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

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

October 2012

Abstract

Financial institutions define their marginal cost of risk on the basis of arbitrarily chosen risk measures. We reverse this approach by calculating the marginal cost for a profit-maximizing firm, and then identifying the risk measure delivering the correct marginal cost. The resulting measure is a weighted average of three components corresponding the drivers of firm capitalization: (1) An external risk measure reflecting regulatory concerns; (2) Value-at-Risk emerging from shareholder concerns; and (3) a novel risk measure that encapsulates counterparty preferences. Our results demonstrate that risk measures used for pricing and performance measurement should be chosen based on economic fundamentals rather than mathematical properties.

JEL classification: G20, G22

Keywords: capital allocation; risk measure; profit maximization; counterparty risk.

*We thank Enrico Biffis, Keith Crocker, Richard Derrig, Anastasia Kartasheva, John Major, Stephen Mildenhall, Giacomo Nocera, Richard Phillips, Ajay Subramanian and seminar participants at the 2012 ASSA Meetings, the 2011 NBER Insurance Project Workshop, the 2012 CAS Spring Meeting, the ICMFE 2011, the ARC 2012, the ARIA 2012 Meeting, GSU, Imperial College, the University of Iowa, LMU, NCCU, Ulm University, and Pennsylvania State University. 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

Financial institutions typically quantify their exposure with risk measures such as Value-at-Risk or Expected Shortfall. The gradients of these measures are used to allocate the firm’s capital to the various risks within its portfolio—a process which effectively determines the marginal cost of risk and thus provides the key inputs for pricing and performance measurement.\(^1\) This practice, however, starts with an arbitrary choice of risk measure, implying that the end products of capital allocation and marginal cost are also arbitrarily defined and lack an economic foundation.

We reverse the sequence of this approach. We start with an economic model of a financial institution with risk-averse counterparties and costly capital, and we show that profit maximization yields an endogenous expression for the marginal cost that can be used for capital allocation. We then derive the risk measure that gives the correct capital allocations and find that this risk measure generally does not adhere to the mathematical axioms typically imposed when selecting risk measures. Our results suggest that the risk measure employed for capital allocation should be tailored to the specific application in view and that conventional risk measures yield inefficient allocations.

We start our analysis with a simplified one-period model in an environment without securities markets but subsequently generalize the results to the case where both the firm and its consumers have access to securities markets and to multiple periods. In the general case, we identify three sources of “discipline” that feed into the marginal cost of risk faced by the firm (and, consequently, the resulting capital allocation). The first stems from a regulatory solvency constraint, which is a familiar feature of the existing literature: Risks are costly in that they force the firm to hold more capital due to regulation. The second derives from the firm’s counterparties: When the firm adds a risk, all of its counterparties are affected and are thus willing to pay less for the firm’s contracts. The final source of discipline stems from the continuation value of the firm: Risks taken on in the current period may lead to bankruptcy of the firm and thus may destroy future profit flows.

\(^1\)For instance, a McKinsey&Company (2011) survey among a “diverse group of 11 leading banks” revealed that the “vast majority of respondents use economic-capital (EC) models”, mostly for “tracking performance of individual business units or portfolios”, although “more sophisticated applications” such as pricing or risk-based strategic decision making also appear in the sample. Similarly, according to the Society of Actuaries (2008), more than 80% of all insurance companies calculated EC or considered the implementation of an EC framework in 2006, where “allocation of capital” and “measure of risk-adjusted performance” were listed as the two primary drivers.
The optimal capital allocation rule is a weighted average of an “external” allocation rule implied by the regulatory constraint (if it binds), an “internal counterparty” allocation rule driven by the institution’s uninsured counterparties, and a “continuation” rule that derives from the firm’s value as a going concern. In the extreme case of no regulation and perfect competition (so that the firm is earning zero economic rents), the allocation rule simply boils down to the “internal counterparty” rule. Another extreme case is a single period model with fully insured counterparties, where the economically optimal allocation follows from the risk measure imposed by regulation. Intermediate cases, however, could feature marginal cost being driven mainly by the “internal counterparty” and/or the “continuation” allocation 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 then investigate the connection between these economically derived capital allocations and risk measures. Specifically, we “reverse-engineer” risk measures whose gradients yield the economically correct capital allocations. Each of the three components of the optimal allocation rule discussed above is connected to a risk measure: The “external” allocation rule is of course connected to a risk measure by definition, as it arises from a risk measure imposed by the regulator. The more interesting finding is that the “internal counterparty” allocation rule can be implemented by a novel 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. Finally, in a multi-period setting, the allocation stemming from the “continuation” value of the firm can be recovered by applying the gradient method to Value-at-Risk (VaR)—which thus arises endogenously in our model. With the possible exception of the “external” measure, these measures are neither coherent nor convex—properties considered by many as imperative.

We derive closed-form expressions for the novel “internal counterparty” allocation rule in two example setups: (1) homogeneous counterparties facing exponentially distributed losses and (2) heterogeneous counterparties facing Bernoulli distributed losses. We compare the resulting allocations to those obtained from Expected Shortfall (ES)—the coherent risk measure currently in favor among many aca-
demics and regulators. We show that ES-based allocations generally fail to weight default outcomes properly. Specifically, in cases where counterparties are strongly risk-averse or where potential losses are large relative to counterparty wealth, ES-based allocations tend to underweight bad outcomes; in cases where counterparties are only weakly risk-averse or where potential losses are relatively small, ES-based allocations tend to overweight bad outcomes. These differences flow from a fundamental difference in the basis for allocation under the “internal counterparty” rule and under ES. The starting point for evaluation of a risk’s impact under ES concerns its share of the institution’s losses in default states, whereas the starting point under the “internal counterparty” rule is the risk’s share of recoveries—as a risk’s impact on recoveries in default states is ultimately what counterparties care about.

This distinction underscores the key point of the paper: The true marginal cost of risk and the associated allocation of capital should flow from the economic context of the problem. A risk measure chosen for its technical properties such as coherence, rather than for the specific economic circumstances, will generally fail to yield correct pricing and efficient allocation of capital from the perspective of its user.

Relationship to the Literature and Organization of the Paper

Formal analysis of the problem of capital allocation based on the gradient of a risk measure appeared in the banking and insurance literatures around the turn of the millennium and was subsequently generalized (see Schmock and Straumann (1999), Myers and Read (2001), Denault (2001), Tasche (2004), Kalkbrener (2005) or Powers (2007), among others). Broadly speaking, these papers start with a differentiable risk measure and end up allocating capital by computing the marginal capital increase required to maintain the risk measure at a threshold value as a particular risk exposure within the portfolio is expanded, an approach referred to as “gradient” allocation or “Euler” allocation.

The technique thus neatly defines the marginal cost of risk if the risk measure is—in one way or another—embedded in the institution’s optimization problem (Zanjani (2002), Stoughton and Zechner (2007)). Unfortunately, excepting highly specialized circumstances, economic theory thus far offers no

---

2 Various papers make a case for ES over VaR (see e.g. Hull (2007)), and both regulation and practice appear to be moving in this direction. For instance, the International Actuarial Association (2004) recommends using ES in a risk based regulatory framework, and ES was embedded in the Swiss Solvency Test and appears to be viewed favorably by US insurance regulators (cf. NAIC (2009)).

3 If institutional preferences are defined by a particular risk-averse utility function of outcomes, a particular risk measure may be implied (see e.g. Föllmer and Schied (2010)). Alternatively, Adrian and Shin (2008) justify using Value-at-Risk in
guidance on the choice of the measure. Perhaps as a consequence, the debate on risk measure selection has largely centered on mathematical properties of the measures (see e.g. Artzner et al. (1999), Föllmer and Schied (2002), or Frittelli and Gianin (2002)). Yet the choice obviously has profound economic consequences, as it ultimately determines how the institution perceives risk.

Other papers derive the marginal cost of risk and capital allocations from the fundamentals of the institution’s profit maximization problem without the imposition of a risk measure. The ensuing results are transparent if complete and frictionless markets are assumed (Phillips, Cummins, and Allen (1998); Ibragimov, Jaffee, and Walden (2010)), although this setting begs the question of why intermediaries would hold capital in the first place. Others have studied incomplete market settings. In particular, Froot and Stein (1998) and Froot (2007) introduce the frictions suggested by Froot, Scharfstein, and Stein (1993) to motivate capital holding and risk management. Their models generate a marginal cost of risk determined by an institution’s portfolio and effective risk aversion (as implied by a concave payoff function and a convex external financing cost). Institution-specific risk pricing and capital allocation is also found by Zanjani (2010) (although in the context of a social planning problem) where risk management is motivated by counterparty risk aversion.

Our paper builds on the incomplete market approaches. Our theoretical foundation features costly external financing, counterparty risk aversion, and regulatory constraints as motivators of risk management and determinants of marginal cost. We recover some familiar results in certain cases, but the general form of capital allocation is multifaceted. The complexity serves as evidence of the force of Froot and Stein’s criticism that allocating capital via arbitrary risk measures is problematic because it is “not derived from first principles to address the objective of shareholder value maximization.” Froot and Stein, however, did not attempt to reconcile risk measure-based approaches with those based on “first principles.” This leaves a gap between financial theory and practice that we close here by extracting capital allocations from marginal cost calculations and then deriving risk measures consistent with the
extracted allocations. This connection between the marginal cost obtained from the fundamentals of the institution’s problem and that obtained from approaches based on risk measures has, to our knowledge, never been explored.

The paper is organized as follows: Section 2 presents the firm’s profit maximization problem in various settings, and we describe how the marginal cost of risk and the allocation of capital within the firm arise as by-products; Section 3 describes the relationship of the resulting allocation rule to risk measures; Section 4 presents our example applications; and finally Section 5 concludes.

2 Profit Maximization and Capital Allocation

We consider the optimization problem of a representative financial institution. We frame our model in terms of an insurance company, and our language reflects this in that we refer to the financial contracts as “insurance coverage” and the counterparties of the institution as “consumers.” The setup obviously fits other institutions providing similar contracts, such as reinsurance companies and private pension plan sponsors—and can be applied with little modification to institutions selling insurance-like contracts (such as credit default swaps) where the main risks emanate from risk in obligations to counterparties. The model can also be adapted to fit other institutions where capital allocation is relevant (such as commercial banks) but where the key risks emanate from the asset side of the balance sheet, by including an additional set of choice variables for investments. The key assumption of the model, however, is that the stakeholders of the institution are exposed to the failure of the institution—and their preferences for solvency drive the motivation for risk management.

