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

May 2013

Abstract

Financial institutions define their marginal cost of risk on the basis of the gradients of arbitrarily chosen risk measures. We reverse this approach by calculating the marginal cost for a profit-maximizing firm with risk-averse counterparties, and then identifying the risk measure delivering the correct marginal cost. The resulting measure is a weighted average of three parts, each corresponding to one of three 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 in default states. 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, Markus Stricker, 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, ETH, GSU, Goethe University Frankfurt, Imperial College, the University of Cincinnati, the University of Copenhagen, the University of Iowa, the University of South Carolina, 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

The question of how to quantify or *measure* the risk of a financial institution has attracted considerable attention over the recent years and has triggered many contributions in various literatures. Among the most influential are a number of papers from the mathematical finance literature that promote an axiomatic approach to risk measurement. The risk measure, defined as a mapping from a set of random variables to the real numbers, is dubbed *coherent* (Artzner et al., 1999) only if it satisfies certain criteria.[1] But do these criteria yield risk measures suitable for every application?

In this paper, we argue that the properties of a risk measure should flow from the economic context of the problem. We demonstrate that the most widely accepted axioms, while potentially appropriate guidelines for risk measures in some applications, are not universally adequate. In particular, they fail when used for risk pricing and performance measurement, which are among the most important applications of risk measures in financial institutions.[2] More precisely, financial institutions use the gradients of risk measures to allocate the firms capital to the various risks within its portfolio—a process which effectively determines the marginal cost of risk and thus provides key inputs for pricing and performance measurement.

We reverse the sequence of this approach. Instead of taking gradients of an arbitrary risk measure to identify the marginal cost of risk, we start with an economic model of a financial institution with risk-averse counterparties in an incomplete market with frictional capital costs. We show that profit maximization in this environment 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 it generally does not adhere to the mathematical axioms. In particular, we show that conventional risk measures satisfying these axioms generally 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.

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

[1] Other authors also designate these axioms to be necessary for a risk measure to be *reasonable* (Cuoco, He, and Isaenko, 2008) or *sensible* (Acerbi and Tasche, 2002), or promulgate different axioms for a risk measure to be *natural* (Kou, Peng, and Heyde, 2012).

[2] 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.

value as a going concern. In the extreme case of no regulation and (close to) 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 academics and regulators.[3] 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.

Various extensions are possible. For instance, the company in our setting has perfect information about its counterparties, and we do not explicitly model the possibility of raising additional capital from capital markets in the multi-period context. We discuss these extensions in more detail in the final section of the paper. However, these extensions do not affect the key point of the paper: The true marginal cost of risk and the associated allocation of capital should flow from the

[3] 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)).

economic context of the problem. Different model setups will yield different risk measures, but 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,[4] economic theory thus far offers no 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; Erel, Myers, and Read, 2013), although this setting begs the question of why intermediaries would hold capital in the first place.[5] 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 bankruptcy,[6] counterparty risk aversion, and regulatory constraints as motivators of risk

[4]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 a model with limited commitment and a specialized risk structure.

[5]Rampini and Viswanathan (2012) provide a rationale for financial intermediary capital in a complete market environment. In their model, it is opportune for intermediaries to hold capital because of different collateral requirements for households and intermediaries arising e.g. from limited enforcement as in Rampini and Viswanathan (2010)—i.e. intermediaries are "collateralization specialists." In contrast, within our incomplete market setting, holding capital is a risk management device that averts default in adverse states of the world.

[6]Bankruptcy costs originate from shareholders not having access to future profits in default states as in Smith and

management and determinants of marginal cost. We recover some familiar results in certain cases,[7] 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 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

Stulz (1985) and Smith, Smithson, and Wilford (1990). As noted by Froot, Scharfstein, and Stein (1993), this produces similar "mechanics" to those obtained when considering a convex cost of external finance as in Froot and Stein (1998).

[7]For instance, in the limiting case of a complete market, our "internal counterparty" allocation rule reduces to the allocation derived in Ibragimov, Jaffee, and Walden (2010), whereas it coincides with the allocation from Zanjani (2010) for the specialized risk structure considered there. In addition, VaR-based allocation (see e.g. Garman (1997) or Kalkbrener (2005)) is recovered in multiperiod settings with fully protected counterparties and no regulation.

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 Φ is a compact choice set. For tractability, we focus on a proportional arrangement, i.e. a linear contract, where the insurer agrees to reimburse q_i per dollar of loss:

$$I_i = I_i(L_i, q_i) = q_i \times L_i. \tag{1}$$

However, generalizations are possible.[8]

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:

$$R_i = \min\left\{I_i(L_i, q_i), \frac{a}{I}\, I_i(L_i, q_i)\right\}. \tag{2}$$

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:

$$e_i = \mathbb{E}\left[R_i\right] = \underbrace{\mathbb{E}\left[R_i\,\mathbf{1}_{\{I<a\}}\right]}_{=e_i^Z} + \underbrace{\mathbb{E}\left[R_i\,\mathbf{1}_{\{I\geq a\}}\right]}_{=e_i^D}.$$

There is a frictional cost—including taxes, agency, and monitoring costs—associated with holding assets in the company. In the spirit of Froot and Stein (1998), we represent the cost as a tax on assets:

$$\tau \times a, \tag{3}$$

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

$$\tau \times \left(a - \mathbb{E}\left[\sum_{i=1}^{N} \min\left\{I_i(L_i, q_i), \frac{a}{I}\, I_i(L_i, q_i)\right\}\right]\right) \tag{4}$$

and this does not change the ensuing allocation result. Without this frictional cost, the problem of default would be trivially solved as there would be no point in burdening the consumers with default risk (i.e. $a = \infty$ would be optimal).

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

$$v_i(a, \mathrm{w}_i - p_i, q_1, \ldots, q_N) = \mathbb{E}\left[U_i\left(\mathrm{w}_i - p_i - L_i + R_i\right)\right], \tag{5}$$

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

The firm then solves:

$$\max_{a,\{q_i\},\{p_i\}} \sum p_i - \sum e_i - \tau a, \tag{6}$$

[8] 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.

subject to participation constraints for each consumer:

$$v_i(a, \mathrm{w}_i - p_i, q_1, ..., q_N) \geq \gamma_i \quad \forall i \tag{7}$$

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

$$s(q_1, ..., q_N) \leq a, \tag{8}$$

where s is imagined to arise from 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 positively homogeneous of degree one.

