallest margin complying with the rules. However, Rudd and Schroeder [13] proved in 1982 that a linear optimization problem (L.P.) modelled the rules adequately, and was sufficient to establish the minimum margin of a portfolio, that is, to evaluate its *risk measure*. It is worth mentioning that given this L.P.-based risk measure, the corresponding allocation problem has been called *linear production game* by Owen [10], see also [5].

For the purpose of the article, we restrict the risk measure to (simplistic) portfolios of calls on the same underlying stock, and all with the same expiry date. This restriction of the SEC rules was used first in [1] as an example of a *non-coherent* risk measure. In the case of a portfolio of calls, the margin is calculated through a representation of the calls by a set of *spread options*, each of which carrying a fixed margin. To obtain a coherent measure of risk, we prove later that it is sufficient to represent the calls by a set of spreads *and* butterfly options.

A.1 Coherent, S.E.C.-like margin calculation

We consider a portfolio consisting of C_P calls at strike price P, where P belongs to a set of strike prices $\mathcal{P} = \{P_{\min}, P_{\min} + 10, \ldots, P_{\max} - 10, P_{\max}\}$. This assumption about the format of the strike prices set $\mathcal{P}$ simplifies the notation. For convenience, we denote the set $\mathcal{P} \setminus \{P_{\min}, P_{\max}\}$ by $\mathcal{P}^-$ and the set $\mathcal{P} \setminus \{P_{\min}, P_{\min} + 10, P_{\max} - 10, P_{\max}\}$ by $\mathcal{P}^{--}$. We also make the simplifying assumption that there are as many long calls as short calls in the portfolio, i.e. $\sum_{P \in \mathcal{P}} C_P = 0$.

We will denote by $\boldsymbol{C_P}$ (in bold) the vector of the C_P parameters, $P \in \mathcal{P}$. While $\boldsymbol{C_P}$ fully describes the *portfolio*, it certainly does not describe the *future value of the portfolio*, which depends on the price of the underlying stock at a future date. We may nevertheless write $\rho(\boldsymbol{C_P})$ for the good reason that the ρ considered here depends by definition *only* on $\boldsymbol{C_P}$, as will be clear below.

We can now define our S.E.C.-like margin requirement. To evaluate the margin (or risk measure) ρ of the portfolio $\boldsymbol{C_P}$, we must replicate its calls with *spreads* and *butterflies*, defined as follows:

Variable	Instrument	Calls equivalent
$S_{H,K}$	Spread, long in H, short in K	One long call at price H, one short call at strike K
B_H^{long}	Long butterfly, centered at H	One long call at $H - 10$, two short calls at H, one long call at $H + 10$
B_H^{short}	Short butterfly, centered at H	One short call at $H - 10$, two long calls at H, one short call at $H + 10$

The variables shall represent the number of each specific instrument. All H and K are understood to be in $\mathcal{P}$, or $\mathcal{P}^-$ for the butterflies; $H \neq K$ for the spreads.

As in the S.E.C. rules, fixed margins are attributed to the instruments used for the replicating portfolio, i.e. spreads and butterflies in our case. Spreads carry a margin of $(H - K)^+ = \max(0, H - K)$; short butterflies are given a margin of 10, while long butterflies require no margin. In simple language, each instrument requires a margin equal to the worst potential loss, or negative payoff, it could yield.

By definition, the margin of a portfolio of spreads and butterflies is the sum of the margins of its components.