To illustrate the main ideas, we will start by considering a greatly simplified one-period model in an environment without securities markets. Subsequently, we generalize the results to the case where both the firm and its consumers have access to securities markets and to multiple periods.

2.1 One-period Model without Securities Markets

Formally, we consider an insurance company that has \( N \) consumers, with consumer \( i \) facing a loss \( L_i \) modeled as a continuous, non-negative, square-integrable random variable on the complete probability
space \((\Omega, \mathcal{F}, \mathbb{P})\) with (joint) density \(f_{L_1, L_2, \ldots, L_N} : \mathbb{R}^N_+ \to \mathbb{R}_+\). 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. For tractability, we focus on a quota share arrangement, i.e. a linear contract, where the insurer agrees to reimburse \(q_i\) per dollar of loss:

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

(1)

However, generalizations are possible.\(^6\)

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:

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

(2)

Accordingly, \(\{I \geq a\} = \{\omega \in \Omega | I(\omega) \geq a\}\) 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:

\[
e_i = \mathbb{E}[R_i] = \mathbb{E}[R_i 1_{\{I < a\}}] + \mathbb{E}[R_i 1_{\{I \geq a\}}] = e_i^P + e_i^D.
\]

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:

\[
\tau \times a,
\]

(3)

\(^6\)For instance, a fixed policy limit as in \(I_i = \min\{L_i, q_i\}\) in conjunction with binary loss distributions also fits our framework, although the lack of differentiability would formally require a separate treatment.
although it is also possible to represent frictional costs as a tax on equity capital, as in:

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

(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:

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

(5)

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

The firm then solves:

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

(6)

subject to participation constraints for each consumer:

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

(7)

and subject to a differentiable solvency constraint imposed by the regulator:

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

(8)

where \( s \) is imagined to arise from, for example, an externally supplied risk measure with a set threshold dictating the requisite capitalization for the firm. As is customary for risk measures (see e.g. the well-known coherence axioms by Artzner et al. (1999)), we assume that \( s \) is positive homogeneous of degree one.

We show in Appendix A that a profit-maximizing firm can implement 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 pricing function. With this assumption in place, the marginal price change at the optimal level of \( q_i \) must satisfy:

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

with \( U'_i = U'_i (w_i - p_i - L_i + R_i) \). The term in brackets represents how the consumer perceives the marginal benefit of additional coverage, which, due to the aforementioned assumption, differs from the true impact of coverage on the utility function by \( E \left[ 1_{\{I \geq a\}} \frac{U'_i}{I^2} I \frac{\partial I_i}{\partial q_i} \right] \).

Appendix B.1 shows that (9) may be rewritten as:

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

where

\[
\tilde{\phi}_i = \frac{\mathbb{E} \left[ 1_{\{I \geq a\}} \sum_k \frac{U'_k}{v'_k} \frac{1}{I^2} I_k \frac{\partial I_i}{\partial q_i} \right]}{\mathbb{E} \left[ 1_{\{I \geq a\}} \sum_k \frac{U'_k}{v'_k} I_k \right]}
\]

The last two terms of (10) imply an allocation of the marginal unit of capital to consumer that “adds up.” More specifically, it can be verified that:

\[
a \times \sum \tilde{\phi}_i q_i = a,
\]

whereas the regulatory constraint “adds up” by the homogeneity assumption:

\[
\sum \frac{\partial s}{\partial q_i} q_i = a.
\]

Thus, the optimal marginal pricing condition (10) can be extended to fully allocate all of the firm’s costs, including the cost of capital:

\[
\sum \frac{\partial p^*_i}{\partial q_i} q_i = \sum \frac{\partial e^Z_i}{\partial q_i} q_i + \left[ P(I \geq a) a + \tau a \right].
\]
Note that the cost of capital as captured in the bracketed term breaks down as:

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

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

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

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

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

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:

\[
\sum_k \frac{\partial v_k}{v_k'} = 0,
\]

so that (10) becomes:

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

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 Denault (2001), Tasche (2004) 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, Constraint (8) is immaterial, so (cf. Eq. (47) in Appendix B.1):

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

meaning that (10) becomes:

\[ \frac{\partial p^*_i}{\partial q_i} = \frac{\partial e^Z_i}{\partial q_i} + \tilde{\phi}_i \times a \times [P(I \geq a) + \tau]. \] (15)

Thus, the marginal cost of risk and the attendant allocation of capital is driven completely by “internal” considerations. Specifically, (11) indicates that the allocation is driven by the time-zero value that consumers place on their anticipated 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, (10) remains in its original form:

\[ \frac{\partial p^*_i}{\partial q_i} = \frac{\partial e^Z_i}{\partial q_i} + \frac{\partial s}{\partial q_i} \left[ P(I \geq a) + \tau - \sum_k \frac{\partial v_k}{\partial a} v'_k \right] + \tilde{\phi}_i \times a \times \left[ \sum_k \frac{\partial v_k}{\partial a} v'_k \right], \] (16)

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

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

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. The extent of this distortion is the key to whether internal counterparty concerns or external regulatory concerns guide
capital allocation. If regulation comes close to replicating the private market outcome:

\[ \mathbb{P}(I \geq a) + \tau \approx \sum_{k} \frac{\partial v_k}{\partial a}, \]

then the second term in (10) will be unimportant relative to the third term, and internal counterparty concerns will dominate. On the other hand, if regulation pushes institutional capitalization well beyond the level that would prevail in the private market:

\[ \mathbb{P}(I \geq a) + \tau \gg \sum_{k} \frac{\partial v_k}{\partial a}, \]

then the second term in (10) overshadows the third, and external regulatory concerns dominate.

### 2.2 Allocation in a Security Market Equilibrium

To keep the setup simple, 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 \( \Omega^{(S)} = \{\omega_1^{(S)}, \ldots, \omega_X^{(S)}\} \) be the set of these states with associated sigma-algebra \( \mathcal{F}^{(S)} \) given by its power set and let \( p_j^{(S)} = \mathbb{P}\left(\{\omega_j^{(S)}\}\right) \) 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 \( \omega_j^{(S)} \), where we assume:

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

This condition allows us to define unique 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:

\[ \bar{\Omega} = \Omega^{(S)} \times \Omega = \{\bar{\omega} = (\omega^{(S)}, \omega) | \omega^{(S)} \in \Omega^{(S)}, \ \omega \in \Omega\}, \]
The marginal cost of risk, risk measures, and capital allocation

\[ \mathcal{F} = \mathcal{F}(S) \lor \mathcal{F}, \text{ and} \]

\[ \Pi \left( \bar{A} \right) = \sum_{j \in \Upsilon_A} p_j^{(S)} \times \Pi \left( A_j \left| \{ \omega_j^{(S)} \} \right. \right) \]

for \( \bar{A} = \bigcup_{j \in \Upsilon_A} \{ \omega_j^{(S)} \} \times A_j \in \bar{F} \) with \( A_j \in \mathcal{F}, j = 1, 2, \ldots, |\Upsilon_A| \).

Our problem now, however, is that the market is no longer complete so that we need a notion of what insurance liabilities are “worth” to the insurer 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, employ the so-called minimal martingale measure:\(^7\)

\[ Q \left( \bar{A} \right) = \sum_{j \in \Upsilon_A} \pi_j \times \Pi \left( A_j \left| \{ \omega_j^{(S)} \} \right. \right), \bar{A} \subseteq \Omega, \]

i.e. \( Q \) is defined by the Radon-Nikodym derivative \( \frac{\partial Q}{\partial P}((\omega_j^{(S)}, \omega)) = \frac{\pi_j}{p_j^{(S)}} \).

Consumer utility now depends on the individual’s chosen security market allocation:

\[ v_i = \mathbb{E}^P \left[ U_i \left( W_i - p_i - L_i + R_i \right) \right] \text{ with } v'_i = \mathbb{E}^P \left[ U'_i \left( W_i - p_i - L_i + R_i \right) \right], \]

where \( W_i \) is \( \mathcal{F}(S) \)-measurable with \( w_{ij} = W_i(\omega_j^{(S)}) \) and \( \sum_j \pi_j w_{ij} = w_i \), whereas \( L_i \) is \( \mathcal{F} \)-measurable. The recovery \( R_i \) 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 \( \bar{\Omega}_j = \{ \bar{\omega} = (\omega^{(S)}, \omega) \mid \omega^{(S)} = \omega_j^{(S)} \} \). We write \( K \) to denote the corresponding\(^7\)

As indicated by Björk and Slinko (2006), the minimal martingale measure “provides us with a canonical benchmark for pricing.” It emerges as the optimal martingale measure given various criteria proposed in the mathematical finance literature if the market for insurance risk is “small” relative to financial markets, i.e. if these risks do not affect the payoff of financial securities (see e.g. Goll and Rüschendorf (2001) or Henderson et al. (2005)), and it has also appears in other settings throughout the finance literature. For instance, the minimal martingale measure coincides with the “hedge-neutral measure” in Basak and Chabakauri (2006), and it arises as the limit of Cochrane and Saá-Requejo price bounds for Itô price processes as shown by Černý (2003). We refer to Föllmer and Schweizer (2010) for a formal definition.
$\mathcal{F}^{(S)}$-measurable random variable. Consumer $i$’s recovery can then be expressed as:

$$R_i = \min \left\{ I_i, \frac{Ka}{I_i} \right\},$$

and the fair valuation of claims is thus:

$$e_i = \mathbb{E}^Q \left[ R_i \right] = \mathbb{E}^Q \left[ R_i \mathbf{1}_{\{I < Ka\}} \right] + \mathbb{E}^Q \left[ R_i \mathbf{1}_{\{I \geq Ka\}} \right].$$

As before, we can now derive the capital allocation according to the company’s marginal cost by working through its optimization problem (see Appendix B.2 for details). We obtain an allocation result similar to that of the previous section: The cost of capital:

$$\mathbb{E}^Q \left[ Ka \mathbf{1}_{\{I \geq Ka\}} + \tau a \right],$$

which now reflects state prices and the company’s asset allocation, can be decomposed according to the marginal costs for each of the individual exposures as:

$$\mathbb{E}^Q \left[ Ka \mathbf{1}_{\{I \geq Ka\}} + \tau a \right] = \sum_i \frac{\partial s}{\partial q_i} q_i \mathbb{E}^Q \left[ K \mathbf{1}_{\{I \geq Ka\}} \right] + \tau - \sum_k \frac{\partial v_k}{\partial a} v'_k \sum_j \pi_j K_j \mathbb{E}^P \left[ \mathbf{1}_{\{I \geq K_j a\}} \right] \mathbb{E}^P \left[ \omega_j \right],$$

(17)

where:

$$\tilde{\phi}_i = \frac{\mathbb{E}^Q \left[ \mathbf{1}_{\{I \geq Ka\}} \sum_k \frac{U'_k}{\mathbb{E}^P[U'_k | \omega_j]} K \frac{I_k}{I} I_k \frac{\partial L_i}{\partial q_i} \right]}{\mathbb{E}^Q \left[ \mathbf{1}_{\{I \geq Ka\}} \sum_k \frac{U'_k}{\mathbb{E}^P[U'_k | \omega_j]} K \frac{I_k}{I} \right]} = \sum_j \pi_j K_j \mathbb{E}^P \left[ \mathbf{1}_{\{I \geq K_j a\}} \sum_k \frac{U'_k}{\mathbb{E}^P[U'_k | \omega_j]} I_k \frac{\partial L_i}{\partial q_i} \right] \frac{\omega'_j}{\omega_j},$$

(18)

Thus, we essentially have the same result as before, although the “internal” allocation rule now only applies in every “branch” of the security market where the incompleteness becomes material. In particular, after adjusting for state prices by conditioning on each “branch,” capital allocation weights are still determined by consumer marginal utility.