The participation constraints contain the parameters γ_i which can be used to incorporate different degrees of competition. For example, γ_i set to reflect uninsured consumer utilities would correspond to a case of pure monopoly, where the monopolist could practice first-degree price discrimination and extract all consumer surplus associated with insurance. At the other end of the spectrum, the γ_i could be set so high as to simulate (close to) perfect competition.

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:

$$\left[\frac{\partial v_i}{\partial q_i} + \mathbb{E}\left[\mathbf{1}_{\{I \geq a\}}\, U_i' \,\frac{a}{I^2}\, I_i \,\frac{\partial I_i}{\partial q_i}\right]\right] - \frac{\partial v_i}{\partial \mathrm{w}} \frac{\partial p_i^*}{\partial q_i} = 0 \tag{9}$$

with $U_i' = U_i'(\mathrm{w}_i - p_i - L_i + R_i)$. 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 $\mathbb{E}\left[\mathbf{1}_{\{I \geq a\}}\, U_i' \,\frac{a}{I^2}\, I_i \,\frac{\partial I_i}{\partial q_i}\right]$.

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

$$\boxed{\frac{\partial p_i^*}{\partial q_i} = \frac{\partial e_i^Z}{\partial q_i} + \frac{\partial s}{\partial q_i}\left[\mathbb{P}(I \geq a) + \tau - \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}\right] + \tilde{\phi}_i \times a \times \left[\sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}\right],} \tag{10}$$

where

$$\boxed{\tilde{\phi}_i = \frac{\mathbb{E}\left[\mathbf{1}_{\{I \geq a\}} \sum_k \frac{U_k'}{v_k'}\, \frac{1}{I^2}\, I_k \,\frac{\partial I_i}{\partial q_i}\right]}{\mathbb{E}\left[\mathbf{1}_{\{I \geq a\}} \sum_k \frac{U_k'}{v_k'}\, \frac{I_k}{I}\right]}.} \tag{11}$$

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 easily verified that:

$$a \times \sum \tilde{\phi}_i\, q_i = a, \tag{12}$$

whereas the regulatory constraint "adds up" by the homogeneity assumption:

$$\sum \frac{\partial s}{\partial q_i} q_i = a. \tag{13}$$

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

$$\sum \frac{\partial p_i^*}{\partial q_i} q_i = \sum \frac{\partial e_i^Z}{\partial q_i} q_i + [\mathbb{P}(I \geq a)\, a + \tau a] = \sum e_i^Z + [\mathbb{P}(I \geq a)\, a + \tau a]\,.$$

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

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

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:

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

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

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

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

Full Coverage by Deposit Insurance and Binding Regulation

If consumers are fully covered by deposit insurance/guaranty funds—whether implicit or explicit—they are indifferent to the capitalization of their financial institution. Mathematically, this means that:

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

so that (10) becomes:

$$\frac{\partial p_i^*}{\partial q_i} = \frac{\partial e_i^Z}{\partial q_i} + \frac{\partial s}{\partial q_i} [\mathbb{P}(I \geq a) + \tau]\,. \tag{14}$$

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.[9]

[9]If deposit insurance were explicit and premiums were charged to financial institutions, then the form of the deposit premium function might also need to be considered in the firm's calculus, particularly if it contained risk penalties (e.g., Cummins (1988)).

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. (49) in Appendix B.1):

$$\sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'} = \left[\mathbb{P}(I \geq a) + \tau\right],$$

meaning that (10) becomes:

$$\frac{\partial p_i^*}{\partial q_i} = \frac{\partial e_i^Z}{\partial q_i} + \tilde{\phi}_i \times a \times \left[\mathbb{P}(I \geq a) + \tau\right]. \tag{15}$$

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. Then, (10) remains in its original form:

$$\frac{\partial p_i^*}{\partial q_i} = \frac{\partial e_i^Z}{\partial q_i} + \frac{\partial s}{\partial q_i}\left[\mathbb{P}(I \geq a) + \tau - \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}\right] + \tilde{\phi}_i \times a \times \left[\sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}\right], \tag{16}$$

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

$$\mathbb{P}(I \geq a) + \tau > \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'},$$

with the interpretation that regulation is forcing the institution to hold assets beyond the level that would be privately efficient from the perspective of serving its counterparties. 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:

$$\mathbb{P}(I \geq a) + \tau \approx \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'},$$

then the second term in (16) 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:

$$\mathbb{P}(I \geq a) + \tau \gg \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'},$$

then the second term in (16) overshadows the third, and external regulatory concerns dominate.

2.2 Allocation in a Securities Market Equilibrium

To this point, we have ignored hedging opportunities associated with securities markets. Such opportunities could be important for the insurer in the course of managing the risks in its liability portfolio, and for consumers in managing the risks associated with insurer default and any residual loss risk that is not insured. Specifically, we now allow both insurers and consumers to invest in (risky) capital market assets with the goals of maximizing insurer value and consumer utility, respectively.

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 formally, 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 $\left(\bar{\Omega}, \bar{\mathcal{F}}, \mathbb{P}\right)$, with:

$$\begin{aligned}
\bar{\Omega} &= \Omega^{(S)} \times \Omega = \left\{\bar{\omega} = (\omega^{(S)}, \omega) \middle| \omega^{(S)} \in \Omega^{(S)},\ \omega \in \Omega\right\}, \\
\bar{\mathcal{F}} &= \mathcal{F}^{(S)} \vee \mathcal{F}, \text{ and} \\
\mathbb{P}\left(\bar{A}\right) &= \sum_{j \in \Upsilon_A} p_j^{(S)} \times \mathbb{P}\left(A_j \middle| \{\omega_j^{(S)}\}\right)
\end{aligned}$$

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*:[10]

$$\mathbb{Q}\left(\bar{A}\right) = \sum_{j \in \Upsilon_A} \pi_j \times \mathbb{P}\left(A_j \middle| \{\omega_j^{(S)}\}\right),\ \bar{A} \subseteq \bar{\Omega},$$

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

[10]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 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 (2000) price bounds for Itô price processes as shown by Černý (2003). We refer to Föllmer and Schweizer (2010) for a formal definition.

