The Marginal Cost of Risk, Risk Measures, and Capital Allocation*

Draft – Please do not cite without the authors’ permission

Daniel Bauer
Department of Risk Management and Insurance, Georgia State University
35 Broad Street. Atlanta, GA 30303. USA

George Zanjani†
Department of Risk Management and Insurance, Georgia State University
35 Broad Street. Atlanta, GA 30303. USA

November 2010

Abstract

The Euler (or gradient) allocation technique defines a financial institution’s marginal cost of a risk exposure via calculation of the gradient of a risk measure evaluated at the institution’s current portfolio position. The technique, however, relies on an arbitrary selection of a risk measure. We reverse the sequence of this approach by calculating the marginal costs of risk exposures for a profit maximizing financial institution with risk averse counterparties, and then identifying a closed-form solution for the risk measure whose gradient delivers the correct marginal costs. We compare the properties of allocations derived in this manner to those obtained through application of the Euler technique to Expected Shortfall.

JEL classification: G20, G22
Keywords: capital allocation; risk measure; profit maximization; counterparty risk aversion.

*The views expressed in this article are those of the authors and do not necessarily reflect the position of Georgia State University.
†Corresponding author. Phone: +1-(404)-413-7464. Fax: +1-(404)-413-7499. E-mail addresses: dbauer@gsu.edu (D. Bauer); gzanjani@gsu.edu (G. Zanjani).
1 Introduction

Merton and Perold (1993) studied the problem of how to allocate a financial institution’s capital to the various risks within its portfolio and deemed it impossible to devise an allocation that could reliably guide decisionmaking in all situations. They constructed examples showing that supramarginal and inframarginal changes to a portfolio, such as those that might be created when entering or exiting a business line, could not be analyzed with a single fixed allocation. This non-existence result, however, offered little consolation to practitioners confronted a real-world allocation problem that simply could not be ignored: When capital is costly to hold, the cost of bearing any individual risk cannot be understood without understanding how much capital is needed to support that individual risk.

Understandably, the subsequent research on capital allocation focused on the margin of the company, since it turned out that allocation methods were analytically valid for analyzing small, marginal changes to a contract portfolio. Work by Schmock and Straumann (1999), Tasche (2000), and Myers and Read (2001), among others, spawned subsequent generalizations (e.g., Kalkbrenner (2005) and Mildenhall (2006)) and a number of applications (e.g., Perold (2005); Stoughton and Zechner (2007)). Broadly speaking, these papers start with a differentiable risk measure and compute the marginal capital increase required to maintain the risk measure at a target value as a particular risk exposure within the portfolio is expanded, an approach we will refer to (as others have) as “gradient” allocation or “Euler” allocation. Under some conditions (see Mildenhall (2004)), the resulting marginal requirements will produce a complete allocation of the capital of the firm. The literature thus provides a practical solution to the real-world problem of capital allocation that seems connected to the marginal cost of risk at the firm level and is often so billed—as “economic” or “based on marginal cost” or “optimal.”

Unfortunately, the connection to economic theory is tenuous. It is true that, if consumer preferences are defined over a particular risk measure (Zanjani (2002)), or if a particular risk measure is assumed to constitute the payoff function in a cooperative game (see Denault (2001); Powers (2007)), the consequent economic analysis points to the gradient allocation principle. Such approaches, however, depend on the choice of risk measure—and it is not clear how economic self interest would guide this choice. While there exist formal connections between certain risk measures and the preferences of a risk-averse firm (see e.g. Föllmer and Schied (2010) or Wächter and Mazzoni (2010)), the standard assumption in economic models is one of a risk-neutral profit-maximizing firm, whose motivation for risk management (to the extent that a motivation exists) derives from some aspect of the cost structure it faces—such as convex costs of taxes or external finance (e.g., Froot and Stein (1998)) or counterparty risk aversion.

Myers and Read (2001) recognized that the argument for using a risk measure to allocate capital depended crucially on institutional context (pp. 550-551) and justified their own choice as applying to a highly specialized situation—where the counterparties of the financial institution are fully covered by a deposit insurance scheme and the regulator possessed some unspecified objective function or budget constraint demanding use of a particular risk measure (the value of the default put option divided by the value of liabilities)—which it would presumably impose on the firms in the industry. Similarly, Tasche also alludes to the role of regulation in justifying risk measure choice (Tasche (2000)) and in motivating a “reasonable” allocation of capital (Tasche
These arguments, however, say nothing about shareholder value or profit maximization beyond the necessity of obeying constraints imposed by regulation. In general, if we take regulatory constraints as given, what is the marginal cost of risk for a profit-maximizing financial institution? And is there a risk measure this institution could use for purposes of pricing and performance measurement that would yield the economically correct allocation of capital? This paper is concerned with these questions.

We consider optimal pricing behavior of a profit-maximizing financial institution with risk-averse counterparties in the presence of a (possibly nonbinding) regulatory constraint tied to a risk measure. We seek the capital allocation rule consistent with the firm’s marginal cost, and we find that the optimal rule depends crucially on the institutional context. As might be expected, in the scenario envisioned by Myers and Read (with fully insured counterparties), the economically optimal allocation follows from the gradient allocation principle as applied to the risk measure imposed by regulation. However, if the counterparties are not fully insured, the optimal allocation rule is not fully determined by the regulatory constraint, even if that constraint is binding.

More specifically, when counterparties are not fully protected, the firm’s marginal cost associated with the risk of a particular counterparty depends on how that risk affects the firm’s other counterparties (and, thus, their willingness to pay for the firm’s contracts)—so the firm must price contracts accordingly. The optimal allocation rule then ends up being a weighted average of an “external” allocation rule implied by the regulatory constraint (if it binds) and an “internal” allocation rule driven by the institution’s uninsured counterparties. In the extreme case of no regulation, the allocation rule simply boils down to the “internal” rule. Intermediate cases, however, could feature marginal cost being driven mainly by the “internal” rule (if the regulatory constraint puts firm capitalization close to the level it would have chosen in the absence of regulation) or the “external” constraint (if regulation forces the firm to hold far more capital than is privately optimal).

We show further that the “internal” allocation rule can be implemented by applying the gradient allocation principle to a particular risk measure—the exponential of a weighted average of the logarithm of portfolio outcomes in states of default, with the weights being determined by the relative values placed on recoveries in the various states of default by the firm’s counterparties. The weights are thus similar in concept to the “spectral” weights proposed by Acerbi (2002) as a means of capturing the “subjective risk aversion” of a financial institution—with the weights being determined endogenously through the process of profit maximization. The risk measure itself is evidently a “tail” risk measure, although the functional transformations ultimately cause it to be non-convex.

Finally, we present numerical results comparing the allocations resulting from the “internal” allocation rule to those arising from the application of the gradient allocation method to Expected Shortfall (ES), showing that ES-based allocations generally fail to weight default outcomes properly. Specifically, in cases where counterparties are strongly risk averse, ES-based allocations tend to underweight bad outcomes; when counterparties are only weakly risk averse, ES-based allocations tend to overweight bad outcomes.

1And if the effect of regulation is simply to verify the choice of a “reasonable” method of capital allocation, then the door is opened to a variety of methods that might be argued to qualify as “reasonable” (see Dhaene et al. (2009) for a recent survey), not all of which conform with the gradient allocation principle.
2 Profit Maximization and Capital Allocation

To illustrate the main ideas, we will start by considering a greatly simplified environment without securities markets and then generalize the results to the case where both the firm and its consumers have access to securities markets (see Appendix D for the generalized treatment).

An insurance company has \( N \) consumers, with consumer \( i \) facing a loss \( L_i \) modeled as a non-negative (square-integrable) random variable on the complete probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Thus, \( L_i(\omega) > 0 \) indicates that consumer \( i \) experienced a loss in the state \( \omega \in \Omega \), while \( L_i(\omega) = 0 \) indicates that she did not.

The firm determines the optimal level of assets \( a \) for the company, as well as levels of insurance coverage for the consumers, with the coverage indemnification level for consumer \( i \) denoted as a function of the loss experienced and a parameter \( q_i \in \Phi \) (where \( \Phi \) is a compact choice set), as in \( I_i(L_i, q_i) \), where we require \( I_i(0, q_i) = 0, i = \{1, 2, \ldots, N\} \). The latter function could take a variety of forms. For example, if indemnification promised to consumer \( i \) amounts to full reimbursement of losses subject to a policy limit \( q_i \), the promised indemnification would be:

\[
I_i = I_i(L_i, q_i) = \min \{L_i, q_i\}.
\]  

(1)

A quota share arrangement, where the insurer agrees to reimburse \( q_i \) per dollar of loss, would be represented as:

\[
I_i = I_i(L_i, q_i) = q_i \times L_i.
\]  

(2)

If a consumer experiences a loss, she claims to the extent of the promised indemnification. If total claims are less than company assets, all are paid in full. If not, all claimants are paid at the same rate per dollar of coverage. The total claims submitted are:

\[
I = I(L_1, L_2, \ldots, L_N, q_1, q_2, \ldots, q_N) = \sum_{j=1}^{N} I_j(L_j, q_j),
\]

and we define the consumer’s recovery as:

\[
R_i = \min \left\{ I_i(L_i, q_i), \frac{a}{I} I_i(L_i, q_i) \right\}.
\]

Accordingly, \( \{I \geq a\} = \{\omega \in \Omega \mid I(\omega) \geq a\} \) denote the states in which the company defaults whereas \( \{I < a\} \) are the solvent states. The expected value of recoveries for the \( i \)-th consumer is whence given by:

\[
e_i = \mathbb{E}[R_i] = \mathbb{E}\left[ R_i 1_{(I < a)} + R_i 1_{(I \geq a)} \right] = e^D_i
\]

(3)

There is a frictional cost—including agency, taxes, and monitoring costs—associated with holding assets in the company. We represent the cost as a tax on assets:

\[
\tau \times a,
\]

although it is also possible to represent frictional costs as a tax on equity capital, as in:

\[
\tau \times \left( a - \mathbb{E}\left[ \sum_{i=1}^{N} \min \left\{ I_i(L_i, q_i), \frac{a}{I} I_i(L_i, q_i) \right\} \right] \right)
\]

(4)
and this does not change the ensuing allocation result.