In the limiting case of a complete market (i.e. the case when $L_i$ and $R_i$ are $\mathcal{F}^{(S)}$-measurable so that
we can write \( l_{ij} = L_i(\omega_j^{(S)}) \) and \( r_{ij} = R_i(\omega_j^{(S)}) \), \( l_{ij}, r_{ij} \in \mathbb{R} \) we obtain:

\[
\tilde{\phi}_i = \frac{\mathbb{E}^Q \left[ 1_{\{I \geq Ka\}} \sum_k K I_k \frac{\partial I_k}{\partial q_i} \right]}{\mathbb{E}^Q \left[ K 1_{\{I \geq Ka\}} \right]} = \frac{\mathbb{E}^Q \left[ K \frac{\partial I}{\partial q_i} \right]}{\mathbb{E}^Q \left[ K 1_{\{I \geq Ka\}} \right]}, \tag{19}
\]

so that:

\[
q_i \times \tilde{\phi}_i \times \mathbb{E}^Q \left[ Ka 1_{\{I \geq Ka\}} \right] = \mathbb{E}^Q \left[ K a \frac{\partial I}{\partial q_i} 1_{\{I \geq Ka\}} \right]
\]

is the fair price of the recovery. This result—where capital is allocated to consumers in proportion to their share of the total market value of recoveries—is the same allocation result as in Ibragimov, Jaffee, and Walden (2010). It is important to note, however, that in the complete market case, purchasing protection from an insurance company with costly capital is inefficient since consumers can hedge insurance risk themselves. \(^8\)

### 2.3 A Multi-Period Version of the Model

In this section, we consider a generalization of the (one-period) setup to multiple periods. Let \( L^t_i \) denote the loss incurred by consumer \( i, i \in \{1, 2, \ldots, N\} \), in period \( t, t \in \{1, 2, \ldots\} \). We assume that \( L^t_i, t > 0 \)—for fixed \( i \)—are independent and identically distributed, and we define the relevant filtration \( \mathcal{F} = (\mathcal{F}_t)_{t \geq 0} \) via \( \mathcal{F}_t = \sigma(L^s_i, i \in \{1, 2, \ldots, N\}, s \leq t) \). The firm determines the optimal level of assets, \( a_t \), in the beginning of each period (i.e. \( a_t \) is \( \mathcal{F} \)-predictable) for a period cost of \( \tau \times a_t \).

Similarly to before, the company chooses \( \mathcal{F} \)-predictable amounts \( q^t_i \) in \( I^t_i = I(L^t_i, q^t_i) = q^t_i \times L^t_i \) and prices \( p^t_i \) at the beginning of the period, and we denote the total claims by \( I^t = \sum_{j=1}^N I^t_j \).

Now the company defaults if \( I^t > a_t \), so that the recovery paid to each consumer is \( R^t_i = \min \{ I^t_i, \frac{a_t}{\tau} I^t_i \} \) and the company shuts down in case of default, i.e. shareholders do not have access

\(^8\)Ibragimov, Jaffee, and Walden (2010) deal with this by assuming “the insurees do not have direct access to the market for risk” whereas the insurer faces a “friction-free complete market for risk.”
to future profits. The consumer’s utility in period $t$ is given by:

$$v_t^i(a_t, w_t^i - p_t^i, q_1^t, \ldots, q_N^t) = \mathbb{E}_{t-1} \left[ U_i(w_t^i - p_t^i - L_t^i + R_t^i) \right],$$

where for simplicity we assume that wealth is homogeneous across periods, i.e. $w_t^i \equiv w_i$.

The company solves:

$$\max_{\{a_t\}, \{q_t^i\}, \{p_t^i\}} V_0 = \sum_{t=1}^{\infty} \mathbb{E} \left[ \mathbb{I}_{\{I_t^j \leq a_1, \ldots, I_t^{t-1} \leq a_{t-1}\}} \times \left( \sum_i p_t^i - \sum_i \mathbb{E}_{t-1}[R_t^i] - \tau a_t \right) \right]$$

(20)

with constraints:

$$v_t^i(a_t, w_t^i - p_t^i, q_1^t, \ldots, q_N^t) \geq \gamma_i \quad \forall i, \forall t,$$

(21)

$$s(q_1^t, \ldots, q_N^t) \leq a_t \quad \forall t.$$  

(22)

Under the assumptions above, it is clear that there exists an optimal stationary policy:

$$(a_t, \{q_t^i\}, \{p_t^i\}) \equiv (a^*, \{q_t^*\}, \{p_t^*\})$$

that solves the Bellman equation:

$$V = \max_{a_t, \{q_t^i\}, \{p_t^i\}} \sum_i p_t^i - \sum_i \mathbb{E}[R_t^i] - \tau a + \mathbb{P}[I_t^j \leq a] \times V$$

(23)

under conditions (21) and (22). Hence, we have a similar program as in the basic setup from Section 2.1 before, where the primary difference is the last term in (23) involving the value function.

Proceeding analogously to before (see Appendix B.3 for details), we obtain the following marginal

---

9 Alternatively, it is possible to allow the distressed company to raise additional funds in the case of default at a higher (or even increasing) cost akin to Froot, Scharfstein, and Stein (1993) and Froot and Stein (1998). Here, we limit our considerations to this simple case and leave the further exploration of alternative settings for future research.

10 Formally, the consumers will form utilities over consumption in multiple periods. In particular, future (random) losses will also be material. Thus, here $U$ should rather be interpreted as a value function (of end-of-period wealth) than as a utility function (of end-of-period consumption).
Thus, akin to the previous sections, again (24) implies an allocation of capital that “adds up” to the cost of capital:

\[
\begin{align*}
\frac{\partial p^*_i}{\partial q_i} &= E\left[ \frac{\partial I_i(L_i^i,q_i)}{\partial q_i} 1\{I^t \leq a\} \right] + V f_I(a) E\left[ \frac{\partial I_i(L_i^i,q_i)}{\partial q_i} | I^t = a \right] \\
+ &\left[ P(I^t > a) + \tau - \sum_k k \frac{\partial a_k}{v_k} - V f_I(a) \right] \frac{\partial s}{\partial q_i} \\
+ &\left[ \sum_k k \frac{\partial a_k}{v_k} \right] \times a \times \left[ E\left[ 1\{I^t > a\} \sum_k k \frac{1}{v_k} I^t \right] \frac{\partial a}{\partial q_i} \right] .
\end{align*}
\]

(24)

Thus, akin to the previous sections, again (24) implies an allocation of capital that “adds up” to the cost of capital:

\[
\begin{align*}
[P(I^t \geq a) a + \tau a] &= \sum_i \tilde{\theta}_i q_i \times [V f_I(a)] + \sum_i \frac{\partial s}{\partial q_i} q_i \times \left[ P(I^t \geq a) + \tau - \sum_k k \frac{\partial a_k}{v_k} - V f_I(a) \right] \\
+ &\sum_i \tilde{\phi} q_i a \times \left[ \sum_k k \frac{\partial a_k}{v_k} \right] .
\end{align*}
\]

In addition to the “external” \( \frac{\partial s}{\partial q_i} \) and “internal counter-party” \( \tilde{\phi} \) allocation rules from before, the allocation now features a third term—\( \tilde{\theta}_i \)—that is associated with the firm’s value as a going concern. In order to obtain insights on the provenance of the corresponding weights, it again is helpful to consider a few specific situations.

**The Limiting Case of Perfect Competition**

In the limiting case of perfect (Bertrand) competition, the firm value \( V \) approaches zero, and, thus, so does the weight associated with the “going concern” allocation \( \tilde{\theta}_i \). Hence, the limiting allocation is

\[
\frac{\partial p^*_i}{\partial q_i} = \frac{\partial e^*_i}{\partial q_i} + \frac{\partial s}{\partial q_i} \times \left[ P(I^t \geq a) + \tau - \sum_k k \frac{\partial a_k}{v_k} \right] + \tilde{\phi} q_i a \times \left[ \sum_k k \frac{\partial a_k}{v_k} \right] .
\]

(25)

i.e. we obtain the same allocation as in the single-period model (10). As before, we could now further break down this setting by distinguishing insured and uninsured consumers and binding and non-
binding regulation to obtain allocations that are fully determined by “external” and “internal counterparty” considerations, respectively. In particular, in this case the remaining two weights in (25) adhere to the same interpretation as in the single period setting.

**Monopolistic Competition, Full Coverage by Deposit Insurance, and Non-Binding Regulation**

In this case, again consumers are indifferent to the capitalization of the firm and there is no external solvency constraint, so that the level—and the allocation—of firm capital is solely determined by the firm’s value as a going concern:

$$\frac{\partial p^*_i}{\partial q_i} = \mathbb{E} \left[ \frac{\partial I_i(L^i_i, q_i)}{\partial q_i} 1_{(I^t \leq a)} \right] + V f_I(a) \mathbb{E} \left[ \frac{\partial I_i(L^i_i, q_i)}{\partial q_i} \right]_{I^t = a}$$

where the latter equality follows from the first order condition for $a$ in the absence of constraints (see Eq. (52) in Appendix B.3).

