Organising the allocation

Capital allocation is an important quantitative methodology that affects a bank's risk-taking behaviours and business strategies. Yadong Li, Marco Naldi, Jeffrey Nisen and Yixi Shi propose a new capital allocation method that offers stability, fairness, computational efficiency and better business incentives, when compared with existing methods

In the wake of the financial crisis of 2008, regulators around the globe started to strengthen regulatory capital requirements, aiming to limit systemic risk by reining in the aggressive risk-taking behaviours of banks. Under the tightened regulations, banks are required to raise more capital than in the precrisis era to support similar business activities. In response to the growing regulatory capital pressure, banks are aggressively repositioning themselves by diverting resources to business areas with the highest return on capital (ROC). This is in contrast to the precrisis era, when resources were primarily allocated according to revenue expectations.

The capital of a bank is computed, reported and regulated at the whole portfolio level; the capital of the whole bank is not a simple sum of the standalone capital of individual business units within the bank due to diversification effects. Therefore, in order to compute the ROC of individual business units within a bank, the total capital of the bank has to be allocated to individual business units in a strictly additive manner. This is why capital allocation has become an important quantitative method that directly affects banks' business strategies.

As illustrated in figure 1, a bank's trading book is typically organised as a tree-like hierarchy, where the nodes of the tree represent trading units such as business divisions, desks and books. The profit and loss (P&L) and risk of a node are the sum of the P&L and the risks of its descendents (subportfolios). The trading book hierarchy of a bank is always well defined and remains largely stable for management, reporting and attribution purposes.

We use $\vec{x} = \sum_i \vec{x}_i$ to denote a bank's whole trading portfolio, where $\vec{x}_i$ represents the notional vector of the ith trading unit's position over the whole universe of tradable instruments. A trading unit could be a business division, a trading desk, a book or even an individual trade at the lowest level. We use $c(\cdot)$ to denote a risk capital cost function for a portfolio, which can be either regulatory or economic capital. With this choice of notation, $c(\vec{x}_i)$ is therefore a business unit's standalone capital, ie, the capital if the business unit was an independent company. In general, the bank's total capital is smaller than the sum of the standalone capital of its trading units, ie, $c(\vec{x}) = c(\sum_i \vec{x}_i) < \sum_i c(\vec{x}_i)$, because of diversification and hedging effects between different business units. The bank's total diversification benefits are therefore $\sum_i c(\vec{x}_i) - c(\vec{x})$.

We use ξ_i to denote the risk capital allocation to the ith trading unit. By definition, the allocations have to sum up to the overall capital of the bank: $c(\vec{x}) = \sum_i \xi_i$. The difference between a trading unit's standalone capital and the allocated capital, $c(\vec{x}_i) - \xi_i$, is the diversification benefit allocated to that unit. Therefore, the allocation methodology, at its core, determines the distribution of diversification benefits among trading units.

One of the primary purposes of the allocated capital ξ_i is to compute the ROC of individual trading units, which is the primary measure for evaluating and comparing their relative performances. The ROC of a trading unit is defined as its revenue over its allocated capital ξ_i.[1] The ROC calculation has to use the allocated capital ξ_i instead of the standalone capital $c(\vec{x}_i)$ in order to capture the hedging benefits. For example, a capital intensive business unit might not have a sustainable ROC when evaluated as a standalone entity, but it might provide valuable hedging benefits to the rest of the bank's portfolio; hence, its ROC could be attractive when evaluated within a bank's whole portfolio using the allocated capital ξ_i.

Given that the ROC is tied to the compensation, reward and resources allocated to a trading unit, the allocated capital is actively managed and optimised by all business units of a bank. Consequently, it is of critical importance to study and understand the business incentives set forth by allocation methodologies. A good capital allocation methodology should promote positive behavioural changes by rewarding the right risk-taking within a bank.

In this article, we first review some commonly used allocation methodologies and highlight their limitations when applied to the risk capital allocation problem of a bank. We then propose a new allocation method, the constrained Aumann-Shapley (CAS) allocation, which addresses these shortcomings. Finally, we conclude by showing some numerical results for value-at-risk attribution using CAS allocation, and we discuss its business incentives in practice.