On the basis of [13], the margin of the portfolio $\boldsymbol{C_P}$ can then be evaluated with the linear optimization problem (SEC-LP):

$$
\begin{aligned}
&\text{minimize} \qquad \sum_{H,K\in\mathcal{P}} (H-K)^+ \, S_{H,K} \;+\; \sum_{H\in\mathcal{P}^-} 10\, B_H^{\text{short}} \\
&\text{subject to} \hspace{20em} \text{(SEC-LP)} \\
&\sum_{K\in\mathcal{P}} S_{P,K} - B_{P-1}^{\text{short}} + 2B_P^{\text{short}} - B_{P+1}^{\text{short}} - \\
&\qquad \sum_{K\in\mathcal{P}} S_{K,P} + B_{P-1}^{\text{long}} - 2B_P^{\text{long}} + B_{P+1}^{\text{long}} = C_P, \quad \forall\, P\in\mathcal{P}^{--} \\
&\sum_{K\in\mathcal{P}} S_{P_{\min},K} - \sum_{K\in\mathcal{P}} S_{K,P_{\min}} - B_{P_{\min}+10}^{\text{short}} + B_{P_{\min}+10}^{\text{long}} = C_{P_{\min}}, \\
&\sum_{K\in\mathcal{P}} S_{P_{\min}+10,K} + 2B_{P_{\min}+10}^{\text{short}} - B_{P_{\min}+20}^{\text{short}} - \\
&\qquad \sum_{K\in\mathcal{P}} S_{K,P_{\min}+10} - 2B_{P_{\min}+10}^{\text{long}} + B_{P_{\min}+20}^{\text{long}} = C_{P_{\min}+10}, \\
&\sum_{K\in\mathcal{P}} S_{P_{\max}-10,K} + 2B_{P_{\max}-10}^{\text{short}} - B_{P_{\max}-20}^{\text{short}} - \\
&\qquad \sum_{K\in\mathcal{P}} S_{K,P_{\max}-10} - 2B_{P_{\max}-10}^{\text{long}} + B_{P_{\max}-20}^{\text{long}} = C_{P_{\max}-10}, \\
&\sum_{K\in\mathcal{P}} S_{P_{\max},K} - \sum_{K\in\mathcal{P}} S_{K,P_{\max}} - B_{P_{\max}-10}^{\text{short}} + B_{P_{\max}-10}^{\text{long}} = C_{P_{\max}}, \\
&\qquad S_{H,K} \ge 0, \quad \forall\, H,K\in\mathcal{P}, H\neq K \\
&\qquad B_H^{\text{long}} \ge 0, \quad \forall\, H\in\mathcal{P}^- \\
&\qquad B_H^{\text{short}} \ge 0, \quad \forall\, H\in\mathcal{P}^-
\end{aligned}
$$

The objective function represent the margin; the equality constraints ensure that the portfolio is exactly replicated.

The margin, or risk measure thus defined is coherent; the proof of this is delayed until section A.4.

Note that despite the unpleasant appearance of the above problem, linear optimization is a well-researched topic, where “large” problems can typically be solved efficiently on laptop computers with standard software.

A.2 Computation of the allocations

Given this risk measure as a linear optimization problem, the Shapley value is easy to compute when the "total portfolio" is divided in a small number of subsets (sub-portfolios). First, the margin of every possible coalition of sub-portfolios is calculated; this step could prove cumbersome if a large number of sub-portfolios was involved. Then, the constructive definition of the Shapley value is used: the margin allocated to the i^{th} sub-portfolio (out of, say, n), is:

$$K_i^{Sh} = \sum_{S \in \mathcal{C}_i} \frac{(s-1)!(n-s)!}{n!} \left(\rho \left(\sum_{i \in S} X_i \right) - \rho \left(\sum_{i \in S/\{i\}} X_i \right) \right),$$

where $\mathcal{C}_i$ represent all coalitions that contain i and $s = |S|$.

The computation of the Aumann-Shapley value is even simpler (for the moment, we make abstraction of the differentiability condition of theorem 5). Note that by working in the scalable player (i.e. scalable sub-portfolio) framework, we implicitely assume that fractions of portfolios are sensible intruments. Recall now that the *per unit* margin allocated to the i^{th} sub-portfolio is

$$k_i^{AS} = \frac{\partial r(\Lambda)}{\partial \lambda_i} \tag{12}$$

where $r(\lambda)$ is the margin required of the sum of all the sub-portfolios i, each scaled by a fraction λ_i/Λ_i, and where "evaluated at Λ" means for the total portfolio; for this example, we simply set Λ to the vector of ones, denoted by e. In vector notation, we write $k^{AS} = \nabla r(\Lambda)$.

