Non-Negative Risk Components

Jeremy Staum
j-staum@northwestern.edu
Department of Industrial Engineering and Management Sciences
Robert R. McCormick School of Engineering and Applied Science
Northwestern University
Evanston, IL 60208-3119, U.S.A.

January 8, 2015

Abstract

We propose two new methods for attributing the risk of a portfolio or system to its components, when it is required to produce non-negative risk components that sum to the risk of the portfolio or system as a whole. One method attributes risk entirely to losses, taking profits for granted. The other method does allow profits in some scenarios to offset losses in other scenarios to some extent, but it not in a way that could yield a negative risk component. The methods are illustrated by applying them to an example of attribution a firm's expected shortfall to business units within the firm and an example of attributing systemic risk to banks. We prove that, under appropriate conditions, the methods proposed have some game-theoretic properties that are desirable for risk attribution.

Key words: cost allocation, cost attribution, non-negative, risk allocation, risk attribution, risk components, systemic risk

1 Introduction

Risk attribution is used for multiple purposes, and the appropriate risk attribution method depends on the purpose. For example, one purpose is to measure the performance of a firm's business lines by return on economic capital. The economic capital required for the firm is based on a risk measure for the firm, and the economic capital for each business line is found by attributing the firm's risk to the business lines (Tasche, 2004, 2008). In a risk audit, the purpose is to find where risk for the firm is generated, so as to follow up with individual business lines about possible adjustments of risk exposures. Similarly, in managing systemic risk, one purpose is to identify firms that are important sources of systemic risk (Drehmann and Tarashev, 2013; Liu and Staum, 2011). Another purpose is to levy systemic risk charges in proportion to each firm's contribution to the cost of insuring the system (Staum, 2012). We focus on risk attribution that yields risk components, which have the property that the system's risk equals the sum of the risk components attributed to the components of the system. This is important in some applications: for example, the sum of the economic capital of the firm's business lines should equal the firm's economic capital, determined by the firm's risk, and the sum of systemic risk charges should equal the cost of insuring the system.

Some risk components may be negative. For example, a business line may decrease the firm's risk if the effect of its profits in reducing the firm's loss in some scenarios outweighs the effect of its losses. Depending on the purpose of risk attribution, it may or may not be acceptable to get negative risk components, or for profits to mask losses in this way at all. A negative systemic risk component can be informative and interesting, because it says that a particular firm does more to reduce systemic risk than to increase it. Tasche (2004) argues that negative economic capital is suitable in performance measurement; small negative mean returns could be a good performance from a business unit whose function is hedging. On the other hand, negative systemic risk charges could be impractical (Staum, 2012) or even dangerous because they create the possibility for extracting profits from the systemic risk management process. When the goal is to identify the sources of risk, the effect of losses should not be masked by the effect of profits in other scenarios. That would hinder the effort to find the firms in the system or business units within the firm that are responsible for large contributions to risk.

This article proposes two methods of generating non-negative risk components using tools of cooperative game theory. The methods are developed using the Shapley and Aumann-Shapley values, which are core concepts for portfolio risk allocation (Denault, 2001; Tasche, 2008). This approach to portfolio risk allocation, grounded in cooperative game theory, uses a risk function that specifies how the participation of the system's components creates risk. This article uses this formalism for systems that are more complicated than portfolios, but the formalism is presented here in the introduction for portfolios to make possible a simple exposition of the methods proposed in this article. The vector $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_n]$ contains the participation levels of the n components in a portfolio. That is, λ_i is a number between 0 and 1 that is multiplied by the weight of component i in the actual portfolio to generate the weight of component i in a new portfolio. Let X_{ij} be the loss of portfolio component i in scenario j. In scenario j, the loss of the new portfolio defined by participation $\boldsymbol{\lambda}$ is $L_j(\boldsymbol{\lambda}) = \sum_{i=1}^n \lambda_i X_{ij}$. The risk function r is given by $r(\boldsymbol{\lambda}) = \rho(\boldsymbol{L}(\boldsymbol{\lambda}))$ where ρ is a risk measure and the random variable $\boldsymbol{L}$ gives the new portfolio's loss in each scenario.

The usual method of producing non-negative risk components is to replace negative risk components by zero and then re-scale the risk components so that they once again sum to the system's risk $\rho(\boldsymbol{L})$. That is, if the original risk components are $\phi_1, \ldots, \phi_n$, then the usual non-negative method replaces ϕ_i by $\rho(\boldsymbol{L}) \max\{0, \phi_i\} / \sum_{j=1}^n \max\{0, \phi_j\}$. The usual non-negative method allows profits in some scenarios to offset losses in other scenarios to some extent. This way of producing non-negative risk components may be useful for applications like systemic risk charges, as in Example 2.

In a similar spirit, we propose the "overall non-negative" method of producing non-negative risk components. To motivate the development of the overall non-negative method, consider risk components that are based on Shapley or Aumann-Shapley values, as is standard. The usual non-negative method applies the Shapley or Aumann-Shapley value to the risk function r, gets risk components that may lack the desired property of non-negativity, and then imposes the desired property afterwards. The overall non-negative method instead modifies the risk function so that the application of the Shapley or Aumann-Shapley value is guaranteed to yield non-negative risk components. Specifically, it uses the greatest non-decreasing function that is dominated by the positive part of r. In a sense, this is the function closest to r that is guaranteed to yield non-negative risk components. This construction gives the overall non-negative method good game-theoretic properties, discussed in Section 4. In particular, one of the advantages of the overall non-negative method over the usual non-negative method is that the former has a property of partial separability that the latter lacks. Table 2, containing the results of various methods applied to Example 2, shows that these two methods can differ significantly and illustrates the issue of partial separability. In this example, the system of three banks can be partially separated into its components: Bank 3 forms a subsystem that has no interactions with the subsystem containing Banks 1 and 2 in terms of the two subsystems' contributions to the probability of a systemic crisis. If the Shapley or Aumann-Shapley values are applied in the usual way, Bank 1 is assigned a negative risk component because it does more to reduce than to increase the probability of a systemic crisis, due to its ability to buffer losses generated by Bank 2 in one scenario. To give Bank 1 a non-negative risk component, one must change the risk components assigned to at least one of the other banks. It is appropriate to change the risk component assigned to Bank 2 and not that assigned to Bank 3 because only Bank 2 interacts with Bank 1. This is what the methods proposed in this article do, but the usual non-negative method reduces the risk component assigned to Bank 3.

Whereas both the usual and overall non-negative methods can allow profits in some scenarios to mask losses in other scenarios, at least to some extent, the "scenario-wise non-negative" method does not. The non-negative cost attributions in Staum (2012, §§ 4.2, 5.2) are a special case of the scenario-wise non-negative method. Here, we develop the scenario-wise non-negative method more fully and in more generality. Like the overall non-negative method, the scenario-wise non-negative method works by modifying the risk function r and gets desirable game-theoretic properties as a result of this construction. Instead of $r(\boldsymbol{\lambda}) = \rho(\boldsymbol{L}(\boldsymbol{\lambda}))$, the scenario-wise non-negative method looks at $\rho(\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda}))$ where, for a portfolio, the loss $\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda}) = \sum_{i=1}^{n} \lambda_i X_{ij}^{+} - \sum_{i=1}^{n} X_{ij}^{-}$, the sum of losses scaled by participation levels and unscaled profits. This amounts to saying that the participation of the system's components is responsible only for losses and not for profits. This makes the scenario-wise non-negative method very different from the usual and overall non-negative methods. It is suitable for different purposes. For example, it is useful for risk audits, as in Example 1.

The methods proposed here are also applicable to cost attribution (see, for example, Moulin and Sprumont, 2007). Cost attribution is fundamental to risk attribution, so we begin by establishing a common framework to describe cost attribution or risk attribution problems.

2 Attribution of Risk or Cost

We want to attribute the risk or cost of a system to the components of the system. In Section 2.2, we consider the Shapley and Aumann-Shapley values, which are prominent methods for attribution in cooperative game theory. These methods have proved to be useful even in cases when one does not regard the system's components as players of a cooperative game. For example, the Aumann-Shapley value is standard for portfolio risk attribution (Tasche, 2008), even though the portfolio arises by design of a portfolio manager, and not as the result of a cooperative game played by the investments in the portfolio. To use the Shapley or Aumann-Shapley values, we must express the risk or cost of a system as a function of the levels of participation of the components.

2.1 Risk or Cost

We need to model the way in which risk or cost depends on the participation of the system's components. Let us focus on risk, but also explain how the same formalism applies to cost. Our model is a function r that maps a participation vector $\boldsymbol{\lambda} \in [0,1]^n$ to the risk $r(\boldsymbol{\lambda})$ of the system in which the participation of component i is λ_i. For the standard methods of cooperative game theory to apply, we must have $r(\mathbf{0}) = 0$, i.e., risk is zero when none of the system's components participate, and that $r(\mathbf{1})$ equals the risk of the actual system.

An important case is when the system can be treated as a portfolio of components that have no interactions in the way their participation affects the systemic loss. This case includes portfolio risk attribution and multi-commodity cost attribution problems, as well as other problems such as systemic risk attribution when the financial system is treated as a portfolio of financial institutions. In this case, we can show in detail how participation affects risk by writing

$$r(\boldsymbol{\lambda}) = \rho(\boldsymbol{L}(\boldsymbol{\lambda})), \quad \boldsymbol{L}(\boldsymbol{\lambda}) = a(\mathbf{X}(\boldsymbol{\lambda})) \text{ where } a(\boldsymbol{X}) = \sum_{i=1}^{n} X_i, \quad \text{and} \quad X_{ij}(\boldsymbol{\lambda}) = \lambda_i X_{ij}, \tag{1}$$

where:

- X_{ij} represents the loss experienced by component i in scenario j. If this is negative, then $-X_{ij}$ represents a gain. In general, we model how the loss $X_{ij}(\boldsymbol{\lambda})$ of each component i in each scenario j depends on $\boldsymbol{\lambda}$, the vector that specifies participation by every component. In the case of a portfolio, the participation λ_i of component i affects only the losses of component i, by providing a linear scaling factor: $X_{ij}(\boldsymbol{\lambda}) = \lambda_i X_{ij}$. In particular, if component i does not participate in the portfolio, i.e., $\lambda_i = 0$, its loss is zero, whereas if it participates fully in the portfolio, i.e., $\lambda_i = 1$, its loss equals its actual loss, X_{ij}.