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

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

where W_i is $\mathcal{F}^{(S)}$-measurable with $\mathrm{w}_{ij} = W_i(\omega_j^{(S)})$ and $\sum_j \pi_j \, \mathrm{w}_{ij} = \mathrm{w}_i$, whereas L_i is $\bar{\mathcal{F}}$-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 $\bar{\Omega}_j = \left\{ \bar{\omega} = (\omega^{(S)}, \omega) \middle| \, \omega^{(S)} = \omega_j^{(S)} \right\}$. We write K to denote the corresponding $\mathcal{F}^{(S)}$-measurable random variable. Consumer i's recovery can then be expressed as:

$$R_i = \min\left\{ I_i, \frac{K \, a}{I} \, I_i \right\},$$

and the fair valuation of claims is thus:

$$e_i = \mathbb{E}^{\mathbb{Q}}\left[R_i\right] = \underbrace{\mathbb{E}^{\mathbb{Q}}\left[R_i \, \mathbf{1}_{\{I < K\,a\}}\right]}_{=e_i^Z} + \underbrace{\mathbb{E}^{\mathbb{Q}}\left[R_i \, \mathbf{1}_{\{I \geq K\,a\}}\right]}_{=e_i^D}.$$

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:

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

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:

$$\boxed{\left[\mathbb{E}^{\mathbb{Q}}\left[K \, a \, \mathbf{1}_{\{I \geq K\,a\}}\right] + \tau \, a\right] = \sum_i \frac{\partial s}{\partial q_i} \, q_i \left[\mathbb{E}^{\mathbb{Q}}\left[K \, \mathbf{1}_{\{I \geq K\,a\}}\right] + \tau - \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}\right] + \sum_i \tilde{\phi}_i \, q_i \, a \times \left[\sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}\right],} \tag{17}$$

where:

$$\boxed{\tilde{\phi}_i = \frac{\mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{I \geq K\,a\}} \sum_k \frac{U_k'}{\mathbb{E}^{\mathbb{P}}[U_k' | \omega^{(S)}]} \frac{K}{I^2} \, I_k \, \frac{\partial I_i}{\partial q_i}\right]}{\mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{I \geq K\,a\}} \sum_k \frac{U_k'}{\mathbb{E}^{\mathbb{P}}[U_k' | \omega^{(S)}]} \, K \, \frac{I_k}{I}\right]} = \frac{\sum_j \pi_j \, K_j \, \mathbb{E}^{\mathbb{P}}\left[\mathbf{1}_{\{I \geq K_j\,a\}} \sum_k \frac{U_k'}{\mathbb{E}^{\mathbb{P}}[U_k' | \omega_j^{(S)}]} \, \frac{I_k}{I} \, \frac{\frac{\partial I_i}{\partial q_i}}{I} \middle| \, \omega_j^{(S)}\right]}{\sum_j \pi_j \, K_j \, \mathbb{E}^{\mathbb{P}}\left[\mathbf{1}_{\{I \geq K_j\,a\}} \sum_k \frac{U_k'}{\mathbb{E}^{\mathbb{P}}[U_k' | \omega_j^{(S)}]} \, \frac{I_k}{I} \middle| \, \omega_j^{(S)}\right]}.} \tag{18}$$

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:

$$\tilde{\phi}_i = \frac{\mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{I \geq K\,a\}} \sum_k \frac{K}{I^2}\, I_k\, \frac{\partial I_i}{\partial q_i}\right]}{\mathbb{E}^{\mathbb{Q}}\left[K\,\mathbf{1}_{\{I \geq K\,a\}}\right]} = \frac{\mathbb{E}^{\mathbb{Q}}\left[K\frac{\frac{\partial I_i}{\partial q_i}}{I}\middle|\, I \geq K\,a\right]}{\mathbb{E}^{\mathbb{Q}}\left[K \middle|\, I \geq K\,a\right]}, \tag{19}$$

so that:

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

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.[11]

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_i^t denote the loss incurred by consumer i, $i \in \{1, 2, \ldots, N\}$, in period t, $t \in \{1, 2, \ldots\}$. We assume that $L_i^t,\ t > 0$—for fixed i—are independent and identically distributed, and we define the relevant filtration $\mathbf{F} = (\mathcal{F}_t)_{t \geq 0}$ via $\mathcal{F}_t = \sigma(L_i^s,\ i \in \{1, 2, \ldots, N\},\ s \leq t)$. The firm determines the optimal level of assets, a_t, in the beginning of each period (i.e. (a_t) is $\mathbf{F}$-predictable) for a period cost of $\tau \times a_t$. Similarly to before, the company chooses $\mathbf{F}$-predictable amounts q_t^i in $I_i^t = I(L_i^t, q_i^t) = q_i^t \times L_i^t$ and prices p_i^t at the beginning of the period, and we denote the total claims by $I^t = \sum_{j=1}^N I_j^t$.

Now the company defaults if $I^t > a_t$, so that the recovery paid to each consumer is $R_i^t = \min\{I_i^t, \frac{a_t}{I^t}\, I_i^t\}$ and the company shuts down in case of default, i.e. shareholders do not have access to future profit flows.[12] The consumer's utility in period t is given by:

$$v_i^t(a_t, \mathrm{w}_i^t - p_i^t, q_1^t, \ldots, q_N^t) = \mathbb{E}_{t-1}\left[U_i(\mathrm{w}_i^t - p_i^t - L_i^t + R_i^t)\right],$$

where for simplicity we assume that wealth is homogeneous across periods, i.e. $\mathrm{w}_i^t \equiv \mathrm{w}_i$.[13]

The company solves:

$$\max_{\{a_t\},\{q_i^t\},\{p_i^t\}} V_0 = \sum_{t=1}^{\infty} \mathbb{E}\left[\mathbf{1}_{\{I^1 \leq a_1, \ldots, I^{t-1} \leq a_{t-1}\}} \times \left(\sum_i p_i^t - \sum_i \mathbb{E}_{t-1}\left[R_i^t\right] - \tau\, a_t\right)\right] \tag{20}$$

[11] 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."

[12] 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.

