Capital Allocation and its Discontents

by Daniel Bauer and George Zanjani

I. Introduction

Few areas of academic inquiry can claim so inauspicious a birth as the theory of capital allocation. Merton and Perold (1993) observed that allocation was generally “not feasible,” while Phillips, Cummins, and Allen (1998) deemed allocation to be “inappropriate” for insurance companies. Yet, despite such dim pronouncements from the halls of Harvard and Wharton, the capital allocation literature blossomed in the first decade of the new millennium.

To a newcomer, the propagation of the literature may be hard to understand. The arguments of Merton and Perold (1993) and Phillips, Cummins, and Allen (1998) were well-founded and continue to resurface from time to time in skeptical articles. Capital allocation continues, however, because of the practical need: Pricing and performance measurement within insurance companies and other financial institutions is not possible in current practice without some allocation of capital---whether implicit or explicit.

This chapter starts by reviewing the rationale for capital allocation, as well as its limitations. Once we have established the justification for allocation, we then review the methods. It is here where the literature becomes diffuse, with many potential approaches to choose from. We focus most of our attention on what can fairly be called the mainstream approach to allocation—the gradient or Euler method (an allocation also implied by game theoretic approaches)—due not only to its widespread acceptance but also because of the respect it pays to the concept of marginal cost. Defining marginal cost, however, is a tricky enterprise—and, after presenting an example from life insurance, we conclude by discussing this weakness of capital allocation approaches along with future directions for research.

II. Allocation Defined

Consider an insurance company with capital $K$, assets $A$, and a set of $N$ exposures denoted by $q_1$ for $i = 1, \ldots, N$. The variable $q_i$ quantifies the extent to which a company is exposed to the $i$-th source of risk, where the sources of risk could be lines of insurance, or individual contracts. The exposures are associated with random claims.

---

1 We thank Nan Zhu for excellent research assistance and Steve Mildenhall for valuable comments on an earlier draft.
2 Department of Risk Management and Insurance, Georgia State University, 35 Broad Street, Atlanta, GA 30303
3 By capital, we are referring to the equity of the firm—or the difference between the value of its assets and its liabilities. To fix ideas, we will adopt a common specification of the difference between the fair value of assets and the expected value of liabilities for examples. However, equity can be calculated in a number of different ways, depending on the accounting treatment accorded to assets and liabilities. As will become clear later, the important consideration for the allocation problem is the source of the costs to the firm, and the source could align with various definitions of capital. Thus, it is important to note that the allocation methods described herein could be applied just as easily to alternative definitions of capital.
An increase in exposure shifts the distribution of the claim random variable so that the resulting distribution has first order stochastic dominance over the former. That is:

$$Pr(I_i(q_i)) \geq z \geq Pr(I_i(q_i)) \geq z \quad \forall z, q_i > q_i$$

As a simple (and ubiquitous) example, imagine the exposure $q_i$ representing an insurance company’s quota share of a customer $i$’s loss $\tilde{L}$, so that

$$I_i(q_i) = q_i \cdot \tilde{L}$$

The variable $L$ represents the aggregate claims for the company, with the sum of the random claims over the sources adding up to the total claim:

$$\sum_{i=1}^{N} I_i = L$$

Actual payments made (denoted by the random variable $X$) differ from the total loss claims made ($L$) because of the possibility of default and can be expressed as:

$$X = \min (L, A)$$

Note we have omitted time in this specification for simplicity, as well as asset risk. This allows us to focus on the essence of the allocation problem with minimal complications and to interpret the allocations adding up over liability risks.

We can also decompose actual payments, where $x_i$ represents the payment delivered to the $i$-th source. Obviously, it must be the case that:

$$\sum_{i=1}^{N} x_i = X$$

The typical assumption in the literature is of equal priority in bankruptcy, so that:

$$x_i = \min (L \cdot \frac{A}{L}, I_i)$$

We will adopt that rule here for simplicity, although alternative rules could conceivably be adopted in the framework to follow. We may now think of allocating capital or assets to the $N$ sources of risk, with $k_i$ representing the capital per unit of exposure allocated to the $i$-th source (and $a_i$ representing a similar quantity for assets). Of course, it must be the case that:
\begin{equation}
\sum_{i=1}^{N} q_i k_i = K
\end{equation}

and

\begin{equation}
\sum_{i=1}^{N} q_i a_i = A
\end{equation}

These relations embody the so-called “adding up” property of an allocation. However, it is important to note that allocating capital or assets differs from allocating losses or actual payments in an important sense. The decomposition of the latter quantities into source-specific pieces is an obvious and unique one—following clearly from the claims made by, or payments made to, the respective sources. Allocating capital or assets, on the other hand, is not obvious and depends—as we will describe later—on the context of the problem (see also Bühlmann (1985) for early related ideas).

III. Why allocate?

\textit{III.a Pricing and performance measurement}

Let \( P_i(q_i) \) represent the premiums collected from the \( i \)-th source, with the amount collected depending on the exposure assumed, so that total premiums collected are

\begin{equation}
\sum_{i=1}^{N} P_i(q_i) = P
\end{equation}

We ignore underwriting expenses and define the total costs faced by the insurer as

\begin{equation}
V(X) + C(A, q_1, ..., q_N)
\end{equation}

where \( V(X) \) represents the fair financial value of the random claims payments, and \( C(A, q_1, ..., q_N) \) represents a frictional financing cost which could originate from tax or agency issues. This latter function can evidently accommodate a variety of different frictional cost assumptions. For example, if capital is regarded as the source of frictional costs, a simple tax \( \tau \) on capital (e.g., Froot and Stein (1998)) could be represented as

\begin{equation}
C(A, q_1, ..., q_N) = \tau(A - E[L]) = \tau K
\end{equation}

One could also imagine frictional costs being represented by a tax on assets:

\begin{equation}
C(A, q_1, ..., q_N) = \tau A
\end{equation}

In any case, fair pricing of insurance implies that:

\begin{equation}
P = V(X) + C(A, q_1, ..., q_N)
\end{equation}

must hold in the aggregate. The key question is how to allocate the requisite aggregate to each of the \( N \) sources, and herein lies a controversy about capital allocation first identified by Phillips, Cummins, and Allen (1998). They studied an environment without frictional costs (i.e., \( C(A, q_1, ..., q_N) \equiv 0 \)) and with complete markets. In this setting, by-line pricing is straightforward:
Thus,  

\[ P_i = V(x_i) \]  

with  

\[ \sum_{i=1}^{N} P_i = P = \sum_{i=1}^{N} V(x_i) = V(\mathbf{X}) \]

The significance of this result is that the fair price for insurance for a given line is evidently independent of any by-source allocation of capital or assets. The valuation of the liabilities associated with the \( i \)-th source depends on total assets as well as the extent of exposure to each of the risk sources:  

\[ x_i(q_1, q_2, ..., q_N, A) \]

but, importantly, it is not necessary to know anything beyond that.\(^4\) This finding underscores an important point about capital allocation. Scholars studying insurance pricing in frictionless markets will find capital allocation unnecessary or arbitrary (Sherris (2006))\(^5\)---and with good reason. There is no need to allocate capital for pricing purposes unless there are frictional costs involved.