In words, we need, for a sub-portfolio i, the rate of change of the margin, when the "presence" of this sub-portfolio varies; this rate of change function is to be evaluated at the point where the total portfolio is present.

Now, the *dual* solution δ obtained automatically when computing the margin of the total portfolio, provides the rates of change of the margin, when the presence of *each specific call* varies. But the amount of calls is a linear function of the amount of sub-portfolios; if there are $|\mathcal{P}|$ different calls (equivalently here, $|\mathcal{P}|$ strike prices), and n sub-portfolios, then an $n \times |\mathcal{P}|$ matrix L maps numbers of calls into sub-portfolios, so that $k^{AS} = L\,\delta$, and

$$K^{AS} = \Lambda .* k^{AS} = e .* L\delta = L\,\delta$$

To put it in one sentence, the Aumann-Shapley allocation is only a matrix product away from the lone evaluation of the margin for the total portfolio.

The question of differentiability remains to be taken care of, and this can be done by considering the paper of Billera and Raanan [5]. (...)

A.3 Numerical examples of coherent allocation

Using the (coherent) risk measure presented above, we can obtain a somewhat more practical feeling of coherent allocation methods, by looking at examples.

For all allocation examples below, the reference "total" portfolio is the same; its values of $C_P, P \in \{10, 20, 30, 40, 50\}$ are:

	C_{10}	C_{20}	C_{30}	C_{40}	C_{50}
Total	-1	-2	8	-7	2

meaning one short call at strike 10, two short calls at strike 20, eight long calls at strike 30, etc. It carries a risk of 40, i.e. $\rho(\boldsymbol{C_{P_1}} + \boldsymbol{C_{P_2}} + \boldsymbol{C_{P_3}}) = 40$.

Consider first the division in three parts of the total portfolio:

	C_{10}	C_{20}	C_{30}	C_{40}	C_{50}
Portfolio 1	-1	0	6	-6	1
Portfolio 2	0	-2	2	0	0
Portfolio 3	0	0	0	-1	1
Total	-1	-2	8	-7	2

Coalitions of portfolios incur margins as follows:

$$\rho(\boldsymbol{C_{P_1}} + \boldsymbol{C_{P_2}}) = 40 \qquad \rho(\boldsymbol{C_{P_1}} + \boldsymbol{C_{P_3}}) = 20 \qquad \rho(\boldsymbol{C_{P_2}} + \boldsymbol{C_{P_3}}) = 30$$
$$\rho(\boldsymbol{C_{P_1}}) = 20 \qquad \rho(\boldsymbol{C_{P_2}}) = 20 \qquad \rho(\boldsymbol{C_{P_3}}) = 10$$

while the Shapley allocation is

$$K_1^{Sh} = 15 \qquad K_2^{Sh} = 20 \qquad K_3^{Sh} = 5$$

and the Aumann-Shapley allocation is

$$K_1^{AS} = 20 \qquad K_2^{AS} = 20 \qquad K_3^{AS} = 0$$

Here, the Shapley value is in the core, and also (strictly) positive.

Consider a second example

	C_{10}	C_{20}	C_{30}	C_{40}	C_{50}
Portfolio 1	-1	0	2	-2	1
Portfolio 2	0	-1	6	-5	0
Portfolio 3	0	-1	0	0	1
Total	-1	-2	8	-7	2