**General Case: Monopolistic Competition, Uninsured Consumers, and Binding Regulation**

Here we obtain (24) and can now identify the three influences on capital allocation. Two are exactly the same as before—with their relative importance determined by how close the regulatory requirement is to the capitalization level chosen by the consumers in an unregulated market. The relative importance of the “new” term $\tilde{\theta}_i$ that derives from the firm’s value as a going concern depends on how different the capitalization would be if regulatory and consumer concerns were immaterial. In particular, if consumer and shareholder considerations in a private (unregulated) market would yield a similar capitalization as imposed by regulation:

$$\mathbb{P}(I^t \geq a) + \tau \approx \sum_k \frac{\partial v_k}{v_k} + V f_I(a),$$

then internal counterparty and continuation value concerns will dominate whereas regulatory concerns will be the key driver otherwise.

Inspecting the form of the regulatory and the shareholder-driven allocation, they are reminiscent of
conventional allocation methods based on the gradients of risk measures. The next section elaborates on these relationships in more detail.

3 Capital Allocation and Risk Measures

The “black board solution” of Section 2 offers little obvious consolation to a practitioner faced with the problem of allocating capital in a real-world setting, as it may seem difficult to implement. Indeed, the popularity of allocation methods based on risk measures, particularly the so-called Euler or Gradient allocation, can be attributed in part to its ease of application. However, there has been little scientific analysis of the question of how to choose risk measures for the purposes of capital allocation. Conventional thinking points to the various mathematical properties of risk measures (e.g., coherence), yet it is not clear whether choices made on such a basis yield economically desirable outcomes (see Gründl and Schmeiser (2007)).

This section aligns our results with those obtained from the Euler method. We first formally introduce the Euler allocation principle. We then derive the risk measures which—if the Euler method were applied to them—would yield our allocation results.

3.1 The Euler Allocation Principle

Although it is not always derived in this way, the Euler method is implied by the maximization of profits subject to a risk measure constraint. To illustrate, assume a company’s profit function $\Pi$ depends on the parameters (volumes) $q_i, 1 \leq i \leq N$, and on capital $a$. Then maximizing profits subject to the risk measure constraint:

$$\rho(q_1, q_2, \ldots, q_N) \chi_\rho \leq a$$

yields:

$$\frac{\partial \Pi}{\partial q_i} = \left( -\frac{\partial \Pi}{\partial a} \right) \chi_\rho \frac{\partial \rho}{\partial q_i}$$

at the optimal values $a^*, q_i^*, 1 \leq i \leq N$. Here $\rho$ is a differentiable risk measure evaluated at the aggregate claims $\sum_{j=1}^N I_j(L_j, q_j)$ and $\chi_\rho$ is an exchange rate that converts risk to capital (which is set at one if
risk is measured in monetary units—i.e. in the case of a monetary risk measure). Hence, for the optimal portfolio, the risk-adjusted marginal return on marginal capital \( \frac{\partial \Pi / \partial q_i}{\chi \rho \times \partial \rho / \partial q_i} \) for each exposure is the same and equals the cost of a marginal unit of capital \( -\frac{\partial \Pi}{\partial a} \). In other words, if the marginal performance of risk \( i \) as measured by its marginal return on marginal risk capital \( \chi \rho \times \frac{\partial \rho}{\partial q_i} \) exceeds (respectively, falls below) the cost of a marginal unit of capital, then increasing (respectively, decreasing) the weight \( q_i \) of that exposure by a small amount improves the overall performance of the portfolio.\(^{11}\) Akin to the results of Tasche (2004) on the suitability of capital allocation principles, this motivates the interpretation of the marginal capital weighted by the corresponding volume \( \chi \rho \times \frac{\partial \rho}{\partial q_i} q_i^* \) as the amount of capital allocated to exposure \( i \). For homogenous risks and risk measure, the allocations to the respective risks “add up” to the entire capital:

\[
\sum_{j=1}^{N} \chi \rho \frac{\partial \rho}{\partial q_j} q_j^* = a^*.
\]

We refer to McNeil, Frey, and Embrechts (2005) and references therein for more details on the Euler principle, and to Denault (2001), Kalkbrener (2005), and Myers and Read (2001) for alternative derivations of the Euler principle based on cooperative game theory, formal axioms, or a contingent claim approach, respectively. Regardless of its provenance, the Euler method’s strength is the feasibility of implementation, as it is only necessary to calculate the partial derivatives of a given risk measure with respect to each exposure evaluated at the current portfolio.

### 3.2 Allocation Based on Risk Measures

We now proceed to derive the risk measures that correspond to the capital allocations derived in Section 2. Each of the three components of capital allocation are associated with a unique risk measure.

The first component of capital allocation—that associated with the external risk measure—maps obviously as the output of the Euler method applied to that external risk measure. To illustrate, when consumers are fully insured by deposit insurance (and perfect competition prevails in the multi-period context), the only relevant constraint on risk-taking derives from the given risk measure \( s \) — so that

capital is fully allocated based on the gradient of \( s \):

\[
\sum_{j=1}^{N} \frac{\partial s}{\partial q_j} q_j = a.
\] (27)

This allocation is also obtained from solving the profit maximization problem underlying the Euler method:

\[
\begin{cases}
\max_{a,\{q_i\}} \Pi(q_1, \ldots, q_N, a) = \sum_k p^*_k(q_k) - \sum_k e_k(q_1, \ldots, q_N, a) - \tau a, \\
 s(q_1, \ldots, q_N) \leq a.
\end{cases}
\] (28)

This optimization problem is not equivalent to the firm’s formal optimization problem (6)-(8), since the optimal premium functions \( p^*_i \) (which are embedded in (28)) require a solution of the original problem. Nevertheless, the new problem (28) yields the same partial solution as the original problem and the same marginal costs at the optimum when using the optimal premium functions as an input.

The allocation component relating to continuation value also maps easily to a familiar risk measure. The capital allocation of the multi-period version of our model in (24)—in contrast to the single period model—features a term relating to the “continuation” value of the firm:

\[
\tilde{\theta}_i = \mathbb{E} \left[ \frac{\partial I_i(L_i, q_i)}{\partial q_i} \mid I = a \right], \ i = 1, \ldots, N.
\]

Now \( \tilde{\theta}_i \) is exactly the derivative of Value-at-Risk (VaR) with confidence level (cf. Property 1 on Page 229 in Gourieroux, Laurent, and Scaillet (2000)):

\[
\tilde{\alpha} = \mathbb{P}(I \leq a).
\] (29)

Thus, in a regime with full deposit insurance and non-binding regulatory requirements, Euler allocation based on \( \rho = \text{VaR}_{\tilde{\alpha}} \):

\[
\tilde{\theta}_i = \frac{\partial}{\partial q_i} \text{VaR}_{\tilde{\alpha}}(I)
\]

yields allocations exactly consistent with (26) from Section 2.3. Although this only holds in specific circumstances, it is more than a curiosity since 1) deposit insurance is prevalent in most developed
banking and (primary) insurance markets and 2) regulatory capital requirements—though common in intermediary markets—frequently do not bind, i.e. solvency ratios exceed the required level (see e.g. Hanif et al. (2010)).

To motivate the measure corresponding to the “internal counterparty” component of capital allocation, consider now a company in a regime without (full) coverage by deposit insurance (but still assume perfect competition in the multi-period context). Then, capitalization will become material to consumers—which mathematically corresponds to the \( \tilde{\phi}_i \)'s emerging in the marginal cost allocations (10)/(24), where:

\[
\tilde{\phi}_i = \frac{\mathbb{E} \left[ \mathbf{1}_{\{I \geq a\}} \sum_k \frac{U_k^i}{v_k^i} \frac{I_k}{I} \frac{\partial I_i}{\partial q_i} \right]}{\mathbb{E} \left[ \mathbf{1}_{\{I \geq a\}} \sum_k \frac{U_k^i}{v_k^i} I_k \right]}.
\]

To align this allocation with the Euler method, introduce the probability measure \( \tilde{\mathbb{P}} \) on \((\Omega, \mathcal{F})\) via its Radon-Nikodym derivative:

\[
\frac{\partial \tilde{\mathbb{P}}}{\partial \mathbb{P}} = \frac{\mathbb{E} \left[ \sum_k \frac{U_k^i}{v_k^i} I_k \mathbf{1}_{\{I \geq a^*\}} \right]}{\mathbb{E} \left[ \sum_k \frac{U_k^i}{v_k^i} I_k \mathbf{1}_{\{I \geq a^*\}} \right]},
\]

(30)

where \( I, U, \) etc. are evaluated at the concurrent (optimal) values \( a^*, p_i^*(q_i^*), q_i^* \), \( 1 \leq i \leq N \). Note that \( \tilde{\mathbb{P}} \) is absolutely continuous with respect to the original probability measure \( \mathbb{P} \) but the measures are not equivalent since under \( \tilde{\mathbb{P}} \) all the probability mass is concentrated in default states. On the set of strictly positive \( \tilde{\mathbb{P}} \)-square integrable random variables:

\[
L^2_+ (\Omega, \mathcal{F}, \tilde{\mathbb{P}}) = \left\{ X \in L^2 (\Omega, \mathcal{F}, \tilde{\mathbb{P}}) \mid X > 0 \tilde{\mathbb{P}} \text{-a.s.} \right\},
\]

we define the risk measure:

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

(31)

Obviously \( \tilde{\rho} \) is monotonic (\( \tilde{\rho}(X) \leq \tilde{\rho}(Y) \) for \( X \leq Y \) a.s.), positive homogenous (\( \tilde{\rho}(aX) = a \tilde{\rho}(X) \), \( a > 0 \)), and it satisfies the constancy condition \( \tilde{\rho}(c) = c \) for \( c > 0 \) (see Frittelli and Gianin (2002) for a discussion of properties of risk measures). However, \( \tilde{\rho} \) is neither translation-invariant nor sub-additive.

\[\text{While its functional form seems similar to that of the so-called entropic risk measure, which has recently gained popularity in the mathematical finance literature (see e.g. Föllmer and Schied (2002) or Detlefsen and Scandolo (2005)), note that the roles of the exponential function and the logarithm are interchanged.}\]
and is therefore not coherent and not convex. In fact, \( \hat{\rho} \) is not even a monetary risk measure and thus may seem ill-suited for use as an external risk measure.\(^{13}\) Nevertheless, in our setting, it is the correct risk measure for internal capital allocation based on the Euler principle.