If, however, frictional costs are present---i.e., if \( C(A, q_1, ..., q_N) > 0 \) ---it becomes necessary to allocate them to lines of business. While \( V(\mathbf{X}) \) decomposes naturally into source-specific components, \( C(A, q_1, ..., q_N) \) may not, and it is this fundamental problem that motivates allocation. While this problem could in principle apply to other types of overhead expense, most recent interest has been directed at the topic of costs relating to capital, in which case the problem of frictional cost allocation ends up boiling down to one of capital allocation.

This is no blackboard curiosity. This is the same problem an actuary confronts when charged with pricing a multiline business to a target return on equity (ROE). The practical manifestation of a “frictional cost of capital” occurs in situations where the target ROE exceeds that implied by the underwriting betas associated with the insurer’s liability risks. Indeed, even the academic literature indicates a significant gap between the insurer’s cost of capital (see Cummins and Phillips (2005)) and the estimated theoretical costs of bearing liability risks (to the extent that these can be measured with any precision at all---see, e.g., Cox and Rudd (1991)). Regardless of the source of the difference, the gap between an insurer’s target ROE and the required rate of return on capital predicted by a model grounded in frictionless financial markets can be thought of as a frictional cost of capital that must be allocated to risks.

**III.b Value Maximization**

Merton and Perold (1993) had a related objective in mind when considering capital allocation. Their interest lay in exploring the feasibility of a capital allocation rule for purposes of making value-

---

4 Note that there is a one-to-one relationship between assets and capital once we are given a set of risk exposures. Thus, we could also have written: \( x_i(q_1, q_2, ..., q_N, K) \).

5 Sherris notes that the default value can be allocated uniquely, but that the extension of this allocation to assets, while potentially appealing, is arbitrary.
maximizing business decisions. To fix ideas, imagine an insurer evaluating its profit function:

\[ \Pi = \sum_{i=1}^{N} P_i(q_i) - V(X(q_1, q_2, \ldots, q_N, A)) - \tau K \]

and wondering if an alternative mix of exposures---for example, exiting a business line or, perhaps, doubling volume in another, would yield an improvement. To answer this question, of course, some connection between exposures and capital is needed (e.g., setting capital according to a risk measure target), and Merton and Perold showed that it was not possible to develop a linear allocation rule (i.e., assigning \( k_i \) units of capital for each unit of exposure to the \( i \)-th source, with \( \sum_{i=1}^{N} q_i k_i = K \)) that would account for the effects of diversification when considering inframarginal or supramarginal changes to the exposure portfolio.

While this finding was generally true, subsequent research would show that allocations could give accurate guidance when considering marginal changes to a portfolio. And it was this insight that spawned much of the literature on the topic that would follow.

IV. Allocation Methods

Many authors have approached the allocation problem from different directions, yet it is reassuring that under certain assumptions---and putting aside relatively minor differences in presentation---most approaches end up in essentially the same place: the so-called Euler or Gradient Principle. For the special case of the allocation according to the so-called covariance principle, the close relationship of the different allocation methods was already pointed out by Urban et al. (2003). Also, Albrecht (2004) recognizes that different approaches lead to the Euler principle.

Of course, the foregoing characterization obscures much nuance and detail. In what follows, we start by introducing the Euler Principle as the most important approach among practitioners and the common ground for many allocation methods. We then attempt to provide the intuition for and the intellectual genesis of the various approaches to the capital allocation problem.

**IV.a The Euler Principle**

The key ingredient for the gradient method is a positively homogeneous risk measure. Formally, a risk measure \( \rho \) is a function mapping the random variable of total claims into a real number and can thus be expressed as:

\[ \rho(L) = \rho(\sum_i l_i(q_i)) \]

although we sometimes directly write \( \rho \) as a function of the exposures as in \( \rho(q_1, q_2, \ldots, q_N) \). If now \( \rho(aL) = a \times \rho(L) \), \( a \geq 0 \), and \( l_i(q_i) = q_i \times l_i \), i.e. if the risk measure and individual loss distributions are (positively) homogeneous, then Euler’s homogeneous function theorem yields:
Herein lies the basis for allocation, with the i-th source receiving a per-unit allocation of capital equivalent to:

\[
\sum_{i=1}^{N} \frac{d\rho}{dq_i} q_i = \rho
\]

and the capital allocations “add up.” In the important special case where capital is determined by the risk measure constraint (i.e., when \( \rho(L) = K \)), the capital allocations correspond to risk allocations: \( k_i = \frac{d\rho}{dq_i} \). But this restriction is not necessary: As can be seen above, scaling by the risk measure in (21) effectively converts “risk shares” into capital shares.

It is important to note that, to this point, all we have described is a mathematical technique. The Euler principle yields an allocation method simply because the prescribed allocations add up. We have not provided a motivation for why one would want to apply the Euler principle, nor any reasoning to guide the choice of risk measure (beyond the requirement that the risk measure be homogeneous). It is at this point that the literature becomes diffuse—-with motivations and guidance depending crucially on the particular context chosen.

This problem is illustrated by one of the seminal papers in the field. Myers and Read (2001) reckon that, given complete markets, default risk can be measured by the default value, i.e. the premium the insurer would have to pay for guaranteeing its losses in the case of a default. They reason further that “sensible” regulation will require companies to maintain the same default value per dollar of liabilities and effectively choose this latter ratio as their risk measure. Myers and Read verify the “adding up” property in this particular case and continue to demonstrate that this observation can be employed to uniquely allocate capital in such a way as to preserve the risk measure target across lines of insurance when considering marginal changes to the risk portfolio.