- The aggregation function a maps a vector $\boldsymbol{X}$ containing losses associated with each component to a scalar $a(\boldsymbol{X})$ describing the resulting loss for the system. The concept of the aggregation function is due to Chen et al. (2013). The notation $\boldsymbol{X}_{\cdot j}$ denotes the column n-vector containing the losses of each component in scenario j. By abuse of notation, $a(\mathbf{X})$ denotes the random variable such that the aggregation function is applied column-by-column, i.e., $(a(\mathbf{X}))_j = a(\boldsymbol{X}_{\cdot j})$ is the systemic loss in scenario j. This is how to interpret $\boldsymbol{L}(\boldsymbol{\lambda}) = a(\mathbf{X}(\boldsymbol{\lambda}))$. In the case of a portfolio, $a(\boldsymbol{X}) = \sum_{i=1}^{n} X_i$, the net loss of the portfolio's components.

- The risk measure ρ maps a random variable $\boldsymbol{L}$ describing the systemic loss to a scalar $r(\boldsymbol{L})$ that describes the risk of this random variable. The notation L_j denotes the systemic loss in scenario j.

For simplicity of exposition, we assume that the random variables are defined on finite probability spaces, but this is not essential to the underlying ideas.

Equation (1) has been described in terms of risk, but it also applies to cost. Let us consider how it applies to a multi-commodity cost attribution problem (see, for example, Koster et al., 1998). The index j of the components of the vector $\boldsymbol{L}$ or the columns of the matrix $\mathbf{X}$ is now interpreted as referring to a commodity instead of a scenario. (Indeed, the Arrow-Debreu security paying \$1 in scenario j can be thought of as one commodity in a financial market.) Interpret X_{ij} as the demand for commodity j by component i; if it is negative, then $-X_{ij}$ represents supply of commodity j by component i. The aggregation function a produces a systemic demand $L_j = a(\boldsymbol{X}_{\cdot j})$ for commodity j. Finally, r is a cost function that maps the systemic demand vector $\boldsymbol{L}$ to a total cost to satisfy that systemic demand.

The single-commodity cost attribution problem is very simple and fundamental; we will return to it in the later exposition. Because there is a single commodity, in Equation (1), the matrix of individual demands $\mathbf{X}$ becomes a vector, and the vector of systemic demands $\boldsymbol{L}$ becomes a scalar.

The systemic cost as a function of participation is $r(\boldsymbol{\lambda}) = \rho(\sum_{i=1}^{n} \lambda_i X_i)$, the cost of satisfying net demand.

Equation (1) is not always applicable. For example, in Liu and Staum (2011), the systemic loss $\boldsymbol{L}$ represents the losses that the financial system imposes on its creditors. In their model, the participation $\boldsymbol{\lambda}$ affects the systemic loss $\boldsymbol{L}(\boldsymbol{\lambda})$ not only through the losses $\mathbf{X}(\boldsymbol{\lambda})$ associated with the firms that are the components of the financial system, but also by affecting the extent to which the firms transmit losses to their creditors. The model cannot be written with $\boldsymbol{L}(\boldsymbol{\lambda}) = a(\mathbf{X}(\boldsymbol{\lambda}))$ and would have to be written as $\boldsymbol{L}(\boldsymbol{\lambda}) = a(\mathbf{X}(\boldsymbol{\lambda}), \boldsymbol{\lambda})$. Also, the aggregation function need not simply add up the individual components' losses or demands. For example, in a multi-commodity cost attribution problem, if one must order an integer number of units of commodity j, then $a(\boldsymbol{X}) = \lceil \sum_{i=1}^{n} X_{ij} \rceil$. Chen et al. (2013) explore several other aggregation functions.

Equation (1) is inessential; as long as the function r is established, one of the methods described in this paper can be applied. However, using Equation (1), which is applicable in many important cases, enables development of multiple methods and provides greater clarity. We consider two risk attribution examples that fit into Equation (1). Because they do, it is only necessary to specify the matrix $\mathbf{X}$ of components' losses in each scenario and the risk measure ρ. The risk components generated by the several methods discussed in this article for the two examples are given in Tables 1 and 2 below.

Example 1 (Risk Audit: Expected Shortfall)**.** *Consider a firm whose profit or loss arises from the profits and losses of multiple business units. The firm's risk is measured as expected shortfall (see, e.g., McNeil et al., 2005, § 2.1) at the* α *level,*

$$\rho(\boldsymbol{L}) = \frac{1}{1-\alpha} \left(\sum_{j: L_j \geq q_\alpha} p_j L_j + q_\alpha \left(1 - \alpha - \sum_{j: L_j \geq q_\alpha} p_j \right) \right),$$

where $q_\alpha = \inf\{q : \sum_{j: L_j \leq q} p_j \geq \alpha\}$ *is the* α*-quantile of loss. The purpose of risk attribution is to say how much of the expected shortfall is generated by each of the business units.*

Consider a particular case in which there are $n = 3$ *units and* $m = 4$ *scenarios, whose probabilities are* $\begin{bmatrix} 90\% & 5\% & 4\% & 1\% \end{bmatrix}$*, such that*

$$\mathbf{X} = \begin{bmatrix} 0 & -180 & -90 & 90 \\ -90 & 0 & 90 & 90 \\ -90 & -90 & -90 & 270 \end{bmatrix} \quad \textit{and} \quad \boldsymbol{L} = \ell(\boldsymbol{X}) = \begin{bmatrix} -180 & -270 & -90 & 450 \end{bmatrix}.$$

The 95% expected shortfall is $18 = (-90 \times 4\% + 450 \times 1\%)/5\%$*.*

Example 2 (Systemic Risk Charges: Probability of a Systemic Crisis)**.** *This simple model of the probability of a systemic crisis treats the financial sector as a portfolio. When the system's loss is* L*, the conditional probability of a systemic crisis is* $\Phi(\beta_0 + \beta_1 L^+)$*, where* β_0 *and* β_1 *are parameters of the model and* Φ *is a monotone increasing, absolutely continuous function bounded above by 1. A net loss for the financial system increases the conditional probability of a systemic crisis, but a net gain (i.e., negative loss) does not decrease the conditional probability of a systemic crisis. The systemic risk is the excess of the unconditional probability of a systemic crisis for this system above the lowest possible such probability,* $\Phi(\beta_0)$*:* $\rho(\boldsymbol{L}) = \sum_{j \in j} p_j \Phi(\beta_0 + \beta_1 L_j^+) - \Phi(\beta_0)$*, where* p_j *is the probability of scenario* j*. It is necessary to normalize the systemic risk measure to get* $r(\mathbf{0}) = 0$

by subtracting $\Phi(\beta_0)$. The purpose of risk attribution is to assign systemic risk charges that are proportional to the risk components, assuming that the cost of insuring the system is proportional to the probability of a systemic crisis.

Consider a particular case in which Φ is the standard normal cumulative distribution function, $\beta_0 = -2.5$, and $\beta_1 = 0.5$. There are $m = 4$ scenarios, whose probabilities are $\begin{bmatrix} 93\% & 3\% & 3\% & 1\% \end{bmatrix}$. The losses of the system components, which are $n = 3$ banks, and the cost, which is the conditional crisis probability, are given by

$$\mathbf{X} = \begin{bmatrix} -1 & -2 & 0 & 1 \\ -1 & 5 & 0 & 4 \\ -1 & 0 & 4 & 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{L} = \ell(\mathbf{X}) = \begin{bmatrix} -3 & 3 & 4 & 5 \end{bmatrix}.$$

The baseline conditional crisis probability is $\Phi(\beta_0) = 0.62\%$ in scenario 1, in which the financial sector does not make a net loss; it is about 16% in scenario 2, 31% in scenario 3, and 50% in scenario 4. The systemic risk is $\rho(\boldsymbol{L}) = 1.86\%$, which is the difference between the unconditional crisis probability of 2.48% and the baseline conditional crisis probability.

2.2 Attribution

Two schemes for attributing risk are of particular interest: the Shapley value, which is common in systemic risk attribution (Drehmann and Tarashev, 2013), and the Aumann-Shapley value, which is standard in portfolio risk attribution (Tasche, 2008).

2.2.1 The Shapley Value

Let $\mathbf{e}_i$ be the vector whose ith component is 1 and whose other components are 0. Let $\boldsymbol{\lambda}(S) = \sum_{i \in S} \mathbf{e}_i$. The incremental risk of the participation of system component i, when those in the set S are already participating, is

$$\Delta_i r(\boldsymbol{\lambda}(S)) = r(\boldsymbol{\lambda}(S \cup \{i\})) - r(\boldsymbol{\lambda}(S)). \tag{2}$$

If formalism (1) applies, then the incremental risk is

$$\Delta_i r(\boldsymbol{\lambda}(S)) = \rho\left(X_i + \sum_{i' \in S} X_{i'}\right) - \rho\left(\sum_{i' \in S} X_{i'}\right), \tag{3}$$

the change in the system's risk when component i is added to a portfolio containing the components in the set S. The Shapley value attributes to participant i the risk

$$\frac{1}{n!} \sum_{S \not\ni i} |S|!(n - |S| - 1)! \Delta_i r(\boldsymbol{\lambda}(S)). \tag{4}$$

It is non-negative if the function r is non-decreasing on $\{0, 1\}^n$, which makes all incremental risks non-negative.

For the single-commodity cost attribution problem, the incremental cost in Equation (3) is non-negative if the cost function ρ is non-decreasing and $X_i \geq 0$. If ρ is increasing and $X_i < 0$, then the Shapley value attributes a negative cost to participant i for its negative demand, i.e., supply, which lowers the cost of the system.

2.2.2 The Aumann-Shapley Value

The Aumann-Shapley value attributes to component i the risk

$$\int_0^1 \frac{\partial r}{\partial \lambda_i}(\gamma \mathbf{1})\,\mathrm{d}\gamma, \tag{5}$$

which is non-negative if the function r is non-decreasing.[1] Suppose that Equation (1) applies, so that $r(\boldsymbol{\lambda}) = \rho(\sum_{i=1}^n \lambda_i \boldsymbol{X}_{i\cdot})$. If the risk measure ρ is suitably differentiable,

$$\int_0^1 \frac{\partial r}{\partial \lambda_i}(\gamma \mathbf{1})\,\mathrm{d}\gamma = \sum_{j\in\Omega} X_{ij} \int_0^1 \frac{\partial \rho}{\partial L_j}(\gamma \boldsymbol{L})\,\mathrm{d}\gamma \tag{6}$$

because the net loss of the portfolio $\boldsymbol{L} = \sum_{i=1}^n \boldsymbol{X}_{i\cdot}$, the sum of the components' losses. That is, the Aumann-Shapley value attributes to component i a risk component that is a weighted average of component i's losses in each scenario, weighted by an average sensitivity of risk to loss in that scenario, where the average is taken as the systemic loss is scaled down from $\boldsymbol{L}$ to 0.