More precisely, define \( \tilde{\chi}_\rho = \frac{a^*}{\hat{\rho}(\sum_{j=1}^N I_j(L_j, q_j^*))} \) as the “exchange rate” between units of risk and capital. Then the application of the Euler principle with profit function \( \Pi(q_1, \ldots, q_N, a) \) as in (28) and risk measure constraint:

\[
\hat{\rho} \left( \sum_{j=1}^N I_j(L_j, q_j) \right) \tilde{\chi}_\rho \leq a,
\]

yields:

\[
\sum_{j=1}^N \tilde{\chi}_\rho \frac{\partial \hat{\rho}}{\partial q_j} q_j^* \left[ \mathbb{P}(I \geq a^*) + \tau \right] = \sum_{j=1}^N \tilde{\chi}_\rho \mathbb{E}^{\hat{\rho}} \left[ \frac{\partial I_j/\partial q_j}{I} \right] \hat{\rho} q_j^* \left[ \mathbb{P}(I \geq a^*) + \tau \right] \tag{33}
\]

\[
= \sum_{j=1}^N \phi_j q_j^* a^* \left[ \mathbb{P}(I \geq a^*) + \tau \right] = a^* \left[ \mathbb{P}(I \geq a^*) + \tau \right],
\]

which is exactly the allocation (15) derived in the previous section.\(^{14}\) While it may appear problematic to use the optimal levels \( a^* \) and \( q_i^* \) in the definition of \( \tilde{\chi}_\rho \) and the Radon-Nikodym derivative \( \frac{\partial \hat{\rho}}{\partial \mathbb{P}} \), note that this is akin to employing the optimal premium functions \( p^* \) for deriving the Euler allocation based on the externally supplied risk measure within the first component described above. Just like there, optimizing \( \Pi \) subject to the risk measure constraint (32) is a simple auxiliary problem that yields the same partial solution and the same capital allocation as the firm’s full optimization problem. Moreover, by inserting the definition of \( \tilde{\chi}_\rho \), we obtain a more familiar representation of the amount of capital allocated to consumer \( i \) (cf. Schmock and Straumann (1999)):

\[
a^* \times q_i^* \times \frac{\frac{\partial \hat{\rho}}{\partial q_i}(q_1^*, \ldots, q_N^*)}{\hat{\rho}(q_1^*, \ldots, q_N^*)}.
\]

Hence, the correct internal capital allocation according to marginal cost can be implemented by the

\(^{13}\)While sub-additivity is subject of an ongoing debate (see e.g. Dhaene et al. (2008) or Kou, Peng, and Heyde (2012)), 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, Frey, and Embrechts (2005)).

\(^{14}\)In a setting with securities markets (Eq. (18)), the risk measure still takes the same form (31) but state prices and allocations to security market states enter the measure transform (30).
Euler principle relying on the risk measure $\tilde{\rho}$.

General institutional circumstances—i.e., binding regulation, imperfect competition, and incomplete deposit insurance—will lead to a risk measure that combines the three foregoing risk measures. In the general case from Section 2.3, the allocation can be rewritten as:

$$
\sum_{j=1}^{N} \frac{\partial s}{\partial q_j} q_j^* \left[ \mathbb{P}(I \geq a^*) + \tau - \sum_{k} \frac{\partial v_k}{\partial a} v'_k - V f_I(a^*) \right] \\
+ \sum_{j=1}^{N} \tilde{\chi}_\rho \frac{\partial \tilde{\rho}}{\partial q_j} q_j^* \left[ \sum_{k} \frac{\partial v_k}{\partial a} v'_k \right] + \sum_{j=1}^{N} \frac{\partial \text{VaR}_{\tilde{\alpha}^*}}{\partial q_j} q_j^* [V f_I(a^*)] = a^* [\mathbb{P}(I \geq a^*) + \tau],
$$

so that:

$$
\sum_{j=1}^{N} \frac{\partial}{\partial q_j} \left( (1 - \zeta_1 - \zeta_2) s + \zeta_1 \tilde{\chi}_\rho \tilde{\rho} + \zeta_2 \text{VaR}_{\tilde{\alpha}^*} \right) q_j^* = a^*,
$$

where:

$$
\zeta_1 = \frac{\sum_k \frac{\partial v_k}{\partial a} v'_k}{\mathbb{P}(I \geq a^*) + \tau},
$$

$$
\zeta_2 = \frac{V f_I(a^*)}{\mathbb{P}(I \geq a^*) + \tau}.
$$

Hence, the Euler principle can still be applied, but the supporting risk measure is a weighted average of the external risk measure $s$, the internal risk measure $\tilde{\rho}$, and VaR. The weights on the latter two measures are determined by the consumers’ and the shareholders’ preferences for capitalization of the company at the margin and are dictated by the considerations outlined in Section 2.3.

### 4 Comparison of Capital Allocation Methods

This section shows how the counterparty-driven internal allocation effectively differs from the allocation based on ES—perhaps the most widely endorsed measure within the academic and practitioner communities—in the context of two single-period examples. In particular, we are interested in how the economic weight assigned to various outcomes and various risks under $\tilde{\rho}$ differs from what would be obtained from applying the Euler technique to ES.
4.1 The Case of Homogeneous Exponential Losses

Assume that there are $N$ identical, independent 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 firm’s optimization problem in the one period model without a regulatory constraint (6)/(7) may be represented as:

$$\max_{a,q,p} \left\{ N \times p - N \times q \times \left[ \frac{1}{\nu} \Gamma_{N-1,\nu} \left( \frac{a}{q} \right) - \frac{\nu^{N-1}}{(N-1)!} e^{-\nu a} \frac{\nu}{\nu - \alpha} \right] - a \times \tilde{\Gamma}_{N,\nu} \left( \frac{a}{q} \right) - \tau \times a \right\}$$

subject to

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

where $\tilde{\Gamma}_{m,b}(x) = 1 - \Gamma_{m,b}(x)$ and $\Gamma_{m,b}(\cdot)$ denotes the cumulative distribution function of the Gamma distribution with parameters $m$ and $b$ (see the Appendix C.1 for the derivation of (34)).

Note that the premiums $p_i \equiv p$ and the coverage amounts $q_i \equiv q$ are equal for all consumers since they are identical. Likewise, any reasonable allocation rule trivially yields identical allocations for each consumer, and particularly:

$$q \tilde{\phi}_i = N^{-1}, \ i = 1, 2, \ldots, N,$$

for the counterparty-driven allocation—rendering a comparison to other allocation methods moot. However, we may analyze how different allocations arrive at this congenerous result, i.e. we can distinguish different allocation methods by comparing the weight they put on different loss states.

For instance, for the allocation based on the Expected Shortfall (ES) according to the Euler princi-
ple, it is well known that (see e.g. McNeil, Frey, and Embrechts (2005)): \(^{15}\)

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

where \(L = \sum_j L_j\), i.e. the Expected Shortfall can be associated with a linear weighting function of the default states.

For the counterparty-driven allocation, on the other hand, we obtain:

\[
\frac{1}{N} \text{ Eq. (11)} = \frac{\mathbb{E} \left[ \left( I_{\{qL \geq a\}} \sum_{j=1}^{N} U' \left( w - p - L_j + a \frac{L_j}{L} \right) \frac{L_j}{L} \right] \right]}{\mathbb{E} \left[ I_{\{qL \geq a\}} \sum_{j=1}^{N} U' \left( w - p - L_j + a \frac{L_j}{L} \right) \frac{L_j}{L} \right]}
= \mathbb{E}^{\tilde{\psi}} \left[ \frac{L_i}{L} \right] = \frac{1}{N} \mathbb{E} \left[ \frac{\partial^p}{\partial \tilde{\psi} L} \right],
\]

(35)

i.e. the weighting function is implied by the measure transform and can be expressed as:

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

(36)

where \(\tilde{c}_{N,\nu,a,q}\) is a constant ensuring that \(\mathbb{E} \left[ \tilde{\psi}(L) \right] = 1.\)

In particular, for the risk measure \(\tilde{\rho}\) evaluated at the aggregate loss \(I\), we have:

\[
\tilde{\rho}(I) = \tilde{\rho}(q L) = \exp \left\{ \mathbb{E} \left[ \tilde{\psi}(L) \log\{q L\} \right] \right\} = \exp \left\{ \mathbb{E} \left[ \tilde{\psi}(L) \log\{q L\} | q L \geq a \right] \right\},
\]

(37)

with \(\tilde{\psi}(l) = P(q L \geq a) \times \tilde{\psi}(l)\). Thus, \(\tilde{\rho}\) in this sense is in fact a tail risk measure, and hence is related to ES. 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). However, while there the risk spectrum purports to encode the “subjective risk aversion of an investor” to justify overweighting bad outcomes, in our setting

\(^{15}\)As in Section 3.2, to compare results we assume that the confidence level is chosen according to the outcome of the firm’s optimization problem, i.e. \(\alpha = P(q L \geq a)\).

\(^{16}\)The derivation of these equations, a closed form solution for \(\tilde{c}_{N,\nu,a,q}\) as well as a representation of \(\tilde{\psi}(\cdot)\) not involving an infinite sum for implementation purposes all are provided in Appendix C.2.
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 a subjectively specified preference function for the firm.

In the absence of weights, the concavity of the logarithmic function will, in the course of the application of the Euler allocation methods, tend to penalize bad outcomes less heavily than ES. In fact, it is evident from (35) that for $\hat{\psi}(\cdot) \equiv 1$, the counterparty-driven allocation will effectively weight all aggregate loss outcomes in excess of the firm’s capital equally, regardless of size. The reason for this is that $\hat{\psi}(\cdot) \equiv 1$ implies that the firm’s counterparties are risk-neutral ($\alpha = 0$ in (36)) and, thus, the value of the firm in all states of default, regardless of how extreme the default, is simply the firm’s assets. At the margin, the counterparties evaluate changes in risk simply from the perspective of how the expected value of recoveries from the firm are affected, and recoveries in mild states of default are weighted no differently from severe ones. This is also the reason why $\tilde{\rho}$ is not sub-additive or translation-invariant: Adding a constant amount to the aggregate loss in high loss states is less precarious than in low loss states because of limited liability.

Under risk aversion, on the other hand, $\hat{\psi}(\cdot) \neq 1$, and counterparties will weight recoveries in severe states of default more heavily than mild ones. In fact, in the current setting $\hat{\psi}(\cdot)$ is increasing and strictly convex, and for all risk aversion levels $\alpha > 0$ there exists a loss level $l_0$ such that the weighting function for the counterparty-driven allocation will exceed that of ES for all loss levels greater than $l_0$.$^{17}$ However, for smaller—and more probable—loss realizations, different shapes are possible rendering either the ES-based allocation or the counterparty-driven allocation to appear more conservative.