However, because their analysis is confined to a particular risk measure, their findings depend on the unstated objectives of the “sensible” regulator and can only be uniquely implemented in a complete markets setting. Moreover, Gründl and Schmeiser (2007) show that the Myers-Read allocation leads to decisions that are sub-optimal from a profit standpoint. While this is not so surprising when one considers that Myers and Read advanced the allocation as being driven by regulation (and not necessarily one consistent with insurer self-interest), the finding underscores the crucial role institutional context plays in risk measure selection. Using a particular risk measure

---

6 The literature often obscures the distinction between “capital” and “assets” in part because in some contexts (e.g., a hedge fund manager seeking to risk-adjust expected returns in a long-only asset portfolio) the distinction is unimportant. The distinction, however, is important for insurance, and from here on we will write about “capital” allocation, although the techniques described could just as easily be used to allocate assets or something else.

7 It is useful here to provide an example verifying that this framework fits insurance applications. An example of the risk measure being defined so that \( \rho(L) = K \) is where capital is being set to satisfy \( \rho(L) = \text{VaR}_\infty(L) - E[L] \) (where \( \text{VaR}_\infty \) represents the Value-at-Risk of the claim distribution at some threshold \( \infty \)). Of course, there is no requirement generally that \( \rho(L) = K \), but this condition will often prevail in real world applications.
may be justified when driven by regulatory fiat but may not necessarily align with economic self-interest.

As we will see, the broader literature has contemplated applying the Euler principle to a wider range of risk measures, but the underlying justification for using any particular measure has always remained murky. We will return to this issue in Section V.

**IV.b Axiomatic Approaches**

Inspired by the axiomatic approach to risk measures by Artzner, Delbaen, Eber, and Heath (1999), Denault (2001) proposes a set of axioms that define a coherent capital allocation principle when $\rho(L) = K$. Aside from the “adding up” property introduced above, he requires a “no undercut” condition, “symmetry”, and “riskless allocation.” Here, the “no undercut” conditions means that no sub-portfolio of risks will require a smaller amount of capital on a stand-alone basis than the aggregated capital allocated to these risk. “Symmetry” means that when adding two risks to any disjoint sub-portfolio results in the same contribution to capital, their allocations must coincide; in other words, risks that are identical relative to all other risks in the portfolio should be treated the same. And, finally, “riskless allocation” means that the allocation of a deterministic “risk”—in excess of its “mean”—is zero (see also Panjer (2002)).

This approach, however, yields an impractical result: In order for a “coherent” allocation to exist, the risk measure must necessarily be linear (see also Buch and Dorfleitner (2008), who show that the problematic axiom is symmetry). The key issue here again traces back to the distinction between marginal and inframarginal changes to the portfolio. More precisely, the axioms above are framed in an indivisible setting where the focus is on (finite) subportfolios of a given (finite) portfolio. This framing effectively requires consideration of inframarginal changes to the total portfolio—a requirement which, as shown in Merton and Perold (1993), leads capital allocation to be an exercise in futility. Denault (2001) shows that this futility can only be overcome through the use of a linear risk measure.

As linear risk measures offer little practical relevance, Denault (2001) moved on to thinking about the more fruitful and practical setting of divisible portfolios and real-numbered portfolio weights—thereby effectively restricting attention to marginal changes in the portfolio. In particular, Denault (2001) proposes a set of five axioms in this divisible setting defining a “fuzzy” coherent allocation principle that exists for any given coherent, differentiable risk measure—and this allocation is given by the Euler principle applied to the supplied risk measure.

A different though slightly more parsimonious and self-contained set of axioms—within the divisible setting—was proposed by Kalkbrener (2005): “Linear aggregation”, which combines the “adding-up” and the “riskless allocation” properties; “Diversification”, which corresponds to the “no undercut” property; and “Continuity”, which means that small changes to the portfolio should only

---

8 A risk measure is called coherent if it satisfies four properties: Monotonicity, sub-additivity, (positive) homogeneity, and translation invariance. See Artzner et al. (1999) for details.
have a small effect on the capital allocated to a subportfolio. He continues to show that these
axioms already uniquely determine the allocation of capital to each subportfolio, and it is given by
(Gâteaux) derivative in the direction of the subportfolio---which is exactly the Euler principle.
Furthermore, the existence of the capital allocation is equivalent to subadditivity and homogeneity
of the underlying risk measure, which in turn are the defining characteristics of a coherent risk
measure relative to a monetary risk measure (see e.g. Föllmer and Schied (2002)).

Although Kalkbrener’s axioms seem intuitive at first glance, Meyers (2005) points out that the
resulting allocation may not yield the appropriate choice from an “economic perspective” that is in
line with the objectives of a profit-maximizing institution (see Meyers (2003) and below). The reason
traces back to an implicit assumption about homogeneity of the underlying loss distributions
embedded in the “linear aggregation” axiom.

This problem with the Euler principle was first identified by Mildenhall (2004), who noted that
actuarial applications often involve inhomogeneous distributions---whose properties change as the
volume is scaled up or down. To illustrate, the return distribution associated with a particular stock
is homogeneous in the sense that the return distribution associated with, say, 100 shares simply
follows from scaling up the distribution associated with a single share by a factor of 100; however,
this is not true of the loss distribution associated with vehicles---where the distribution of losses
associated with 100 vehicles is not (except in the case of perfect correlation across vehicles) a simple
scaling up of the loss distribution associated with a single vehicle.

That said, even with inhomogeneous distributions, gradient allocation methods can be resurrected
after generalization that takes into account “volumetric diversification” by adjusting the structure of
the underlying vector space (Mildenhall (2006)). However, note that capital allocated by this
generalized gradient principle may not “add up”.