Review of allocation methodologies

Arguably the simplest of all allocation methods is the standalone pro rata allocation, also known as standalone allocation, in which a business unit's allocation is proportional to its standalone risk capital:

$$\xi_i = \frac{c(\vec{x})}{\sum_i c(\vec{x}_i)} c(\vec{x}_i)$$

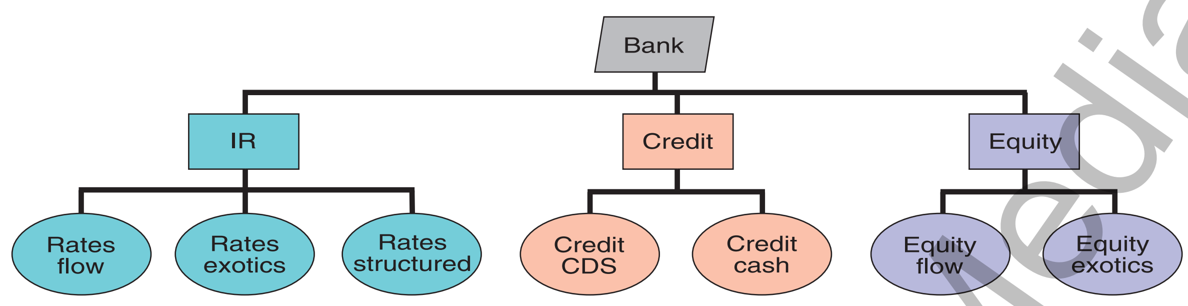
and $c(\vec{x}) / \sum_i c(\vec{x}_i)$ represents a normalisation factor to ensure additivity, $c(\vec{x}) = \sum_i \xi_i$.

The Shapley allocation, which originated from co-operative game theory, is another well-known allocation method. It has been rigorously proven that the Shapley allocation is the only fair allocation of the surplus from a collaborative game among players, under an axiomatic definition of fairness (Shapley 1953). Each player's Shapley allocation equals the average of their incremental contributions to the surplus, where the average is taken over all possible permutations of players. In our context, the 'surplus' is the risk capital and the 'players' are the business units.

We use a stylised example of replacement cost (RC) to highlight the key features of different allocation methods. The RC is one of the capital components in the Basel III leveraged balance sheet (Basel Committee on Banking Supervision 2014); it is defined as the total present value

[1] *If the allocated capital is negative, ROC is not well defined; however, it is uncommon in practice for a business unit to have negative allocated capital.*

1 A bank's trading book hierarchy



A. Replacement cost allocation

Allocation method	A	B	C	Total
Standalone pro rata	4	8	0	12
Shapley	10	16	−14	12
Euler/Aumann-Shapley	12	24	−24	12
C-Shapley/CAS	12	15	−15	12

B. Shapley allocation calculation

Permutations	Cumulative RC	A	B	C
A=12, B=24, C=−24	12, 36, 12	12	24	−24
A=12, C=−24, B=24	12, 0, 12	12	12	−12
B=24, C=−24, A=12	24, 0, 12	12	24	−24
C=−24, B=24, A=12	0, 0, 12	12	0	0
B=24, A=12, C=−24	24, 36, 12	12	24	−24
C=−24, A=12, B=24	0, 0, 12	0	12	0
Averages		**10**	**16**	**−14**

C. Associativity

	Bottom up		Top down	
Allocation method	**A**	**B+C**	**A**	**B+C**
Standalone pro rata	4	8	12	0
Shapley	10	2	12	0
Euler/Aumann-Shapley	12	0	12	0
C-Shapley/CAS	12	0	12	0

(PV) of all trades against a given counterparty floored at 0, ie, $\mathrm{RC}(\vec{y}) = \max(\vec{y}\cdot\vec{v}, 0)$, where $\vec{y}$ is the portfolio's notional vector, and $\vec{v}$ is the price vector of all the tradable instruments on the market. The dot product of $\vec{y}\cdot\vec{v}$ is therefore the portfolio's total PV. The RC captures the cost to replace the counterparty's portfolio in the event of its default.

Suppose there are only three trades in a bank's portfolio facing a given counterparty, with PVs of ($\mathrm{A} = 12$, $\mathrm{B} = 24$, $\mathrm{C} = -24$); the bank's total RC to the counterparty is therefore $\max(\mathrm{A} + \mathrm{B} + \mathrm{C}, 0) = 12$. Table A shows the resulting standalone and Shapley allocations to the three trades as well as the other allocation methods we will introduce later. Table B shows all six permutations of the three trades underpinning the Shapley allocation and each trade's corresponding incremental contribution permutation-wise.

One disadvantage of Shapley allocation is its computational cost: it is certainly not feasible to exhaustively enumerate the massive number of possible permutations in any real-world capital allocation problem. A Monte Carlo simulation is required to implement Shapley allocation in practice, where only a small fraction of all possible permutations are sampled. The computational cost of the Monte Carlo implementation of Shapley allocation for large portfolios can be prohibitively expensive, which is one of the factors that has prevented its wide adoption.