[13] 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).

with constraints:

$$v_i^t(a_t, w_i^t - p_i^t, q_1^t, \dots, q_N^t) \;\geq\; \gamma_i \;\forall i,\; \forall t, \tag{21}$$

$$s(q_1^t, \dots, q_N^t) \;\leq\; a_t \;\forall t. \tag{22}$$

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

$$(a_t, \{q_i^t\}, \{p_i^t\}) \equiv (a^*, \{q_i^*\}, \{p_i^*\})$$

that solves the Bellman equation:

$$V = \max_{a,\{q_i\},\{p_i\}} \sum_i p_i - \sum_i \underbrace{\mathbb{E}[R_i^t]}_{e_i} - \tau\, a + \mathbb{P}[I^t \leq a] \times V \tag{23}$$

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

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

$$\boxed{\begin{aligned} \frac{\partial p_i^*}{\partial q_i} \;=\; & \mathbb{E}\left[\frac{\partial I_i(L_i^t, q_i)}{\partial q_i}\, 1_{\{I^t \leq a\}}\right] + [V\, f_I(a)]\; \tilde{\theta}_i + \left[\mathbb{P}\left(I^t > a\right) + \tau - \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'} - V\, f_I(a)\right]\; \frac{\partial s}{\partial q_i} \\ & + \left[\sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}\right] \times a \times \tilde{\phi}_i, \end{aligned}} \tag{24}$$

where

$$\tilde{\theta}_i = \mathbb{E}\left[\left.\frac{\partial I_i(L_i^t, q_i)}{\partial q_i}\right| I^t = a\right] \text{ and } \tilde{\phi}_i = \frac{\mathbb{E}\left[\left(1_{\{I^t > a\}} \sum_k \frac{U_k'}{v_k'} \frac{I_k^t}{I^t}\right) \frac{\frac{\partial I_i^t}{\partial q_i}}{I^t}\right]}{\mathbb{E}\left[1_{\{I^t > a} \sum_k \frac{U_k'}{v_k'} \frac{I_k^t}{I^t}\right]}.$$

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

$$\begin{aligned} \left[\mathbb{P}(I^t \geq a)\, a + \tau\, a\right] \;=\; & \sum_i \tilde{\theta}_i\, q_i \times [V\, f_I(a)] + \sum_i \frac{\partial s}{\partial q_i}\, q_i \times \left[\mathbb{P}(I^t \geq a) + \tau - \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'} - V\, f_I(a)\right] \\ & + \sum_i \tilde{\phi}\, q_i\, a \times \left[\sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}\right]. \end{aligned}$$

In addition to the "external" ($\frac{\partial s}{\partial q_i}$) and "internal counterparty" ($\tilde{\phi}_i$) 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

$$\frac{\partial p_i^*}{\partial q_i} = \frac{\partial e_i^Z}{\partial q_i} + \frac{\partial s}{\partial q_i} \times \left[\mathbb{P}\left(I^t \geq a\right) + \tau - \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}\right] + \tilde{\phi}_i \times a \times \left[\sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}\right], \tag{25}$$

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.

Imperfect 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:

$$\begin{aligned}\frac{\partial p_i^*}{\partial q_i} &= \mathbb{E}\left[\frac{\partial I_i(L_i^t, q_i)}{\partial q_i} 1_{\{I^t \leq a\}}\right] + V\, f_I(a)\, \mathbb{E}\left[\frac{\partial I_i(L_i^t, q_i)}{\partial q_i}\middle| I^t = a\right] \\ &= \mathbb{E}\left[\frac{\partial I_i(L_i^t, q_i)}{\partial q_i} 1_{\{I^t \leq a\}}\right] + \left[\mathbb{P}\left(I^t > a\right) + \tau\right] \times \mathbb{E}\left[\frac{\partial I_i(L_i^t, q_i)}{\partial q_i}\middle| I^t = a\right], \end{aligned} \tag{26}$$

where the latter equality follows from the first order condition for a in the absence of constraints (see Eq. (54) in Appendix B.3).

General Case: Imperfect Competition, Uninsured Consumers, and Binding Regulation

Here we obtain (24), and we 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:

$$\mathbb{P}(I^t \geq a) + \tau \approx \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'} + V\, f_I(a),$$

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 popularity of risk measures for the purpose of calculating the marginal cost of risk and allocating capital can be attributed in part to their ease of application. However, there has been little scientific analysis of the question of how to choose said risk measures. Conventional thinking points to various mathematical properties (e.g., coherence or convexity), yet it is not clear whether choices made on such a basis yield economically desirable outcomes (see Gründl and Schmeiser (2007)).

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

3.1 The Gradient Allocation Principle

The gradient method follows from the maximization of profits subject to a risk measure constraint. To illustrate, assume a company's profit function Π depends on the indemnification parameters q_i, $1 \leq i \leq N$, and on capital a. Then maximizing profits subject a *risk measure constraint*:

$$\begin{cases} \max_{a,\{q_i\}} \Pi(q_1, \ldots, q_N, a) \\ \rho\,(q_1, q_2, \ldots, q_N)\ \chi_\rho \leq a \end{cases} \tag{27}$$

immediately yields:

$$\frac{\partial \Pi}{\partial q_i} = \left(-\frac{\partial \Pi}{\partial a}\right) \chi_\rho \, \frac{\partial \rho}{\partial q_i} \tag{28}$$

at the optimum. Here ρ is a differentiable risk measure evaluated at the aggregate claims $\sum_{j=1}^{N} I_j$ and χ_ρ is an exchange rate that converts risk to capital (which is set to 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 $-\frac{\partial \Pi}{\partial a}$.

This motivates the interpretation of the marginal capital weighted by the corresponding volume $\chi_\rho \, \frac{\partial \rho}{\partial q_i} \, q_i$ as the amount of capital allocated to exposure i (see e.g. Tasche (2004) or the section "Economic Justification of the Euler Principle" in McNeil, Frey, and Embrechts (2005)). In particular, for *homogeneous* risks and a *homogeneous* risk measure, the allocations to the respective risks "add up" to the entire capital:

$$\sum_{j=1}^{N} \chi_\rho \, \frac{\partial \rho}{\partial q_j} \, q_j = a. \tag{29}$$

We refer to Denault (2001), Kalkbrener (2005), and Myers and Read (2001) for alternative derivations of the gradient allocation principle (29) based on cooperative game theory, formal axioms, or a contingent claim approach, respectively. Regardless of its provenance, its 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 Economic Allocation Based on Risk Measures

In Section 2, we derive the marginal cost of risk and corresponding capital allocations in the context of an economic model of the firm within different settings. We now proceed to align these economic allocation results with risk measures. As before, it is useful to consider specific institutional circumstances, where only certain parts of our general allocation rule prevail.

Full Coverage by Deposit Insurance and Binding Regulation in the One-Period Model[14]

As already indicated in Section 2.1, in this case the only relevant constraint on risk-taking derives from the external risk measure s, and therefore the marginal cost of risk—and the attendant capital allocation—is indeed completely determined by its gradient.