To 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 (34) with different risk aversion levels. The properties are as expected: $a$, $p$, and $q$ all are increasing in risk aversion. Figure 1 plots the weighting function $\hat{\psi}$ against the linear weighting function associated with the Expected Shortfall.

---

$^{17}$It is worth noting that $\hat{\psi}(\cdot)$ generally is not necessarily convex; it is possible to construct examples where the shape is strictly concave or even linear.
**Exponential Losses: Parametrizations**

<table>
<thead>
<tr>
<th>Nr.</th>
<th>N</th>
<th>ν</th>
<th>τ</th>
<th>α</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.

for varying risk aversion levels. We find two qualitatively different shapes: 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. Thus, the counterparty-driven allocation can be deemed more conservative than the ES-based allocation in this case. For the low risk aversion level, on the other hand, \( \hat{\psi} \) crosses the linear weighting function twice; in this case, the weighting function within the new risk measure \( \tilde{\rho} \) puts more mass on low and extremely high loss states, while the weights are smaller for intermediate to high loss states. Thus, it can be deemed less conservative than ES in this case.

Therefore, when relying on the one-size-fits-all solution of Expected Shortfall for the purpose of internal capital allocation, the loss-specific weights may be too conservative or not conservative enough, depending on the situation. In contrast, the risk measure \( \tilde{\rho} \) from (37) provides a more nuanced answer that can be tailored to the specific circumstances.

### 4.2 The Case of Heterogenous Bernoulli Losses

To analyze the impact of the different weighting of loss states on the allocation of capital, we now consider a setup with heterogeneous consumers. Assume there are \( m \) groups of consumers, where each group \( i \) contains \( N_i \) identical consumers with wealth level \( w_i \) and utility function \( U_i(\cdot) \) that face independent, Bernoulli distributed losses \( l_i \) occurring with a probability \( \pi_i, i = 1, \ldots, m \). The partic-
The marginal cost of risk, risk measures, and capital allocation

...ipation 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 optimization problem in the one period model without a regulatory constraint (6)/(7) can then be conveniently set up by observing 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 consumer in group \(i\):

\[ q_i \hat{\phi}_i = \hat{c} \sum_{k_1=0}^{N_1} \cdots \sum_{k_i=1}^{N_i} \cdots \sum_{k_m=0}^{N_m} \left( \begin{array}{c} N_1 \\ k_1 \end{array} \right) \cdots \left( \begin{array}{c} N_i - 1 \\ k_i - 1 \end{array} \right) \cdots \left( \begin{array}{c} N_m \\ k_m \end{array} \right) \\
\times \pi_1^{k_1} \cdots \pi_i^{k_i} \cdots \pi_m^{k_m} (1 - \pi_1)^{N_1-k_1} \cdots (1 - \pi_i)^{N_i-k_i} \cdots (1 - \pi_m)^{N_m-k_m} \]

of the Bernoulli setup.
the risk aversion level, which is due to a combination of two effects. On the one hand, more risk-averse consumers with differing loss exposure and risk aversion levels. Specifically, we assume consumers in the first group have wealth level \( w_1 = 7 \), CARA preferences with an absolute risk aversion level \( \alpha_1 = 1.75 \), and face a loss of size \( l_1 = 2 \) occurring with a probability of \( \pi_1 = 10\% \). Consumers in the second group also have wealth \( w_2 = 7 \), CARA preferences, and face a loss with an occurrence probability of \( \pi_2 = 10\% \), but we vary their risk aversion level \( \alpha_2 \) and the size of the loss \( l_2 \). Figure 2 shows the optimal capital level, \( a \), and the optimal coverage amount for consumers in group 2, \( q_2 \), as a function of \( \alpha_2 \) and \( l_2 \) for group sizes \( N_1 = N_2 = 25 \) and capital cost \( \tau = 5\% \).

Unsurprisingly, Panel 2(a) shows that the optimal capital level is increasing in the loss level—which of course is due to an increase in the expected loss. Moreover, we find that capital is also increasing in the risk aversion level, which is due to a combination of two effects. On the one hand, more risk-averse
consumers are—ceteris paribus—more worried about nonperformance of their insurance contract and thus prefer a higher level of capital. On the other hand, as is evident from Panel 2(b), higher risk aversion levels prompt the consumers in group 2 to increase their coverage level $q_2$, which in turn increases expected losses.\footnote{For the participation amount of consumers in group 1, $q_1$, we observe the opposite relationship, i.e. it is decreasing in group 2 loss and risk aversion levels. The reason here is that—ceteris paribus—recoveries are larger and insurance is more expensive when the capital level is higher.}

To demonstrate the influence of these effects on capital allocation, Figure 3 plots the difference in the proportion of capital allocated to group 2 between the counterparty-based allocation rule, $q_2 \bar{\phi}_2$, and the ES allocation rule, $a_{ES}^\alpha$, as a function of $l_2$ and $\alpha_2$. From the figure, the functional relationship appears to be non-smooth, which—aside from numerical errors—is due to genuine discontinuities emerging from the discreteness of the resulting aggregate loss distribution. To better illustrate the underlying relationships, we therefore include a $5 \times 5$-polynomial fit as well as contour lines based on the smoothed surface.

To interpret the results, it is important to notice that the figure displays differences in allocations—the absolute allocation levels vary much more across the different loss and risk aversion levels for both allocation methods. For instance, for a risk aversion level of $\alpha_2 = 1.75$ and a loss level $l_2 = 6.5$, the proportion of capital allocated to consumers in group 2 based on ES is about 86.6% and about 87.5% for the counterparty-based allocation principle, whereas for a risk aversion level $\alpha_2 = 1.75$ and loss
$l_2 = 2$, both allocations clearly yield 50% since the two consumer groups are identical in this case. In particular, the difference between the two allocations is zero in the latter case as is also evident from the corresponding contour line.

At first glance, the differences of about -1% for relatively low loss and risk aversion levels and about 2% for relatively large loss and risk aversion levels—resulting in an overall range of about 3% across all combinations—appear rather small. However, it is important to note that with Bernoulli distributions and group sizes of $N_i = 25$, diversification effects are very pronounced so that capital allocations to each group are largely driven by their respective contributions to expected losses. Smaller group sizes yield effects that are qualitatively similar but considerably larger; however, the aforementioned discontinuities are also much more pronounced due to the discreteness of the loss distribution and impede clear presentation of the primary effects.

Figure 3: Difference in allocation of capital to group 2 between the counterparty-based allocation rule and the ES allocation rule $(q_2 \tilde{\phi}_2 - a^{ES}_2/a_2)$ for group 2 loss levels $l_2$ between between 2 and 6.5 and group 2 risk aversion levels $\alpha_2$ between 0.25 and 2.95.

As before, we find that the ES based allocations are too conservative—i.e. they put too much weight on the group with the high risk consumers in group 2—if those consumers are not very risk averse relative to group 1 consumers and the difference in loss sizes is relatively small. On the other hand, if
group 2 consumers are relatively risk averse and face relatively large losses, the ES-based allocations are not conservative enough—i.e. the counterparty-based principle allocates more capital to the high risk group than ES.

5 Conclusion

The gradient allocation principle prescribes an allocation of capital to risks in the company’s portfolio according to their marginal costs as defined via a given risk measure. However, risk measure selection is a thorny issue that can be resolved only through careful consideration of institutional context—a problem implicitly recognized in the early literature on capital allocation (Myers and Read (2001), Tasche (2004)—both of whom ultimately fall back on regulation as motivating the choice of risk measure) but since overlooked in favor of a focus on mathematical properties of allocation rules and risk measures. Instead of starting with a risk measure, this paper starts with primitive assumptions and calculates the marginal cost of risk from the perspective of a profit-maximizing firm with risk averse counterparties. We then take the additional step of identifying the risk measure whose gradient yields allocations consistent with marginal cost.

The resulting composite measure is a weighted average of three risk measures corresponding to the interests of the different parties concerned with the capitalization of the firm: the regulator, shareholders, and counterparties. The first risk measure relating to regulation is specified exogenously as a part of the company’s optimization problem, while the risk measures corresponding to shareholder and counterparty concerns arise endogenously in our setting. Shareholder concerns yield VaR whereas counterparty concerns yield a novel risk measure, which can be expressed as the exponential of the expected value of the logarithm of the portfolio outcomes under an alternative probability measure which reflects weights for counterparty preferences in default states. The composite risk measure will not generally obey the principles of convexity or coherence. Nevertheless, it is the only one that yields the appropriate allocation of capital for the profit-maximizing firm in our setting.

To obtain these theoretically economically correct allocations, one needs considerable information to choose the appropriate weights as well as the probability transform encapsulated in the novel risk
measure. That said, conventional risk measures also rely on selections for parameter values—such as the confidence level in VaR or ES, or the weighting function in spectral risk measures—so from the standpoint of practice the proposed measure starts from an economically sound foundation and offers no greater level of complication than is already present. Moreover, as is evident from our numerical applications, the parameters are structural, i.e. they correspond fundamental quantities, so they may be calibrated according to the application in view.

We compare the allocations obtained in a single period setting without regulation to those obtained from the gradient of ES—the coherent risk measure currently favored by many academics and regulators. We show that ES may underweight or overweight severe states of default, depending on the nature of customer risk aversion. This raises the interesting possibility that a transition away from a system of regulation based on risk measure-based solvency assessment to one based on market (counterparty) discipline will not necessarily mitigate the oft-lamented failure of financial institutions to penalize “tail” risk.

While the calculations in this paper are done from the perspective of a profit-maximizing firm, one could also contemplate the calculus of a regulator or social planner. In specific cases, the calculus will be similar. For example, a regulator without responsibility for unpaid losses (i.e., if no deposit insurance scheme exists) but in a context where counterparties are uninformed will view risk similarly to the profit-maximizing firm. However, a regulator responsible for unpaid losses would presumably have to consider the value of unpaid losses in selecting a risk measure as well as other issues—such as bankruptcy costs not internalized by private firms and the production cost technology associated with deposit insurance—that would determine the optimal level of capitalization for financial institutions as well as the social cost of risk. These issues are of course complex and well beyond the scope of this paper, but they are intriguing areas for future research.
Appendix

A  Implementation of the Firm’s Optimization Problem 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 (in Appendix B.1) that the marginal price increase associated with coverage must satisfy (10) 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

\[
\tilde{v}_i \left( w_i - p_i^*(q_i); q_i, \tilde{a}, \tilde{q}_1, \ldots, \tilde{q}_N \right) = \mathbb{E} \left[ U_i \left( w_i - p_i^*(q_i) - L_i + \tilde{R}_i \right) \right],
\]

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.  