IV.c Game Theoretic Approaches

The Shapley value is a concept from cooperative game theory that assigns each player a unique share of
the cost that adheres to several axioms (Shapley, 1953). Although it therefore also technically
constitutes an axiomatic approach, a distinction from the previous section is useful since the axioms
here are more general and not tied to the specifics of capital allocation. In fact, the idea to rely on this
concept for other allocation problems in insurance such as cost or risk allocation already occurs in
Lemaire (1984) and Mango (1998). However, it is again Denault (2001) who formalized the application in
the context of the capital allocation problem by aligning the general axioms with his specific allocation
axioms introduced in the previous subsection. Thus, as pointed out there, the direct application of the
Shapley value proves disappointing as it yields a linear risk measure. However, as the game-theoretic
analogue to moving from the indivisible to the divisible portfolio case in the previous subsection, relying
on the theory of fuzzy cooperative games introduced by Aubin (1981) proves to be more practical and
fruitful.
To illustrate the main idea, assume the cost functional $c$ of a cooperative game is defined via the risk measure $\rho$:

$$
(22) \quad c(q_1, q_2, ..., q_N) = \rho(q_1, q_2, ..., q_N)
$$

Then the core of the fuzzy game is defined as (see also Tsanakas and Barnett (2003))

$$
(23) \quad C = \{(k_1, k_2, ..., k_N) | c(q_1, q_2, ..., q_N) = \sum k_i q_i & c(u) \geq \sum k_i u_i, u \in [0, q_1] \times ... \times [0, q_N] \}
$$

Hence, the core consists of allocations such that for each (fractional) subportfolio the aggregated per-unit costs increase, which is a generalization of the “no-undercut” rule from above. It now turns out that if the cost function is subadditive, positively homogeneous, and differentiable---which is equivalent to requiring these properties from the underlying risk measure and homogeneous loss distributions---the core consists of a single allocation only, namely that implied by the Euler method (cf. Aubin (1981)). In particular, in this case the allocations coincide with the so-called Aumann-Shapley values, which satisfy axioms stemming from different backgrounds (see Aumann and Shapley (1974), Billera and Heath (1982), or Mirman and Tauman (1982)):

$$
(24) \quad k_i = \frac{\partial}{\partial u_i} \int_0^1 c(\gamma u) \, dy \bigg|_{u_j=q_j \forall j}
$$

Now obviously if $c(\gamma u) = \gamma \times c(u)$, i.e. if the underlying risk measure is homogeneous, this expression immediately reduces to the gradient allocation (20)/(21). In general, the Aumann-Shapley value aggregates the marginal contributions of each “slice” of a risk factor $i$, $[\gamma u_i, (\gamma + dyu_i), \gamma \in (0,1)]$, when the risk portfolio is uniformly expanded.

The Aumann-Shapley value thus also serves as a starting point for generalizations. Specifically, to cope with the problem of inhomogeneous loss distributions, Powers (2007) demonstrates that although the Euler principle will not apply, the Aumann-Shapley value can be used for the risk-allocation problem. Similarly, it may offer a solution if the underlying risk measure does not satisfy the homogeneity condition. For instance, Tsanakas (2009) shows how to allocate capital with convex risk measures, although the absence of homogeneity is shown to potentially produce an incentive for infinite fragmentation of portfolios. The intuition for this rather undesirable feature are risk aggregation penalties within inhomogeneous convex risk measures.\(^\text{10}\)

**IV.d Economic Approaches: Profit Maximization**

---

\(^9\) Tsanakas (2004) shows that for distortion risk measures, the gradient allocation also pertains if one allows for nonlinear portfolios, which give rise to a so-called non-atomic core.

\(^{10}\) A risk measure is called convex if it satisfies three properties: Monotonicity, translation invariance, and convexity. See Föllmer and Schied (2002) for details.
The relationship of the Euler principle to profit maximization was sensed in early work by Tasche (2004)\(^{11}\) and Schmock and Straumann (1999). More specifically, Tasche (2004) calls a capital allocation suitable for performance measurement if it satisfies the following property: If the marginal performance of risk \(i\) as measured by its return on allocated risk capital exceeds (respectively, falls below) the company’s total risk-adjusted return of capital (RAROC)---i.e. its return per unit of risk \(\frac{\Pi}{\rho(L)}\)---then increasing (respectively, decreasing) the exposure by a small amount improves the overall performance of the portfolio. The authors then continue to show that the only suitable allocation is given by the Euler principle. Similarly, Schmock and Straumann (1999) call an allocation consistent if all individual risk-adjusted returns are equal to the optimal risk-adjusted company return, which again yields Euler.

More formally, Zanjani (2002) derives the gradient solution in the context of a profit maximization problem where the firm’s policyholder/counterparty preferences are defined over a measure of the overall portfolio risk of the institution. Similarly, Stoughton and Zechner (2007) arrive at a gradient based allocation by framing the institution’s profit maximization problem with a capitalization constraint tied to a risk measure. To illustrate, consider the optimization problem

\[
(25) \quad \max_{k,q_1,\ldots,q_N} \left\{ \frac{\sum_{i=1}^{N} P_i(q_i) - V(X(q_1, q_2, \ldots, q_N, K)) - C_k(K)}{\Pi} \right\}
\]

subject to

\[
(26) \quad \rho(q_1, q_2, \ldots, q_N) \times \vartheta_X \leq K
\]

where \(\vartheta_X\) is an exchange rate that converts risk to capital, which is often chosen to be unity if risk is measured in monetary units. After eliminating Lagrange multipliers from the optimality conditions associated with this problem, one obtains:

\[
(27) \quad \frac{\partial \Pi}{\partial q_i} = \left( -\frac{\partial \Pi}{\partial K} \right) \times \vartheta_X \times \frac{\partial \rho}{\partial q_i}
\]

at the optimal exposures and capital level. Hence, for the optimal portfolio, the risk adjusted marginal return \(\frac{\partial \Pi/\partial q_i}{\vartheta_X \times \partial \rho/\partial q_i}\) for each exposure \(i\) is the same and equals the cost of a marginal unit of capital \(-\frac{\partial \Pi}{\partial K}\). In this context, the appeal of the allocation produced by the gradient method is its consistency with marginal cost (see also Meyers (2003)), and for this reason, the gradient method is often claimed to be “economic” in nature. However, it is again important to stress that any economic content flows from the imposition of a risk measure constraint (2), and that this imposition may well be arbitrary.