For the single-commodity cost attribution problem,

$$\frac{\partial r}{\partial \lambda_i}(\boldsymbol{\lambda}) = \frac{\partial}{\partial \lambda_i}\rho\left(\sum_{i'=1}^n \lambda_{i'} X_{i'}\right) = X_i \rho'\left(\sum_{i'=1}^n \lambda_{i'} X_{i'}\right),$$

so the Aumann-Shapley value attributes to participant i the cost

$$X_i \int_0^1 \rho'\left(\gamma \sum_{i'=1}^n X_{i'}\right)\,\mathrm{d}\gamma = \frac{X_i}{\sum_{i'=1}^n X_{i'}}\rho\left(\sum_{i'=1}^n X_{i'}\right).$$

This is known as "average-cost pricing," because $\rho\left(\sum_{i'=1}^n X_{i'}\right) / \sum_{i'=1}^n X_{i'}$ is the average cost per unit of demand. It is non-negative if cost is non-negative and $X_i \geq 0$. If $X_i < 0$, the Aumann-Shapley cost attribution is negative as a reward for supply.

This example, leading to an interpretation of the Aumann-Shapley value as average-cost pricing, illustrates how explicit formulae for the Aumann-Shapley value can give insight about what it does. For this reason, we will emphasize explicit formulae, although they are not needed for computational purposes. It is possible to approximate Equation (5) or (6) by numerical integration and by replacing the partial derivatives with finite-difference approximations.

Example 1 In this example, in Equation (6) we have

$$\frac{\partial \rho}{\partial L_j}(\gamma \boldsymbol{L}) = 1\{L_j \geq q_\alpha\}\frac{p_j}{1-\alpha}$$

[1] Strictly speaking, Equation (5) applies if r is differentiable at almost every point on the "diagonal" $\{\gamma\mathbf{1} : \gamma \in [0,1]\}$ and the function that maps $\gamma \in [0,1]$ to $r(\gamma\mathbf{1})$ is absolutely continuous. The treatment of the case in which Equation (5) is unusable due to non-differentiability goes back to Mertens (1988). Boonen et al. (2012) and Haimanko (2001) provide a treatment of the issue for the specific applications of risk attribution and cost attribution, respectively. The same considerations apply to the Aumann-Shapley values derived later in this paper.

unless $q_\alpha = L_j = L_{j'}$ for some $j' \neq j$, in which case differentiability fails. The partial derivatives do not depend on γ because expected shortfall is positively homogeneous. The Aumann-Shapley value for participant i is

$$\frac{1}{1-\alpha} \sum_{j:L_j \geq q_\alpha} p_j X_{ij}.$$

The Shapley and Aumann-Shapley values allocate negative risk components to Units 1 and 3 (Table 1). The main explanation for this is that the profits (negative losses) of these units in scenario 3 outweigh their positive losses in scenario 4, and scenarios 3 and 4 are the tail scenarios for the real system.

Example 2 Applying Equation (6) to Example 2, we compute

$$\frac{\partial \rho}{\partial L_j}(\gamma \boldsymbol{L}) = 1\{L_j > 0\} p_j \beta_1 \Phi'(\beta_0 + \beta_1 \gamma L_j^+).$$

This yields an Aumann-Shapley value for component i of

$$\sum_{j:L_j>0} p_j X_{ij} \int_0^1 \beta_1 \Phi'(\beta_0 + \beta_1 \gamma L_j^+) \, \mathrm{d}\gamma = \sum_{j:L_j>0} p_j \frac{X_{ij}}{L_j^+} \Big(\Phi(\beta_0 + \beta_1 L_j^+) - \Phi(\beta_0) \Big).$$

In each scenario in which the portfolio has a net loss, which increases the crisis probability, each bank is attributed a fraction of the increase in crisis probability, proportional to the bank's contribution to the net loss. This contribution is positive for banks that generate losses but negative for those that generate profits, which can lead to negative risk components. The Shapley and Aumann-Shapley values allocate a risk component to Bank 1 which is negative (Table 2). The main explanation for this is that what Bank 1 does to reduce the conditional crisis probability in scenario 2 (which has probability $p_2 = 3\%$) by generating the profit $-X_{12} = 2$ outweighs what it does to increase the conditional crisis probability in scenario 4 (which has probability $p_4 = 1\%$) by generating the loss $X_{14} = 1$.

3 Non-Negative Attribution of Risk or Cost

We present two methods for non-negative attribution of risk or cost. Recall that risk attribution can be interpreted as cost attribution in a multi-commodity problem, where the commodities are Arrow-Debreu securities, each representing wealth in one scenario only (Section 2.1). The two methods are based on a single fundamental idea: to attribute cost to system components only insofar as their participation increases cost, and to take their participation for granted when it decreases cost. In the single-commodity cost attribution problem, this idea means attributing cost to demand (losses) only, while taking supply (profits) for granted. The "scenario-wise non-negative" method of risk attribution applies the fundamental idea to losses in each scenario separately, and then assesses the implications for risk. That is, it takes all profits for granted, and attributes risk by considering losses only. The "overall non-negative" method of risk attribution applies the fundamental idea directly to risk: it takes the participation of a system component for granted when it decreases risk, and attributes risk to system components insofar as their participation increases risk. Thus, it may ignore the losses of a system component when their contribution to risk is outweighed by the negative contribution to risk of the system component's profits.

Tables 1 and 2 contain the systemic risk components in Examples 1 and 2 generated by the Shapley and Aumann-Shapley values for the usual method (Section 2.2), the usual non-negative method (Section 1), the scenario-wise non-negative method (Section 3.1), and the overall non-negative method (Section 3.2). These tables support two main conclusions about the advantages that the proposed methods have over the usual non-negative method. The scenario-based non-negative method can assign large positive risk components where the other methods assign values that are negative or zero. This makes it advantageous when the goal is to detect contributions that losses make to systemic risk without allowing them to be masked by profits. However, sometimes it is desirable to compute risk components in a way that allows profits in some scenarios to offset losses in other scenarios, at least to some extent. The usual and overall non-negative methods do this. They may coincide, as in Example 1, or differ significantly, as in Example 2. In the latter example, the overall non-negative method assigns to Bank 3 its stand-alone risk of 0.91%, which is the amount by which it increases the conditional crisis probability in scenario 3. As argued in Section 1, this is an example of how the property of partial separability is appropriate. Because Bank 3 neither contributes to nor reduces the conditional crisis probability in the other scenarios, and Banks 1 and 2 have zero loss in scenario 3, there is no interaction between Bank 3 and the other banks in their contributions to systemic risk. This makes it desirable to assign 0.91% as the risk component for Bank 3. It is a disadvantage of the usual non-negative method that it fails to do so.

Table 1: Risk Components in Example 1: Expected Shortfall.

		Risk Components		
Method	Value	Unit 1	Unit 2	Unit 3
Usual	Shapley	-18	54	-18
Usual	Aumann-Shapley	-54	90	-18
Usual Non-Negative	Shapley	0	18	0
Usual Non-Negative	Aumann-Shapley	0	18	0
Overall Non-Negative	Shapley	0	18	0
Overall Non-Negative	Aumann-Shapley	0	18	0
Scenario-Based Non-Negative	Shapley	6	6	6
Scenario-Based Non-Negative	Aumann-Shapley	2	10	6

We assume henceforth that r satisfies the standard conditions that $r(\mathbf{0}) = 0$ and $r(\mathbf{1})$ is the system's cost or risk, which is non-negative. If it is negative, then it is impossible to produce non-negative cost components.

3.1 Scenario-Wise Non-Negative Method

Suppose that $r = \rho \circ \boldsymbol{L}$, as in Equation (1). Further suppose that ρ is non-decreasing, as is often assumed in risk measurement theory,[2] that $\rho(\mathbf{0}) = 0$, and that $\boldsymbol{L}(\mathbf{0}) = \mathbf{0}$. We transform the loss $\boldsymbol{L}$

[2] An exception is the deviation measures of Rockafellar et al. (2006), which include such risk measures as standard deviation.

Table 2: Risk Components in Example 2: Systemic Crisis Probability.

		Risk Components		
Method	Value	Bank 1	Bank 2	Bank 3
Usual	Shapley	-0.41%	1.36%	0.91%
Usual	Aumann-Shapley	-0.21%	1.16%	0.91%
Usual Non-Negative	Shapley	0	1.11%	0.74%
Usual Non-Negative	Aumann-Shapley	0	1.04%	0.82%
Overall Non-Negative	Shapley	0.01%	0.94%	0.91%
Overall Non-Negative	Aumann-Shapley	0%	0.95%	0.91%
Scenario-Based Non-Negative	Shapley	0.10%	0.85%	0.91%
Scenario-Based Non-Negative	Aumann-Shapley	0.10%	0.85%	0.91%

into a non-decreasing function $\boldsymbol{L}^{\uparrow}$ scenario by scenario, then replace r by $r^{\uparrow\uparrow} = 0 \vee (\rho \circ \boldsymbol{L}^{\uparrow})$. The function $\boldsymbol{L}^{\uparrow}$ is given by

$$L_j^{\uparrow}(\boldsymbol{\lambda}) = \inf_{\boldsymbol{\lambda}'} \left\{ L_j(\boldsymbol{\lambda}') : \boldsymbol{\lambda} \leq \boldsymbol{\lambda}' \leq \mathbf{1} \right\}. \tag{7}$$

The result is that we get a non-decreasing function $r^{\uparrow\uparrow} = 0 \vee (\rho \circ \boldsymbol{L}^{\uparrow})$ to use in the Shapley and Aumann-Shapley values. Because $\boldsymbol{L}^{\uparrow}(\mathbf{1}) = \boldsymbol{L}$ and $\rho(\boldsymbol{L}) \geq 0$, $r^{\uparrow\uparrow}(\mathbf{1}) = \rho(\boldsymbol{L})$. Because $\boldsymbol{L}^{\uparrow}(\mathbf{0}) \leq \boldsymbol{L}(\mathbf{0}) = \mathbf{0}$, $\rho(\mathbf{0}) = 0$, and ρ is non-decreasing, $\rho(\boldsymbol{L}^{\uparrow}(\mathbf{0})) \leq 0$ and therefore $r^{\uparrow\uparrow}(\mathbf{0}) = 0$.

In the particular case of a portfolio, as in Equation (1), $L_j(\boldsymbol{\lambda}) = \sum_{i=1}^{n} \lambda_i X_{ij}$, so

$$L_j^{\uparrow}(\boldsymbol{\lambda}) = \sum_{i=1}^{n} \lambda_i X_{ij}^{+} - \sum_{i=1}^{n} X_{ij}^{-}. \tag{8}$$

Then $\rho(\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda}(S))) = \rho(\sum_{i \in S} \boldsymbol{X}_{i\cdot}^{+} - \sum_{i=1}^{n} \boldsymbol{X}_{i\cdot}^{-})$ is the risk of a portfolio exposed to the profits of all the components and the losses of the components in the set S.