We denote the premium charged to the consumer \( i \) as \( p_i \), and consumer utility may be expressed as:

\[
v_i(a, w_i - p_i, q_1, ..., q_N) = \mathbb{E} \left[ U_i \left( w_i - p_i - L_i + R_i \right) \right],
\]

where \( w_i \) denotes consumer \( i \)'s wealth, and we write \( v_i' (\cdot) = \frac{\partial}{\partial w_i} v_i (\cdot) \).

The firm then solves:

\[
\max_{a, (q_i), (p_i)} \sum p_i - \sum e_i - \tau a,
\]

subject to participation constraints for each consumer:

\[
v_i(a, w_i - p_i, q_1, ..., q_N) \geq \gamma_i \quad \forall i
\]

and subject to a differentiable solvency constraint imposed by the regulator

\[
s(q_1, ..., q_N) \leq a,
\]

where \( s \) is imagined to arise from, for example, a risk measure with a set threshold dictating the requisite capitalization for the firm.

In order to ensure differentiability of the objective function, it is necessary to impose conditions on the distributions of the loss random variables. For instance, for a quota share arrangement (2), the objective function will be continuously differentiable if the \( L_i \) are jointly continuously distributed (see Appendix A for details). For discrete loss distributions, on the other hand, if \( I_i (L_i, \cdot) \) is differentiable, it is evident that the objective function is only piecewise differentiable. In this case, the analytic complications presented by the “kinks” may be overcome by considering the one-sided derivatives similarly to Zanjani (2010). In what follows, with little loss of generality, we simply assume that the solution lies in the differentiable region of the objective function. Let \( \lambda_k \) be the Lagrange multiplier associated with the participation constraint (7) for consumer \( k \), and let \( \xi \) the multiplier associated with (8). The first order conditions for an interior solution are then:

\[
[q_i] \quad - \sum_k \frac{\partial e_k}{\partial q_i} + \sum_k \lambda_k \frac{\partial v_k}{\partial q_i} - \frac{\partial s}{\partial q_i} \xi = 0,
\]

\[
[a] \quad - \sum_k \frac{\partial e_k}{\partial a} - \tau + \sum_k \lambda_k \frac{\partial v_k}{\partial a} + \xi = 0,
\]

\[
[p_i] \quad 1 - \lambda_i \frac{\partial v_i}{\partial w} = 0.
\]

We show in Appendix B that a profit-maximizing firm will be able to achieve the optimum by offering each consumer a smooth and monotonic premium schedule, where consumer \( i \) is free to choose any level of \( q_i \) desired. We denote the variable premium as \( p_i^* (q_i) \) and consider its construction under the assumption that each consumer is a “price taker” and ignores the impact of her own purchase at the margin on the level of recoveries in states of default. This assumption is discussed in Zanjani (2010), who followed the transportation economics literature on congestion pricing (Keeler and Small (1977)) by using the assumption when calculating the optimal consumer pricing function.\(^2\) With this assumption in place, the marginal price change at the optimal level

\(^2\) The assumption is ubiquitous (although often implicit rather than stated) within many fields of economics. Its importance here is that, without it, the marginal cost associated with each consumer’s risk will feature a cost related to an implicit allocation of capital that will not “add up” across consumers to the total capital of the firm.
or


depth previously referenced:

Second, we require the “adding up” property on the regulatory constraint—which is satisfied under conditions previously referenced:

\[
\sum \frac{\partial s}{\partial q_i} q_i = a.
\]
When the foregoing conditions hold, the optimal marginal pricing condition (13) can be extended to fully allocate all of the firm’s costs, including the cost of capital:

$$\sum \frac{\partial p^*_i}{\partial q_i} q_i = \sum \frac{\partial e^Z_i}{\partial q_i} q_i + [P(I \geq a) a + \tau a].$$

Note that the cost of capital as captured in the bracketed term breaks down as:

$$\left[ \sum_k \frac{\partial e_k}{\partial a} a + \tau a \right] = \sum_i \frac{\partial s}{\partial q_i} q_i \times \left[ P(I \geq a) + \tau - \sum_k \frac{\partial v_k}{\partial a} v'_k \right] + \sum_i \tilde{\phi}_i q_i a \times \left[ \sum_k \frac{\partial v_k}{\partial a} \right].$$

So for any one individual consumer, their capital allocation has two components. The first derives from an “internal” marginal cost—driven by the cross-effects of consumers on each other:

$$\tilde{\phi}_i q_i a \times \left[ \sum_k \frac{\partial v_k}{\partial a} v'_k \right]$$

and the second originates from an “external” marginal cost imposed by regulators:

$$\frac{\partial s}{\partial q_i} q_i \times \left[ P(I \geq a) + \tau - \sum_k \frac{\partial v_k}{\partial a} v'_k \right].$$

It is useful at this point to consider several different institutional scenarios.

**Full Coverage by Deposit Insurance and Binding Regulation**

If consumers are fully covered by deposit insurance, they are indifferent to the capitalization of their financial institution. Mathematically, this means that

$$\sum_k \frac{\partial v_k}{\partial a} v'_k = 0,$$

so that (13) becomes:

$$\frac{\partial p^*_i}{\partial q_i} = \frac{\partial e^Z_i}{\partial q_i} + \frac{\partial s}{\partial q_i} [P(I \geq a) + \tau].$$

Thus, the marginal cost of risk, and the attendant allocation of capital, is completely determined by the gradient of the binding regulatory constraint. This is the world of Myers and Read, Tasche, and others involved in the development of the gradient allocation principle. In this world, the marginal cost of risk is indeed completely determined by an arbitrarily chosen risk measure.

**No Deposit Insurance and Non-Binding Regulation**

At the opposite extreme is the case of an unregulated market with no deposit insurance. Here, $$\xi = 0$$, so (cf. Eq. (10)):

$$\sum_k \frac{\partial v_k}{\partial a} v'_k = [P(I \geq a) + \tau],$$
meaning that (13) becomes:

\[ \frac{\partial p^*_i}{\partial q_i} = \frac{\partial e^Z_i}{\partial q_i} + \tilde{\phi}_i a \times [\mathbb{P}(I \geq a) + \tau]. \]

Thus, the marginal cost of risk and the attendant allocation of capital is driven completely by “internal” considerations. Specifically, (14) indicates that the allocation is driven by the time-zero value that affected consumers place on their contingent claims on recoveries in the various states of default.

**General Case: Uninsured Consumers and Binding Regulation**

In general, we may imagine the case where both of the considerations isolated above—an “external” regulatory constraint, and “internal” concerns driven by counterparty preferences—are influencing the marginal cost of risk. In this case, (13) remains in its original form:

\[ \frac{\partial p^*_i}{\partial q_i} = \frac{\partial e^Z_i}{\partial q_i} + \frac{\partial s}{\partial q_i} \left[ \mathbb{P}(I \geq a) + \tau - \sum_k \frac{\partial v_k}{\partial a} v'_k \right] + \tilde{\phi}_i a \times \left[ \sum_k \frac{\partial v_k}{\partial a} v'_k \right], \]

but we are now able to see more clearly the two influences on capital allocation. When the regulatory constraint binding, we know that:

\[ \mathbb{P}(I \geq a) + \tau > \sum_k \frac{\partial v_k}{\partial a} v'_k, \]

with the interpretation that regulation is forcing the institution to hold assets beyond the level that would be privately efficient from the perspective of serving its counterparties. However, the extent of this distortion is key to identifying whether internal counterparty concerns or external regulatory concerns guide capital allocation. If regulation comes close to replicating the private market outcome:

\[ \mathbb{P}(I \geq a) + \tau \approx \sum_k \frac{\partial v_k}{\partial a} v'_k, \]

then the second term in (13) will be unimportant relative to the third term, and internal counterparty concerns will dominate. On the other hand, if regulation has the effect of pushing institutional capitalization well beyond the level that would prevail in the private market:

\[ \mathbb{P}(I \geq a) + \tau \gg \sum_k \frac{\partial v_k}{\partial a} v'_k, \]

then the second term in (13) will overshadow the third term, and external regulatory concerns will dominate.

### 3 Implementation of Internal Capital Allocation

The discussion in the preceding section introduced the notion that internal counterparty concerns will in general influence capital allocation and may be the most important influence in a profit-maximizing firm, even in cases where a regulatory constraint binds. This naturally leads one to
ask 1) how counterparty-driven allocation effectively differs from the now-commonplace methods based on arbitrary risk measures, and 2) if the former approach is easily implemented.

In view of the latter question, notice that

$$\tilde{\phi}_i = \mathbb{E} \left[ \frac{\sum_k \frac{U_k}{v_k} I_k 1_{\{I \geq a\}} \frac{\partial q_i}{\partial \tilde{q}_i}}{\mathbb{E} \left[ \sum_k \frac{U_k}{v_k} I_k 1_{\{I \geq a\}} \right]} \right]$$

$$= \mathbb{E} \left[ \frac{\sum_k \frac{U_k}{v_k} I_k 1_{\{I \geq a\}} \frac{\partial \log \{I_1(L_1, \tilde{q}_1) + \ldots + I_N(L_N, \tilde{q}_N)\}\{\tilde{q}_j = q_j\}}{\partial \tilde{q}_i} \right]$$

$$= \frac{\partial}{\partial \tilde{q}_i} \mathbb{E}_\tilde{P} \left[ \log \left\{ \sum_j I_j(L_j, \tilde{q}_j) \right\} \right]_{\tilde{q}_j = q_j} = \frac{\partial}{\partial \tilde{q}_i} \mathbb{E}_\tilde{P} \left[ \log \{I\} \right]_{\tilde{q}_j = q_j}, \quad (18)$$

where $\bar{I} = I(L_1, \ldots, L_N, \tilde{q}_1, \ldots, \tilde{q}_N)$ and the probability measure $\tilde{P}$ is given via its Radon-Nikodym derivative

$$\frac{\partial \tilde{P}}{\partial P} = \frac{\sum_k \frac{U_k}{v_k} I_k 1_{\{I \geq a\}}}{\mathbb{E} \left[ \sum_k \frac{U_k}{v_k} I_k 1_{\{I \geq a\}} \right]}$$

for fixed policy limits $\{q_j\}$ and assets $a$. Hence, marginal capital is allocated to consumers according to

$$a_i = a \phi_i q_i = a \frac{q_i}{\rho(\bar{I})} \frac{\partial}{\partial \tilde{q}_i} \rho(\bar{I})_{\tilde{q}_j = q_j}, \quad i = 1, \ldots, N,$$

where the risk measure $\rho$ is defined by

$$\rho(X) = \exp \left\{ \mathbb{E}_\tilde{P} \left[ \log \{X\} \right] \right\}.$$

In other words, capital can be allocated according to the gradient approach or the Euler capital allocation principle based on the positively homogeneous risk measure $\rho$ (see Schmock and Straumann (1999) or Section 6.3 in McNeil et al. (2005) for details). Aside from positive homogeneity, $\rho$ obviously is monotonic and satisfies the constancy condition of risk measures (see Frittelli and Gianin (2002) for a discussion of properties of risk measures).