---

\(^{11}\) Although the final version was published in 2004, all the important ideas are already contained in a working paper version entitled “Risk Contributions and Performance Measurement” from 2000.
An alternative economic foundation for capital allocation—that does not rely on the imposition of a risk measure—is offered by Zanjani (2010), who shows that economic capital allocation can be derived in a social planning problem when starting from primitive assumptions on risk and preferences. Bauer and Zanjani (2011) extend this line of reasoning by establishing the allocation consistent with economic self-interest—i.e. that is again resulting from profit maximization—but they explicitly take the preferences of the underlying counterparties into account. This is achieved by attaching additional participation constraints to the problem (25) and keeping—or dropping—constraint (26), which is interpreted as an exogenously supplied solvency constraint from a regulator. The key idea is that it is not only external regulation but also the counterparties’ preferences for capitalization that drive the company’s capital allocation. The resulting allocation then is a weighted average of the Euler method applied to the exogenous risk measure and an internal allocation rule, where the weight depends on how much the imposed capital constraint differs from the level of capital held in an unregulated regime.

IV.e Alternatives to the Gradient Method based on Risk Measurement

There are a number of alternative approaches based on risk measures that do not necessarily land at the Euler principle. One idea that first occurred in Dhaene, Goovaerts, and Kaas (2003) and was extended by Laeven and Goovaerts (2004) and Dhaene, Tsanakas, Valdez, and Vanduffel (2011) is to derive allocations based on an optimization procedure. The idea is to choose an allocation such that the deviation of the individual losses and the allocated capital are maximally “close”. Specifically, Laeven and Goovaerts propose the optimization problem:

\[
\min_{k_1, k_2, \ldots, k_N} \rho \left( \sum_{i=1}^{N} (l_i(q_i) - q_i k_i)^+ \right) \quad \text{s.t.} \quad \sum q_i k_i = K
\]

(28) to identify an allocation \((k_1, k_2, \ldots, k_N)\), whereas Dhaene et al. (2011) consider the program:

\[
\min_{k_1, k_2, \ldots, k_N} \sum_{i=1}^{N} q_i E \left[ \theta_i D \left( \frac{l_i(q_i)}{q_i} - k_i \right) \right] \quad \text{s.t.} \quad \sum q_i k_i = K
\]

(29) where \(D\) is a (distance) measure and \(\theta_i\) are weighting random variables with \(E[\theta_i] = 1\).

While the former yields rather complex allocation rules, the approach by Dhaene et al. (2011)—when choosing \(D\) and \(\theta_i\) adequately—gives rise to various allocation methods proposed in the literature. For instance, for \(D(x) = x^2\) and \(K = \sum E[\theta_i l_i(q_i)]\), they arrive at so-called weighted risk capital allocations \(k_i = E[\theta_i l_i(q_i)]\) studied in detail by Furman and Zitikis (2008). For other choices, they uncover an array of other allocation principles, including several that can also be derived from the application of the Euler principle. Thus, while not unambiguously collapsing to the Euler principle, the approaches are yet again related.

V. Conceptual Issues with the Choice of Risk Measure

With few exceptions, most of the capital allocation approaches start with the choice of a risk measure. This fairly reflects the weight of the academic literature. It also reflects revealed preference among practitioners for the tractability of gradient methods applied to risk measures.
Academics may concern themselves with all sorts of esoterica when analyzing capital allocation, but when it comes to actually implementing capital allocation, gradient methods applied to risk measures are hard to beat.

However, while ease of implementation may justify the special attention paid to approaches based on risk measures, it does not necessarily imply coherence of logic. We have yet to establish which risk measure is appropriate and, on a deeper level, whether it makes sense to be guided by risk measures at all.

To illustrate the latter point, is not clear why coherence of an allocation method should be specified via a risk measure as in Kalkbrener (2005), nor why the risk measure should specify the cost function of the cooperative game from Denault (2001), nor why the policyholders in Zanjani (2002) would assess company quality with a risk measure, nor why the bank in Stoughton and Zechner (2007) would constrain itself with a risk measure. As noted earlier, a risk measure constraint will not necessarily help a firm improve profitability (e.g., Gründl and Schmeiser (2007)).

Of course, one could argue that risk measure constraints are driven by the dictates of rating agencies (on whom uninformed consumers rely for assessments of creditworthiness) or regulators, who set standards for companies via risk measure-based analytics. But such an argument inevitably leads one to question whether rating agencies and regulators should be using risk measures in this way. Does a regulator serve the public interest by setting standards with a risk measure, and, if so, which is the correct risk measure?

For the gradient method at least, much rides on the answer. The gradient method’s claim of superiority rests entirely on the propriety of the risk measure. Without economic justification for the risk measure, it is not clear that the gradient method, despite its mathematical elegance, offers a superior allocation.

Thus, it is not surprising that much attention in the current discourse on capital allocation pertains to the choice of risk measure. Capital allocation scholars have largely joined other risk scholars in embracing so-called “tail” risk measures with the consequence that risk measures similar in concept to Expected Shortfall (ES) are rapidly gaining favor among academics and practitioners. Proponents of ES stress the theoretical appeal of coherence (Artzner et al. (1999)) and the practical appeal of weighting tail events more heavily than VaR. The weighting of tail events can be taken further with the enhancement of a spectral weighting function (Acerbi (2002)), a transformation that preserves coherence.

It is an open question, however, whether coherence offers a one-size-fits-all guide to risk measure selection. To illustrate, the capital allocation yielded in the economic model of the insurance

---

12 Aside from minor subtleties in the case of non-continuous distributions, the Expected Shortfall is identical to the Tail Value at Risk (TVaR) or Conditional Tail Expectation (CTE). We will treat them as synonyms for the purpose of this chapter.
company used in Bauer and Zanjani (2011) can only be implemented through the application of the gradient technique to a particular risk measure:

\[
\dot{\rho}(X) = \exp\{E^Q[\log\{X\}])
\]

where \(Q\) is a probability measure that shifts the entire probability mass to default states and includes weights determined by relative values placed on recoveries by the firm’s policyholders. This risk measure, however, is not coherent.

The deeper point here is that, even if one accepts the inevitability of using risk measures for capital allocation, different foundational assumptions may point the way to different risk measures with different mathematical properties. This suggests that the appropriate risk measure for capital allocation based on the gradient method may depend very much on context.

VI. An Example from Life Insurance

To illustrate the use of the capital allocation methods introduced above, we consider a life insurance company selling three product lines: Term life insurance contracts with constant annual premium payments, endowment insurance contracts with constant annual premium payments, and life annuities. More specifically, we analyze the allocation problem for the stylized insurance company introduced in Zhu and Bauer (2011b), the portfolio of which is detailed in Table 1: Portfolio of the insurance company.