Alternatively, we could also specify

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

where the consumer is cognizant of her own coverage but “scales” the company according to her own coverage level and this would not change the presentation in the main text. Again, the important point is that the consumer expects the same recovery per dollar of coverage in default states independent of her choice of \( q_i \).
and in addition the new constraint:
\[
\tilde{q}_i \in \arg \max_{q_i} \tilde{v}_i(w_i - p^*_i(q_i), q_i; \tilde{a}, \tilde{q}_1, ..., \tilde{q}_N), \ \forall i.
\] (42)

Equation (42) 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 \(\hat{a}_i\), \(\{\hat{q}_i\}\), and \(\{\hat{p}_i\}\) 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 \(p^*_i(\hat{q}_i) = \hat{p}_i\) and induces consumers to choose the original solution:
\[
\hat{q}_i \in \arg \max_{q_i} \hat{v}_i(w_i - p^*_i(q_i), q_i; \hat{a}, \hat{q}_1, ..., \hat{q}_N), \ \forall i.
\]

The following lemma shows that this function exists.

**Lemma A.1.** Suppose \(\hat{a}_i\), \(\{\hat{q}_i\}\), and \(\{\hat{p}_i\}\) 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. \(p^*_i(\hat{q}_i) = \hat{p}_i\).
2. \(\hat{q}_i \in \arg \max_{q_i} \tilde{v}_i(w_i - p^*_i(q_i), q_i; \hat{a}, \hat{q}_1, ..., \hat{q}_N)\).

**Proof.** Start by noting that it is evident that the constraints (7) all bind. Note further that the function of \(x\):
\[
g(x) = \tilde{v}_i(w_i - x, 0; \hat{a}, \hat{q}_1, ..., \hat{q}_N)
\]
is monotonically decreasing and, hence, invertible, so that we may uniquely define:
\[
p^*_i(0) = g^{-1}(\gamma_i),
\] (43)
which obviously satisfies:
\[
\tilde{v}_i(w_i - p^*_i(0), 0; \hat{a}, \hat{q}_1, ..., \hat{q}_N) = \gamma_i.
\]

Furthermore, let \(p^*_i(\cdot)\) be a solution to the differential equation (initial value problem):
\[
\frac{\partial p^*_i(x)}{\partial x} = \frac{\partial}{\partial x} \tilde{v}_i(w_i - p^*_i(x), x; \hat{a}, \hat{q}_1, ..., \hat{q}_N), \quad p^*_i(0) = g^{-1}(\gamma_i),
\] (44)
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.\(^{22}\)

\(^{22}\)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.
Moving on, by construction we know that:

\[
\tilde{v}_i(w_i - p_1^*(q_i), q_i; \hat{a}, \hat{q}_1, ..., \hat{q}_N) = \gamma_i + \int_0^{q_i} \left[ \frac{\partial}{\partial q_i} \tilde{v}_i(w_i - p_1^*(x), x; \hat{a}, \hat{q}_1, ..., \hat{q}_N) - \frac{\partial}{\partial w} \tilde{v}_i(w_i - p_1^*(x), x; \hat{a}, \hat{q}_1, ..., \hat{q}_N) \times \frac{\partial p_1^*(x)}{\partial x} \right] dx
\]

In particular,

\[
\tilde{v}_i(w_i - p_1^*(\hat{q}_i), \hat{q}_i; \hat{a}, \hat{q}_1, ..., \hat{q}_N) = \gamma_i,
\]

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

\[p_1^*(\hat{q}_i) = \hat{p}_i,\]

proving the first part of the lemma. Moreover, (45) directly implies that:

\[\hat{q}_i \in \arg \max_{q_i} \tilde{v}_i(w_i - p_1^*(q_i), q_i; \hat{a}, \hat{q}_1, ..., \hat{q}_N),\]

proving the second part.

\[\square\]

### B Identities in Section 2

#### B.1 Derivation of Equation (10)

Let \(\lambda_k\) be the Lagrange multiplier associated with the participation constraint (7) for consumer \(k\), and let \(\xi\) the multiplier associated with (8). Then, the first order conditions for an interior solution of Problem 6 are:\(^{23}\)

\[
[q_i] - \sum_k \frac{\partial e_k}{\partial q_i} + \sum_k \lambda_k \frac{\partial v_k}{\partial q_i} - \frac{\partial s}{\partial q_i} \xi = 0, \quad (46)
\]

\[
[a] - \sum_k \frac{\partial e_k}{\partial a} - \tau + \sum_k \lambda_k \frac{\partial v_k}{\partial a} + \xi = 0, \quad (47)
\]

\[
[p_i] 1 - \lambda_i \frac{\partial v_i}{\partial w} = 0. \quad (48)
\]

Using (46) and (48), we see that (9) may be rewritten as:

\[
\frac{\partial p_1^*}{\partial q_i} \overset{(48)}{=} \frac{1}{v'_i} \left[ \frac{\partial v_i}{\partial q_i} + \mathbb{E} \left( \mathbf{1}_{\{I \geq a\}} \frac{U'_i}{I} \frac{\partial I_i}{\partial q_i} \right) \right]
\]

\[
\overset{(46)}{=} \sum_k \frac{\partial e_k}{\partial q_i} + \frac{\partial s}{\partial q_i} \xi - \sum_{k \neq i} \frac{\partial p_k}{\partial q_i} \frac{\partial v_k}{v'_k} + \mathbb{E} \left( \mathbf{1}_{\{I \geq a\}} \frac{U'_i}{I} \frac{\partial I_i}{\partial q_i} \right) \frac{\partial I_i}{\partial q_i}.
\]

\(^{23}\)Despite the apparent kinks, the objective function and the constraints are differentiable for continuous distributions as can be easily verified by an application of the Leibniz rule.
or, simplifying and using (47):
\[
\frac{\partial p_i^*}{\partial q_i} = \frac{\partial e_i^Z}{\partial q_i} + \frac{\partial s}{\partial q_i} \left[ \sum_k \frac{\partial e_k}{\partial a} + \tau - \sum_k \frac{\partial w_k}{\partial a} \right] + \mathbb{E} \left[ 1_{\{I \geq a\}} \sum_k \frac{U_k'}{v_k'} I_k \frac{\partial I_i}{\partial q_i} \right].
\]

Moving on:
\[
\frac{\partial p_i^*}{\partial q_i} = \frac{\partial e_i^Z}{\partial q_i} + \frac{\partial s}{\partial q_i} \left[ P(I \geq a) + \tau - \sum_k \frac{\partial v_k}{\partial a} v_k' \right] + \mathbb{E} \left[ 1_{\{I \geq a\}} \sum_k \frac{U_k'}{v_k'} I_k \frac{\partial I_i}{\partial q_i} \right] \times \sum_k \frac{\partial w_k}{\partial a} v_k' \times a
\]

and, therefore, (10).

**B.2 Derivation of Equation (18)**

In the setting of Section 2.2, the firm’s problem becomes:

\[
\max_{a,\{q_i\},\{p_i\},\{K_j\},\{w_{ij}\}} \sum p_i - \sum e_i - \tau a,
\]

subject to:

\[
\begin{align*}
v_i & \geq \gamma_i, \\
s(q_1,\ldots,q_N) & \leq a, \\
\sum_j \pi_j K_j & = 1, \\
\sum_j \pi_j w_{ij} & = w_i.
\end{align*}
\]

In addition to a new set of optimality conditions connected with \{K_j\} and \{w_{ij}\}, we have the same set of first order conditions:

\[
\begin{align*}
[q_i] & - \sum_k \frac{\partial e_k}{\partial q_i} + \sum_k \lambda_k \frac{\partial v_k}{\partial q_i} - \frac{\partial s}{\partial q_i} \xi = 0, \\
[a] & - \sum_k \frac{\partial e_k}{\partial a} - \tau + \sum_k \lambda_k \frac{\partial v_k}{\partial a} + \xi = 0, \\
[p_i] & 1 - \lambda_i \frac{\partial v_i}{\partial w} = 0.
\end{align*}
\]

The first order condition for \{w_{ij}\} is:

\[
[w_{ij}] \lambda_i \frac{\partial v_i}{\partial w_{ij}} - \eta_i \pi_j = 0
\]

\[
\Leftrightarrow \lambda_i p_j^{(S)} \mathbb{E}[U_i'(W_i - p_i - L_i + R_i)|\omega_j^{(S)}] - \eta_i \pi_j = 0,
\]

(49)
where \( \{ \eta_i \} \) are the Lagrange multipliers for the individual wealth constraints. Since

\[
0 = \sum_j \left( \lambda_i p_j^{(S)} \mathbb{E}^P \left[ U_i'(W_i - p_i - L_i + R_i) \omega_j^{(S)} \right] - \eta_i \pi_j \right)
\]

\[
= \lambda_i \frac{\partial v_i}{\partial w} - \eta_i,
\]

with \( [p_i] \) we obtain \( \eta_i \equiv 1 \). Thus, we also have:

\[
\frac{\pi_j}{p_j^{(S)}} = \frac{\mathbb{E}^P \left[ U_i'(W_i - p_i - L_i + R_i) \omega_j^{(S)} \right]}{\mathbb{E}^P \left[ U_i'(W_i - p_i - L_i + R_i) \right]}.
\]

(50)

As before, we seek a pricing function satisfying:

\[
\left[ \frac{\partial v_i}{\partial q_i} + \mathbb{E}^P \left[ 1_{\{i \geq K \}} U_i' \frac{K a}{I^2} I_i \frac{\partial I_k}{\partial q_i} \right] - \frac{\partial v_i}{\partial w} \right] = 0.
\]

Proceeding analogously to Appendix B.1, we arrive at the marginal pricing condition associated with a decentralized implementation:

\[
\frac{\partial p_i^*}{\partial q_i} = \sum_k \frac{\partial e_k}{\partial q_i} + \frac{\partial s}{\partial q_i} \xi - \sum_{k \neq i} \frac{\partial a_k}{v_k} \eta_k + \mathbb{E}^P \left[ 1_{\{i \geq K \}} U_i' \frac{K a}{I^2} I_i \frac{\partial I_k}{\partial q_i} \right].
\]