On the other hand, we notice that its functional form is, at first glance, similar to that of the so-called entropic risk measure, which has recently gained increasing popularity in the mathematical finance literature (see e.g. Föllmer and Schied (2002) or Dhaene et al. (2008)). However, the roles of the exponential function and the logarithm are interchanged. While this modification yields positive homogeneity (in contrast to the entropic risk measure) and thus allows for an immediate application of the Euler principle, $\rho$ neither is translation-invariant nor sub-additive, and therefore is not coherent and not convex. Thus, while $\rho$ may have useful properties for the purpose of internal capital allocation at the firm level, it will not necessarily perform well in other applications.\(^3\)

\(^3\)While sub-additivity is subject of an ongoing debate (see e.g. Heyde et al. (2006) or Dhaene et al. (2008)), at least translation invariance is generally deemed adequate for an external risk measure and even “necessary for the risk-capital interpretation [...] to make sense” (see p. 239 in McNeil et al. (2005)).
Furthermore, this discussion of the general properties of $\tilde{\rho}$ misses the point that the risk measure—and particularly the weighting implied by the change of measure—were tailored for the measurement of a specific random variable and its neighborhood, namely the aggregate loss $I$. More specifically, we have

$$\tilde{\rho}(I) = \exp \left\{ \mathbb{E}^{\tilde{\mathbb{P}}}[\log \{I\}] \right\} = \exp \left\{ \mathbb{E} \left[ \frac{\partial \tilde{\mathbb{P}}}{\partial \mathbb{P}} \log \{I\} \right] \right\} = \exp \left\{ \mathbb{E} \left[ \mathbb{E} \left[ \frac{\partial \tilde{\mathbb{P}}}{\partial \mathbb{P}} I \log \{I\} \mid I \geq a \right] \right] \right\}$$

which illustrates that $\tilde{\rho}$ in this sense is in fact a tail risk measure, and hence is related to the Expected Shortfall.

Here, the weights $\tilde{\psi}(\cdot)$ perform a role similar to the risk spectrum within the so-called spectral risk measures as introduced in Acerbi (2002). According to the author, the “subjective risk aversion of an investor [or a regulator] can be encoded” in this function, which may justify over-weighting bad outcomes, but he does not provide guidance on how to choose an explicit form. To close this gap, Dowd et al. (2008) provide some ad-hoc examples whereas Sriboonchitta et al. (2010) and Wächter and Mazzoni (2010) attempt to establish theoretical links of the risk spectrum to the preferences of the user of the risk measure—or, depending on its application, the preferences of an external supplier such as a regulator—by relying on results from robust statistics and the dual theory of choice, respectively. In contrast, in our setting the weights represent an adjustment to objective probabilities based on the value placed by claimants on recoveries in various states of default. Thus, the pivotal characteristics for our weights lie in the primitives of the firm’s profit maximization problem (namely, the preferences of counterparties)—which ultimately determine the overall choice of capitalization as well as the values consumers place on state contingent recoveries—rather than in an arbitrarily specified concave preference function for the firm, which will generally fail to capture limited liability.

The answer to the question of how counterparty-driven allocation effectively differs from the allocation based on common risk measures such as the Expected Shortfall lies in the form of the weights. For instance, if a specification satisfied

$$\frac{\partial \tilde{\mathbb{P}}}{\partial \mathbb{P}} = \tilde{\psi}(I) = \text{const} \times I \times 1_{\{I \geq a\}},$$

Equation (18) implies that the resulting counterparty-driven allocation would then be identical to a gradient allocation based on the Expected Shortfall (for a suitably chosen level—see the next section for details). Similarly, it is conceivable that other specifications may result in qualitatively different outcomes, in either direction. The next section sheds more light on these issues by considering a variety of situations (including one which satisfies (19)).
4 Comparison of Capital Allocation Methods

In this section, we consider the question of the practical and numerical implications of allocating capital based on the method discussed in the previous sections. In particular, we compare the results to the more widely used risk measures of Expected Shortfall. We illustrate in the context of several examples.

4.1 The Case of Exponential Losses

Assume that there are \( N \) identical consumers with wealth level \( w \) in a regime with non-binding regulation that face independent, Exponentially distributed losses \( L_i \sim \text{Exp}(\nu) \), \( 1 \leq i \leq N \). Assume further that all consumers exhibit a constant absolute risk aversion of \( \alpha < \nu \), and that their participation constraint is given by the autarky level

\[
\gamma = \gamma_i = \mathbb{E} [U(w - L_i)] = -e^{-\alpha w} \frac{\nu}{\nu - \alpha}.
\]

Then, the optimization problem (6)/(7)/(8) may be written as

\[
\begin{align*}
\max_{a,q,p} & \quad N \times p - N \times q \times \left[ \frac{1}{p} \Gamma_{N-1,\nu} \left( \frac{a}{q} \right) - \nu^{N-1} \left( \frac{a}{q} \right)^{N-1} \left( \frac{1}{p} + \frac{1}{q} \right) \right] \\
\text{subject to} & \quad \gamma \leq e^{-\alpha (w-p)} \left\{ \nu \frac{\Gamma_{N-1,\nu} \left( \frac{a}{q} \right) - \nu^{N-1} \left( \frac{a}{q} \right)^{N-1} \left( \frac{1}{p} + \frac{1}{q} \right)}{\nu - \alpha} + \sum_{k=0}^{\infty} \frac{a^k}{(N-1+k)!} \left( \frac{\nu}{1-\nu} \right)^k \right\},
\end{align*}
\]

where \( \Gamma_{m,b}(x) = 1 - \text{Exp}(x) \) and \( \Gamma_{m,b}(\cdot) \) denotes the cumulative distribution function of the Gamma distribution with parameters \( m \) and \( b \) (see the Appendix C for the derivation of (20)).

For the allocation of capital to the individual consumers, we trivially obtain

\[
q \hat{\phi}_i = N^{-1}, \ i = 1, 2, \ldots, N,
\]

which is the same for any risk measure. More specifically,

\[
q \hat{\phi}_i \quad \text{Eq.(14)} \quad \mathbb{E} \left[ \mathbf{1}_{\{qL \geq a\}} \mathbb{E} \left[ \sum_{j=1}^{N} U \left( w - p - L_j + a \frac{L_j}{L} \right) \left| L \right| \right] \right] = \frac{1}{N} \mathbb{E} \left[ \hat{\psi}(L) \mid qL \geq a \right],
\]

where \( L = \sum_j L_j \),

\[
\hat{\psi}(l) = \mathbf{1}_{\{q \geq a\}} \hat{c}_{N,\nu,\alpha,a,q} \sum_{k=0}^{\infty} \frac{(k+1) (\alpha l - a)^k}{(N+k)!},
\]

and \( \hat{c}_{N,\nu,\alpha,a,q} \) is a constant ascertaining \( \mathbb{E} \left[ \hat{\psi}(L) \mid qL \geq a \right] = 1 \). For the risk measure \( \bar{\rho} \), on the other hand, we obtain

\[
\bar{\rho}(I) = \bar{\rho}(qL) = \exp \left\{ \mathbb{E} \left[ \hat{\psi}(I) \log \{I \} \mid I \geq a \right] \right\} = \exp \left\{ \mathbb{E} \left[ \hat{\psi}(L) \log \{qL \} \mid qL \geq a \right] \right\},
\]
i.e. $\hat{\psi}(\cdot)$ is the corresponding weighting function.\footnote{The derivation of these equations, a closed form solution for $\hat{c}_{N,\nu,\alpha,a,q}$ as well as a representation of $\hat{\psi}(\cdot)$ not involving an infinite sum for implementation purposes all are provided in Appendix C.} Hence, the risk measure $\bar{\rho}$ in this case naturally accounts for risk aversion ($\alpha$) as well as for diversification effects ($N$).

For the allocation based on the Expected Shortfall (ES) according to the Euler principle, it is well known that (see e.g. Schmock and Straumann (1999) or Dhaene et al. (2009)):\footnote{Here, we always assume that the confidence level is chosen corresponding to the counterparty-driven allocation, namely $P(qL \geq a)$ in this case.}

$$a_i = \frac{q \mathbb{E}[L_i | qL \geq a]}{\mathbb{E}[I | I \geq a]} = \frac{\mathbb{E}[\mathbb{E}[L_i | L | qL \geq a]]}{\mathbb{E}[L | qL \geq a]} = \frac{1}{N} \mathbb{E}[\text{const} \times L | qL \geq a],$$

i.e. the Expected Shortfall can be associated with a linear weighting function of the loss states. Since $\hat{\psi}(\cdot)$ is increasing and strictly convex for all risk aversion levels $\alpha > 0$, there always exists a loss level $l_0$ such that the weighting function for the counterparty-driven allocation will be higher for all loss levels greater than $l_0$. In this sense, the allocation based on $\bar{\rho}$ always appears more conservative in the current setting. However, we also see that for fixed parameters,

$$\hat{\psi}(l) \rightarrow (P(qL \geq a))^{-1} > \frac{a}{\mathbb{E}[qL | qL \geq a]}, \quad N \rightarrow \infty,$$

which is the left end-point for the Expected Shortfall weighting function. Similarly, for $\alpha = 0$, we obtain $\hat{\psi}(l) \equiv (P(qL \geq a))^{-1}$, i.e. a flat weighting function. Thus, for large enough companies or risk-neutral consumers, the weight on relatively low loss levels will always be higher for the counterparty-driven allocation, rendering it to appear less conservative.