<table>
<thead>
<tr>
<th>Table 1: Portfolio of the insurance company</th>
</tr>
</thead>
<tbody>
<tr>
<td>term life, face value 100,000.00</td>
</tr>
<tr>
<td>age(term)</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>30/20</td>
</tr>
<tr>
<td>35/15</td>
</tr>
<tr>
<td>40/10</td>
</tr>
<tr>
<td>45/5</td>
</tr>
</tbody>
</table>

We assume that the required capital is calculated in a one-year mark-to-market approach as

\[
K = \rho(L^{TEA})
\]

where the “loss” is defined as

\[
L^{TEA} = p(0,1) \times (V_1 - A_1)
\]

\(p(0,1)\) denotes the price of a one-year zero coupon bond, and \(A_1\) and \(V_1\) denote the values of the insurer’s assets consisting of all premiums paid and liabilities at time 1, respectively. Here it is assumed that the insurer allocates assets into 1, 3, 5, and 10-year UK government bonds as well as an equity index (FTSE) at equal proportions, which are modeled via an extended Black-Scholes model with Vasicek stochastic interest rates calibrated to UK data between 6/1998 and 6/2008. We
refer to Zhu and Bauer (2011b) for the model dynamics and parameters. In addition, we consider
the required capital for companies that only have one or two lines of business, with corresponding
losses denoted by $L^{TE}$ for a company with term-life and endowment business only, $L^{TA}$ for a
company with term-life and annuity business only, $L^{EA}$ for a company with endowment and annuity
business only, $L^{T}$ for a company with term-life business only, $L^{E}$ for a company with endowment
business only, and finally---by $L^{A}$ for a company with annuity business only.

For evaluating the insurance liabilities---in addition to the stochastic interest rate model---we need
to make an assumption about the evolution of mortality. Following Zhu and Bauer (2011b), we use
two different approaches: First, we assume mortality evolves deterministically, so that
(unsystematic) mortality risk only comes to effect in that the number of deaths within each cohort is
sampled from a Binomial distribution—which a known mortality probability given via the
corresponding generation life table for the England and Wales general male population compiled
based on the Lee-Carter method. In what follows, we refer to this as the “deterministic case”.
Second, in addition to unsystematic mortality risk as above, we consider aggregate (or systematic)
mortality risk by sampling the generational life table at time $1$, i.e. we allow mortality rates to evolve
stochastically over the year. Specifically, we use the forward mortality factor model introduced in
Zhu and Bauer (2011a). We refer to this as the “stochastic case”.

Based on each sampled scenario—that is a combination of stock return, interest rate, generational
life table, and death counts for each cohort—we can then evaluate $A_{1}$ and $V_{1}$ (we refer to Zhu and
Bauer (2011b) for further details). Since the assets at time zero consist of the premiums only, which
—for each business line—are proportional to the corresponding insured amounts, and since the
same is true for the liabilities, $L^{TEA}$ can be represented as a linear combination of random variables
corresponding to each business line with the amounts as the weights, i.e. the loss distributions are
homogeneous in this case. More specifically, we can write

$$L^{TEA} = \text{FaceVal}^{T} \times \bar{L}^{T} + \text{FaceVal}^{E} \times \bar{L}^{E} + \text{FaceVal}^{A} \times \bar{L}^{A}$$

(33)

where $\text{FaceVal}^{i}$ is the face value in business line $i \in \{T, E, A\}$ and $\bar{L}^{i}$ is the corresponding
normalized loss for a face value of 1. Obviously, we have $L^{i} = \text{FaceVal}^{i} \times \bar{L}^{i}$, $i \in \{T, E, A\}$. Thus,
given a homogeneous risk measure, we can evaluate capital allocations via the Euler principle:

$$q_{i}k_{i} = \text{FaceVal}^{i} \times \frac{\partial \rho(L^{TEA})}{\partial \text{FaceVal}^{i}} \approx \text{FaceVal}^{i} \times \frac{\rho(L^{TEA+\Delta L^{i}}) - \rho(L^{TEA})}{\Delta_{i}}$$

(34)

where we choose $\Delta_{i} = 1\% \times \text{FaceVal}^{i}$, $i \in \{T, E, A\}$.

Table 2: Results for the Deterministic Mortality Case and Table 3: Results for the Stochastic Mortality
Case display our results for the deterministic and the stochastic mortality case, respectively. To keep
the presentation concise, we only provide point estimates based on 100,000 simulations—we refer to
Zhu and Bauer (2011b) for corresponding standard errors. In particular, we calculate capital
allocations for four risk measures: (i) The standard deviation risk measure with parameter $a = 2$,
\( (35) \quad \rho(X) = E[X] + a \times \text{StDev}[X] \)

(ii) Value-at-Risk at the 99% level estimated by the empirical quantile; (iii) Expected Shortfall at the 90% level, that is

\( (36) \quad \rho(X) = E[X | X \geq \text{VaR}_{90\%}(X)] \)

where again the value at risk is estimated by the empirical quantile; and (iv) Expected Shortfall at the 99% level. All of these risk measures are homogeneous so that we may apply the Euler principle. Formally, the standard deviation risk measure (35) yields the so-called covariance allocation principle,

\( (37) \quad k_i = E[\tilde{U}] + a \times \frac{\text{Cov}(\tilde{L}^T \text{L}^T \text{E} \text{A})}{\text{StDev}(\text{L}^T \text{E} \text{A})} \)

Value-at-risk gives

\( (38) \quad k_i = E[\tilde{U} | \text{L}^T \text{E} \text{A} = \text{VaR}_\alpha(\text{L}^T \text{E} \text{A})] \)

and the Expected Shortfall (36) yields (cf. Section 6.3 in McNeil et al. (2005) for derivations of these formulas)

\( (39) \quad k_i = E[\tilde{U} | \text{L}^T \text{E} \text{A} \geq \text{VaR}_\alpha(\text{L}^T \text{E} \text{A})] \)

whereas our calculation via (34) only provides an easy-to-calculate approximation. In particular, our allocations from Table 2: Results for the Deterministic Mortality Case and Table 3: Results for the Stochastic Mortality Case shown in the first lines for each considered risk measure do not perfectly add up to the capital \( K \) calculated according to (31)---which is shown in the last column of the third line for each risk measure. However, with the possible exception of Value-at-Risk, the approximation error is small and the calculated allocations are close to the “adjusted” allocations calculated based on \( K \) and the proportions shown in the respective second lines.