Coalitions of portfolios incur margins as follows:

$$\rho(\boldsymbol{C_{P_1}} + \boldsymbol{C_{P_2}}) = 30 \qquad \rho(\boldsymbol{C_{P_1}} + \boldsymbol{C_{P_3}}) = 50 \qquad \rho(\boldsymbol{C_{P_2}} + \boldsymbol{C_{P_3}}) = 20$$

$$\rho(\boldsymbol{C_{P_1}}) = 20 \qquad \rho(\boldsymbol{C_{P_2}}) = 10 \qquad \rho(\boldsymbol{C_{P_3}}) = 30$$

while the Shapley allocation is

$$K_1^{Sh} = 20 \qquad K_2^{Sh} = 0 \qquad K_3^{Sh} = 20$$

and the Aumann-Shapley allocation is

$$K_1^{AS} = 20 \qquad K_2^{AS} = 10 \qquad K_3^{AS} = 10$$

Again, the Shapley allocation is in the core. This time, a player is allocated a capital of 0.

A third example is

	C_{10}	C_{20}	C_{30}	C_{40}	C_{50}
Portfolio 1	-1	-1	4	-2	0
Portfolio 2	0	-1	4	-3	0
Portfolio 3	0	0	0	-2	2
Total	-1	-2	8	-7	2

Coalitions of portfolios incur margins as follows:

$$\rho(\boldsymbol{C_{P_1}} + \boldsymbol{C_{P_2}}) = 40 \qquad \rho(\boldsymbol{C_{P_1}} + \boldsymbol{C_{P_3}}) = 30 \qquad \rho(\boldsymbol{C_{P_2}} + \boldsymbol{C_{P_3}}) = 10$$

$$\rho(\boldsymbol{C_{P_1}}) = 30 \qquad \rho(\boldsymbol{C_{P_2}}) = 10 \qquad \rho(\boldsymbol{C_{P_3}}) = 20$$

while the Shapley allocation is

$$K_1^{Sh} = 26.66 \qquad K_2^{Sh} = 6.66 \qquad K_3^{Sh} = 6.66$$

and the Aumann-Shapley allocation is

$$K_1^{AS} = 30 \qquad K_2^{AS} = 10 \qquad K_3^{AS} = 0$$

This time, the Shapley value, while strictly positive, is *not* in the core, i.e. it allows undercut.

A.4 Proof of the coherence of the measure

We prove here that the risk measure ρ obtained through (SEC-LP) is *coherent*, as defined in [1]: a risk measure is coherent if it satisfies the four properties subadditivity, degree one homogeneity, translation invariance, and monotonicity. We prove each in turn, below.

1) Subadditivity: For any two portfolios $\boldsymbol{C_P^*}$ and $\boldsymbol{C_P^{**}}$,

$$\rho\left(\boldsymbol{C_P^*} + \boldsymbol{C_P^{**}}\right) \;\leq\; \rho(\boldsymbol{C_P^*}) + \rho(\boldsymbol{C_P^{**}})$$

Proof: If solving (SEC-LP) with $\boldsymbol{C_P^*}$ as right-hand side of the equality constraints yields a solution $\boldsymbol{S}^*$, and solving with $\boldsymbol{C_P^{**}}$ yields a solution $\boldsymbol{S}^{**}$, then $\boldsymbol{S}^* + \boldsymbol{S}^{**}$ is a feasible solution for the (SEC-LP) with $\boldsymbol{C_P^*} + \boldsymbol{C_P^{**}}$ as right-hand side. Subadditivity follows directly, given the linearity of the objective function.

2) Degree one homogeneity: For any $\lambda \geq 0$ and any portfolio $\boldsymbol{C_P}$,

$$\rho(\lambda \boldsymbol{C_P}) = \lambda \rho(\boldsymbol{C_P})$$

Proof: This is again a direct consequence of the linear optimization nature of ρ, as $\lambda \boldsymbol{S}$ is a solution of the (SEC-LP) with $\lambda \boldsymbol{C_P}$ as right-hand side of the constraints, provided $\boldsymbol{S}$ is a solution of the (SEC-LP) with $\boldsymbol{C_P}$ as right-hand side. Note that not only the proof, but the very definition of homogeneity requires that we allow fractions of calls to be sold and bought.