The incremental risk used to compute the Shapley value is

$$\Delta_i r^{\uparrow\uparrow}(\boldsymbol{\lambda}(S)) = \begin{cases} \rho(\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda}(S \cup \{i\}))) - \rho(\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda}(S))) & \text{if } \rho(\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda}(S))) \geq 0 \\ 0 & \text{if } \rho(\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda}(S \cup \{i\}))) < 0 \\ \rho(\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda}(S \cup \{i\}))) & \text{otherwise} \end{cases}. \tag{9}$$

The risk component attributed by the Aumann-Shapley value to component i is $\int_0^1 \frac{\partial r^{\uparrow\uparrow}}{\partial \lambda_i}(\gamma \mathbf{1})\, \mathrm{d}\gamma$. If ρ and $\boldsymbol{L}^{\uparrow}$ are suitably differentiable, then this Aumann-Shapley value is given by

$$\int_0^1 \frac{\partial r^{\uparrow\uparrow}}{\partial \lambda_i}(\gamma \mathbf{1})\, \mathrm{d}\gamma = \int_{\gamma^*}^1 \left(\nabla \rho(\boldsymbol{L}^{\uparrow}(\gamma \mathbf{1})) \right) \frac{\partial}{\partial \lambda_i} \boldsymbol{L}^{\uparrow}(\gamma \mathbf{1})\, \mathrm{d}\gamma. \tag{10}$$

where $\gamma^* = \inf \left\{ \gamma : \rho(\boldsymbol{L}^{\uparrow}(\gamma \mathbf{1})) > 0) > 0 \right\}$. Recall that we assumed $r(\mathbf{0}) = 0$ and $r(\mathbf{1}) \geq 0$, which implies that $0 \leq \gamma^* \leq 1$. In the case of a portfolio, using Equation (8), the risk component is

$$\left(\int_{\gamma^*}^1 \left(\nabla \rho \left(\gamma \sum_{i'=1}^{n} \boldsymbol{X}_{i'\cdot}^{+} - \sum_{i'=1}^{n} \boldsymbol{X}_{i'\cdot}^{-} \right) \right) \mathrm{d}\gamma \right) \boldsymbol{X}_{i\cdot}^{+}. \tag{11}$$

Assuming that the number of scenarios in Ω is finite, this can be written more concretely as

$$\sum_{j\in\Omega}\left(\int_{\gamma^*}^{1}\frac{\partial\rho}{\partial L_j}\left(\gamma\sum_{i'=1}^{n}\boldsymbol{X}_{i'\cdot}^{+}-\sum_{i'=1}^{n}\boldsymbol{X}_{i'\cdot}^{-}\right)\mathrm{d}\gamma\right)X_{ij}^{+}. \tag{12}$$

This weighted sum of the losses of component i has the interpretation of average-cost pricing. A numerical approximation to this is

$$\sum_{j\in\Omega}\left(\frac{1}{N}\sum_{k=1}^{N}1\left\{\rho\left(\frac{k-1}{N}\sum_{i'=1}^{n}\boldsymbol{X}_{i'\cdot}^{+}-\sum_{i'=1}^{n}\boldsymbol{X}_{i'\cdot}^{-}\right)>0\right\}\frac{\partial\rho}{\partial L_j}\left(\frac{k-1}{N}\sum_{i'=1}^{n}\boldsymbol{X}_{i'\cdot}^{+}-\sum_{i'=1}^{n}\boldsymbol{X}_{i'\cdot}^{-}\right)\right)X_{ij}. \tag{13}$$

The partial derivative can be replaced by a finite-difference approximation if need be.

Example 1 The risk components for Example 1 are reported in Table 1. In this example, $\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda})=\begin{bmatrix}-180 & -270 & 90\lambda_2-180 & 90\lambda_1+90\lambda_2+270\lambda_3\end{bmatrix}$. The scenario-wise non-negative Shapley value allocates equal amounts of risk to all three banks. This happens because $r^{\uparrow\uparrow}(\boldsymbol{\lambda}(S)) = 0\vee\rho(\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda}(S)))=0$ for any proper subset $S\subset\{1,2,3\}$. When two or fewer banks are in the subset, by adding the participation of a third bank only in scenarios in which it earns a profit, the resulting loss vector $\boldsymbol{L}^{\uparrow}$ has so much profit in some of its tail scenarios that its expected shortfall is not positive. To compute the scenario-wise non-negative Aumann-Shapley values, we make the following observations. The risk $\rho(\boldsymbol{L}^{\uparrow}(\gamma\mathbf{1}))$ is positive for $\gamma>\gamma^*=8/9$. For all $\boldsymbol{\lambda}$ sufficiently near the diagonal $\{\gamma\mathbf{1}:\gamma\in[0,1]\}$, scenarios 3 and 4 belong to the 5% tail of worst losses, and there are no ties among scenarios, which would pose difficulties for differentiation. The gradient of ρ near the diagonal is then $\begin{bmatrix}0 & 0 & 0.8 & 0.2\end{bmatrix}$, because scenario 3 is four times as likely as scenario 4. Also using the sensitivities of $\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda})=\begin{bmatrix}-180 & -270 & 90\lambda_2-180 & 90\lambda_1+90\lambda_2+270\lambda_3\end{bmatrix}$, Equation (12) gives an Aumann-Shapley value of

$$\int_{8/9}^{1}\left(0.8\times\begin{bmatrix}0 & 90 & 0\end{bmatrix}+0.2\times\begin{bmatrix}90 & 90 & 270\end{bmatrix}\right)\mathrm{d}\gamma=\begin{bmatrix}2 & 10 & 6\end{bmatrix}.$$

The ratio of the risk component of Unit 3 to that of Unit 1 is much higher than in other schemes. Here the ratio is three because the loss in scenario 4 of Unit 3 is three times as large as that of Unit 1.

Example 2 The risk components for Example 2 are reported in Table 2. Table 3 contains the systemic loss $L_j^{\uparrow}$ for scenarios $j=1-4$, and the resulting risk measures, which are used to compute the scenario-wise non-negative Shapley value. Compared to the overall non-negative Shapley value, the risk component of Bank 1 is larger. This is because the scenario-wise non-negative Shapley value recognizes the contribution of Bank 1 to increasing the conditional crisis probability in scenario 4, but takes for granted the effect of Bank 1 in reducing the conditional crisis probability in other scenarios. Applying Equation (12) to $\rho(\boldsymbol{L})=\sum_{j\in\Omega}p_j\Phi(\beta_0+\beta_1L_j^{+})-\Phi(\beta_0)$, we get an Aumann-Shapley value for component i of

$$\begin{aligned}&\sum_{j\in\Omega}\left(\int_{\gamma^*}^{1}p_j\beta_1\Phi'\left(\beta_0+\beta_1\left(\gamma\sum_{i'=1}^{n}X_{i'j}^{+}-\sum_{i'=1}^{n}X_{i'j}^{-}\right)\right)\mathrm{d}\gamma\right)X_{ij}^{+}\\&=\sum_{j\in\Omega}p_j\left(\Phi(\beta_0+\beta_1L_j^{+})-\Phi(\beta_0)\right)\frac{X_{ij}^{+}}{\sum_{i'=1}^{n}X_{i'j}^{+}}.\end{aligned}$$

In each scenario in which the portfolio has a net loss, the conditional crisis probability is elevated, and each bank that generates a loss (not a profit) is attributed a fraction of the increase in conditional crisis probability, in proportion to the bank's loss.

Table 3: Computation of the Scenario-Wise Non-Negative Shapley Value in Example 2.

Set S of Banks	$\emptyset$	$\{3\}$	$\{2\}$	$\{2,3\}$	$\{1\}$	$\{1,3\}$	$\{1,2\}$	$\{1,2,3\}$
Scenario 1: Loss $L_1^{\uparrow}(\boldsymbol{\lambda}(S))$	-3	-3	-3	-3	-3	-3	-3	-3
Scenario 2: Loss $L_2^{\uparrow}(\boldsymbol{\lambda}(S))$	-2	-2	3	3	-2	-2	3	3
Scenario 3: Loss $L_3^{\uparrow}(\boldsymbol{\lambda}(S))$	0	4	0	4	0	4	0	4
Scenario 4: Loss $L_4^{\uparrow}(\boldsymbol{\lambda}(S))$	0	0	4	4	1	1	5	5
Excess crisis probability $\rho(\boldsymbol{L}^{\uparrow})$	0%	0.91%	0.76%	1.67%	0.02%	0.92%	0.95%	1.86%

3.2 Overall Non-Negative Method

As we saw from the theoretical analysis of the Shapley and Aumann-Shapley values (Section 2.2), they would give us non-negative risk components if the function r, which maps participation of the components to risk, were non-decreasing. It may fail to be non-decreasing when the participation of a component lowers the system's risk by generating profits. The overall non-negative method of risk attribution avoids this problem by working with the greatest non-decreasing function dominated by $0 \vee r$, given by

$$r^{\uparrow}(\boldsymbol{\lambda}) = \inf_{\boldsymbol{\lambda}'} \left\{ 0 \vee r(\boldsymbol{\lambda}') : \boldsymbol{\lambda} \leq \boldsymbol{\lambda}' \leq \mathbf{1} \right\} = 0 \vee \inf_{\boldsymbol{\lambda}'} \left\{ r(\boldsymbol{\lambda}') : \boldsymbol{\lambda} \leq \boldsymbol{\lambda}' \leq \mathbf{1} \right\} \tag{14}$$

for $\boldsymbol{\lambda} \in [0,1]^n$. The reason to work with $0 \vee r$ instead of r is that this ensures the condition $r^{\uparrow}(\mathbf{0}) = 0$, needed to apply the Shapley and Aumann-Shapley values. As the following proposition shows, the function $r^{\uparrow}$ is suitable to use in risk attribution, and even inherits the often-desirable property of convexity, if r has it.