### Exponential Losses: Parametrizations

<table>
<thead>
<tr>
<th>Nr.</th>
<th>$N$</th>
<th>$\nu$</th>
<th>$\tau$</th>
<th>$\alpha$</th>
<th>$w$</th>
<th>$a$</th>
<th>$p$</th>
<th>$q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>2.0</td>
<td>0.050</td>
<td>0.25</td>
<td>3.0</td>
<td>1.4663</td>
<td>0.2598</td>
<td>0.5713</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>2.0</td>
<td>0.050</td>
<td>1.25</td>
<td>3.0</td>
<td>4.0036</td>
<td>0.7401</td>
<td>0.9494</td>
</tr>
</tbody>
</table>

Table 1: Parametrizations of the Exponential Losses model.

To further analyze this relationship, in Table 1 we present two parametrizations of the setup and the corresponding optimal parameters $a$, $p$, and $q$ as solutions of the program (20). The properties of the solutions are as may be expected: $a$, $p$, and $q$ all are increasing in risk aversion. Figure 1 now shows the weighting function $\hat{\psi}$ as opposed to the linear weighting function associated with the Expected Shortfall for varying risk aversion levels. We find two qualitatively different shapes:\footnote{Analyses with respect to other parameters such as company size $N$ or the expected loss $1/\nu$ show similar results.} For the high risk aversion level, $\hat{\psi}$ crosses the linear weighting function once from below; thus, in this case, relatively lower loss states are weighted more heavily for the allocation based on the Expected Shortfall, whereas the weighting is higher for the counter-party driven allocation in high...
loss states. For the low risk aversion level, \( \hat{\psi} \) crosses the linear weighting function twice; in this case, the weighting function within the new risk measure \( \bar{\rho} \) puts more mass on low and extremely high loss states, while the weights are smaller for moderate to high loss states.

![Figure 1: Weighting function \( \hat{\psi} \) for varying risk aversion parameter \( \alpha \).](image)

Hence, when relying on the Expected Shortfall for the purpose of internal capital allocation, the loss-specific weights may be too conservative or not conservative enough, depending on, among other factors, company size or the risk aversion level. These considerations are naturally taken into account by the risk measure \( \bar{\rho} \).

### 4.2 The Case of Homogenous Bernoulli Losses

Again, we consider \( N \) identical consumers with wealth level \( w \) in a regime with non-binding regulation whose preferences are given by the (same) smooth utility function \( U(\cdot) \). However, in contrast to the previous section, we now assume that the consumers face Bernoulli distributed losses \( L_i \), \( 1 \leq i \leq N \), with loss level \( l \) and loss probability \( \pi \). Their participation constraint once again is given by the autarky level

\[
\gamma = \gamma_i = \mathbb{E}[U(w - L_i)].
\]
In this case, the optimization problem (6)/(7)/(8) takes the form

\[
\max_{a,q,p} \left\{ N \times \left( p - \pi \times \sum_{k=0}^{N-1} \left( \frac{N-1}{k} \right) \pi^k (1-\pi)^{N-1-k} \times \left[ qI_{\{k<\frac{N}{q}\}} + \frac{p}{k+1} I_{\{k\geq\frac{N}{q}\}} \right] - \tau \times a \right\} 
\]

subject to
\[
\gamma \leq (1-\pi) U(w-p) + \pi \times \sum_{k=0}^{N-1} \left( \frac{N-1}{k} \right) \pi^k (1-\pi)^{N-1-k} \times \\
U(w-p - (1-q)l) I_{\{k<\frac{N}{q}\}} + U(w-p - l + \frac{a}{N}) I_{\{k\geq\frac{N}{q}\}} \right].
\]

For the allocation to the individual consumer, similarly to the previous section, we obtain
\[
q \tilde{\phi}_i = \frac{1}{N} \mathbb{E} \left[ \frac{1}{\mathbb{E} \left[ I_{\{\Gamma \geq \frac{a}{q}\}} \right]} U'(w-p-l+\frac{a}{|\Gamma|}) \right] = \frac{1}{N},
\]
where $|\Gamma|$ denotes the number of total losses. And the risk measure now takes the form
\[
\tilde{\rho}(I) = \tilde{\rho}(qI \mid \Gamma) = \exp \left\{ \mathbb{E} \left[ \tilde{\psi}(I) \log \{I\} \mid I \geq a \right] \right\} = \exp \left\{ \mathbb{E} \left[ \tilde{\psi}(|\Gamma|) \log \{qI \mid |\Gamma|\} \mid |\Gamma| \geq \frac{a}{qI} \right] \right\},
\]
whereas for the Expected Shortfall based allocation, again similarly to Section 4.1,
\[
\frac{a_i}{a} = \frac{q \mathbb{E} [L_i \mid qL \geq a]}{\mathbb{E} [I \mid I \geq a]} = \frac{1}{N} \mathbb{E} \left[ \text{const} \mid |\Gamma| \mid |\Gamma| \geq \frac{a}{qI} \right],
\]
so again it can be associated with a linear weighting function.

Hence, in this case, the weighting function for $\tilde{\rho}$ is a composition of the marginal utility and a reciprocal function. As such, it will always be increasing if consumers are risk averse, and again we obtain a flat allocation for the risk neutral case. For $w \geq p + l$, a sufficient condition for the concavity of $\tilde{\psi}$ is a level of relative prudence smaller than two (see e.g. Kimball (1990) for the concept of relative prudence). However, for high levels of relative prudence, a convex shape is possible. Thus, again, it is not immediately clear how the counterparty-driven allocation compares to the linear weighting implied by the Expected Shortfall measure.

Table 2 now displays several parametrizations in this setup for different Constant Absolute Risk Aversion (CARA) and Constant Relative Risk Aversion (CRRA) utility functions, where the optimal parameters $a$, $p$, and $q$ are determined as solutions of program (22); again, they exhibit the expected relationships. Figure 2 now shows the weighting function $\tilde{\psi}$ for parametrizations 1 and 2 (CARA utility) as well as the corresponding weighting function for the Expected Shortfall (ES) and risk-neutral consumers ($\alpha = 0$). While in contrast to the previous section $\tilde{\psi}$ is concave in both cases, again we obtain two qualitatively different relationships: For the low risk aversion, $\tilde{\psi}$ crosses the linear weighting function from above, implying a less conservative state-specific allocation; for the strong risk averser, on the other side, the crossing is from below, ensuing a more conservative configuration. So yet again, risk aversion appears to play an important role.

For the CRRA case, the contract and firm parameters were chosen so that wealth is close to premium level plus losses, since for $w = p + l$, $\tilde{\psi}$ will be concave (convex) if and only if relative
prudence is smaller (greater) than two. In particular, for log-utility, where relative prudence is constant and equals two, the weighting will be linear and, therefore, identical to the Expected Shortfall weights. This can also be seen from Figure 3, where the weighting functions for different relative risk aversion levels $\gamma$ are plotted: For log-utility and $w \approx p + l$, the ES weighting and $\bar{\psi}$ are roughly identical. In contrast, for the lower constant relative risk aversion—and consequently a lower level of relative prudence—the crossing is from above and the shape is concave (panel (b)). For the higher level of constant relative risk aversion and whence more prudent consumers, the crossing is from above and the shaper of $\bar{\psi}$ is convex (panel (c)).

Thus, we find that for the shape of the weighting function $\bar{\psi}$ and accordingly for the comparison of the resulting statewise allocation with that implied by the Expected Shortfall, not only risk aversion but—at a second order—also other characteristics of the consumers’ utility functions are relevant. In particular, we can give a positive answer to the existence question raised towards the end of Section 3: Indeed, there exist special cases where the weighting is linear so that the counterparty-driven allocation and the Expected Shortfall based allocation are identical, although in general—of course—they will differ.

### 4.3 The Case of Heterogenous Bernoulli Losses

Similarly to the last section, we consider consumers that face Bernoulli distributed losses. However, in contrast to the previous setup, we now allow for heterogeneity in consumer preferences as well as in the losses. More specifically, we assume that there are $m$ groups of consumers, where group $i$ contains $N_i$ identical consumers with wealth level $w_i$ and utility function $U_i(\cdot)$ that face independent losses $l_i$ occurring with a probability $\pi_i$, $i = 1, \ldots, m$. The participation constraint again is given by their autarky levels:

$$\gamma_i = \mathbb{E} [U_i(w_i - L_i)] = \pi_i U_i(w_i - l_i) + (1 - \pi_i) U_i(w_i).$$

The observation that for $\gamma = 1$ we obtain the Expected Shortfall is a curious analogy to the so-called power spectral risk measures from Dowd et al. (2008), who show that these measures degenerate into the Expected Loss for $\gamma = 1$, and that there is a qualitative difference between the case $\gamma < 1$ and $\gamma > 1$.  

---

#### Homogeneous Bernoulli Losses: Parametrizations

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Type</th>
<th>$N$</th>
<th>$l$</th>
<th>$\pi$</th>
<th>$\tau$</th>
<th>$\alpha/\gamma$</th>
<th>$w$</th>
<th>$a$</th>
<th>$p$</th>
<th>$q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CARA</td>
<td>100</td>
<td>3.00</td>
<td>0.10</td>
<td>0.05</td>
<td>0.1</td>
<td>5.0</td>
<td>24.8970</td>
<td>0.2727</td>
<td>0.9762</td>
</tr>
<tr>
<td>2</td>
<td>CARA</td>
<td>100</td>
<td>3.00</td>
<td>0.10</td>
<td>0.05</td>
<td>1.0</td>
<td>5.0</td>
<td>40.4654</td>
<td>1.0585</td>
<td>0.9969</td>
</tr>
<tr>
<td>3</td>
<td>CRRA</td>
<td>10</td>
<td>3.25</td>
<td>0.45</td>
<td>0.05</td>
<td>0.5</td>
<td>5.0</td>
<td>12.9492</td>
<td>1.2893</td>
<td>0.8704</td>
</tr>
<tr>
<td>4</td>
<td>CRRA</td>
<td>10</td>
<td>3.25</td>
<td>0.45</td>
<td>0.05</td>
<td>1.0</td>
<td>5.0</td>
<td>16.5461</td>
<td>1.6882</td>
<td>0.9350</td>
</tr>
<tr>
<td>5</td>
<td>CRRA</td>
<td>10</td>
<td>3.25</td>
<td>0.45</td>
<td>0.05</td>
<td>2.0</td>
<td>5.0</td>
<td>19.5488</td>
<td>2.1840</td>
<td>0.9693</td>
</tr>
</tbody>
</table>

Table 2: Parametrizations of the Bernoulli model.
The marginal cost of risk, risk measures, and capital allocation

Figure 2: Weighting function $\tilde{\psi}$ for varying absolute risk aversion parameter $\alpha$; parametrizations 1 and 2 (CARA utility).