The replacement cost example also reveals a common flaw in both standalone and Shapley allocation methods when applied to banks' risk capital allocation. Consider the case in which B and C are two trades of the same desk, and A belongs to another desk. The allocation to the desk $\mathrm{B} + \mathrm{C}$ can be computed in two ways: the first is a bottom-up approach that sums up the allocations to its two trades, as given in table A; the second is a top-down approach that treats $\mathrm{B} + \mathrm{C}$ as an atomic unit and directly applies the allocation method to the two desks with PVs of ($\mathrm{A} = 12$, $\mathrm{B} + \mathrm{C} = 0$). If these two approaches yield identical results, we say that the allocation algorithm is associative. Table C shows that neither the standalone pro rata nor the Shapley allocation are associative.

Associativity is of critical importance for banks' risk capital allocation: it ensures that the allocation to a trading unit remains unchanged, regardless of how the trades are booked or grouped within the unit. Without associativity, an allocation method is vulnerable to potential manipulations via trade rebookings. For instance, considering the bottom-up, standalone pro rata allocation in table C, the manager of desk $\mathrm{B} + \mathrm{C}$ could reduce their overall allocation from 8 to 0 by rebooking B and C into a single trade, without changing the economics of their positions.

Fairness is a strong theoretical advantage of the Shapley allocation. In search of a suitable risk capital allocation method for banks, we took the approach of modifying the Shapley allocation to enforce associativity, while preserving its fairness advantages. One known method to make Shapley allocation associative is to define the allocation to be the sum of allocations to the smallest granularity possible, thus removing the ambiguity between top down and bottom up. The result of this approach is the Aumann-Shapley allocation (Aumann & Shapley 1974; Denault 2001).

The Aumann-Shapley allocation is a continuous limit of the Shapley allocation, where the allocation units are trades with infinitesimal notional amounts. The Aumann-Shapley allocation is commonly expressed as the allocation price $\vec{u}$, which is the allocation per unit notional amount:

$$\vec{u} = \int_0^1 \frac{\partial c(q\vec{x})}{\partial(q\vec{x})}\,\mathrm{d}q \tag{1}$$

where q is a scalar between 0 and 1, $q\vec{x}$ represents a fraction of the bank's whole portfolio and the partial derivative is a short-hand notation for the gradient of:

$$\frac{\partial c(q\vec{x})}{\partial(q\vec{x})} = \nabla c(q\vec{x}) = \left(\frac{\partial c(q\vec{x})}{\partial(qx_1)}, \frac{\partial c(q\vec{x})}{\partial(qx_2)}, \dots, \frac{\partial c(q\vec{x})}{\partial(qx_n)}\right)$$

D. Incentives for risk reduction		
	A=12, B=−10, RC=2	**A=12, B=−14, RC=0**
Standalone pro rata	A=2, B=0	A=0, B=0
Aumann-Shapley/Euler	A=12, B=−10	A=0, B=0
Shapley/C-Shapley/CAS	A=7, B=−5	A=6, B=−6

which represents the marginal contribution of an infinitesimal trade to the risk capital when we hold a fraction q of the whole bank's portfolio $(q\vec{x})$. The Aumann-Shapley allocation to the ith business unit is therefore the dot product $\xi_i = \vec{u} \cdot \vec{x}_i$.

A key characteristic of the Aumann-Shapley allocation is its total disregard for the trading books' organisational information, as it only depends on the bank's total portfolio vector $\vec{x}$. No matter how the bank's trading book is organised, the allocation price for a given instrument is identical under the Aumann-Shapley allocation.

The additivity of the Aumann-Shapley allocation can be proven as:

$$\sum_i \xi_i = \sum_i \vec{u} \cdot \vec{x}_i = \vec{u} \cdot \vec{x} = \int_0^1 \frac{\partial c(q\vec{x})}{\partial (q\vec{x})} \cdot \vec{x}\, \mathrm{d}q$$
$$= \int_0^1 \frac{\partial c(q\vec{x})}{\partial (q\vec{x})} \cdot \frac{\partial (q\vec{x})}{\partial q}\, \mathrm{d}q$$
$$= \int_0^1 \mathrm{d}c\,(q\vec{x}) = c(\vec{x}) - c(\vec{0}) = c(\vec{x})$$

The last equality is because the risk capital of an empty portfolio is zero.

The Aumann-Shapley allocation is constructed to be associative; this directly follows the associativity of the vector summation, because the allocation is linear with respect to the position vector $\vec{x}_i$.