Proposition 1. *The function $r^{\uparrow}$ is non-decreasing and satisfies $r^{\uparrow}(\mathbf{0}) = 0$ and $r^{\uparrow}(\mathbf{1}) = r(\mathbf{1})$. If r is convex, then $r^{\uparrow}$ is convex.*

Proof. The first part is clear from the construction of $r^{\uparrow}$. Assume that r is convex. Then the function $r^{+} = 0 \vee r$ is convex. For any $\alpha \in [0,1]$ and $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in [0,1]^n$,

$$\begin{aligned}
r^{\uparrow}(\alpha\boldsymbol{\lambda}_1 + (1-\alpha)\boldsymbol{\lambda}_2) &= \inf_{\boldsymbol{\lambda}'} \{ r^{+}(\boldsymbol{\lambda}') : \alpha\boldsymbol{\lambda}_1 + (1-\alpha)\boldsymbol{\lambda}_2 \leq \boldsymbol{\lambda}' \leq \mathbf{1} \} \\
&= \inf_{\boldsymbol{\lambda}_1', \boldsymbol{\lambda}_2'} \{ r^{+}(\alpha\boldsymbol{\lambda}_1' + (1-\alpha)\boldsymbol{\lambda}_2') : \boldsymbol{\lambda}_1 \leq \boldsymbol{\lambda}_1' \leq \mathbf{1}, \boldsymbol{\lambda}_2 \leq \boldsymbol{\lambda}_2' \leq \mathbf{1} \} \\
&\leq \inf_{\boldsymbol{\lambda}_1', \boldsymbol{\lambda}_2'} \{ \alpha r^{+}(\boldsymbol{\lambda}_1') + (1-\alpha) r^{+}(\boldsymbol{\lambda}_2') : \boldsymbol{\lambda}_1 \leq \boldsymbol{\lambda}_1' \leq \mathbf{1}, \boldsymbol{\lambda}_2 \leq \boldsymbol{\lambda}_2' \leq \mathbf{1} \} \\
&= \alpha \inf_{\boldsymbol{\lambda}_1'} \{ r^{+}(\boldsymbol{\lambda}_1') : \boldsymbol{\lambda}_1 \leq \boldsymbol{\lambda}_1' \leq \mathbf{1} \} + (1-\alpha) \inf_{\boldsymbol{\lambda}_2'} \{ r^{+}(\boldsymbol{\lambda}_2') : \boldsymbol{\lambda}_2 \leq \boldsymbol{\lambda}_2' \leq \mathbf{1} \} \\
&= \alpha r^{\uparrow}(\boldsymbol{\lambda}_1) + (1-\alpha) r^{\uparrow}(\boldsymbol{\lambda}_2),
\end{aligned}$$

which shows that $r^{\uparrow}$ is convex. $\square$

If Equation (1) applies, then Equation (14) can be written as

$$r^{\uparrow}(\boldsymbol{\lambda}) = 0 \vee \inf_{\boldsymbol{\lambda}'} \left\{ \rho\left(\sum_{i=1}^{n} \lambda'_i \boldsymbol{X}_{i\cdot} \right) : \boldsymbol{\lambda} \leq \boldsymbol{\lambda}' \leq 1 \right\}.$$

The infimum has the interpretation of the minimal risk attainable starting from the system with participation $\boldsymbol{\lambda}$ and "hedging" to reduce risk, but only by increasing participation of some components.

The incremental risk used to compute the Shapley value is

$$\Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S)) = \min_{S'} \left\{ 0 \vee r(\boldsymbol{\lambda}(S')) : i \in S', S \subset S' \right\} - \min_{S'} \left\{ 0 \vee r(\boldsymbol{\lambda}(S')) : S \subseteq S' \right\}.$$

If Equation (1) applies, then this becomes

$$\Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S)) = \min_{S'} \left\{ 0 \vee \rho\left(\sum_{i' \in S'} \boldsymbol{X}_{i'\cdot} \right) : i \in S', S \subset S' \right\} - \min_{S'} \left\{ 0 \vee \rho\left(\sum_{i' \in S'} \boldsymbol{X}_{i'\cdot} \right) : S \subseteq S' \right\}.$$

It is zero if component i is included in an "optimal hedge" for the system that has participation by the components in the set S, i.e., if adding the participation of a set of components including i is a way to minimize risk or reduce it to zero.

The Aumann-Shapley value attributes to component i a risk component of $\int_0^1 \frac{\partial r^{\uparrow}}{\partial \lambda_i}(\gamma \mathbf{1})\, \mathrm{d}\gamma$. The partial derivative $\frac{\partial r^{\uparrow}}{\partial \lambda_i}(\gamma \mathbf{1})$ is zero if the constraint $\lambda'_i \geq \lambda_i$ is not binding optimally in Equation (14). In that case, for the system in which all participation levels are γ, there is an optimal hedge that increases the participation of component i. That is, the participation of a component is taken for granted when increasing it reduces risk. It is possible to approximate the Aumann-Shapley value by numerical integration, finite-difference approximations of partial derivatives, and solving Equation (14) approximately with a derivative-free nonlinear optimization algorithm. For example, one may approximate the risk attributed to component i by

$$\frac{1}{N} \sum_{k=1}^{N} \frac{1}{h} \left(\widehat{r^{\uparrow}} \left(\frac{k-1}{N} \mathbf{1} + h \mathbf{e}_i \right) - \widehat{r^{\uparrow}} \left(\frac{k-1}{N} \mathbf{1} \right) \right), \tag{15}$$

where N is the number of quadrature points, h is the finite difference, and $\widehat{r^{\uparrow}}(\boldsymbol{\lambda})$ is the approximate solution to Equation (14).

Example 1 The risk components in Example 1 by the overall non-negative method are reported in Table 1. The Shapley value is zero for Unit 3 because it has negative expected shortfall in a portfolio by itself, and does not increase expected shortfall when added to any portfolio containing some of the other business units. The Shapley value is zero for Unit 1 even though it has positive expected shortfall in a portfolio by itself. This is because any increase in expected shortfall resulting from adding Unit 1 to a portfolio can be eliminated by also adding Unit 3. Table 4 details the computation of the Shapley value. The entries on the last line of the table show that incremental risk is positive only when adding Unit 2 to the set S. Thus, the Shapley value attributes all of the risk to Unit 2. Next, we turn to the Aumann-Shapley value. In this example, for all $\boldsymbol{\lambda} \in [0,1]^3$ sufficiently near the diagonal $\{\gamma \mathbf{1} : \gamma \in [0,1]\}$,

$$\begin{aligned} r^{\uparrow}(\boldsymbol{\lambda}) &= 0 \vee r\left(\begin{bmatrix}1 & \lambda_2 & 1\end{bmatrix}^{\top}\right) \\ &= 0 \vee \rho\left(\begin{bmatrix}-(90+90\lambda_2) & -270 & -180+90\lambda_2 & 360+90\lambda_2\end{bmatrix}\right) \\ &= 0 \vee ((-180+90\lambda_2) \times 4\% + (360+90\lambda_2) \times 1\%)/5\% \\ &= 0 \vee (-72+90\lambda_2) \\ &= (90\lambda_2 - 72)^{+}. \end{aligned}$$

For $\boldsymbol{\lambda}$ near the diagonal, the gradient is $\nabla r^{\uparrow}(\boldsymbol{\lambda}) = \begin{bmatrix}0 & 90 & 0\end{bmatrix}$ if $\lambda_2 > 0.8$ and $\mathbf{0}$ if $\lambda_2 < 0.8$. The Aumann-Shapley value is $\int_0^1 \nabla r^{\uparrow}(\gamma \mathbf{1})\,\mathrm{d}\gamma = \int_{0.8}^1 \begin{bmatrix}0 & 90 & 0\end{bmatrix}\,\mathrm{d}\gamma = \begin{bmatrix}0 & 18 & 0\end{bmatrix}$. It is zero for Unit 1 and Unit 3 because the contribution they make to reducing the portfolio loss in scenario 2 outweighs the contribution they make to increasing it in scenario 4.

Table 4: Computation of the Overall Non-Negative Shapley Value in Example 1.

Set S of Business Units	$\emptyset$	$\{3\}$	$\{2\}$	$\{2,3\}$	$\{1\}$	$\{1,3\}$	$\{1,2\}$	$\{1,2,3\}$
Expected Shortfall $r(\boldsymbol{\lambda}(S))$	0	-90	90	72	18	0	36	18
Optimal $S' \supseteq S$	$\emptyset$	$\{3\}$	$\{1,2,3\}$	$\{1,2,3\}$	$\{1,3\}$	$\{1,3\}$	$\{1,2,3\}$	$\{1,2,3\}$
$r^{\uparrow}(\boldsymbol{\lambda}(S)) = 0 \vee r(\boldsymbol{\lambda}(S'))$	0	0	18	18	0	0	18	18

Example 2 The risk components in Example 2 by the overall non-negative method are reported in Table 2. The Shapley value is positive for Bank 1, because Bank 1 by itself increases the probability of a systemic crisis. However, it is small, because adding Bank 1 to a system including one or both of the other banks does not increase the crisis probability. The Aumann-Shapley values reported in Table 2 were approximated using Equation (15) with $N = 10^3$ and $h = 10^{-4}$, and computing $\widehat{r^{\uparrow}}$ with the Hooke-Jeeves algorithm as implemented in the R package `dfoptim`.

4 Game Theory

Cooperative game theory studies the suitability for various purposes of cost allocation schemes, such as the usual Shapley and Aumann-Shapley values. This is done by proving that, under appropriate conditions, the cost allocations have or do not have certain properties, e.g., of fairness, or of producing incentives for cooperation. See, e.g., Moulin and Sprumont (2007) for an introduction. A similar approach has been taken in risk attribution (Denault, 2001; Kalkbrener, 2005). There are many properties that are investigated in cooperative game theory. Here we investigate three important properties in relation to the scenario-wise and overall non-negative risk attribution methods. In Section 4.1, we study a separability property, which says that if one system component has no interactions with the other components of the system, then the risk attributed to it equals the risk generated by that component in isolation. In Section 4.2, we study a diversification property, which says that the risk attributed to any component as part of a portfolio or system cannot be more than the risk generated by that component in isolation. This property is considered desirable in portfolio risk attribution. In Section 4.3, we study a monotonicity property, which says that the risk attributed to a system component cannot decrease if the losses generated by that component increase. If this property were to be violated, then the risk attribution method could provide bad incentives. For example, it would be problematic if a firm could lower its systemic risk charge by taking worse risks that increase the losses it would incur in some scenarios.

4.1 Separability

In cooperative game theory, there is a focus on cost attribution methods, such as the Shapley and Aumann-Shapley values, that satisfy a separability axiom (Moulin and Sprumont, 2007). The separability axiom says that if $r(\boldsymbol{\lambda}) = \sum_{i=1}^{n} r_i(\lambda_i)$, then the risk component attributed to system component i is $r_i(1)$ for all $i = 1, \ldots, n$. This axiom is not useful in our investigation of methods of generating non-negative risk components. In general, a method that satisfies the separability axiom cannot be guaranteed to generate non-negative risk components. If we consider only non-negative risk functions r, then the separability axiom is compatible with avoiding negative risk components. However, it is so easy to satisfy the separability axiom for non-negative risk functions that it tells us almost nothing about which methods of generating non-negative risk components are attractive. Any method that coincides with the usual Shapley or Aumann-Shapley value when these are non-negative satisfies the separability axiom for non-negative risk functions. Consequently, we consider a separability property that is more difficult to satisfy. Let $\boldsymbol{\lambda}_{-i}$ represent the vector formed from $\boldsymbol{\lambda}$ by deleting the ith component, λ_i.