The optimization problem (6)/(7)/(8) can then be easily set up by noticing that the number of losses in the different groups follow independent Binomial$(N_i, \pi_i)$ distributions.

For the counterparty-based allocation, we obtain for each group $i$

$$q_i \tilde{\phi}_i = \tilde{c} \sum_{k_1}^{N_1} \cdots \sum_{k_i}^{N_i} \cdots \sum_{k_m}^{N_m} \left( \frac{N_1}{k_1} \right) \cdots \left( \frac{N_i}{k_i} \right) \cdots \left( \frac{N_m}{k_m} \right)$$

$$\times \pi_1^{k_1} \cdots \pi_i^{k_i} \cdots \pi_m^{k_m} (1 - \pi_1)^{N_1-k_1} \cdots (1 - \pi_i)^{N_i-k_i} \cdots (1 - \pi_m)^{N_m-k_m}$$

$$\times \left\{ \sum_{j=1}^{m} k_j U'(w_j - p_j - l_j + q_j l_j) \frac{a}{k_1 q_1 l_1 + \cdots + k_m q_m l_m} \right\}$$

$$\times \frac{k_i q_i l_i}{k_1 q_1 l_1 + \cdots + k_m q_m l_m},$$

where $\tilde{c}$ is a constant such that $\sum_i q_i \tilde{\phi}_i = 1$. Thus, while the analytical form of the weights $\tilde{\psi}(I) = E \left[ \frac{\partial \tilde{\phi}}{\partial P} \mid I \right]$ is less transparent in this case, again we notice that they immediately depend on the marginal utilities of recoveries in various states of default. For the allocation based on the
The marginal cost of risk, risk measures, and capital allocation

Figure 3: Weighting function $\bar{\psi}$ for varying relative risk aversion parameter $\gamma$; parametrizations 3, 4, and 5 (CRRA utility).
Expected Shortfall, on the other hand, we obtain

\[
\frac{a_i}{a} = \text{const} \sum_{k_1}^{N_1} \cdots \sum_{k_i}^{N_i} \cdots \sum_{k_m}^{N_m} \binom{N_1}{k_1} \cdots \binom{N_i}{k_i} \cdots \binom{N_m}{k_m} \times \pi_i \cdots \pi_{m_i} \cdots \pi_{m_m} (1 - \pi_1)^{N_1 - k_1} \cdots (1 - \pi_i)^{N_i - k_i} \cdots (1 - \pi_m)^{N_m - k_m} \times 1_{\{k_1 q_1 l_1 + \cdots + k_m q_m l_m \geq a\}}(k_i q_i l_i),
\]

i.e. it is of a similar form as (23) but now 1) does not contain the adjustment based on the marginal utilities \(\frac{\partial \tilde{P}}{\partial P}\), and 2) the state-specific loss for consumer \(i\), \((k_i q_i l_i)\), is not scaled by the aggregate loss—which, in the counterparty-driven allocation, is a consequence of the proportional partitioning of the recoveries in states of default.

To assess the consequences of these adjustments, we consider the case of identical group sizes \(N_i = N\), identical CARA preferences and wealth levels throughout the population, identical loss probabilities \(\pi_i = \pi\), but differing loss levels. More specifically, in Table 3, we present model parametrizations for a setup with \(m = 3\) groups, whose members face loss levels \(l_1 = 1\), \(l_2 = 2\), and \(l_3 = 3\), respectively. For the first five parametrizations, the only difference in the assumed parameters is the group size but remarkably the influence on premiums \(\{p_i\}\) and the choice parameters \(\{q_i\}\) is marginal, and may be completely attributable to numerical inaccuracies. Since the participation constraint binds, it appears that the solution features an adjustment of the asset level such that the utility level is the same in all cases. In particular, since the resulting asset level is concave in the group size due to obvious diversification effects, this implies that the companies’ monopoly rents increase disproportionate to the firm size.

### Heterogenous Bernoulli Losses: Parametrizations

<table>
<thead>
<tr>
<th>Nr.</th>
<th>(\alpha)</th>
<th>(N)</th>
<th>(\pi)</th>
<th>(\tau)</th>
<th>(w)</th>
<th>(a)</th>
<th>(p_1)</th>
<th>(p_2)</th>
<th>(p_3)</th>
<th>(q_1)</th>
<th>(q_2)</th>
<th>(q_3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>2</td>
<td>0.1</td>
<td>0.01</td>
<td>5</td>
<td>5.37</td>
<td>0.16</td>
<td>0.49</td>
<td>1.06</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>5</td>
<td>0.1</td>
<td>0.01</td>
<td>5</td>
<td>8.84</td>
<td>0.16</td>
<td>0.49</td>
<td>1.06</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>10</td>
<td>0.1</td>
<td>0.01</td>
<td>5</td>
<td>13.63</td>
<td>0.16</td>
<td>0.49</td>
<td>1.06</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>15</td>
<td>0.1</td>
<td>0.01</td>
<td>5</td>
<td>17.86</td>
<td>0.16</td>
<td>0.49</td>
<td>1.06</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>25</td>
<td>0.1</td>
<td>0.01</td>
<td>5</td>
<td>25.76</td>
<td>0.16</td>
<td>0.49</td>
<td>1.06</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>6</td>
<td>1.25</td>
<td>10</td>
<td>0.1</td>
<td>0.01</td>
<td>5</td>
<td>14.07</td>
<td>0.18</td>
<td>0.60</td>
<td>1.31</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 3: Parametrizations of the heterogeneous Bernoulli model with CARA preferences and loss levels \((l_1, l_2, l_3) = (1, 2, 3)\).

Figure 4 shows the counterparty-based allocation of capital to the three groups for parametrizations 1 through 5 as well as the corresponding allocations based on the Expected Shortfall (ES). At first glance, the two graphs look roughly identical. More specifically, for small group sizes, both of them entail a disproportionate allocation of capital to the high loss group \((l = 3)\), whereas for an increasing group size, both allocations become more and more linear. This clearly is a
consequence of diversification: For small group sizes, and individual loss in the high loss group is closely linked to a company default, whereas for extremely large group sizes, individual losses in either group and company defaults become essentially independent. Taking into account that positive contributions for either allocation (23) or (24) are only accrued in states where 1) the individual in view incurs a loss and 2) the company defaults, the similarity of the graphs becomes more perspicuous. In particular, this effect appears to dominate other influences stemming from, for instance, risk aversion.

Figure 4: Allocation for varying company sizes $N$; parametrizations 1-5.

To demonstrate that such effects are present nonetheless when adjusting the focus, Figure 5 illustrates the allocation of capital to the low ($l = 1$) and high ($l = 3$) loss group for parametrization 6 and varying levels of absolute risk aversion. Obviously, the allocation based on the expected shortfall (ES) is independent of consumer preferences. For the counterparty-driven allocation, on the other hand, we find a decreasing and slightly concave dependence for the low loss group (panel (a)), and an increasing and slightly convex relationship for the high loss group (panel (b)). In particular, there exist risk aversion levels for which either of the allocations yield higher/lower charges for the high/low loss group, although in this case already moderate levels of risk aversion render the counterparty-driven allocation to penalize the high loss group more heavily.

Figure 5: Allocation for varying absolute risk aversion parameter $\alpha$; parametrization 6.
5 Conclusion

The early literature on capital allocation recognized that risk measure selection was a thorny issue that could be resolved only through careful consideration of institutional context. The subsequent literature on capital allocation celebrated, refined, and justified (in mathematical terms) the technique of Euler allocation technique while ignoring the institutional context, with the consequence that the demon of arbitrary risk measure selection has never been exorcised. The demon, however, cannot and should not be ignored: It is well known (see Basak and Shapiro (2001)) that risk measure selection has a profound influence on an organization’s perception of the cost of risk, and, consequently, on the systemic risks to which society is exposed.

This paper identifies the risk measure consistent with the marginal cost of risk from the perspective of a profit-maximizing firm with risk averse counterparties. Surprisingly, this risk measure is neither coherent nor convex; nevertheless, it is the only one that yields the appropriate allocation of capital for the profit-maximizing firm. At first glance, it would seem that spectral risk measures such as spectral ES could be designed to deliver the appropriate results due to flexibility in the design of the spectral weighting function—yet we have shown that this is not necessarily the case because ES-based allocation could either be underweighting or overweighting severe states of default, depending on the nature of customer risk aversion, and a spectral weighting function as traditionally defined will only be able to address the former problem.

The reason for the difference in allocations derives from a fundamental difference in the source of marginal cost in an economic model based on counterparty risk aversion and one based on the imposition of ES or spectral ES. In the economic model, marginal cost derives from the impact of the expansion of a particular risk exposure on the recoveries of the counterparties to the firm, which is determined by the assets of the firm and the value that counterparties place on those assets in various states of default. ES-based allocation, on the other hand, is driven by loss outcomes rather than recoveries.

Finally, we have shown that the marginal cost of risk to the firm depends on the risk preferences of its counterparties, even in cases where it faces a binding regulatory constraint—so much so that the influence of the regulatory constraint on capital allocation could be dominated by the influence of counterparty risk preferences. Even in less extreme circumstances, however, we argue that the risk preferences of counterparties must generally be considered in the capital allocation process, and for this purpose we recommend an approach based on a newly introduced risk measure.