Table 2: Results for the Deterministic Mortality Case

<table>
<thead>
<tr>
<th>Deterministic Mortality</th>
<th>Term Insurance</th>
<th>Endowment Insurance</th>
<th>Annuities</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{StDev} )</td>
<td>Cap Alloc.</td>
<td>22,624</td>
<td>1,519,072</td>
<td>5,162,791</td>
</tr>
<tr>
<td>Percentage</td>
<td>0.34%</td>
<td>22.66%</td>
<td>77.01%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Cap Alloc. (adj.)</td>
<td>22,614</td>
<td>1,518,365</td>
<td>5,160,389</td>
<td>6,701,368</td>
</tr>
<tr>
<td>Stand Alone</td>
<td>227,136</td>
<td>2,165,157</td>
<td>5,352,843</td>
<td>7,745,136</td>
</tr>
<tr>
<td>Infr. Increase</td>
<td>19,068</td>
<td>1,330,541</td>
<td>4,509,569</td>
<td>5,859,178</td>
</tr>
<tr>
<td>( 99% \text{ VaR} )</td>
<td>Cap Alloc.</td>
<td>12,769</td>
<td>2,077,348</td>
<td>5,444,049</td>
</tr>
<tr>
<td>Perc</td>
<td>0.17%</td>
<td>27.57%</td>
<td>72.26%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Cap Alloc. (adj.)</td>
<td>13,249</td>
<td>2,155,460</td>
<td>5,648,752</td>
<td>7,817,461</td>
</tr>
<tr>
<td>Stand Alone</td>
<td>337,442</td>
<td>2,692,282</td>
<td>5,908,572</td>
<td>8,938,296</td>
</tr>
</tbody>
</table>

15
### Infr. Increase

<table>
<thead>
<tr>
<th>Cap Alloc.</th>
<th>16,240</th>
<th>1,880,233</th>
<th>5,095,335</th>
<th>6,991,808</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perc</td>
<td>0.36%</td>
<td>26.34%</td>
<td>73.30%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Cap Alloc. (adj.)</td>
<td>20,668</td>
<td>1,512,383</td>
<td>4,207,893</td>
<td>5,740,944</td>
</tr>
<tr>
<td>Stand Alone</td>
<td>224,516</td>
<td>2,007,531</td>
<td>4,372,433</td>
<td>6,604,480</td>
</tr>
<tr>
<td>Infr. Increase</td>
<td>17,555</td>
<td>1,351,388</td>
<td>3,710,757</td>
<td>5,079,700</td>
</tr>
</tbody>
</table>

### 99% ES

<table>
<thead>
<tr>
<th>Cap Alloc.</th>
<th>27,157</th>
<th>2,456,872</th>
<th>6,611,156</th>
<th>9,095,185</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perc</td>
<td>0.30%</td>
<td>27.01%</td>
<td>72.99%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Cap Alloc. (adj.)</td>
<td>27,149</td>
<td>2,456,105</td>
<td>6,609,092</td>
<td>9,092,346</td>
</tr>
<tr>
<td>Stand Alone</td>
<td>384,072</td>
<td>3,121,017</td>
<td>6,832,316</td>
<td>10,337,405</td>
</tr>
<tr>
<td>Infr. Increase</td>
<td>23,549</td>
<td>2,234,172</td>
<td>5,942,845</td>
<td>8,200,566</td>
</tr>
</tbody>
</table>

Comparing the proportional allocations in line 2 for each risk measure in the deterministic case (Table 2: Results for the Deterministic Mortality Case), we find that all principles yield rather similar allocations with a relatively small weight on the term business and the majority of capital allocated to the annuity business. In particular, the allocations based on 90% ES and 99% ES are very close. Similarly, we observe rather congenerous allocations for all considered risk measures in the stochastic case, with again the ES based allocations being very similar. One minor difference between the allocations is the increased weight put on the annuities line within the covariance allocation relative to the other, tail-based risk measures. This indicates that the combination of risk factors driving annuities plays a major role over the entire domain of the distribution—and in some ranges possibly even more so than in the tails.

**Table 3: Results for the Stochastic Mortality Case**

<table>
<thead>
<tr>
<th>StDev</th>
<th>Term Insurance</th>
<th>Endowment Insurance</th>
<th>Annuities</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cap Alloc.</td>
<td>-105,602</td>
<td>1,276,893</td>
<td>6,272,827</td>
<td>7,444,118</td>
</tr>
<tr>
<td>Percentage</td>
<td>-1.42%</td>
<td>17.15%</td>
<td>84.27%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Cap Alloc. (adj.)</td>
<td>-105,544</td>
<td>1,276,192</td>
<td>6,269,383</td>
<td>7,440,031</td>
</tr>
<tr>
<td>Stand Alone</td>
<td>375,020</td>
<td>2,182,978</td>
<td>6,519,600</td>
<td>9,077,598</td>
</tr>
<tr>
<td>Infr. Increase</td>
<td>-113,693</td>
<td>1,064,411</td>
<td>5,172,510</td>
<td>6,123,228</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>99% VaR</th>
<th>Term Insurance</th>
<th>Endowment Insurance</th>
<th>Annuities</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cap Alloc.</td>
<td>-54,416</td>
<td>1,550,635</td>
<td>7,424,958</td>
<td>8,921,177</td>
</tr>
<tr>
<td>Perc</td>
<td>-0.61%</td>
<td>17.38%</td>
<td>83.23%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Cap Alloc. (adj.)</td>
<td>-54,067</td>
<td>1,540,694</td>
<td>7,377,356</td>
<td>8,863,983</td>
</tr>
<tr>
<td>Stand Alone</td>
<td>467,622</td>
<td>2,706,795</td>
<td>7,668,752</td>
<td>10,843,169</td>
</tr>
<tr>
<td>Infr. Increase</td>
<td>-194,183</td>
<td>1,385,223</td>
<td>6,059,950</td>
<td>7,250,990</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>90% ES</th>
<th>Term Insurance</th>
<th>Endowment Insurance</th>
<th>Annuities</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cap Alloc.</td>
<td>-106,402</td>
<td>1,253,615</td>
<td>5,355,359</td>
<td>6,502,572</td>
</tr>
<tr>
<td>Perc</td>
<td>-1.64%</td>
<td>19.28%</td>
<td>82.36%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Cap Alloc. (adj.)</td>
<td>-106,328</td>
<td>1,252,742</td>
<td>5,351,625</td>
<td>6,498,039</td>
</tr>
</tbody>
</table>