Definition 1. *The risk attribution scheme satisfies the* partial separability property *if* $r(\boldsymbol{\lambda}) = r_i(\lambda_i) + r_{-i}(\boldsymbol{\lambda}_{-i})$ *for some functions* r_i *and* r_{-i} *and all* $\boldsymbol{\lambda} \in [0, 1]^n$ *implies that the risk component attributed to system component* i *is* $r_i(1)$.

The property of partial separability says that if component i has no interaction with other components, i.e., it is additively separable from the other components, then its risk component should equal its stand-alone risk $r(\mathbf{e}_i) = r_i(1)$. The methods for generating non-negative risk components that we have considered can be interpreted as redistributing the credit for risk reduction away from negative risk components. The property of partial separability says that this credit should not be distributed to a system component that has no interactions with the other components.

The property of partial separability has perhaps not been emphasized in the literature on additive cost allocation methods because, given that the axiom of additivity holds, separability and partial separability are both equivalent to the dummy axiom. See Moulin and Sprumont (2007, § 3). The methods of generating non-negative risk components considered in this article are not additive.

In Table 2, we see that the usual non-negative Shapley and Aumann-Shapley values violate the property of partial separability in Example 2. Next we prove that the overall and scenario-wise non-negative Shapley and Aumann-Shapley values have the property of partial separability, under some conditions that hold in Example 2.

Proposition 2. *If* r *is non-negative, then the overall non-negative Shapley and Aumann-Shapley values satisfy the partial separability property.*

Proof. Suppose that r is non-negative and $r(\boldsymbol{\lambda}) = r_i(\lambda_i) + r_{-i}(\boldsymbol{\lambda}_{-i})$ for some functions r_i and r_{-i}

and all $\boldsymbol{\lambda} \in [0,1]^n$. Then

$$r^{\uparrow}(\boldsymbol{\lambda}) = \inf_{\lambda'_i} \left\{ r_i(\lambda_i) : \lambda_i \leq \lambda'_i \leq 1 \right\} + \inf_{\boldsymbol{\lambda}'_{-i}} \left\{ r_{-i}(\boldsymbol{\lambda}_{-i}) : \boldsymbol{\lambda}_{-i} \leq \boldsymbol{\lambda}'_{-i} \leq \mathbf{1} \right\}.$$

Because the Shapley and Aumann-Shapley values satisfy the property of partial separability, they assign to component i the risk component $r^{\uparrow}(\mathbf{e}_i) = r_i(1)$. □

Proposition 3. *If Equation (1) holds and ρ is non-negative, then the scenario-wise non-negative Shapley and Aumann-Shapley values satisfy the partial separability property.*

Proof. Suppose that Equation (1) holds, ρ is non-negative, and $r(\boldsymbol{\lambda}) = r_i(\lambda_i) + r_{-i}(\boldsymbol{\lambda}_{-i})$ for some functions r_i and r_{-i} and all $\boldsymbol{\lambda} \in [0,1]^n$. Then $r_i(\lambda_i) = \rho_i(\lambda_i X_i)$ and $r_{-i}(\boldsymbol{\lambda}_{-i}) = \rho_{-i}(\sum_{i' \neq i} \lambda_{i'} X_{i'})$ for some non-negative risk measures ρ_i and ρ_{-i}. Then $r^{\uparrow\uparrow}(\boldsymbol{\lambda}) = \rho_i(\lambda_i X_i^+ - X_i^-) + \rho_{-i}(\sum_{i' \neq i}(\lambda_{i'} X_{i'}^+ - X_{i'}))$. Because the Shapley and Aumann-Shapley values satisfy the property of partial separability, they assign to component i the risk component $r^{\uparrow\uparrow}(\mathbf{e}_i) = \rho_i(\lambda_i X_i) = r_i(1)$. □

4.2 Diversification

If the function r is convex and positively homogeneous, then a system that combines the participation of multiple components has cost bounded above by the sum of the costs of the systems in which each of the components participates alone. Thus, it seems appropriate for the risk attribution method to obey the "diversification" (Kalkbrener, 2005) or "stand-alone" property (Moulin and Sprumont, 2007) that the risk allocated to any component i is bounded above by its stand-alone risk $r(\mathbf{e}_i)$. In Equation (1), $r(\mathbf{e}_i) = \rho(\boldsymbol{X}_{i\cdot})$, the risk measure applied to the losses generated by component i alone.

Definition 2. *Let ϕ_i represent the risk component attributed to system component i when performing risk attribution on the function r. The risk attribution scheme satisfies the* diversification property *if, for all $i = 1, \ldots, n$, $\phi_i \leq r(\mathbf{e}_i)$.*

In this section, we prove that the overall and scenario-wise non-negative Shapley and Aumann-Shapley values have the diversification property, under some conditions.

4.2.1 Overall Non-Negative Method

Proposition 4. *If r is convex and positively homogeneous, then the overall non-negative Shapley and Aumann-Shapley values satisfy the diversification property.*

Proof. First, consider the Shapley value. The incremental risk

$$\Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S)) \leq \min_{S'} \{ r(\boldsymbol{\lambda}(S')) : i \in S', S \subset S' \} - \min_{S'} \{ r(\boldsymbol{\lambda}(S') : S \subseteq S' \}.$$

Let T be a set satisfying $r(\boldsymbol{\lambda}(S \cup T)) = \min_{S'}\{r(\boldsymbol{\lambda}(S') : S \subseteq S'\}$, i.e., an optimal hedge for the set S. Then $r(\boldsymbol{\lambda}(S')) : i \in S', S \subset S'\} \leq r(\boldsymbol{\lambda}(S \cup \{i\} \cup T))$ because the optimal hedge for $S \cup \{i\}$ is at least as good as T. Because r is sublinear, $r(\boldsymbol{\lambda}(S \cup \{i\} \cup T)) \leq r(\boldsymbol{\lambda}(S \cup T)) + r(\boldsymbol{\lambda}(\{i\}))$. Therefore $\Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S)) \leq r(\boldsymbol{\lambda}(\{i\}))$.

Next, consider the Aumann-Shapley value, which is an average of

$$\frac{\partial r^{\uparrow}}{\partial \lambda_i}(\gamma \mathbf{1}) = \lim_{h \downarrow 0} \frac{1}{h} \left(r^{\uparrow}(\gamma \mathbf{1} + h \mathbf{e}_i) - r^{\uparrow}(\gamma \mathbf{1}) \right)$$

over γ ranging from 0 to 1. It suffices to prove that the partial derivative is bounded above by $r(\mathbf{e}_i)$ for all $\gamma \in [0, 1]$. It follows from Equation (14) that

$$\frac{\partial r^{\uparrow}}{\partial \lambda_i}(\gamma \mathbf{1}) \leq \lim_{h \downarrow 0} \frac{1}{h} \left(\inf_{\boldsymbol{\lambda}'} \{ r(\boldsymbol{\lambda}') : \gamma \mathbf{1} + h \mathbf{e}_i \leq \boldsymbol{\lambda}' \leq \mathbf{1} \} - \inf_{\boldsymbol{\lambda}'} \{ r(\boldsymbol{\lambda}') : \gamma \mathbf{1} \leq \boldsymbol{\lambda}' \leq \mathbf{1} \} \right).$$

Choose any $\epsilon > 0$ and let $\boldsymbol{\lambda}^*$ be such that $r(\boldsymbol{\lambda}^*) < \inf_{\boldsymbol{\lambda}'} \{ r(\boldsymbol{\lambda}') : \gamma \mathbf{1} \leq \boldsymbol{\lambda}' \leq \mathbf{1} \} + \epsilon$. For $h \in (0, 1 - \gamma)$, we have

$$\inf_{\boldsymbol{\lambda}'} \{ r(\boldsymbol{\lambda}') : \gamma \mathbf{1} + h \mathbf{e}_i \leq \boldsymbol{\lambda}' \leq \mathbf{1} \} \leq r(\boldsymbol{\lambda}^* + h \mathbf{e}_i).$$

Because r is sublinear, this is less than or equal to $r(\boldsymbol{\lambda}^*) + h r(\mathbf{e}_i)$. Therefore, for any $\epsilon > 0$,

$$\inf_{\boldsymbol{\lambda}'} \{ r(\boldsymbol{\lambda}') : \gamma \mathbf{1} + h \mathbf{e}_i \leq \boldsymbol{\lambda}' \leq \mathbf{1} \} - \inf_{\boldsymbol{\lambda}'} \{ r(\boldsymbol{\lambda}') : \gamma \mathbf{1} \leq \boldsymbol{\lambda}' \leq \mathbf{1} \} < \epsilon + h r(\mathbf{e}_i).$$

This provides an upper bound on $\frac{\partial r^{\uparrow}}{\partial \lambda_i}(\gamma \mathbf{1})$ of $r(\mathbf{e}_i)$, which was what was needed. □

Corollary 1. *If Equation (1) holds with the risk measure ρ being convex and positively homogeneous, then r is convex and positively homogeneous, and the overall non-negative Shapley and Aumann-Shapley values satisfy the diversification property.*

Proof. In Equation (1), $r = \rho \circ \boldsymbol{L}$ and $\boldsymbol{L}$ is linear. Therefore r is convex and positively homogeneous if ρ is. Now apply Prop. 4. □

4.2.2 Scenario-Wise Non-Negative Method

The scenario-wise method for non-negative risk components works especially well when, in Equation (1), the risk measure ρ is a shortfall risk measure (Staum, 2013) and also convex and positively homogeneous. Related concepts, which could also fit well with non-negative risk components, appear in Cont et al. (2013) and Koch-Medina et al. (2013). The mathematical finance literature on risk measures is split as to sign convention. Here we have adopted the sign convention of Chen et al. (2013), that $\rho(\boldsymbol{L})$ is the risk of a system whose loss is $\boldsymbol{L}$, where $L_j > 0$ if there is a loss in scenario j and $L_j < 0$ if there is a profit in scenario j. However, other articles, including Staum (2013), represent the risk of a system whose loss is $\boldsymbol{L}$ as $\rho(-\boldsymbol{L})$. Translating the definition of Staum (2013) to account for the difference in sign convention, a shortfall risk measure ρ is

- normalized, meaning that $\rho(\mathbf{0}) = 0$,
- non-negative, meaning that $\rho(\boldsymbol{L}) \geq 0$ for all $\boldsymbol{L}$,
- non-decreasing, meaning that $\rho(\boldsymbol{L}) \geq \rho(\boldsymbol{L}')$ for all $\boldsymbol{L}$ and $\boldsymbol{L}'$ such that $\boldsymbol{L} \geq \boldsymbol{L}'$, and
- excess-invariant, meaning that ρ depends only on losses and not on profits: $\rho(\boldsymbol{L}) = \rho(\boldsymbol{L}^+)$ for all $\boldsymbol{L}$.