Simplifying, we obtain:

\[
\frac{\partial p_i^*}{\partial q_i} = \sum_k \frac{\partial e_k}{\partial q_i} + \frac{\partial s}{\partial q_i} \left[ \frac{\partial e_k}{\partial a} + \tau - \sum_k \frac{\partial a_k}{v_k} \right] + \mathbb{E}^P \left[ 1_{\{i \geq K \}} \sum_k \frac{U_k}{v_k} \frac{K a}{I^2} I_k \frac{\partial I_k}{\partial q_i} \right] - \sum_k \frac{\partial a_k}{v_k} \xi - \xi \frac{\partial \eta_k}{v_k} + \mathbb{E}^P \left[ 1_{\{i \geq K \}} \sum_k \frac{U_k}{v_k} \frac{K a}{I^2} I_k \frac{\partial I_k}{\partial q_i} \right] \times \sum_k \frac{\partial a_k}{v_k} \times \alpha
\]

\[
= \sum_k \frac{\partial e_k}{\partial q_i} + \frac{\partial s}{\partial q_i} \left[ \mathbb{E}^Q \left[ K 1_{\{i \geq K \}} \right] + \tau - \sum_k \frac{\partial a_k}{v_k} \right] + \mathbb{E}^Q \left[ 1_{\{i \geq K \}} \sum_k \frac{U_k}{v_k} \frac{K a}{I^2} I_k \frac{\partial I_k}{\partial q_i} \right] \times \sum_k \frac{\partial a_k}{v_k}
\]

\[
= \sum_k \frac{\partial e_k}{\partial q_i} + \frac{\partial s}{\partial q_i} \left[ \mathbb{E}^Q \left[ K 1_{\{i \geq K \}} \right] + \tau - \sum_k \frac{\partial a_k}{v_k} \right] + \tilde{\omega}_i \times \alpha \times \left[ \sum_k \frac{\partial a_k}{v_k} \right],
\]

i.e. (17), where:

\[
\tilde{\omega}_i = \frac{\mathbb{E}^P \left[ 1_{\{i \geq K \}} \sum_k \frac{U_k}{v_k} \frac{K a}{I^2} I_k \frac{\partial I_k}{\partial q_i} \right]}{\mathbb{E}^P \left[ 1_{\{i \geq K \}} \sum_k \frac{U_k}{v_k} K I_k \right]} \quad (50)
\]

\[
= \frac{\mathbb{E}^Q \left[ 1_{\{i \geq K \}} \sum_k \frac{U_k}{v_k} \frac{K a}{I^2} I_k \frac{\partial I_k}{\partial q_i} \right]}{\mathbb{E}^Q \left[ 1_{\{i \geq K \}} \sum_k \frac{U_k}{v_k} K I_k \right]},
\]

i.e. (18). In the limiting case of a complete market, clearly \( U_i'(w_{ij} - p_i - l_{ij} + r_{ij}) \) is \( \mathcal{F}^{(S)} \)-measurable, so that we immediately obtain (19).
B.3 Derivation of Equation (24)

The first order conditions of program (23) are:

\[
[q_i] - \sum_k \frac{\partial e_k}{\partial q_i} - V E \left[ \frac{\partial I_i(L_i, q_i) \mathbf{1}_{\{I=a\}}}{\partial q_i} \right] dt + \sum_k \lambda_k \frac{\partial v_k}{\partial q_i} - \xi \frac{\partial s}{\partial q_i} = 0, 
\]

\[
[a] - \sum_k \frac{\partial e_k}{\partial a} - \tau + \sum_k \lambda_k \frac{\partial v_k}{\partial a} + \xi + V f_I(a) = 0, 
\]

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

where—as before—\( \lambda_i \) and \( \xi \) denote the Lagrange multipliers for the first set of conditions and the second condition, respectively, and \( f_I \) denotes the cumulative density function of \( I \).

Again, we assume 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, so that the marginal price change at the optimal level of \( q_i \) must satisfy:

\[
\frac{\partial v_i}{\partial w} \frac{\partial p_i^*}{\partial q_i} = \frac{\partial v_i}{\partial q_i} + E \left[ 1_{\{I>a\}} \frac{a}{(I)^2} I_i \frac{\partial I_i}{\partial q_i} U_i' \right]. 
\]

Therefore:

\[
\frac{\partial p_i^*}{\partial q_i} = \frac{1}{\frac{\partial v_i}{\partial w}} \left[ \frac{\partial v_i}{\partial q_i} + E \left[ 1_{\{I>a\}} \frac{a}{(I)^2} I_i \frac{\partial I_i}{\partial q_i} U_i' \right] \right] 
\]

\[
= \sum_k \frac{\partial e_k}{\partial q_i} + \xi \frac{\partial s}{\partial q_i} + V E \left[ \frac{\partial I_i(L_i, q_i) \mathbf{1}_{\{I=a\}}}{\partial q_i} \right] dt - \sum_{k \neq i} \frac{\partial v_k}{\partial q_i} v_k' + E \left[ 1_{\{I>a\}} \frac{a}{(I)^2} I_i \frac{\partial I_i}{\partial q_i} U_i' \right] 
\]

\[
= \frac{\partial e_k}{\partial q_i} + \left[ \mathbb{P}(I \geq a) + \tau - \sum_k \frac{\partial e_k}{\partial a} - V f_I(a) \right] \frac{\partial s}{\partial q_i} + V f_I(a) E \left[ \frac{\partial I_i}{\partial q_i} I = a \right] 
\]

\[
+ E \left[ 1_{\{I>a\}} \sum_k U_k' v_k' \frac{a}{(I)^2} I_k \frac{\partial I_i}{\partial q_i} \right] 
\]

and since:

\[
\sum_k \frac{\partial e_k}{\partial a} v_k' = E \left[ 1_{\{I>a\}} \sum_k U_k' I_k v_k' \right], 
\]

we obtain (24).

\(^{24}\)For simplicity, we omit the “\( t \)” super- and subscripts in case no ambiguity arises.
C Identities in Section 4

C.1 Derivation of Equation (34)

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

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

For part ii., note that \(L_{-N} + L_N\) is \(\text{Beta}(1, N - 1)\) distributed independent of \(L_{-N} + L_N \sim \text{Gamma}(N, \nu)\). Hence, part ii. can be written as

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

For part i., we have

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

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

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

Therefore, since all consumers are identical, the objective function (6) takes the form displayed in (34). For condition (7), on the other hand, we have

\[
V = V_N = \mathbb{E} \left[ U (w - p - L_N + R_N) \right] = \mathbb{E} \left[ U (w - p - (1 - q) L_N) 1_{\{q(L_{-N} + L_N) \leq a\}} \right] + \mathbb{E} \left[ U (w - p - L_N + a \frac{L_N}{L_{-N} + L_N}) 1_{\{q(L_{-N} + L_N) \geq a\}} \right]
\]

For part i., we obtain

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

For part \( ii. \), note that

\[ \begin{align*}
\mathbb{E} \left[ U \left( w - p - ((L-N) + L_N) - a \right) \frac{L_N}{L - N + L_N} \right]_{(L-N + L_N) \geq a} \\
= \ & \int_0^1 \int_0^\infty e^{-\alpha(w-p-(l-a)y)} I_{\{l \geq a/q\}} \frac{\nu^N}{(N-1)!} l^{N-1} e^{-\nu l} (N-1)(1-y)^{N-2} \, dl \, dy \\
= \ & e^{-\alpha(w-p)} \int_0^{\infty} \frac{\nu^N}{(N-1)!} e^{-\nu l} l^{N-1} \left[ \int_0^1 e^{(\alpha(l-a))y} (N-1)(1-y)^{N-2} \, dy \right] \frac{mgf_{Beta(1,N-1)}(\alpha(l-a))}{l^{\alpha(l-a)}} \\
= \ & e^{-\alpha(w-p)} \int_0^{\infty} \frac{\nu^N}{(N-1)!} e^{-\nu l} l^{N-1} \sum_{k=0}^{\infty} \frac{(\alpha(l-a))^k}{(N-1+k)!} \frac{\alpha^k}{l^{\alpha(l-a)}} (N+k-j-1)! \\
& \quad \times \int_{a/q}^{\infty} \frac{\nu^{N+k-j}}{(N+k-j-1)!} e^{-\nu(l-a)} (l-a)^{N+k-j-1} \, dl \\
= \ & e^{-\alpha(w-p)} \sum_{k=0}^{\infty} \frac{(N-1)!}{(N-1+k)!} \frac{(av)^j}{j!} e^{-\nu a} \frac{\alpha^k}{l^{\alpha(l-a)}} (N+k-j-1)! \frac{(N+k-j-1)!}{(N-1-j)!} \Gamma_{N-j+k,\nu}\left(\frac{a}{q}\right).
\end{align*} \]

### C.2 Derivation of Equation (36)

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

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

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

\[ f_{\frac{L_j}{L}, \frac{L_N}{L}}(x, y) = (1-x-y)^{N-3} (N-2) (N-1) 1_{(x,y \geq 0, x+y \leq 1)}, \quad j \neq N. \]

Whence, for part \( i. \),

\[ \begin{align*}
\mathbb{E} \left[ U \left( w - p - (L - a) \frac{L_N}{L} \right) \frac{L_N-1}{L} \frac{L_N}{L} \right].
\end{align*} \]
whereas for part ii.,

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

so that

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

For the denominator,

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

\[ q \hat{\phi}_i = \frac{1}{N} \mathbb{E} \left[ I\{q > a\} \right. \right.

\left. \sum_{k=0}^{\infty} \frac{(k+1)(\alpha(L-a))^k}{(N+k)!} e^{-\alpha a} \sum_{j=0}^{N-1} \frac{(\alpha N_j + j)!}{j!} \frac{(N + k - j - 1)!}{(N - j - 1)!} \Gamma_{N+k-j,\nu} \left( a \left( \frac{1 - q}{q} \right) \right) \right] = \hat{\psi}(L)

For implementation purposes, the numerator can be expressed as

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

Finally,

\[ \hat{\psi}(I) = \hat{\psi}(q) = \mathbb{E} \left[ \frac{\partial \hat{\psi}}{\partial \theta} \bigg|_{\theta = \alpha(L-a)} \right] \]

\[ = \text{const} \mathbb{E} \left[ \sum_{j=1}^{N} \alpha \exp \left\{ - \alpha \left( w - p - L_j + a \frac{L_j}{L} \right) \right\} \frac{L_j}{L} \right] \]

\[ = \text{const} \mathbb{E} \left[ c^{\alpha(L-a)/L} \frac{L_j}{L} \right] \]

\[ = \text{const} \frac{\partial}{\partial \alpha} \text{mgf}_{\text{Beta}(1, N-1)}(x) \bigg|_{x = \alpha(L-a)} = \hat{\psi}(L), \]

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


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