This can also be derived in the formal setting of the gradient allocation principle introduced in 3.1. In the context of our model, the supporting optimization problem (27) takes the form:

$$\begin{cases} \max_{a,\{q_i\}} \Pi^* \left(q_1, \ldots, q_N, a\right) = \sum_k p_k^*(q_k) - \sum_k e_k(q_1, \ldots, q_N, a) - \tau\, a, \\ s(q_1, \ldots, q_N) \leq a, \end{cases} \tag{30}$$

and the solution (28) takes the form:

$$\left[\frac{\partial p_i^*}{\partial q_i} - \underbrace{\sum_k \frac{\partial e_k}{\partial q_i}}_{=\frac{\partial e_i^Z}{\partial q_i}}\right] = \left[\tau + \underbrace{\sum_k \frac{\partial e_k}{\partial a}}_{=\mathbb{P}(I \geq a)}\right] \frac{\partial s}{\partial q_i},$$

i.e. exactly (14), the relationship derived in Section 2. As before, capital is fully allocated due to the assumed properties on s :

$$\sum_{j=1}^{N} \frac{\partial s}{\partial q_j} q_j = a. \tag{31}$$

However, it is important to note that the optimization problem (30) is *not* equivalent to the firm's formal maximization problem (6)-(8), particularly since the derivation of the optimal premium functions embedded in (30) requires a solution of the original problem. Rather, the optimization problem in the context of the gradient principle (27)/(30) is a simpler auxiliary problem that yields the same marginal cost at the optimum—and thus can be used for capital allocation.

Full Coverage by Deposit Insurance and Non-Binding Regulation

In this case, risk-taking is only hindered by shareholder concerns about profit flows in future periods. Consider the auxiliary optimization problem (27) with our profit function Π^* and a Value-at-Risk constraint with confidence level $\alpha^* = \mathbb{P}(I \leq a)$. Again, it turns out that the solution to this auxiliary problem (28) lines up precisely with the pricing condition from our economic model (cf.

[14]Or full deposit insurance, binding regulation, and (close to) perfect competition in the multi-period context.

Eq. (26)):[15]

$$\begin{aligned} \frac{\partial p_i^*}{\partial q_i} - \frac{\partial e_i^Z}{\partial q_i} &= \big[\tau + \mathbb{P}\left(I^t > a\right)\big] \times \mathbb{E}\left[\frac{\partial I_i(L_i^t, q_i)}{\partial q_i}\middle| I^t = a\right] \\ \Leftrightarrow \frac{\partial \Pi^*}{\partial q_i} &= \underbrace{\left[\tau + \sum_k \frac{\partial e_k}{\partial a}\right]}_{=-\frac{\partial \Pi^*}{\partial a}} \times \frac{\partial}{\partial q_i}\mathrm{VaR}_{\alpha^*}(I). \end{aligned}$$

Hence, in this case, the marginal cost of risk and resulting capital allocations can also be calculated based on a risk measure, VaR, which thus arises endogenously within our framework. Again, it is worth noting that (27) is *not* equivalent to the original maximization problem (6)-(8) but it is a simplified optimization problem that requires the (fixed) inputs p^* and α^*. However, based on optimal choices of these inputs, again this simplification yields an easy-to-implement prescription to calculating the marginal cost of risk and for allocating capital: Just take the derivative of VaR at the current portfolio position in the direction of the exposure.

Finally, we note that although a VaR-based allocation only emerges in these 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 frequently exceed the required level (see e.g. Hanif et al. (2010)).

Uninsured Consumers and Non-Binding Regulation in the One-Period Model[16]

In this case, capitalization becomes material to consumers—which mathematically corresponds to the $\tilde{\phi}_i$'s emerging in the marginal cost allocations, where:

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

To align the allocation with the gradient method, we introduce the probability measure $\mathbb{P}^*$ on $(\Omega, \mathcal{F})$ via its Radon-Nikodym derivative:

$$\frac{\partial \mathbb{P}^*}{\partial \mathbb{P}} = \frac{\sum_k \frac{U_k'}{v_k'} \frac{I_k}{I} \mathbf{1}_{\{I \geq a\}}}{\mathbb{E}\left[\sum_k \frac{U_k'}{v_k'} \frac{I_k}{I} \mathbf{1}_{\{I \geq a\}}\right]}, \tag{32}$$

where I, U, etc. are evaluated at the optimum. Note that $\mathbb{P}^*$ is absolutely continuous with respect to the original probability measure—i.e. only events that have positive probability under $\mathbb{P}$ have positive probability under $\mathbb{P}^*$—but the measures are not equivalent since under $\mathbb{P}^*$ all the probability mass is concentrated in default states. On the set of strictly ($\mathbb{P}^*$-almost surely) positive

[15] See e.g. Gourieroux, Laurent, and Scaillet (2000) for a formal derivation of the gradient of Value-at-Risk.

[16] Or uninsured consumers, non-binding regulation, and (close to) perfect competition in the multi-period context.

square-integrable random variables $L^2_+(\Omega, \mathcal{F}, \mathbb{P}^*)$, that is on the set of square-integrable random variables that are positive in default states, we define the risk measure:

$$\boxed{\tilde{\rho}(X) = \exp\left\{\mathbb{E}^{\mathbb{P}^*}\left[\log\{X\}\right]\right\}.} \tag{33}$$

Moreover, define

$$\chi^*_\rho = \frac{a}{\tilde{\rho}\left(\sum_{j=1}^N I_j(L_j, q_j)\right)},$$

as the "exchange rate" between units of risk and capital, again evaluated at the optimum. Then the marginal cost of risk and attendant capital allocations in this case can also be calculated based on a risk measure—namely $\tilde{\rho}$.

More precisely, consider the gradient allocation problem (27) with our profit function Π^* and the risk measure constraint based on $\tilde{\rho}$ and χ^*_ρ. Then the solution (28) reads:

$$\begin{aligned}\left[\frac{\partial p_i^*}{\partial q_i} - \frac{\partial e_i^Z}{\partial q_i}\right] &= \left[\tau + \mathbb{P}(I \geq a)\right] \chi^*_\rho \, \frac{\partial \tilde{\rho}(I)}{\partial q_i} \\ &= \left[\tau + \mathbb{P}(I \geq a)\right] a \, \underbrace{\mathbb{E}^{\mathbb{P}^*}\left[\frac{\partial I_i / \partial q_i}{I}\right]}_{=\tilde{\phi}_i},\end{aligned} \tag{34}$$

i.e. exactly Equation (15), the relationship derived in Section 2.[17] Akin to p^* and α^* in the previous cases, $\mathbb{P}^*$ and χ^*_ρ are (fixed) input parameters in the simple auxiliary problem delivering the correct marginal cost. In particular, by inserting χ^*_ρ, we find that the attendant capital allocation to consumer i indeed takes the familiar form of a so-called *proportional allocation* (see e.g. Schmock and Straumann (1999) or Dhaene et al. (2009)) based on the homogeneous risk measure $\tilde{\rho}$:

$$a \times \frac{q_i \frac{\partial \tilde{\rho}(I)}{\partial q_i}}{\tilde{\rho}(I)} = a \times \frac{q_i \frac{\partial \tilde{\rho}(I)}{\partial q_i}}{\sum_{j=1}^N q_j \, \frac{\partial \tilde{\rho}(I)}{\partial q_i}}.$$