Appendix

A On the Differentiability of the Objective Function

To keep the presentation concise, we limit our attention to the objective function only and solely consider the case of $N = 2$ consumers under a quota share agreement; similar considerations apply for the constraints and the case of $N > 2$ consumers. Denote the corresponding losses by $L_1$ and $L_2$ with corresponding (joint) probability density function $f(x, y)$. Then the objective function (6)
can be written as

\[
\sum_i p_i - \sum_i \mathbb{E} \left[ \min \left\{ q_i L_i, \frac{a}{q_1 x_1 + q_2 x_2} q_i x_i \right\} \right] - \tau a
\]

Obviously, this function is (continuously) differentiable with respect to \( p_i \). For the differentiability with respect to \( a \) and \( q_i \), problems may only arise for the summands of the second term. Without loss of generality, we focus on the first summand. Here,

\[
\int_0^\infty \int_0^\infty \min \left\{ q_1 x_1, \frac{a}{q_1 x_1 + q_2 x_2} q_1 x_1 \right\} f(x_1, x_2) \, dx_1 \, dx_2 = \int_0^\infty \int_0^\infty \frac{a}{q_1 x_1 + q_2 x_2} q_1 x_1 f(x_1, x_2) \, dx_1 \, dx_2.
\]

By the Leibniz rule, after some simple calculus the derivatives with respect to \( a, q_1, \) and \( q_2 \) are thus given by

\[
\frac{\partial}{\partial a} = \int_0^\infty \int_0^a \frac{q_1 x_1}{q_1 x_2 + q_2 x_2} f(x_1, x_2) \, dx_1 \, dx_2
\]

\[
\frac{\partial}{\partial q_1} = \int_0^\infty \int_0^a \frac{a q_2 x_2}{(q_1 x_2 + q_2 x_2)^2} f(x_1, x_2) \, dx_1 \, dx_2, \quad \text{and}
\]

\[
\frac{\partial}{\partial q_2} = - \int_0^\infty \int_0^a \frac{a q_1 x_1}{(q_1 x_2 + q_2 x_2)^2} f(x_1, x_2) \, dx_1 \, dx_2,
\]

respectively, which are obviously continuous for \( a, q_1, q_2 > 0 \).

**B. Implementation of the Solution to (6)/(7)/(8) via a Premium Schedule**

In the text we consider the solution of maximizing (6) subject to (7) and (8). We claim further that—if the consumer acts as a “price taker” with respect to the recovery rates offered by the company within the various states of default—that the company can implement the optimum by offering a smooth and monotonically increasing premium schedule that allows each consumer to freely choose the level of coverage desired for the premium indicated by the schedule. It is subsequently shown that the marginal price increase associated with coverage must satisfy (13) when evaluated at the optimum. It remains to be shown that this premium schedule exists and can be used to implement the optimum.

A complication arises in modeling the consumer as a price-taker with free choice of coverage level. To introduce the consumer’s ignorance of his own influence on recoveries, we define price schedule described above as \( p^*_i(\cdot) \) and modify the original utility function to

\[
\tilde{v}_i \left( w_i - p^*_i(q_i), q_i; \bar{a}, \bar{q}_1, \ldots, \bar{q}_N \right) = \mathbb{E} \left[ U_i \left( w_i - p^*_i(q_i) - L_i + \tilde{R}_i \right) \right], \quad (25)
\]
where
\[
\tilde{R}_i = \tilde{R}_i (q_i; \tilde{a}, \tilde{q}_1, \ldots, \tilde{q}_N) = \min \left\{ I_i(L_i, q_i), \frac{\tilde{a}}{\sum_{j=1}^{N} I_j(L_j, \tilde{q}_j)} I_i(L_i, q_i) \right\}.
\]

The idea here is to fix recovery rates by fixing the quantities \( \tilde{a} \) and \( \{\tilde{q}_i\} \), leaving the consumer with the free choice of \( q_i \)—but with the caveat that this choice does not influence recovery rates.

The firm’s objective function is identical to the previous one, except that 1) the firm now specifies a price function rather than a single price point, and 2) the firm fixes the recovery rates for purposes of consumer incentive compatibility by choosing \( \tilde{a} \) and \( \{\tilde{q}_i\} \) instead of the “true” levels of \( a \) and \( \{q_i\} \):
\[
\max_{\tilde{a}, \{p_i^*(\cdot)\}, \{\tilde{q}_i\}} \left\{ \sum p_i^*(\tilde{q}_i) - \sum e_i - \tau \tilde{a} \right\}.
\]
(26)

The firm still faces the previous constraints (7) and (8),
\[
\begin{align*}
v_i (\tilde{a}, w_i - p_i^*(\tilde{q}_i), \tilde{q}_1, \ldots, \tilde{q}_N) & \geq \gamma_i, \\
(\tilde{q}_1, \ldots, \tilde{q}_N) & \leq a,
\end{align*}
\]
and in addition the new constraint:
\[
\tilde{q}_i \in \arg \max_{q_i} \tilde{v}_i (w_i - p_i^*(q_i), q_i ; \tilde{a}, \tilde{q}_1, \ldots, \tilde{q}_N), \ \forall i.
\]
(27)

Equation (27) is an incentive compatibility constraint requiring the choice of coverage level to be consistent with the consumer optimizing, given her perception of own utility (which ignores her own impact on recovery rates) and the selected pricing function.

It is evident that the firm’s profits under this maximization can be no better than those achieved under the original program (maximizing (6) subject to (7) and (8)), since we have simply added another constraint and choosing the premium schedule at different points than \( \tilde{q}_i \) is immaterial to the company’s profits. It is therefore clear that, given optimal choices \( \tilde{a} \), \( \{\tilde{q}_i\} \), and \( \{\tilde{p}_i\} \) to the original program, the firm would maximize profits under the new setup if it could choose those same asset and coverage levels and find a pricing function \( p_i^*(\cdot) \) that both satisfies \( p_i^*(\tilde{q}_i) = \tilde{p}_i \) and induces consumers to choose the original solution:
\[
\hat{q}_i \in \arg \max_{q_i} \tilde{v}_i (w_i - p_i^*(q_i), q_i ; \tilde{a}, \hat{q}_1, \ldots, \hat{q}_N), \ \forall i.
\]

The following lemma shows that this function exists.

Lemma B.1. Suppose \( \tilde{a} \), \( \{\tilde{q}_i\} \), and \( \{\tilde{p}_i\} \) are the optimal choices maximizing (6) subject to (7) and (8). Then, for each \( i \), there exists a smooth, monotonically increasing function \( p_i^*(\cdot) \) satisfying:

1. \( p_i^*(\hat{q}_i) = \tilde{p}_i \).

2. \( \hat{q}_i \in \arg \max_{q_i} \tilde{v}_i (w_i - p_i^*(q_i), q_i ; \tilde{a}, \hat{q}_1, \ldots, \hat{q}_N) \).

Proof. Start by noting that it is evident that the constraints (7) all bind. Note further that the function of \( x \)
\[
g(x) = \tilde{v}_i (w_i - x, 0 ; \tilde{a}, \hat{q}_1, \ldots, \hat{q}_N)
\]
is monotonically decreasing and, hence, invertible, so that we may uniquely define:

\[ p_i^*(0) = g^{-1}(\gamma_i), \]  

(28)

which obviously satisfies

\[ \tilde{v}_i(w_i - p_i^*(0), 0; \hat{a}, \hat{q}_1, \ldots, \hat{q}_N) = \gamma_i. \]

Furthermore, let \( p_i^*(\cdot) \) be a solution to the differential equation (initial value problem)\(^8\)

\[ \frac{\partial p_i^*(x)}{\partial x} = \frac{\partial}{\partial w} \tilde{v}_i(w_i - p_i^*(x), x; \hat{a}, \hat{q}_1, \ldots, \hat{q}_N) \cdot p_i^*(0) = g^{-1}(\gamma_i), \]  

(29)

on the compact choice set for \( q_i \). Due to Peano’s Theorem, we are guaranteed existence of such a function and that it is smooth. Moreover, since \( \frac{\partial \tilde{v}_i}{\partial w}, \frac{\partial \tilde{v}_i}{\partial q_i} > 0 \), we know that the function is monotonically increasing.

Moving on, by construction we know that:

\[
\begin{align*}
\tilde{v}_i(w_i - p_i^*(q_i), q_i; \hat{a}, \hat{q}_1, \ldots, \hat{q}_N) &= \gamma_i + \int_0^{q_i} \left[ \frac{\partial}{\partial q_i} \tilde{v}_i(w_i - p_i^*(x), x; \hat{a}, \hat{q}_1, \ldots, \hat{q}_N) - \frac{\partial}{\partial w} \tilde{v}_i(w_i - p_i^*(x), x; \hat{a}, \hat{q}_1, \ldots, \hat{q}_N) \times \frac{\partial p_i^*(x)}{\partial x} \right] dx \\
&= \gamma_i + 0, \quad q_i > 0.
\end{align*}
\]

(30)

In particular,

\[ \tilde{v}_i(w_i - p_i^*(\hat{q}_i), \hat{q}_i; \hat{a}, \hat{q}_1, \ldots, \hat{q}_N) = \gamma_i, \]

which, since it is evident that the constraints (7) all bind in the original optimization, can be true if and only if

\[ p_i^*(\hat{q}_i) = \hat{p}_i, \]

proving the first part of the lemma. Moreover, (30) directly implies that

\[ \hat{q}_i \in \arg \max_{q_i} \tilde{v}_i(w_i - p_i^*(q_i), q_i; \hat{a}, \hat{q}_1, \ldots, \hat{q}_N). \]

proving the second part. \( \square \)

\section{C Identities in Section 4.1}

**Derivation of Equation (20)**

For consumer \( N, L_N \sim \text{Exp}(\nu) \) and the loss incurred by “the other” consumers is \( L_{-N} = \sum_{i=1}^{N-1} L_i \sim \text{Gamma}(N-1, \nu) \). Then

\[ e = e_N = \mathbb{E} \left[ q L_N \mathbf{1}_{\{q(L_{-N} + L_N) < a\}} \right] + a \mathbb{E} \left[ \frac{q L_N}{q (L_{-N} + L_N)} \mathbf{1}_{\{q(L_{-N} + L_N) \geq a\}} \right]. \]

\(^8\)Here, \( \frac{\partial \tilde{v}_i}{\partial w} \) and \( \frac{\partial \tilde{v}_i}{\partial q_i} \) denote the derivatives with respect to the first and the second argument of \( \tilde{v}_i \), respectively.
For part ii., note that \( \frac{L_N}{L_{-N} + L_N} \) is Beta(1, \( N - 1 \)) distributed independent of \( L_{-N} + L_N \sim \text{Gamma}(N, \nu) \). Hence, part ii. can be written as

\[
a \mathbb{P} \left( L_{-N} + L_N \geq \frac{a}{q} \right) = a \Gamma_{N, \nu} \left( \frac{a}{q} \right) N^{-1}.
\]