If the risk capital cost function is homogeneous in trade notional to the first order, ie, $c(q\vec{x}) = qc(\vec{x})$, and differentiable, then the Aumann-Shapley allocation reduces to the Euler allocation (Tasche 2008):

$$\vec{u} = \int_0^1 \frac{\partial c(q\vec{x})}{\partial (q\vec{x})}\, \mathrm{d}q = \int_0^1 \frac{\partial c(\vec{x})}{\partial \vec{x}}\, \mathrm{d}q = \frac{\partial c(\vec{x})}{\partial \vec{x}} \qquad (2)$$

This applies to the case of replacement cost, which is a homogeneous cost function:

$$\vec{u} = \frac{\partial \mathrm{RC}(\vec{x})}{\partial \vec{x}} = \frac{\partial \max(\vec{x} \cdot \vec{v}, 0)}{\partial \vec{x}} = \mathbf{1}(\vec{x} \cdot \vec{v} > 0)\vec{v}$$

where $\mathbf{1}(\cdot)$ is the indicator function. The Aumann-Shapley (or equivalently Euler) allocation of RC to the individual business unit or trade is therefore $\xi_i = \vec{u} \cdot \vec{x}_i = \mathbf{1}(\vec{x} \cdot \vec{v} > 0)\vec{v} \cdot \vec{x}_i$, ie, the allocation of RC is the trade's PV if the bank's overall portfolio against the counterparty is in the money, and zero otherwise. Tables A and C confirm that the Aumann-Shapley allocation is indeed additive and associative in the case of RC. Another advantage of the Aumann-Shapley allocation is its computational efficiency, as it only requires partial derivatives and the integral in (1); there is no need to run an expensive Monte Carlo simulation.

Despite these desirable properties, the Aumann-Shapley allocation could lead to counterintuitive results or even perverse incentives in practice. Let us consider another example of the RC allocation, as shown in table D, where two desks' positions offset each other against a counterparty: desk A's total mark to market (MtM) is 12, and desk B's total MtM is -10, for a total RC of $\max(12-10, 0) = 2$. Using Aumann-Shapley/Euler, the allocations to the desks are equal to their PV of (A $= 12$, B $= -10$). To encourage the reduction in overall capital, a bank often establishes a cost of capital measure, so that the desk with positive capital allocation (A in our example) effectively compensates the desk with negative allocation (B in our example). Now, let us suppose that desk B further reduces its MtM to the same counterparty to -14, expecting a bigger compensation; the Aumann-Shapley allocation, however, becomes (A $= 0$, B $= 0$), thus discouraging such a capital reduction initiative by B.

Such counterintuitive behaviour in the Aumann-Shapley allocation is due to its negligence of banks' organisational structures; therefore, when applied to a discrete business unit, it may not yield sensible results in practice, even though it is theoretically fair for infinitesimal trades.

Constrained Aumann-Shapley

In the previous section, we reviewed several well-known allocation methodologies and showed that none of them are satisfactory when applied to the risk capital allocation of a bank.

A closer look at the permutations behind the Shapley allocation in table B reveals the associativity is broken by a single permutation of (C $= -24$, A $= 12$, B $= 24$), which gives B + C an allocation of 12; this does not match the allocation of zero from the top-down method. This permutation, however, would not be admissible under a bank's organisational constraints. In practice, trades cannot be moved freely across desks and business boundaries within a bank. For example, it is not possible to move a credit default swap (CDS) trade from a credit desk to an interest rate desk without violating the desks' business mandates and management protocols. We observe that if we remove the two permutations (highlighted in yellow) that mix the positions of the two desks in table B, then the Shapley allocation would in fact become associative.

Observing that the flaws in the Shapley and Aumann-Shapley allocations are both caused by disregarding the bank's organisational constraints, we propose the following modification to the Shapley allocation.

Constrained Shapley (C-Shapley) allocation: the allocation of risk capital to a trade or a portfolio equals its average incremental contribution, taken over all possible permutations of trades that are admissible under the organisational constraints.

The organisational constraints for admissible permutations can be captured by two simple rules:

- all permutations are admissible for nodes with the same parent;
- a branch (ie, a node and all of its descendents) has to be permuted as a whole.

For example, in figure 1, the permutation of rates exotics, credit cash, equity flow, etc, is not an admissible permutation because it mixes the desks belonging to different parent businesses.

Let us denote by 'leaf nodes' the smallest trading units in the organisational hierarchy, where the P&L, risk and capital are independently managed. Those nodes roughly correspond to the lowest-level books in a bank, of which there can be thousands in a large global investment bank. The manager of a leaf node usually has the authority to take various trading actions, including rebooking, entering/exiting trades and hedging, etc. In the fictitious hierarchy illustrated in figure 1, the ovals are the leaf nodes.

It is straightforward to show that the C-Shapley allocation is both additive and associative; the associativity is ensured by the constraint of admissible permutation. Tables A, C and D also include the C-Shapley allocation results, which are more sensible than other methods in all the examples considered; for example, this allocation produces correct incentives for risk reduction in table D.