Before we continue with the consideration of more general institutional circumstances, we note that $\tilde{\rho}$ does *not* satisfy the mathematical properties that are typically imposed when selecting risk measures such as *coherence* (Artzner et al., 1999) or *convexity* (Föllmer and Schied, 2002). While $\tilde{\rho}$ obviously is monotonic ($\tilde{\rho}(X) \leq \tilde{\rho}(Y)$ for $X \leq Y$ a.s.), positively homogeneous ($\tilde{\rho}(a\,X) = a\,\tilde{\rho}(X)$, $a > 0$), and satisfies the constancy condition $\tilde{\rho}(c) = c$ for $c > 0$, it is neither translation-invariant nor sub-additive (see e.g. Frittelli and Gianin (2002) for a discussion of properties of risk measures). While these latter properties are viewed by many as desirable, the important point is that they do not follow from the economic context of the problem.

General Case: Imperfect Competition, Uninsured Consumers, and Binding Regulation

General institutional circumstances lead to a risk measure that combines the three foregoing risk measures. Specifically, with (24) we have the following expression for the marginal cost for con-

[17] In a setting with securities markets (see Eq. (18)), the risk measure still takes the same form (33) but state prices and allocations to security market states enter the measure transform (32).

sumer i at the optimum:

$$
\begin{aligned}
\left[\frac{\partial p_i^*}{\partial q_i} - \frac{\partial e_i^Z}{\partial q_i}\right] &= \left[V\, f_I(a)\right]\frac{\partial \mathrm{VaR}_{\alpha^*}}{\partial q_i} + \left[\sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}\right]\tilde{\chi}_\rho^*\,\frac{\partial \tilde{\rho}}{\partial q_i} \\
&\quad + \left[\mathbb{P}(I \geq a) + \tau - \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'} - V\, f_I(a)\right]\frac{\partial s}{\partial q_i} \qquad (35) \\
&= \left[\mathbb{P}(I \geq a) + \tau\right] \times \frac{\partial}{\partial q_i}\left((1-\zeta_1^* - \zeta_2^*)\ s + \zeta_1^*\,\tilde{\chi}_\rho^*\,\tilde{\rho} + \zeta_2^*\,\mathrm{VaR}_{\alpha^*}\right)
\end{aligned}
$$

where:

$$
\zeta_1^* = \frac{\sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}}{\mathbb{P}(I \geq a) + \tau} \qquad \text{and} \qquad \zeta_2^* = \frac{V\, f_I(a)}{\mathbb{P}(I \geq a) + \tau},
$$

so that we have the following expression for the resulting capital allocation:

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

Hence, the gradient principle still applies and we can envision the supporting problem (27) with a risk measure constraint based on the weighted average of the external risk measure s, the internal risk measure $\tilde{\rho}$, and VaR. Just as the risk measure parameters α^*, $\mathbb{P}^*$, and χ_ρ^*, the *exogenous* weights ζ_1^* and ζ_2^* follow from the formal maximization problem (6)-(8), and they depend on regulators', consumers', and shareholders' preferences for capitalization of the company at the margin.

Alternatively, it is also possible to think of the auxiliary problem (27) subject to three different risk measure constraints: one based on s, one based on $\chi_\rho^*\,\tilde{\rho}$, and one based on VaR_{α^*}. The solution then satisfies:

$$
\begin{aligned}
\frac{\partial \Pi^*}{\partial q_i} &= \left[-\frac{\partial \Pi^*}{\partial a} - (\beta_1 + \beta_2)\right]\frac{\partial s}{\partial q_i} + \left[\beta_1\right]\chi_\rho^*\,\frac{\partial \tilde{\rho}}{\partial q_i} + \left[\beta_2\right]\frac{\partial \mathrm{VaR}_{\alpha^*}}{\partial q_i} \\
\Leftrightarrow \frac{\partial p_i^*}{\partial q_i} - \frac{\partial e_i^Z}{\partial q_i} &= \left[\mathbb{P}(I \geq a) + \tau - (\beta_1 + \beta_2)\right]\frac{\partial s}{\partial q_i} + \left[\beta_1\right]\chi_\rho^*\,\frac{\partial \tilde{\rho}}{\partial q_i} + \left[\beta_2\right]\frac{\partial \mathrm{VaR}_{\alpha^*}}{\partial q_i},
\end{aligned}
$$

where β_1 and β_2 are the Lagrange multipliers from the constraints associated with $\chi_\rho^*\,\tilde{\rho}$ and VaR_{α^*}, respectively. Comparing the solution to (35), it is clear that we have:

$$
\beta_1 = \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'} \qquad \text{and} \qquad \beta_2 = V\, f_I(a).
$$

Therefore, when considering multiple constraints, we still obtain the correct marginal cost and now the weights on the drivers of firm capitalization—corresponding to regulatory, shareholder, and counterparty concerns—also follow from the auxiliary problem.

The foregoing shows that it is indeed *possible* to derive correct marginal costs and resulting capital allocations based on risk measures. Casting the "right" inputs, such as the confidence level α^* or the weighting function $\mathbb{P}^*$, is of course complex (and formally requires a solution of the "full" optimization problem of the firm). On the other hand, determining inputs such as confidence levels and *spectral* weighting functions for risk measures (see Acerbi (2002)) is already a problem for practitioners. Economic analysis of the profit maximization problem yields 1) the correct *form* of the risk measure for allocating capital consistent with optimal pricing and performance measurement; and 2) the input parameters in terms of fundamental economic quantities that can be estimated. In our particular setting, the correct form is a weighted average of three risk measures, at least two of which (VaR and $\tilde{\rho}$) generally do not adhere to the axioms of coherence or convexity.