For part ii., we have

\[
q \mathbb{E} \left[ L_N 1_{\{q(L_{-N} + L_N) < a\}} \right] = q \int_0^\infty \int_0^\infty 1_{\{i + i < a/q\}} l \nu e^{-\nu l} \frac{\nu^{N-1}}{(N-2)!} i^{N-2} e^{-i \nu} dl di
\]

\[
= q \frac{\nu^N}{(N-2)!} \int_0^a \int_0^{a-i} i e^{-\nu l} dl i^{N-2} e^{-i \nu} di
\]

\[
= q \frac{\nu^N}{(N-2)!} \int_0^a \int_0^{a-i} i e^{-\nu l} dl i^{N-2} e^{-i \nu} di
\]

\[
= q \frac{\nu^N}{(N-2)!} \int_0^a i^{N-2} e^{-i \nu} di - q \frac{\nu^{N-1}}{(N-2)!} e^{-\nu a/q} \left[ \left( \frac{a}{q} + 1 \right) \int_0^{a/q} i^{N-2} di - \int_0^{a/q} i^{N-1} di \right]
\]

\[
= q \frac{\nu}{\nu} \Gamma_{N-1, \nu} (a/q) - q \frac{\nu^{N-1}}{(N-1)!} e^{-\nu a/q} \left( \frac{a}{q} \right)^{N-1} \left[ \frac{1}{\nu} + \frac{a}{N q} \right] .
\]

Therefore, since all consumers are identical, the objective function (6) takes the form displayed in (20). For condition (7), on the other hand, we have

\[
V = V_N = \mathbb{E} \left[ U \left( w - p - L_N + R_N \right) \right] = \mathbb{E} \left[ U \left( w - p - \left( 1 - q \right) L_N \right) 1_{\{q(L_{-N} + L_N) < a\}} \right] \tag{part i.}
\]

\[
+ \mathbb{E} \left[ U \left( w - p - L_N + a \frac{L_N}{L_{-N} + L_N} \right) 1_{\{q(L_{-N} + L_N) \geq a\}} \right] \tag{part ii.}
\]

For part i, we obtain

\[
\mathbb{E} \left[ U \left( w - p - \left( 1 - q \right) L_N \right) 1_{\{q(L_{-N} + L_N) < a\}} \right] = - \int_0^\infty \int_0^\infty 1_{\{i + i < a/q\}} e^{-\alpha(w - p - \left( 1 - q \right) l)} \nu e^{-\nu l} \frac{\nu^{N-1}}{(N-2)!} i^{N-2} e^{-i \nu} dl di
\]

\[
= e^{-\alpha(w-p)} \int_0^{a/q} \frac{1}{\nu - \alpha(1-q)} \left[ 1 - e^{-\alpha/(1-q) \nu} - i(1-q) \nu + i \right] \frac{\nu^{N-1}}{(N-2)!} i^{N-2} e^{-i \nu} di
\]

\[
= e^{-\alpha(w-p)} \left\{ \frac{1}{\nu - \alpha(1-q)} \Gamma_{N-1, \nu} (a/q) - \frac{e^{-\alpha/(1-q) \nu} - i(1-q) \nu + i}{((1-q) \alpha)^{N-1}} \Gamma_{N-1, (1-q) \alpha} \left( \frac{a}{q} \right) \right\} .
\]
For part \(ii\), note that

\[
\begin{align*}
\mathbb{E} \left[ U \left( w - p - \left( (L_m - N) + L_N \right) - a \right) \frac{L_N}{L_N - L_m} \right] _{q(L_m - N) + L_N \geq a} & = \int_0^1 \int_0^\infty e^{-\alpha(w-aL_m)} I_{\{\geq a\}} \nu_N \frac{L_N}{(N-1)!} \int_0^1 e^{(\alpha - t)(N-1)1 - y} (1 - y)^N - 2 dy dl \ dy \\
& = -e^{-\alpha(w-aL_m)} \int_0^{\infty} \frac{\nu_N}{(N-1)!} \sum_{t=0}^{\infty} (\alpha - t)^k \frac{(N + k - j - 1)!}{(N - 1 + k)!} \Gamma_{N - j + k, \nu} \left( \nu \left( \frac{1 - q}{q} \right) \right) \\
& = -e^{-\alpha(w-aL_m)} \sum_{t=0}^{\infty} \sum_{j=0}^{N-1} (N-1)! \frac{(\nu)_{j+1}}{j!} e^{-\nu a} \left( \frac{\alpha}{\nu} \right) ^k \left( \frac{(N + k - j - 1)!}{(N - 1 + k)!} \Gamma_{N - j + k, \nu} \left( a \left( \frac{1 - q}{q} \right) \right) \\
\end{align*}
\]

**Derivation of Equation (21)**

Similar to the previous part, for consumer \(N\) with \(L = \sum_{i=1}^{N} L_i\):

\[
\begin{align*}
\mathbb{E} \left[ \sum_{j=1}^{N} U' \left( w - p - L_j + a \frac{L_j L_N}{L} \right) \frac{L_j L_N}{L} \right] & = \sum_{j=1}^{N-1} \mathbb{E} \left[ U' \left( w - p - (L - a) \frac{L_j L_N}{L} \right) \frac{L_j L_N}{L} \right] + \mathbb{E} \left[ U' \left( w - p - (L - a) \frac{L_N}{L} \right) \left( \frac{L_N}{L} \right) ^2 \right] . \\
\end{align*}
\]

Note that \(\frac{L_j}{L}, \frac{L_N}{L} \sim \text{Beta}(1, N - 1)\) and for the joint distribution

\[
d_f \frac{L_j}{L}, \frac{L_N}{L} (x, y) = (1 - x - y)^{N-3} (N - 2) (N - 1) I_{\{x, y \geq 0, x + y \leq 1\}}, j \neq N.
\]
Whence, for part i.,

\[
\mathbb{E}\left[ U' \left( w - p - (L - a) \frac{L_{N-1}}{L} \right) \frac{L_{N-1} L_N}{L} \mid L \right] \\
= \alpha e^{-\alpha(w-p)} \int_0^1 \int_0^{1-x} e^{\alpha(L-a)x} x \, y \, (N-1) \, (N-2) \, (1-x-y)^{N-3} \, dy \, dx \\
= \alpha e^{-\alpha(w-p)} \int_0^1 e^{\alpha(L-a)x} \int_0^{1-x} y \, (N-1) \, (N-2) \, (1-x-y)^{N-3} \, dy \, dx \\
= \alpha e^{-\alpha(w-p)} \beta(2, N) \int_0^1 e^{\alpha(L-a)x} \frac{1}{\beta(2, N)} x \, (1-x)^{N-1} \, dx \\
= \alpha e^{-\alpha(w-p)} \frac{1}{N(N+1)} (N+1)! \sum_{k=0}^{\infty} \frac{(k+1)}{(N+k+1)!} (\alpha(L-a))^k,
\]

whereas for part ii.,

\[
\mathbb{E}\left[ U' \left( w - p - (L - a) \frac{L_N}{L} \right) \left( \frac{L_N}{L} \right)^2 \mid L \right] \\
= \alpha e^{-\alpha(w-p)} \mathbb{E}\left[ \exp \left\{ \alpha(L-a) \frac{L_N}{L} \right\} \left( \frac{L_N}{L} \right)^2 \mid L \right] \\
= \alpha e^{-\alpha(w-p)} (N-1)! \sum_{k=0}^{\infty} \frac{(k+1)(k+2)}{(N+k+1)!} (\alpha(L-a))^k,
\]

so that

\[
\mathbb{E}\left[ \sum_{j=1}^{N} u' \left( w - p - L_j + a \frac{L_j}{L} \right) \frac{L_j L_N}{L} \mid L \right] \\
= \alpha e^{-\alpha(w-p)} (N-1)! \sum_{k=0}^{\infty} \frac{(N-1)(k+1)}{(N+k+1)!} (\alpha(L-a))^k + \frac{(k+1)(k+2)}{(N+k+1)!} (\alpha(L-a))^k
\]
For the denominator,

\[ \mathbb{E} \left[ 1_{\{q \geq \alpha\}} \sum_{j=1}^{N} \left( w - p - L_j - a \frac{L_j}{L} \right) \frac{L_j}{L} \right] \]

\[ = \mathbb{E} \left[ 1_{\{q \geq \alpha\}} N \alpha e^{-\alpha(w-p)} (N-1)! \sum_{k=0}^{\infty} \frac{(k+1) \left( \alpha(L-a) \right)^k}{(N+k)!} \right] \]

\[ = N! \alpha e^{-\alpha(w-p)} \sum_{k=0}^{\infty} \frac{k+1}{(N+k)!} \alpha^k \int_{a/q}^{\infty} \left( \frac{\nu^N}{(N-1)!} (l-a)^k l^{N-1} e^{-\nu l} \right) dl \]

\[ = N! \alpha e^{-\alpha(w-p)} \sum_{k=0}^{\infty} \left( \frac{\alpha}{\nu} \right)^k \frac{k+1}{(N+k)!} e^{-\nu a} \sum_{j=0}^{N-1} \frac{(a \nu)^j}{j!} \frac{(N+k-j-1)!}{(N-1)!} \left( \Gamma_{N+k-j,\nu} \left( a \left( \frac{1-q}{q} \right) \right) \right) \]

Hence,

\[ q \phi_i = \frac{1}{N} \mathbb{E} \left[ 1_{\{q \geq \alpha\}} \sum_{k=0}^{\infty} \frac{(k+1) \left( \alpha(L-a) \right)^k}{(N+k)!} \right] \]

\[ = \psi(L) \]