In Example 1, ρ is not excess-invariant, although it has the other properties of a shortfall risk measure. In Example 2, ρ is a shortfall risk measure.

Proposition 5. *If Equation (1) holds and ρ is a convex, positively homogeneous shortfall risk measure, then the scenario-wise non-negative Shapley and Aumann-Shapley values satisfy the diversification property.*

Proof. By Lemma 1 below, the Shapley or Aumann-Shapley value attributes to component i a risk component that equals $\sum_{j\in\Omega}\mu_j X_{ij}^+$ for some vector μ such that $\sum_{j\in\Omega}\mu_j X_{ij}^+ \leq \rho(\boldsymbol{X}_{i\cdot}^+)$. By excess-invariance, $\rho(\boldsymbol{X}_{i\cdot}^+) = \rho(\boldsymbol{X}_{i\cdot})$. $\square$

Lemma 1. *If Equation (1) holds and ρ is non-decreasing, convex, and positively homogeneous, then there exists a set $\mathcal{M}$ such that $\rho(\boldsymbol{L}) = \sup\{\sum_{j\in\Omega}\mu_j L_j : \mu \in \mathcal{M}\}$ for all $\boldsymbol{L}$, and the risk component i generated by the scenario-wise non-negative Shapley or Aumann-Shapley value equals $\sum_{j\in\Omega}\mu_j X_{ij}^+$ for some $\mu \in \mathcal{M}$.*

Proof. The risk component comes from the Shapley or Aumann-Shapley value applied to $r^{\uparrow\uparrow} = 0 \vee (\rho \circ \boldsymbol{L}^{\uparrow}) = \rho \circ \boldsymbol{L}^{\uparrow}$, because the shortfall risk measure is non-negative. Following Ruszczyński and Shapiro (2006), Staum (2013, Thm. 5.1) shows that $\rho(\boldsymbol{L}) = \sup\{\sum_{j\in\Omega}\mu_j L_j : \mu \in \mathcal{M}\}$, where $\mathcal{M}$ is the set of subgradients of ρ at $\boldsymbol{0}$, and $\mathcal{M}$ is convex.

As a preliminary to analyzing the Shapley and Aumann-Shapley values, let $\rho'(\boldsymbol{L};\boldsymbol{H})$ denote the directional derivative of ρ at the point $\boldsymbol{L}$ in the direction $\boldsymbol{H}$, i.e.,

$$\rho'(\boldsymbol{L};\boldsymbol{H}) = \lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}\left(\rho(\boldsymbol{L}+\epsilon\boldsymbol{H}) - \rho(\boldsymbol{L})\right).$$

The directional derivative equals the inner product of $\boldsymbol{H}$ with a subgradient of ρ at $\boldsymbol{L}$, i.e., it is $\sum_{j\in\Omega}H_j\mu_j$ for some μ that is a subgradient of ρ at $\boldsymbol{L}$ (Rockafellar, 1970, Thm. 23.4). Because ρ is convex and positively homogeneous, a subgradient at any point $\boldsymbol{L}$ is a subgradient at $\boldsymbol{0}$. (This claim can be seen as follows. Let μ be a subgradient of ρ at $\boldsymbol{L}$, meaning that for all $\boldsymbol{L}'$, $\rho(\boldsymbol{L}') \geq \rho(\boldsymbol{L}) + \sum_{j\in\Omega}(L_j' - L_j)\mu_j$. Defining $\boldsymbol{L}'' = \boldsymbol{L}' - \boldsymbol{L}$, the inequality can be rewritten as $\rho(\boldsymbol{L}+\boldsymbol{L}'') - \rho(\boldsymbol{L}) \geq \sum_{j\in\Omega}L_j''\mu_j$. Together with subadditivity of ρ and $\rho(\boldsymbol{0}) = 0$, this inequality implies $\rho(\boldsymbol{L}'') - \rho(\boldsymbol{0}) \geq \sum_{j\in\Omega}L_j''\mu_j$. That is, μ is a subgradient of ρ at $\boldsymbol{0}$.)

First, consider the Aumann-Shapley value, which is

$$\left(\int_0^1 \partial\rho\left(\gamma\sum_{i'=1}^n \boldsymbol{X}_{i'\cdot}^+ - \sum_{i'=1}^n \boldsymbol{X}_{i'\cdot}^-\right)\mathrm{d}\gamma\right)\boldsymbol{X}_{i\cdot}^+,$$

where $\partial\rho(\boldsymbol{L})$ represents the subgradient of ρ at $\boldsymbol{L}$ such that the directional derivative $\rho'(\boldsymbol{L};\boldsymbol{X}_{i\cdot}^+) = \sum_{j\in\Omega}X_{ij}^+(\partial\rho(\boldsymbol{L}))_j$. As stated, $\partial\rho(\boldsymbol{L})$ is in $\mathcal{M}$ for every $\boldsymbol{L}$. Therefore, the Aumann-Shapley value is the product of $\boldsymbol{X}_{i\cdot}^+$ with an average of elements of $\mathcal{M}$. Because $\mathcal{M}$ is convex, this average is also in $\mathcal{M}$.

Finally, consider the Shapley value. The incremental risk of adding the participation of component i to the set S is

$$\begin{aligned}
\Delta_i r^{\uparrow\uparrow}(\boldsymbol{\lambda}(S)) &= \rho\left(\boldsymbol{X}_{i\cdot}^{+} + \sum_{i' \in S} \boldsymbol{X}_{i'\cdot}^{+} - \sum_{i'=1}^{n} \boldsymbol{X}_{i'\cdot}^{-}\right) - \rho\left(\sum_{i' \in S} \boldsymbol{X}_{i'\cdot}^{+} - \sum_{i'=1}^{n} \boldsymbol{X}_{i'\cdot}^{-}\right) \\
&= \int_0^1 \rho'\left(\gamma \boldsymbol{X}_{i\cdot}^{+} + \sum_{i' \in S} \boldsymbol{X}_{i'\cdot}^{+} - \sum_{i'=1}^{n} \boldsymbol{X}_{i'\cdot}^{-}; \boldsymbol{X}_{i'\cdot}^{+}\right) \mathrm{d}\gamma \\
&= \left(\int_0^1 \partial\rho\left(\gamma \boldsymbol{X}_{i\cdot}^{+} + \sum_{i' \in S} \boldsymbol{X}_{i'\cdot}^{+} - \sum_{i'=1}^{n} \boldsymbol{X}_{i'\cdot}^{-}\right) \mathrm{d}\gamma\right) \boldsymbol{X}_{i\cdot}^{+},
\end{aligned}$$

where the second equality follows from Cor. 24.2.1 of Rockafellar (1970). Therefore, the incremental risk is the product of $\boldsymbol{X}_{i\cdot}^{+}$ with an average of elements of $\mathcal{M}$. Because the weights on the incremental risks in Equation (4) are non-negative and sum to one, the Shapley value is also the product of $\boldsymbol{X}_{i\cdot}^{+}$ with an average of elements of $\mathcal{M}$. Just as for the Aumann-Shapley value, the conclusion follows from convexity of $\mathcal{M}$. □

4.3 Monotonicity

Friedman and Moulin (1999) describe a property called "demand monotonicity." Recall that loss in our context of risk attribution is analogous to demand in the context of cost allocation. To define a corresponding monotonicity property in our context, we need to expand the notation for risk in Equation (1) to depend explicitly on the loss matrix $\mathbf{X}$: write

$$r(\boldsymbol{\lambda}; \mathbf{X}) = \rho\left(\sum_{i=1}^{n} \lambda_i \boldsymbol{X}_{i\cdot}\right). \tag{16}$$

In words, a risk attribution scheme has the monotonicity property if the risk component attributed to a system component is non-decreasing in the loss generated by that system component.

Definition 3. *Let ϕ_i and $\tilde{\phi}_i$ represent the risk component attributed to system component i when performing risk attribution on the function $r(\cdot; \mathbf{X})$ and on the function $r(\cdot; \tilde{\mathbf{X}})$, respectively. The risk attribution scheme satisfies the* monotonicity property *if, for all $i = 1, \ldots, n$, $\tilde{\phi}_i \geq \phi_i$ whenever $\tilde{\mathbf{X}} - \mathbf{X}$ is non-negative and is zero in every row except the ith row.*

Friedman and Moulin (1999) prove that the Shapley value has their demand monotonicity property, but the Aumann-Shapley value does not have it, in general. In this section, we prove that the scenario-wise and overall non-negative Shapley values have our monotonicity property, under some conditions.

Proposition 6. *If Equation (1) holds and ρ is non-decreasing, then the scenario-wise non-negative Shapley value has the monotonicity property.*

Proof. Define $L_j^{\uparrow}(\boldsymbol{\lambda}; \mathbf{X}) = \sum_{i=1}^{n} \lambda_i X_{ij}^{+} - \sum_{i=1}^{n} X_{ij}^{-}$. Focus on a particular component i and let $\tilde{\mathbf{X}}$ be such that $\tilde{\boldsymbol{X}}_{i\cdot} \geq \boldsymbol{X}_{i\cdot}$ and $\tilde{\boldsymbol{X}}_{i'\cdot} = \boldsymbol{X}_{i'\cdot}$ for $i' \neq i$. For any scenario j, the difference

$$L_j^{\uparrow}(\boldsymbol{\lambda}; \tilde{\mathbf{X}}) - L_j^{\uparrow}(\boldsymbol{\lambda}; \mathbf{X}) = \begin{cases} \lambda_i(\tilde{X}_{ij}^{+} - X_{ij}^{+}) & \text{if } \tilde{X}_{ij}, X_{ij} \geq 0 \\ \lambda_i \tilde{X}_{ij}^{+} + X_{ij}^{-} & \text{if } \tilde{X}_{ij} \geq 0, X_{ij} \leq 0 \\ X_{ij}^{-} - \tilde{X}_{ij}^{-} & \text{if } \tilde{X}_{ij}, X_{ij} \leq 0 \end{cases}$$

is non-negative, non-decreasing in λ_i, and has no dependence on $\lambda_{i'}$ for $i' \neq i$. Consider a set S that does not contain i. Then $\rho(\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}})) = \rho(\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda}(S); \mathbf{X}))$ and $\rho(\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda}(S \cup \{i\}); \tilde{\mathbf{X}})) \geq \rho(\boldsymbol{L}^{\uparrow}(\boldsymbol{\lambda}(S \cup \{i\}); \mathbf{X}))$. Therefore the incremental risk used to compute the Shapley value $\Delta_i r^{\uparrow\uparrow}(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) \geq \Delta_i r^{\uparrow\uparrow}(\boldsymbol{\lambda}(S); \mathbf{X})$. Thus, the scenario-wise non-negative Shapley value for component i is no less when the loss matrix is $\tilde{\mathbf{X}}$ than when it is $\mathbf{X}$. □