4 Comparison of Capital Allocation Methods

This section compares approaches for calculating the marginal cost of risk and allocating capital. Since there are many papers that consider VaR-based allocations and their relationship to other methods (see, among many others, Garman (1997) or Kalkbrener (2005)), we focus on the counterparty-driven internal allocation corresponding to the *new* risk measure $\tilde{\rho}$. We analyze how this 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 gradient 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 no 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 $\alpha < \nu$, and that their participation constraint is given by the autarky level:

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

Then, the firm's optimization problem in the one period model (6)/(7) may be represented as:

$$\begin{cases} \max_{a,q,p}\Big\{N \times p - N \times q \times \left[\frac{1}{\nu}\Gamma_{N-1,\nu}\left(\frac{a}{q}\right) - \frac{\nu^{N-1}}{(N-1)!}e^{-\nu\frac{a}{q}}\left(\frac{a}{q}\right)^{N-1}\left(\frac{1}{\nu} + \frac{1}{N}\frac{a}{q}\right)\right] \\ \qquad\qquad -a \times \bar{\Gamma}_{N,\nu}\left(\frac{a}{q}\right) - \tau \times a\Big\} \\ \text{subject to} \\ \gamma \leq e^{-\alpha(\mathrm{w}-p)}\Big\{\frac{\nu}{\nu-(1-q)\alpha}\left[\Gamma_{N-1,\nu}\left(\frac{a}{q}\right) - \frac{e^{-\frac{a}{q}(\nu-(1-q)\alpha)}\nu^{N-1}}{((1-q)\alpha)^{N-1}}\,\Gamma_{N-1,(1-q)\alpha}\left(\frac{a}{q}\right)\right] \\ \qquad + \sum_{k=0}^{\infty}\left(\frac{\alpha}{\nu}\right)^k\frac{(N-1)!}{(N-1+k)!}e^{-\nu\,a}\sum_{j=0}^{N-1}\frac{(a\,\nu)^j}{j!}\frac{(N-j+k-1)!}{(N-j-1)!}\bar{\Gamma}_{N-j+k,\nu}\left(a\left(\frac{1-q}{q}\right)\right)\Big\}, \end{cases} \tag{36}$$

where $\bar{\Gamma}_{m,b}(x) = 1 - \Gamma_{m,b}(x)$ and $\Gamma_{m,b}(\cdot)$ denotes the cumulative distribution function of the Gamma distribution with parameters m and b (see Appendix C.1 for the derivation of (36)).

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 ES with confidence level $\alpha^* = \mathbb{P}(q\,L \geq a)$ according to the gradient principle, it is well known that (see e.g. McNeil, Frey, and Embrechts (2005)):

$$\frac{1}{N} = \frac{q_i\,\frac{\partial \mathrm{ES}_{\alpha^*}(I)}{\partial q_i}}{\mathrm{ES}_{\alpha^*}(I)} \quad = \quad \frac{q\,\mathbb{E}\left[L_i|\,q\,L \geq a\right]}{\mathbb{E}\left[I|\,I \geq a\right]} = \frac{\mathbb{E}\left[\mathbb{E}[L_i|L]|\,q\,L \geq a\right]}{\mathbb{E}\left[L|\,q\,L \geq a\right]} = \frac{1}{N}\,\mathbb{E}\left[\text{const} \times L|\,q\,L \geq a\right],$$

where $L = \sum_j L_j$, i.e. 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:

$$\begin{aligned} \frac{1}{N} \quad &\overset{\text{Eq. (11)}}{=} \quad \frac{\mathbb{E}\left[\left[\mathbf{1}_{\{qL\geq a\}} \sum_{j=1}^{N} U'\left(\mathrm{w} - p - L_j + a\frac{L_j}{L}\right)\frac{L_j}{L}\right]\frac{L_i}{L}\right]}{\mathbb{E}\left[\mathbf{1}_{\{qL\geq a\}} \sum_{j=1}^{N} U'\left(\mathrm{w} - p - L_j + a\frac{L_j}{L}\right)\frac{L_j}{L}\right]} \\ &= \quad \mathbb{E}^{\mathbb{P}^*}\left[\frac{L_i}{L}\right] = \frac{1}{N}\mathbb{E}\Bigg[\underbrace{\mathbb{E}\left[\frac{\partial \mathbb{P}^*}{\partial \mathbb{P}}\middle|\, L\right]}_{\psi^*(L)}\Bigg], \end{aligned} \tag{37}$$

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

$$\boxed{\psi^*(l) = \mathbf{1}_{\{ql\geq a\}}\,\hat{c}_{N,\nu,\alpha,a,q}\,\sum_{k=0}^{\infty}\frac{(k+1)\,(\alpha(l-a))^k}{(N+k)!},} \tag{38}$$

where $\tilde{c}_{N,\nu,\alpha,a,q}$ is a constant ensuring that $\mathbb{E}\left[\psi^*(L)\right] = 1$.[18]

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

$$\boxed{\tilde{\rho}(I) = \tilde{\rho}(q\,L) = \exp\left\{\mathbb{E}\left[\psi^*(L)\,\log\{qL\}\right]\right\} = \exp\left\{\mathbb{E}\left[\hat{\psi}^*(L)\,\log\{qL\}\middle|\,qL \geq a\right]\right\},} \tag{39}$$

with $\hat{\psi}^*(l) = \mathbb{P}(qL \geq a) \times \psi^*(l)$. Thus, $\tilde{\rho}$ in this sense is in fact a *tail risk measure*, and hence is related to ES. The weights $\hat{\psi}^*(\cdot)$ perform a role similar to the *risk spectrum* within the so-called *spectral risk measures* 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 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 gradient allocation method, tend to penalize bad outcomes *less* heavily than ES.

[18] The derivation of (38), a closed-form solution for $\hat{c}_{N,\nu,\alpha,a,q}$, and a representation of $\hat{\psi}^*(\cdot)$ not involving an infinite sum for implementation purposes are provided in Appendix C.2.

In fact, it is evident from (37) 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 (38)) and, thus, the value placed on the recoveries 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 in terms of its effect on recoveries 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 in 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.[19] However, for smaller—and more probable—loss realizations, different shapes are possible rendering either ES-based allocation or the counterparty-driven allocation to be "more conservative."