The manager of a leaf node can and often does make trading and hedging decisions at arbitrary notional amounts. Therefore, Aumann-Shapley's abstraction of infinitesimal trades is appropriate for capital allocation within a leaf node. In addition, the continuous extension brings much needed computational efficiency to the C-Shapley allocation, which is similar to the computational gain of the Aumann-Shapley allocation over the Shapley allocation. This motivates us to combine the C-Shapley with the Aumann-Shapley allocation, and the resulting allocation method is referred to as the constrained Aumann-Shapley (CAS) allocation.

Considering the allocation to leaf node B under a given admissible permutation, we use $\vec{x}_{\mathrm{A}}$ to denote the union of all the leaf nodes' portfolios preceding the leaf node B in this permutation. Conditioned on the preceding portfolio $\vec{x}_{\mathrm{A}}$, the allocation price for trades in node B is:

$$\vec{u}(\vec{x}_{\mathrm{B}} \mid \vec{x}_{\mathrm{A}}) = \int_0^1 \frac{\partial c(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}})}{\partial(q\vec{x}_{\mathrm{B}})}\,\mathrm{d}q \tag{3}$$

where $c(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}})$ is the risk capital of the portfolio A plus a fraction q of the leaf node B, with $0 < q < 1$. Taking the expectation over all admissible permutations of the leaf nodes (ie, all possible $\vec{x}_{\mathrm{A}}$), we arrive at the unconditional allocation price of the trades in leaf node B:

$$\vec{u}_{\mathrm{B}} = \mathbb{E}[\vec{u}(\vec{x}_{\mathrm{B}} \mid \vec{x}_{\mathrm{A}})] \tag{4}$$

Here, the expectation is the same as the arithmetic average. The allocation to the leaf node B as a whole is therefore the dot product of $\xi_{\mathrm{B}} = \vec{u}_{\mathrm{B}} \cdot \vec{x}_{\mathrm{B}}$.

Unlike the Aumann-Shapley allocation, the allocation price of CAS captures the full organisational constraints; the same instrument, depending on which leaf node it belongs to, has different allocation prices under the CAS method. This is a compelling advantage of CAS, as it recognises that the same instrument could play different roles in different trading units. For example, an interest rate (IR) swap hedge in the credit desk should have a different allocation price from the same IR swap used by the rates desk to take a speculative bet, as the former decreases the firm's overall risk and the latter increases it.

The CAS allocation method is additive and associative by construction: additivity can be proven by noting that trade level allocations within a leaf node always add up to its C-Shapley allocation under any admissible permutation:

$$\begin{aligned}\vec{u}(\vec{x}_{\mathrm{B}} \mid \vec{x}_{\mathrm{A}}) \cdot \vec{x}_{\mathrm{B}} &= \int_0^1 \frac{\partial c(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}})}{\partial(q\vec{x}_{\mathrm{B}})} \cdot \vec{x}_{\mathrm{B}}\,\mathrm{d}q \\ &= \int_0^1 \frac{\partial c(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}})}{\partial(q\vec{x}_{\mathrm{B}})} \cdot \frac{\partial(q\vec{x}_{\mathrm{B}})}{\partial q}\,\mathrm{d}q \\ &= \int_0^1 \mathrm{d}c\,(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}}) = c(\vec{x}_{\mathrm{A}} + \vec{x}_{\mathrm{B}}) - c(\vec{x}_{\mathrm{A}})\end{aligned} \tag{5}$$

The associativity of CAS directly follows that of the C-Shapley allocation, because the CAS and C-Shapley allocations are identical for all nodes in the trading book hierarchy; they only differ in the allocation to individual trades within leaf nodes.

Similar to the C-Shapley allocation, the CAS allocation can be implemented by a Monte Carlo simulation, where only a small fraction of the admissible permutations of leaf nodes are sampled randomly. For each admissible permutation, the conditional unit price given in (3) can be computed via numerical integration, by consecutively adding small portions of portfolio B to the preceding portfolio A. Since CAS does not require Monte Carlo sampling of individual trades' permutations, it is much more efficient than the C-Shapley allocation to the trade level.

The efficiency of the Monte Carlo simulation of the CAS allocation can be improved much further if the integrand in (3) bears the following separable form:

$$\frac{\partial c(\vec{x}_{\mathrm{A}} + q \cdot \vec{x}_{\mathrm{B}})}{\partial(q\vec{x}_{\mathrm{B}})} = S\vec{w}(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}}) \tag{6}$$

where S is a matrix that captures the set of characteristics pertaining to the individual instruments, such as trade type, maturity, reference entity or historical P&L vectors; $\vec{w}(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}})$ is a weighting vector that only depends on the features of the aggregated portfolio. The $S\vec{w}(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}})$ is the matrix product between S and vector $\vec{w}$. When (6) holds, the unconditional CAS allocation price in (4) is also separable:

$$\begin{aligned}\vec{u}_{\mathrm{B}} = \mathbb{E}[\vec{u}(\vec{x}_{\mathrm{B}} \mid \vec{x}_{\mathrm{A}})] &= \mathbb{E}\biggl[\int_0^1 \frac{\partial c(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}})}{\partial(q\vec{x}_{\mathrm{B}})}\,\mathrm{d}q\biggr] \\ &= \mathbb{E}\biggl[\int_0^1 S\vec{w}(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}})\,\mathrm{d}q\biggr] \\ &= S\mathbb{E}\biggl[\int_0^1 \vec{w}(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}})\,\mathrm{d}q\biggr]\end{aligned} \tag{7}$$