3) Translation invariance: For any portfolio $\boldsymbol{C_P}$ and all $\alpha \in \mathbb{R}$,

$$\rho(\boldsymbol{C_P} + \alpha r_f) \;=\; \rho(\boldsymbol{C_P}) - \alpha \tag{13}$$

where r_f is the rate of return on a reference, riskless investment.

Proof: There is little prove here; we rather need to define the behaviour of ρ in the presence of cash (represented by the odd expression $\boldsymbol{C_P} + \alpha r_f$), and naturally choose (13) to do so.

4) Monotonicity: For any two portfolios $\boldsymbol{C_P^*}$ and $\boldsymbol{C_P^{**}}$ such that the future worth of $\boldsymbol{C_P^*}$ is always less than or equal to that of $\boldsymbol{C_P^*}$,

$$\rho(\boldsymbol{C_P^*}) \;\geq\; \rho(\boldsymbol{C_P^{**}})$$

Before proving monotonicity, let us first introduce the values V_P, for $P \in \{P_{\min}+10, \ldots, P_{\max}, P_{\max}+10\}$, which represents the future payoff, or worth, of the portfolio for the (future) prices P of the underlying. (Again, we write $\boldsymbol{V_P}$ to denote the vector of all V_P's) The components of $\boldsymbol{V_P}$ are completely determined by the number of calls in the portfolio:

$$V_P \;=\; \sum_{p=P_{\min}}^{P-10} C_p(P-p) \qquad \forall\, P \in \{P_{\min}+10, \ldots, P_{\max}, P_{\max}+10\}$$

which is alternatively written $\boldsymbol{V_P} = M\boldsymbol{C_P}$, with the square, invertible matrix M:

$$M \;=\; \begin{bmatrix} 10 & 0 & 0 & \cdots \\ 20 & 10 & 0 & \cdots \\ 30 & 20 & 10 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

The antecedent of the monotonicity property is, of course, the componentwise $\boldsymbol{V_P^*} \leq \boldsymbol{V_P^{**}}$.

We will also use the following lemma:

Lemma 4 *Under subadditivity, two equivalent formulations of monotonicity are, for any three portfolios* $\boldsymbol{C_P}$, $\boldsymbol{C_P^*}$ *and* $\boldsymbol{C_P^{**}}$*:*

$$\boldsymbol{V_P^*} \;\leq\; \boldsymbol{V_P^{**}} \;\Longrightarrow\; \rho\,(M^{-1}\boldsymbol{V_P^*}) \;\geq\; \rho\,(M^{-1}\boldsymbol{V_P^{**}})$$

and

$$0 \;\leq\; \boldsymbol{V_P} \;\Longrightarrow\; \rho\,(M^{-1}\boldsymbol{V_P}) \;\leq\; 0$$

Proof: The upper condition is sufficient, as it implies

$$\rho(0) \ \geq \ \rho(M^{-1}\boldsymbol{V_P}),$$

and $\rho(0) = 0$ from the very structure of (SEC-LP). The upper condition is necessary, as

$$\begin{aligned}\rho\,(M^{-1}\boldsymbol{V_P^{**}}) &= \rho\,\left(M^{-1}(\boldsymbol{V_P^{*}} + (\boldsymbol{V_P^{**}} - \boldsymbol{V_P^{*}})\right) \\ &\leq \rho\,(M^{-1}\boldsymbol{V_P^{*}}) + \rho\,\left(M^{-1}(\boldsymbol{V_P^{**}} - \boldsymbol{V_P^{*}})\right) \\ &\leq \rho\,(M^{-1}\boldsymbol{V_P^{*}}).\end{aligned}$$

□

Proof of monotonicity: As a consequence of the above lemma, it is sufficient to prove that if a portfolio of calls *always* has nonnegative future payoff, then its associated margin is nonpositive.