For implementation purposes, the numerator can be expressed as

\[ \sum_{k=0}^{\infty} \frac{(k+1) t^k}{(N+k)!} \bigg|_{t=\alpha(L-a)} = \frac{\partial}{\partial t} \left[ \sum_{k=0}^{\infty} \frac{t^{k+1}}{(N+k)!} \right]_{t=\alpha(L-a)} = \frac{\partial}{\partial t} \left[ t^{-(N-1)} \sum_{k=N}^{\infty} \frac{t^k}{k!} \right]_{t=\alpha(L-a)} \]

\[ = \left. \frac{\partial}{\partial t} \left[ t^{-(N-1)} \left( e^t - \sum_{k=0}^{N-1} \frac{t^k}{k!} \right) \right] \right|_{t=\alpha(L-a)} \]

\[ = -(N-1) t^{-N} \left( e^t - \sum_{k=0}^{N-1} \frac{t^k}{k!} \right) + t^{-(N-1)} \left( e^t - \sum_{k=0}^{N-1} \frac{t^{k-1}}{k!} \right) \bigg|_{t=\alpha(L-a)} \]

\[ = \left( e^t - \sum_{k=0}^{N-2} \frac{t^k}{k!} \right) \times \left( t^{-(N+1)} - t^{-N} (N-1) \right) = \left( \frac{(N-1)}{(N-1)!} \right) t^{-1} \bigg|_{t=\alpha(L-a)} \]
Finally,

\[
\tilde{\psi}(I) = \tilde{\psi}(qL) = \mathbb{E} \left[ \frac{\partial \tilde{P}}{\partial p} \bigg| L \right] \\
= \text{const} \mathbb{E} \left[ \sum_{j=1}^{N} \alpha \exp \left\{ -\alpha \left( w - p - L_j + a \frac{L_j}{L} \right) \right\} \frac{L_j}{L} \right] \\
= \text{const} \cdot N \mathbb{E} \left[ e^{\alpha(L-a) L_j/L} \right] \\
= \text{const} \cdot \frac{\partial}{\partial x} \text{mgf}_{\text{Beta}(1,N-1)}(x) \bigg|_{x=\alpha(L-a)} = \tilde{\psi}(L),
\]

since \( \mathbb{E} \left[ \tilde{\psi}(I) \right] = 1. \)

\section*{D Allocation in a Security Market Equilibrium}

To keep the setup as simple as possible, we limit our considerations to a one-period market with a finite number of securities \( (M) \), each security with potentially distinct payoffs in \( X \) states and assume that the risk-free rate is zero. More specifically, let \( \Omega^{(S)} = \{ \omega^{(S)}_1, \ldots, \omega^{(S)}_X \} \) be the set of these states with associated sigma-algebra \( \mathcal{F}^{(S)} \) given by its power set and let \( p_j^{(S)} = \mathbb{P} \left( \{ \omega^{(S)}_j \} \right) \) denote the associated physical probabilities. Let then \( D \) be the \( M \times X \) matrix with \( D_{ij} \) describing the payoff of the \( i \)th security in state \( \omega^{(S)}_j \), where we assume

\[
\text{span}(D) = \mathbb{R}^X.
\]

This condition allows us to define state prices, consistent with the absence of arbitrage within the securities market, denoted by \( \pi_j, j = 1, \ldots, X \). Thus, any arbitrary menu of securities-market-sub-state-contingent consumption can be purchased at time zero. However, it would be misleading to characterize markets as complete, since \( \Omega^{(S)} \) does not provide a complete description of the “states of the world”. Instead, we characterize the full probability space as \( (\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}) \), with

\[
\bar{\Omega} = \Omega^{(S)} \times \Omega = \left\{ (\omega, \omega') \in \Omega^{(S)} \times \omega \in \Omega \right\}, \\
\bar{\mathcal{F}} = \mathcal{F}^{(S)} \vee \mathcal{F}, \quad \text{and} \\
\mathbb{P} \left( A \right) = \sum_{j \in \mathcal{I}_A} p_j^{(S)} \times \mathbb{P} \left( A_j \big| \omega_j^{(S)} \right)
\]

for \( \bar{A} = \bigcup_{j \in \mathcal{I}_A} \{ \omega_j^{(S)} \} \times A_j \in \bar{\mathcal{F}} \) with \( A_j \subseteq \Omega. \)

Our problem now, however, is that the market is no longer complete so that we need a notion of what insurance liabilities are “worth” when they cannot be hedged completely. We make the assumption that the insurance market is “small” relative to the securities market and, for purposes of valuing insurance liabilities, define the probability measure

\[
\mathbb{Q} \left( A \right) = \sum_{j \in \mathcal{I}_A} \pi_j \times \mathbb{P} \left( A_j \big| \omega_j^{(S)} \right), \quad \bar{A} \subseteq \bar{\Omega},
\]
i.e. it given by the Radon-Nikodym derivative \( \frac{\partial Q}{\partial P}((\omega^{(S)}_j, \omega)) = \frac{\pi'_j}{\rho'_j} \).

Consumer utility now depends on the individual’s chosen security market allocation and may be expressed as
\[
v_i = \mathbb{E}^P\left[ U_i (W_i - p_i - L_i + R_i) \right] \quad \text{with} \quad v'_i = \mathbb{E}^P\left[ U'_i (W_i - p_i - L_i + R_i) \right],
\]
where \( W_i \) is \( \mathcal{F}^{(S)} \)-measurable with \( \sum_j \pi_j W_i \left\{ \omega_j^{(S)} \right\} = w_i \) whereas \( L_i \) — as before — is \( \mathcal{F} \)-measurable. The recovery \( R_i \), on the other hand, now depends both on insurance loss activity as well as portfolio decisions made within the insurance company. To elaborate on this, the budget constraint of the insurance company may be expressed as
\[
a = \sum_j \pi_j K_j a \Rightarrow 1 = \sum_j \pi_j K_j,
\]
where \( K_j a \) reflects consumption purchased in the securities market state \( \omega_j^{(S)} \) or — more precisely — in the states of the world \( \Omega_j = \left\{ \bar{\omega} = (\omega^{(S)}, \omega) \big| \omega^{(S)} = \omega_j^{(S)} \right\} \). We write \( K \) to denote the corresponding \( \mathcal{F}^{(S)} \)-measurable random variable. Consumer \( i \)'s recovery can then be expressed as
\[
R_i = \min \left\{ I_i, \frac{K a}{I} I_i \right\}
\]
and the fair valuation of claims is thus
\[
e_i = \mathbb{E}^Q\left[ R_i \right] = \mathbb{E}^Q\left[ R_i 1_{\{I < K a\}} \right] + \mathbb{E}^Q\left[ R_i 1_{\{I \geq K a\}} \right].
\]
Hence, the firm’s problem becomes
\[
\max_{a, \{q_i\}, \{p_i\}, \{K_j\}} \sum_i p_i - \sum e_i - \tau a,
\]
subject to
\[
\begin{align*}
v_i &\geq \gamma_i, \\
\sum_i \pi_j K_j &\geq 1.
\end{align*}
\]

In addition to a new set of optimality conditions connected with \( \{K_j\} \), we have the same set of first order conditions (as before, we sacrifice technical rigor by assuming a solution in a smooth part of the function):
\[
\begin{align*}
[q_i] &\quad - \sum_k \frac{\partial e_k}{\partial q_i} + \sum_k \lambda_k \frac{\partial v_k}{\partial q_i} - \frac{\partial s}{\partial q_i} \xi = 0, \\
[a] &\quad - \sum_k \frac{\partial e_k}{\partial a} - \tau + \sum_k \lambda_k \frac{\partial v_k}{\partial a} + \xi = 0, \\
[p_i] &\quad 1 - \lambda_i \frac{\partial v_i}{\partial w} = 0.
\end{align*}
\]
As before, we seek a pricing function satisfying:

\[
\left[ \frac{\partial v_i}{\partial q_i} + \mathbb{E}^P \left[ \mathbf{1}_{\{I \geq K a\}} U'_i \frac{K a}{I^2} I_i \frac{\partial I_i}{\partial q_i} \right] \right] - \frac{\partial v_i}{\partial w} \frac{\partial p^*_i}{\partial q_i} = 0.
\]

Proceeding analogously to Section 2, we arrive at the marginal pricing condition associated with a decentralized implementation:

\[
\frac{\partial p^*_i}{\partial q_i} = \sum_k \frac{\partial e_k}{\partial q_i} + \frac{\partial s}{\partial q_i} \xi - \sum_{k \neq i} \frac{\partial v_k}{\partial q_i} + \mathbb{E}^P \left[ \mathbf{1}_{\{I \geq K a\}} U'_i \frac{K a}{I^2} I_i \frac{\partial I_i}{\partial q_i} \right] v'_i.
\]

Simplifying, we obtain

\[
\frac{\partial p^*_i}{\partial q_i} = \frac{\partial e^Z_i}{\partial q_i} + \frac{\partial s}{\partial q_i} \left[ \sum_k \frac{\partial e_k}{\partial a} + \tau - \sum_k \frac{\partial v_k}{\partial a} v'_k \right] + \mathbb{E}^Q \left[ K \mathbf{1}_{\{I \geq K a\}} \right] + \tau - \sum_k \frac{\partial v_k}{\partial a} v'_k + \mathbb{E}^P \left[ \mathbf{1}_{\{I \geq K a\}} U'_i \frac{K a}{I^2} I_k \frac{\partial I_k}{\partial q_i} \right] \sum_k \frac{\partial v_k}{\partial a} v'_k \times a + \tilde{\phi}_i \times a \times \left[ \sum_k \frac{\partial v_k}{\partial a} \right],
\]

where

\[
\tilde{\phi}_i = \frac{\mathbb{E}^P \left[ \mathbf{1}_{\{I \geq K a\}} U'_i \frac{K a}{I^2} I_k \frac{\partial I_k}{\partial q_i} \right]}{\mathbb{E}^P \left[ \mathbf{1}_{\{I \geq K a\}} \sum_k U'_i \frac{K a}{I^2} I_k \right]}.
\]

Hence, we essentially obtain the same result as before—with the only differences worth noting being that 1) the marginal cost of capital, \( \mathbb{E}^Q \left[ K \mathbf{1}_{\{I \geq K a\}} \right] + \tau \), as well as the component of marginal cost deriving from claimant recoveries in solvent company states, \( \frac{\partial e^Z_i}{\partial q_i} \), now reflect state prices derived from the security markets, and 2) the “weights” used in the internal capital allocation formula now include a factor \( K \), reflecting the fact that available company resources may vary across default states due to company asset allocation. Importantly, note that the capital allocation weights are still determined by customer marginal utility, rather than state prices—although the latter may well be connected in some way to the former due to customer asset allocation choices.

### References