Since the instrument-specific matrix appears as a separable factor out of the expectation in (7), we only need to run Monte Carlo simulation to compute the average weight $\vec{w}_{\mathrm{B}} := \mathbb{E}[\int_0^1 \vec{w}(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}})\,\mathrm{d}q]$, using just the portfolio-level information. With (7), we no longer need to track individual trades in the Monte Carlo simulation; therefore, the order of computational and storage costs reduces from the number of trades to the number of leaf nodes (books), often leading to a thousandfold performance gain in practice.

The CAS allocation is generic, and it is applicable to many risk capital allocation problems. Moreover, the vast majority of risk capital metrics are in the separable form of (6), allowing for very efficient numerical implementation. For example, all variations of VAR, such as incremental risk charge (IRC) and credit valuation adjustment (CVA) VAR, are separable; so are the expected shortfall (ES) and sensitivity-based approach (SBA) under the upcoming *Fundamental Review of the Trading Book* (FRTB) regulation (Basel Committee 2016).

Application to VAR

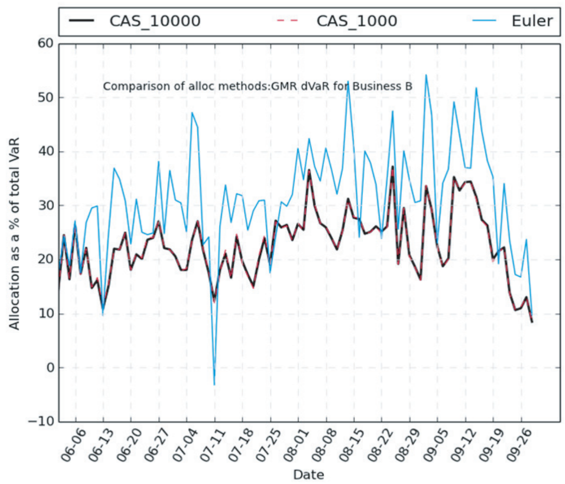
VAR is the most common risk capital metric in practice. Since VAR, being a quantile function, is positive 1-homogeneous, its Aumann-Shapley allocation is equivalent to the Euler allocation (often referred to as the conditional VAR (CoVAR) method). In this section, we present the numerical results of VAR allocation using CAS and compare them with the Euler allocation.

In the case of VAR, the integrand in (3) can be expressed in the following conditional expectation form (Gourieroux, Laurent & Scaillet 2000):

$$\frac{\partial \mathrm{VAR}(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}})}{\partial(q\vec{x}_{\mathrm{B}})} = \mathbb{E}[\vec{\delta v} \mid (\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}}) \cdot \vec{\delta v} = \mathrm{VAR}(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}})] \tag{8}$$

where $\vec{\delta v}$ is the P&L vector of all the tradable instruments and should not be confused with the price vector $\vec{v}$ for the RC. Typically, (8) is estimated using the kernel method (Epperlein & Smillie 2006), which yields

2 GMR DVAR allocations over time



a convenient separable form of (6). Specifically, under the kernel method, the matrix S comprises the P&L vectors of all the tradable instruments, and $\vec{w}(\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}})$ is the Gaussian kernel weights for the portfolio $\vec{x}_{\mathrm{A}} + q\vec{x}_{\mathrm{B}}$. By taking advantage of the performance gain from the separability, the numerical implementation of the CAS is extremely efficient for VAR. It takes less than a minute to allocate the VAR of Barclays Investment Bank to millions of individual trades, using a single PC.