---

13 See also Asimit et al. (2012) for related theoretical results on the limiting behavior of ES-based allocations as the confidence level $\alpha$ approaches 1.
One potentially surprising observation is that while in the deterministic case the capital contributions of each line are still positive, this is no longer the case under stochastic mortality. More precisely, although the total required capital increases for each risk measure relative to the deterministic case, the term life insurance block now is allocated a negative amount of capital. The intuition is of course given by natural hedging effects between the different business lines: While liabilities from endowments and—especially—life annuities increase when survival probabilities increase systematically, the term life insurance liabilities decrease since fewer term policyholders are going to decease. The resulting negative dependence between the term life profits-and-losses and the remainder of the portfolio then turns out to be beneficial for the insurer.\(^\text{14}\)

However, it is necessary to point out that the interpretation of negative allocations is not immediately clear when considering them in the context of the risk-adjusted return on capital (RAROC). A low—or even negative—RAROC now seems desirable from the insurer’s perspective since it implies only a small loss—or even a gain—on a business line together with a substantial capital relief, whereas a high RAROC implies large losses relative to the (possibly minor) capital relief. This challenge was already noted in Tasche (2004), who points out that in this case the RAROC should be interpreted as “the profit of a counterparty and should therefore—from the investor’s [insurer’s] point of view—be hold as small as possible.” In the present setting, one can for instance think of term-policyholders as counterparties providing the company with a hedge for which they want to be compensated—but of course the insurer wants this compensation to be as small as possible. Since we are evaluating risks by their RAROC, the largest acceptable return to be paid to the counterparties is the target return. Thus, pricing according to a given positive target RAROC is still possible, and in this case the company will be happy to sell term life contracts under par—i.e. it is willing to incur a loss on the line since it benefits the company overall.

Of course this will not be the case for a company only offering term life insurance. To illustrate, in the fourth line for each considered risk measure of Table 2: Results for the Deterministic Mortality Case and Table 3: Results for the Stochastic Mortality Case, we provide the economic capital for single line companies, \(\rho(l^i), i \in \{T, E, A\}\). Not surprisingly, we find that in all cases the stand-alone capital exceeds the allocated capital in the enterprise setting and, consequently, that the sum of the capital for the single-line companies exceeds the required capital for the multiline company due to diversification benefits.

---

\(^{14}\) See also Powers (2007) in this context, who notes that “a negative value […] simply means that the presence of member \(i\) serves to decrease (i.e., offset some portion of) the total portfolio’s risk.”

---

<table>
<thead>
<tr>
<th></th>
<th>Stand Alone</th>
<th>Infr. Increase</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Infr. Increase</td>
<td>-112,988</td>
<td>1,068,368</td>
<td>4,404,135</td>
<td>5,359,515</td>
<td></td>
</tr>
<tr>
<td>99% ES Cap Alloc.</td>
<td>-177,992</td>
<td>1,958,468</td>
<td>8,610,591</td>
<td>10,391,067</td>
<td></td>
</tr>
<tr>
<td>Perc</td>
<td>-1.71%</td>
<td>18.85%</td>
<td>82.87%</td>
<td>100.00%</td>
<td></td>
</tr>
<tr>
<td>Cap Alloc. (adj.)</td>
<td>-177,930</td>
<td>1,957,784</td>
<td>8,607,585</td>
<td>10,387,439</td>
<td></td>
</tr>
<tr>
<td>Stand Alone</td>
<td>544,508</td>
<td>3,149,934</td>
<td>8,957,039</td>
<td>12,651,481</td>
<td></td>
</tr>
<tr>
<td>Infr. Increase</td>
<td>-188,271</td>
<td>1,674,377</td>
<td>7,123,448</td>
<td>8,609,554</td>
<td></td>
</tr>
</tbody>
</table>
It is important to bear in mind, however, that the calculated capital contributions are to be understood at the margin. As shown by Merton and Perold (1993), capital allocation is generally unfeasible at the inframarginal level. To illustrate, the fifth line for each considered risk measure in Table 2: Results for the Deterministic Mortality Case and Table 3: Results for the Stochastic Mortality Case provides the inframarginal capital increase when a company with two business lines decides to enter the remaining, third business line. For instance, for calculating the inframarginal capital increase for the term business, we calculate

\[(40) \quad \rho(L^{TEA}) - \rho(L^{EA})\]

with similar equations for the endowment and annuity lines. We find that the inframarginal increase always is smaller than the capital allocated within the Euler principle. The intuition is that each dollar increase of the new business line enjoys the full diversification benefits of the existing lines when the portfolio is changed inframarginally, whereas the Euler principle relies on a uniform (marginal) extension of the entire portfolio (see also Equation (24) for the Aumann-Shapley value and the following discussion). In particular, we find that the resulting increases in capital do not add up to the total capital.\(^{15}\)

VII. Conclusion

Initial skepticism about the exercise of capital allocation was grounded in the notion of frictionless markets, where allocation is unnecessary and arbitrary. Once frictions are introduced, however, allocation becomes well-defined, at least at the margin of the risk portfolio. The predominant allocation technique relies on calculating the gradient of a chosen risk measure of the portfolio.

In the end, allocation methods—regardless of provenance—must be judged on their ability to give an accurate picture of marginal cost of risk. The literature has firmly established that a fixed allocation makes sense when considering marginal changes to the risk portfolio but fails when considering infra- or supra- marginal changes, so the best we can hope for is that the allocation “gets it right” on the margin.

A remaining problem, however, is that marginal cost often ends up being defined by the way we allocate capital—-as opposed to the reverse, where marginal cost dictates how we allocate capital. The first approach is easy but self-referential in its justification. The second has its own pitfalls in that a great deal of information may be required to assess the “true” marginal costs of the firm. Going forward, we regard the challenge for the capital allocation literature as being to connect the mathematical techniques of allocation with the real operating objectives and constraints faced by the institution under consideration.

VIII. References

\(^{15}\) While “adding-up” in principle would be delivered by the Shapley value, as indicated in Section IV.b and IV.c, a “coherent” capital allocation via the Shapley value would require a linear risk measure.