Proposition 7. *If Equation (1) holds and ρ is non-decreasing, then the overall non-negative Shapley value has the monotonicity property.*

Proof. Define r_0 and T_0 by

$$r_0(\boldsymbol{\lambda}(S); \mathbf{X}) = \min_{S'} \left\{ \rho\left(\sum_{i' \in S'} \boldsymbol{X}_{i'\cdot} \right) : S \subseteq S' \right\} = \rho\left(\sum_{i' \in S} \boldsymbol{X}_{i'\cdot} + \sum_{i' \in T_0} \boldsymbol{X}_{i'\cdot} \right). \tag{17}$$

For $i \notin S$, define r_i and T_i by

$$r_i(\boldsymbol{\lambda}(S); \mathbf{X}) = \min_{S'} \left\{ \rho\left(\sum_{i' \in S'} \boldsymbol{X}_{i'\cdot} \right) : i \in S', S \subset S' \right\} = \rho\left(\boldsymbol{X}_{i\cdot} + \sum_{i' \in S} \boldsymbol{X}_{i'\cdot} + \sum_{i' \in T_i} \boldsymbol{X}_{i'\cdot} \right), \tag{18}$$

so that $\Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \mathbf{X}) = r_i(\boldsymbol{\lambda}(S); \mathbf{X}) - r_0(\boldsymbol{\lambda}(S); \mathbf{X})$. Let $\tilde{\mathbf{X}}$ be such that $\tilde{\boldsymbol{X}}_{i\cdot} \geq \boldsymbol{X}_{i\cdot}$ and $\tilde{\boldsymbol{X}}_{i'\cdot} = \boldsymbol{X}_{i'\cdot}$ for $i' \neq i$. Because ρ is non-decreasing,

$$r_0(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) \geq r_0(\boldsymbol{\lambda}(S); \mathbf{X}) \quad \text{and} \quad r_i(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) \geq r_i(\boldsymbol{\lambda}(S); \mathbf{X}). \tag{19}$$

Because Equation (18) has an extra constraint compared to Equation (17),

$$r_i(\boldsymbol{\lambda}(S); \mathbf{X}) \geq r_0(\boldsymbol{\lambda}(S); \mathbf{X}) \quad \text{and} \quad r_i(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) \geq r_0(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}). \tag{20}$$

We have

$$\Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \mathbf{X}) = \left\{ \begin{array}{rll} r_i(\boldsymbol{\lambda}(S); \mathbf{X}) - r_0(\boldsymbol{\lambda}(S); \mathbf{X}) & \text{if } r_0(\boldsymbol{\lambda}(S); \mathbf{X}) \geq 0 & \text{(case i)} \\ r_i(\boldsymbol{\lambda}(S); \mathbf{X}) & \text{if } r_0(\boldsymbol{\lambda}(S); \mathbf{X}) < 0 \leq r_i(\boldsymbol{\lambda}(S); \mathbf{X}) & \text{(case ii)} \\ 0 & \text{if } r_i(\boldsymbol{\lambda}(S); \mathbf{X}) < 0 & \text{(case iii)} \end{array} \right. .$$

By Equation (4), it suffices to prove that $\Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) \geq \Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \mathbf{X})$ for all i and $S \not\ni i$. Before analyzing the three cases, we establish the key inequality

$$r_i(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) - r_0(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) \geq r_i(\boldsymbol{\lambda}(S); \mathbf{X}) - r_0(\boldsymbol{\lambda}(S); \mathbf{X}). \tag{21}$$

Suppose that $r_i(\boldsymbol{\lambda}(S); \mathbf{X}) = r_0(\boldsymbol{\lambda}(S); \mathbf{X})$. Then Inequality (21) holds because of Inequality (20). Otherwise, i.e., supposing $r_i(\boldsymbol{\lambda}(S); \mathbf{X}) > r_0(\boldsymbol{\lambda}(S); \mathbf{X})$, we must have that T_0 in Equation (17) does not contain i. Let $\tilde{T}_0$ represent the value that T_0 takes on when $\tilde{\mathbf{X}}$ is substituted for $\mathbf{X}$ in Equation (17). Because of the properties of $\tilde{\mathbf{X}}$, $\tilde{T}_0$ also does not contain i. That is, if it is not optimal to hedge S using i when the loss matrix is $\mathbf{X}$, then it cannot be optimal to hedge S using i when the loss matrix is $\tilde{\mathbf{X}}$, which makes component i worse. Therefore $r_0(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) = r_0(\boldsymbol{\lambda}(S); \mathbf{X})$. This, together with Inequality (19), implies Inequality (21). This establishes Inequality (21).

- Case i: Inequality (19) implies that $r_0(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) \geq r_0(\boldsymbol{\lambda}(S); \mathbf{X}) \geq 0$, so $\Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) = r_i(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) - r_0(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}})$. By Inequality (21), $\Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) \geq \Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \mathbf{X})$.

- Case ii: Suppose that $r_0(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) < 0$. Then $\Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) = r_i(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) \geq r_i(\boldsymbol{\lambda}(S); \mathbf{X}) = \Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \mathbf{X})$, where the inequality is due to Inequality (19). Otherwise, i.e., supposing $r_0(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) \geq 0$, it follows that

$$\begin{aligned}
\Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) &= r_i(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) - r_0(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) \\
&\geq r_i(\boldsymbol{\lambda}(S); \mathbf{X}) - r_0(\boldsymbol{\lambda}(S); \mathbf{X}) \\
&\geq r_i(\boldsymbol{\lambda}(S); \mathbf{X}) = \Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \mathbf{X}),
\end{aligned}$$

  where the first inequality is Inequality (21).

- Case iii: $\Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) \geq 0 = \Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \mathbf{X})$.

In all cases, $\Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \tilde{\mathbf{X}}) \geq \Delta_i r^{\uparrow}(\boldsymbol{\lambda}(S); \mathbf{X})$, which is what needed to be shown. □

5 Conclusion

We introduced two methods for generating non-negative risk components with the Shapley value or Aumann-Shapley value: the scenario-wise non-negative method and the overall non-negative method. When applied to examples in firm-wide risk management and systemic risk management, they yielded very different risk components from each other and from the usual method of generating non-negative risk components. Like the usual non-negative method, the overall non-negative method allows profits in some scenarios to offset losses in other scenarios in computing a risk component. The scenario-wise non-negative method does not allow profits in some scenarios to mask losses in other scenarios, so it is useful for detecting all sources of risk. We showed that, under appropriate conditions, the risk components generated by the two proposed methods have desirable properties of partial separability, diversification, and monotonicity. In an example, we saw that the usual non-negative method lacks the property of partial separability. This disadvantage of the usual non-negative method is a reason to choose the overall non-negative method instead.

It would be valuable to have further research on other game-theoretic properties of the proposed methods, whether in a general setting or a specialized setting for a particular application. At this point, four methods for generating non-negative risk components are under consideration: the overall and scenario-wise non-negative methods, applied with either the Shapley or Aumann-Shapley values. This situation contrasts with that in portfolio risk attribution, where (if negative risk components are allowed), a consensus emerged that applying the Aumann-Shapley value to the risk function defined in Equation (1) is the single best method (Denault, 2001; Kalkbrener, 2005; Tasche, 1999, 2008). Would further considerations, perhaps appropriate to a particular application, support a unique choice of the best method for generating non-negative risk components?

Acknowledgements

We thank Eric Friedman for suggesting references and an anonymous referee for constructive suggestions that led to substantial improvements in the article.

References

Boonen, T., De Waegenaere, A., Norde, H., November 2012. A generalization of the Aumann-Shapley value for risk allocation problems. CentER discussion paper 2012-091, Tilburg University.

Chen, C., Iyengar, G., Moallemi, C. C., 2013. An axiomatic approach to systemic risk. Management Science 59 (6), 1373–1388.

Cont, R., Deguest, R., He, X. D., 2013. Loss-based risk measures. Statistics and Risk Modeling 30, 133–167.

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

Drehmann, M., Tarashev, N., 2013. Measuring the systemic importance of interconnected banks. Journal of Financial Intermediation 22 (4), 586–607.

Friedman, E., Moulin, H., 1999. Three methods to share joint costs or surplus. Journal of Economic Theory 87, 275–312.

Haimanko, O., 2001. Cost sharing: the nondifferentiable case. Journal of Mathematical Economics 35, 445–4623.

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

Koch-Medina, P., Moreno-Bromberg, S., Munari, C., 2013. Surplus-invariant capital adequacy tests and their risk measures.
URL http://ssrn.com/abstract=2348590

Koster, M., Tijs, S., Borm, P., 1998. Serial cost sharing methods for multi-commodity situations. Mathematical Social Sciences 36 (1), 229–242.

Liu, M., Staum, J., January 2011. Systemic risk components and deposit insurance premia. Working paper 2011-03, Center for Financial Research, Federal Deposit Insurance Corporation.
URL http://ssrn.com/abstract=1982131

McNeil, A. J., Frey, R., Embrechts, P., 2005. Quantitative Risk Management. Princeton University Press, Princeton, New Jersey.

Mertens, J.-F., 1988. The Shapley value in the non differentiable case. International Journal of Game Theory 17 (1), 1–65.

Moulin, H., Sprumont, Y., 2007. Fair allocation of production externalities: recent results. Revue d'Économie Politique 117 (1), 7–36.

Rockafellar, R. T., 1970. Convex Analysis. Princeton University Press, Princeton, New Jersey.

Rockafellar, R. T., Uryasev, S., Zabarankin, M., 2006. Generalized deviations in risk analysis. Finance and Stochastics 10, 51–74.

Ruszczyński, A., Shapiro, A., 2006. Optimization of convex risk functions. Mathematics of Operations Research 31, 433–452.

Staum, J., 2012. Systemic risk components and deposit insurance premia. Quantitative Finance 12 (4), 651–662.

Staum, J., 2013. Excess invariance and shortfall risk measures. Operations Research Letters 41, 47–53.

Tasche, D., 1999. Risk contributions and performance measurement, working paper, Technische Universität München.

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

Tasche, D., 2008. Capital allocation to business units and sub-portfolios: the Euler principle, available at arXiv:0708.2542v3.