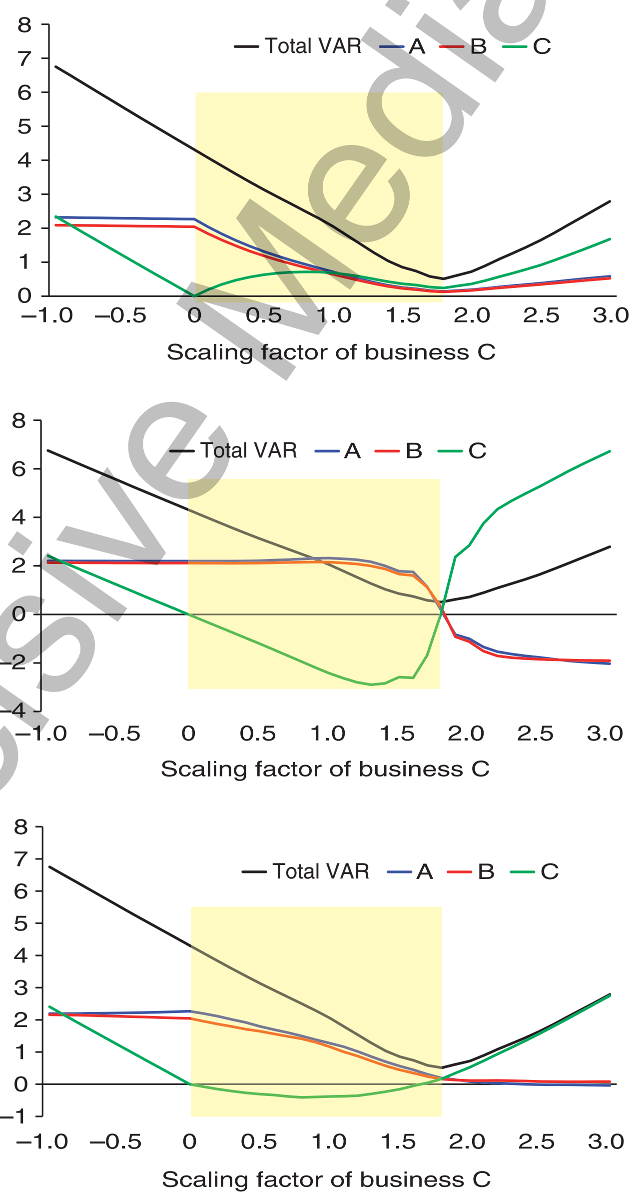
The CAS allocation of VAR is expected to be more stable than the Euler allocation, as the CAS allocation is effectively an average of an integral (as in (4)), while the Euler allocation is a derivative. Figure 2 is a real-world example of general market risk (GMR) DVAR allocation of a business unit in Barclays. The result is shown as a percentage of the firm's total VAR. The CAS allocations are computed using Monte Carlo simulation with 1,000 and 10,000 paths; the results suggest that simulation with 1,000 paths achieved excellent convergence. In addition, figure 2 confirms superior stability of the CAS allocation over the Euler allocation. For example, this business unit's Euler allocation plummeted from 50% to below 0% of the firm's overall GMR DVAR within a few days in early July 2013, while the CAS allocation only experienced a 10% swing during the same period. The better stability of CAS over Euler is also observed for other risk capital measures, such as ES.

Business incentives

As discussed earlier, ROC has become the primary measure of business units' performances within a bank. Under such a paradigm, a business unit is more focused on managing its own risk capital allocation than on the firm's overall risk. There is little incentive for a business unit to actively reduce the firm's overall risk capital unless doing so also reduces its own capital allocation. Therefore, it is extremely important for an allocation method to give incentives for individual business units to act collectively toward the reduction of the firm's overall risk and capital.

We use a stylised VAR allocation example to study the incentives from the three allocation methods we examined: standalone pro rata, Aumann-Shapley/Euler and C-Shapley/CAS. We consider a bank consisting of three businesses [A, B, C]. The joint return distribution of the businesses follows a multivariate normal with standard normal marginals and pairwise

3 Business incentives of allocation methods: top, standalone; middle, Aumann-Shapley/Euler; bottom, C-Shapley/CAS



A scaling factor < 0, $= 0$ or > 0 implies a portfolio that is short, not invested and long in business C, respectively.

correlations of $\mathrm{corr}(\mathrm{A}, \mathrm{B}) = 0.65$, $\mathrm{corr}(\mathrm{A}, \mathrm{C}) = \mathrm{corr}(\mathrm{B}, \mathrm{C}) = -0.9$ (ie, business C acts as a strong hedge to businesses A and B). Holding A and B unchanged, we scale business C by different factors to mimic its growth and downsize. Though an analytical solution exists for Euler allocation for this simple example, we used numerical solutions for all three allocation methods to mimic a realistic setting in practice.

Figure 3 shows the bank's overall VAR and the businesses' allocations in response to the scaling of C. The yellow areas, spanning from a zero scaling factor to the minimum of the firm's total VAR, highlight the range where growing business C reduces the overall VAR of the bank. To the left of this region, the bank is taking the opposite position in C (negative scaling factor). To the right of the region, business C is scaled up so much that it becomes the dominant risk contributor to the bank; therefore, continuing its growth no longer reduces risk.