A close look at (SEC-LP) shows that the margin assigned to the portfolio will be nonpositive (in fact, zero), if and only if a nonnegative, feasible solution of (SEC-LP) exists in which all spreads variables $S_{H,K}$ with $H > K$ and all short butterflies variables B^{short} have value zeRO. This means that there exists a nonnegative solution to:

$$\forall\, P \in \mathcal{P}^{--},\quad \sum_{\substack{K\in\mathcal{P}\\ P<K}} S_{P,K} - \sum_{\substack{K\in\mathcal{P}\\ K<P}} S_{K,P} + B^{\text{long}}_{P-1} - 2B^{\text{long}}_{P} + B^{\text{long}}_{P+1} = \ C_P,$$

$$\sum_{\substack{K\in\mathcal{P}\\ P_{\min}<K}} S_{P_{\min},K} - \sum_{\substack{K\in\mathcal{P}\\ K<P_{\min}}} S_{K,P_{\min}} + B^{\text{long}}_{P_{\min}+10} = C_{P_{\min}},$$

$$\sum_{\substack{K\in\mathcal{P}\\ P_{\min}+10<K}} S_{P_{\min}+10,K} - \sum_{\substack{K\in\mathcal{P}\\ K<P_{\min}+10}} S_{K,P_{\min}+10} - 2B^{\text{long}}_{P_{\min}+10} + B^{\text{long}}_{P_{\min}+20} = \ C_{P_{\min}+10},$$

$$\sum_{\substack{K\in\mathcal{P}\\ P_{\max}-10<K}} S_{P_{\max}-10,K} - \sum_{\substack{K\in\mathcal{P}\\ K<P_{\max}-10}} S_{K,P_{\max}-10} - 2B^{\text{long}}_{P_{\max}-10} + B^{\text{long}}_{P_{\max}-20} = \ C_{P_{\max}-10},$$

$$\sum_{\substack{K\in\mathcal{P}\\ P_{\max}<K}} S_{P_{\max},K} - \sum_{\substack{K\in\mathcal{P}\\ K<P_{\max}}} S_{K,P_{\max}} + B^{\text{long}}_{P_{\max}-10} = C_{P_{\max}},$$

$$\forall\, H, K \in \mathcal{P},\ H < K,\quad S_{H,K} \geq 0,$$

$$\forall\, H \in \mathcal{P}^{-},\quad B^{\text{long}}_{H} \geq 0,$$

Perhaps more clearly, this is a linear system $Ax = \boldsymbol{C_P},\ \ x \geq 0$ where A has the form

$$A \ = \ \begin{bmatrix} 1 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 & -2 & 1 & \cdots & 0 \\ 0 & -1 & \cdots & 0 & 1 & -2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & -2 \\ 0 & 0 & \cdots & -1 & 0 & 0 & \cdots & 1 \end{bmatrix},$$

and x is the appropriate vector of spreads and butterflies variables. We obtain a new, equivalent system of equations $MAx = M\boldsymbol{C_P} = \boldsymbol{V_P}$ by pre-multiplying by the invertible matrix M introduced above. Recall now that we have made the assumption that the portfolio contains as many short calls as long calls, i.e. $\sum_{P \in \mathcal{P}} C_P = 0$. Thus, we need only prove that there exists a nonnegative solution to the system

$$MA\,x = \boldsymbol{V_P} \quad \text{whenever} \quad \boldsymbol{V_P} \geq 0 \quad \text{and} \quad e^t M^{-1} \boldsymbol{V_P} = 0$$

(e^t is a row vector of 1's). A simple observation of MA shows that its columns span the same subspace as the set of columns

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}, \cdots \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

Observing furthermore that

$$e^t M^{-1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

so that any $\boldsymbol{V_P}$ satisfying $e^t M^{-1} \boldsymbol{V_P} = 0$ has identical last two components, the right-hand side of $MA\,x = \boldsymbol{V_P}$ can always be expressed as a nonnegative linear combination of the columns of MA.