In the yellow region, allocations to C are mostly negative under C-Shapley/CAS and Aumann-Shapley/Euler, meaning that A and B have to compensate for C offsetting their risks. As such, C possesses a strong incentive to scale up. In contrast, the standalone pro rata method produces a positive allocation to C in the yellow region, incorrectly incentivising C to cut back.

Another interesting observation from figure 3 relates to how the hedging benefit is shared among all businesses. Under C-Shapley/CAS, the growth of business C and the resulting reduction in the bank's VAR in the yellow region is accompanied by a simultaneous reduction in the allocations to A and B. On the contrary, the hedging benefit is largely monopolised by C under Aumann-Shapley/Euler: A's and B's allocations remain unchanged for the majority of the yellow region. Given that the hedging benefit is the fruition of collaborations among multiple business units, it is fairer and more sensible to distribute the hedging benefits among all participating business units. A side effect of Aumann-Shapley/Euler's attributing the hedging benefits only to the risk-reducing business (C in our example) is a more dramatic change in allocations when the roles of hedge and primary risk switch, as can be seen near the upper bound of the yellow region under the Aumann-Shapley/Euler allocation. The resulting allocation scheme is very sensitive to the portfolio compositions, and it also poses difficulties in risk and capital management in practice. In contrast, the C-Shapley/CAS allocations change smoothly across the full range of scaling.

Also notable from figure 3 is that both C-Shapley/CAS and Aumann-Shapley/Euler recognise C as the main driver of the bank's risk to the right of the yellow region: most of the increase in the total VAR is allocated to C, while the allocations to A and B remain largely unchanged. The C-Shapley/CAS allocation outperforms the Aumann-Shapley/Euler, as its allocation to A and B turns flat much faster to the right of the yellow region. In this regard, the standalone pro rata allocation is counterintuitive, as it redistributes the increase of the bank's VAR among all three businesses, when in fact C is the sole driver of the total risk.

We used a normal return distribution solely for ease of exposition. Similar analyses were performed using skewed, heavy-tailed return distributions as well as those from real trading portfolios, and the aforementioned behaviours are very much consistent.

Conclusion

The C-Shapley/CAS allocation explicitly captures the organisational constraints of a bank's trading book; these are overlooked by several well-known allocation methodologies, but they are of critical importance for the risk capital allocation in a bank. As a result, the C-Shapley/CAS allocation scheme is well behaved and has desirable properties, such as stability and associativity. The associativity of CAS ensures that no business unit can manipulate its capital charge by creatively rebooking its trades. CAS also provides the right incentive structure for individual business units to reduce the firm's overall risk.

The proposed CAS methodology is computationally efficient for most common types of risk capital metrics, which adds to its appeal as a potential standard allocation method for all risk capital metrics within a bank, including the banking book capital. In addition, CAS is applicable to the allocation of other types of operating costs in a bank, such as the cost associated with liquidity reserve or the initial margin. The CAS allocation is also a good candidate for multi-strategy buy-side firms to allocate diversification benefits among different trading strategies, and to evaluate their returns per unit risk with the inclusion of diversification benefits.

Although there are good reasons to believe the better business incentive offered by CAS could result in better resource allocation, and, thus, better returns for banks' shareholders, directly providing evidence for such effects is difficult. We leave this for future research. ■

Yixi Shi currently works for Magnetar Capital, and the other authors work for Barclays Corporate and Investment Bank. The views expressed in this article are the authors' own; they do not necessarily represent those of their current or previous employers.
We thank Yury Korsky, Barry Mcquaid, Pierre Micottis, Simon O'Callaghan, Vladimir Piterbarg, Ariye Shater, Jonathan Taylor and Katherine Hartl for numerous insightful discussions. We are also grateful to two anonymous referees for useful comments.
Email: yadong.li@barclays.com, marco.naldi@barclays.com, jeffrey.nisen@barclays.com, yixi.shi@magnetar.com.

REFERENCES

Aumann R and L Shapley, 1974
Values of Non-Atomic Games
Princeton University Press

Basel Committee on Banking Supervision, 2014
Basel III leverage ratio framework and disclosure requirements
Report, Bank for International Settlements

Basel Committee on Banking Supervision, 2016
Minimum capital requirement for market risk
Report, Bank for International Settlements

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

Epperlein E and A Smillie, 2006
Cracking VAR with kernels
Risk August, pages 70–74

Gourieroux C, JP Laurent and O Scaillet, 2000
Sensitivity analysis of value at risk
Journal of Empirical Finance 7(3–4), pages 225–245

Shapley L, 1953
A value for N-person games
In *Contributions to the Theory of Games*, volume 2, pages 307–318
Annals of Mathematics Studies, volume 28
Princeton University Press

Tasche D, 2008
Capital allocation to business units and sub-portfolios: the Euler principle
In *Pillar II in the New Basel Accord: The Challenge of Economic Capital*, pages 423–453
Risk Books