□

B Proof of the nonnegativity condition

The following result was given in section 6 and is proved here.

Theorem 6 *A sufficient condition for a nonnegative, "no undercut", full allocation to exist is:*

$$\forall\, \lambda \in \mathbb{R}_+^N, \quad \rho\left(\sum_{i\in N} X_i\right) \min_{i\in N}\{\lambda_i\} \;\leq\; \rho\left(\sum_{i\in N} \lambda_i X_i\right)$$

Proof: Let us first define the "coalitions vectors" $e^M \in \mathbb{R}^n$ as

$$e_i^M \;=\; \begin{cases} 1 & i \in M \\ 0 & i \notin M \end{cases} \qquad \forall M \subset N$$

A nonnegative, "no undercut", full allocation K exists when

$$\begin{aligned} &\exists\, K \in \mathbb{R}^N \text{such that} \\ &\qquad K^t e^M \leq \rho\left(\sum_{i\in M} X_i\right) \qquad \forall\, M \subsetneq N \\ &\qquad K^t e^N = \rho\left(\sum_{i\in N} X_i\right) \\ &\qquad K \geq 0 \end{aligned} \tag{14}$$

Using Farkas's lemma, this is equivalent to

$$\begin{aligned} &\sum_{M\subsetneq N} e^M y_M + e^N y_N \geq\, 0, \quad \forall\, y_M \in \mathbb{R}_+,\, \forall\, M \subsetneq N \\ &\qquad \Longrightarrow \quad \sum_{M\subsetneq N} \rho\left(\sum_{i\in M} X_i\right) y_M + \rho\left(\sum_{i\in N} X_i\right) y_N \geq\, 0 \end{aligned} \tag{15}$$

which in turn is equivalent to

$$\begin{aligned} &y_N \geq -\sum_{M\ni i} y_M, \quad \forall\, y_M \geq 0,\, \forall\, M \subsetneq N,\, \forall\, i \in N, \\ &\qquad \Longrightarrow \quad \sum_{M\subsetneq N} \rho(\sum_{i\in M} X_i) y_M \geq -\rho\left(\sum_{i\in N} X_i\right) y_N \end{aligned} \tag{16}$$

Now, using the homogeneity and the subadditivity of ρ,

$$\begin{aligned}
\sum_{M \subsetneq N} \rho\left(\sum_{i \in M} X_i\right) y_M &= \sum_{M \subsetneq N} \rho\left(y_M \sum_{i \in M} X_i\right) \\
&\geq \rho\left(\sum_{M \subsetneq N}\left(y_M \sum_{i \in M} X_i\right)\right) \\
&= \rho\left(\sum_{i \in N}\left(X_i \sum_{M \ni i} y_M\right)\right)
\end{aligned}$$

Therefore, a sufficient condition for (14) (or (15) or (16)) to hold, is

$$y_N \geq -\sum_{M \ni i} y_M, \quad \forall\, y_M \geq 0,\ \forall\, M \subsetneq N,\ \forall\, i \in N,$$
$$\Longrightarrow \quad \rho\left(\sum_{i \in N}\left(X_i \sum_{M \ni i} y_M\right)\right) \geq \rho\left(\sum_{i \in N} X_i\right)(-y_N)$$

Finally, using the definition $\lambda_i \triangleq \sum_{M \ni i} y_M$, we can write the *sufficient* condition for (14)

$$\forall\, \lambda_i \geq 0,\ i \in N, \quad \rho\left(\sum_{i \in N} \lambda_i X_i\right) \geq \rho\left(\sum_{i \in N} X_i\right)\left(\min_{i \in N} \lambda_i\right)$$

□

Note that in the last step, we also used $\rho\left(\sum_{i \in N} X_i\right) \geq 0$, a *necessary* condition for the existence of a nonnegative, "no undercut", full allocation; checking $y_N = 1,\ y_M = 0\ \ \forall\, M \subsetneq N$ in (15) shows this point.