Exponential Losses: Parametrizations

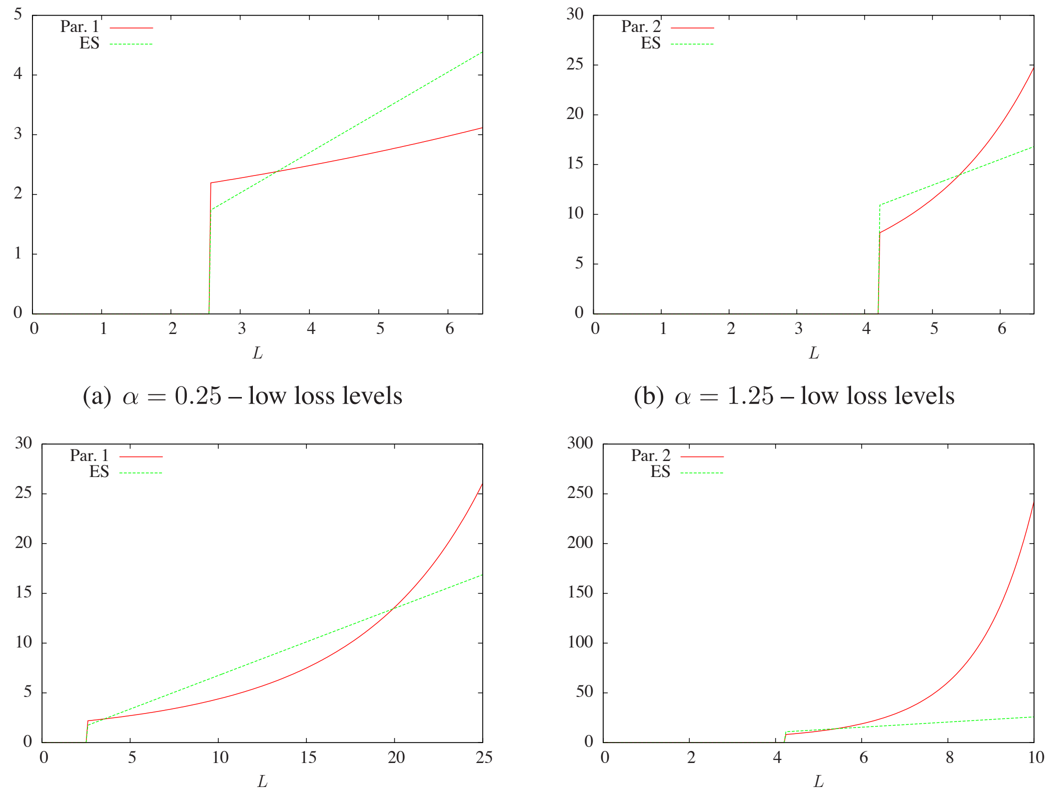
Nr.	N	ν	τ	α	w	a	p	q
1	5	2.0	0.050	0.25	3.0	1.4663	0.2598	0.5713
2	5	2.0	0.050	1.25	3.0	4.0036	0.7401	0.9494

Table 1: Parametrizations of the Exponential Losses model.

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 (36) with different risk aversion levels. The properties are as expected: a, p, and q all are increasing in risk aversion. Figure 1 plots the corresponding weighting function $\hat{\psi}^*$ against the linear weighting function associated with Expected Shortfall. We find two qualitatively different shapes:[20] 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 Expected Shortfall, whereas the weighting is higher for the counterparty driven allocation in high loss states. Thus, the counterparty-driven allocation can be deemed more conservative than 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.

[19] It is worth noting that $\hat{\psi}^*(\cdot)$ is not necessarily convex for alternative preference specifications; indeed, it is possible to construct examples where the shape is strictly concave or even linear.

[20] Analyses with respect to other parameters such as company size N or the expected loss $1/\nu$ show similar results.



(a) $\alpha = 0.25$ – low loss levels

(b) $\alpha = 1.25$ – low loss levels

(c) $\alpha = 0.25$ – high loss levels

(d) $\alpha = 1.25$ – high loss levels

Figure 1: Weighting function $\hat{\psi}^*$ for varying risk aversion parameter α.

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 (39) 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 π_i, $i = 1, \ldots, m$.[21] The participation constraints again are given by their autarky levels:

$$\gamma_i = \mathbb{E}\left[U_i(\mathrm{w}_i - L_i)\right] = \pi_i\, U_i(\mathrm{w}_i - l_i) + (1 - \pi_i)\, U_i(\mathrm{w}_i).$$

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

[21] While a discrete loss distribution does not formally fit our setup from Section 2, as indicated in Footnote 8, generalizations are possible at the expense of conciseness in the presentation. We accept this slight inconsistency due to the tractability of the Bernoulli setup.

independent Binomial(N_i, π_i) distributions.

For the counterparty-based allocation, we obtain for each consumer in group i:

$$\begin{aligned} q_i\,\tilde{\phi}_i &= \tilde{c} \sum_{k_1=0}^{N_1} \ldots \sum_{k_i=1}^{N_i} \ldots \sum_{k_m=0}^{N_m} \binom{N_1}{k_1} \ldots \binom{N_i-1}{k_i-1} \ldots \binom{N_m}{k_m} \\ &\quad \times \pi_1^{k_1} \ldots \pi_i^{k_i} \ldots \pi_m^{k_m} (1-\pi_1)^{N_1-k_1} \ldots (1-\pi_i)^{N_i-k_i} \ldots (1-\pi_m)^{N_m-k_m} \\ &\quad \times \underbrace{\left[\mathbf{1}_{\{k_1 q_1 l_1 + \ldots + k_m q_m l_m \geq a\}} \left\{ \sum_{j=1}^{m} \frac{k_j\, U'\left(\mathrm{w}_j - p_j - l_j + q_j\, l_j\, \frac{a}{k_1 q_1 l_1 + \ldots + k_m q_m l_m}\right) q_j\, l_j}{v_j' \times (k_1 q_1 l_1 + \ldots + k_m q_m l_m)} \right\} \right]}_{=\text{const} \times \frac{\partial \mathbb{P}^*}{\partial \mathbb{P}}} \\ &\quad \times \frac{q_i\, l_i}{k_1 q_1 l_1 + \ldots + k_m q_m l_m}, \end{aligned} \tag{40}$$

where $\tilde{c}$ is a constant such that $\sum_i N_i\, q_i\, \tilde{\phi}_i = 1$. Thus, while the analytical form of the weights $\psi^*(I) = \mathbb{E}\left[\frac{\partial \mathbb{P}^*}{\partial \mathbb{P}} \middle| I\right]$ is less transparent in this case, again we notice that they immediately depend on the marginal utilities of recoveries in various states of default. For the allocation based on Expected Shortfall with confidence level α^*, on the other hand, we obtain for each consumer in group i:

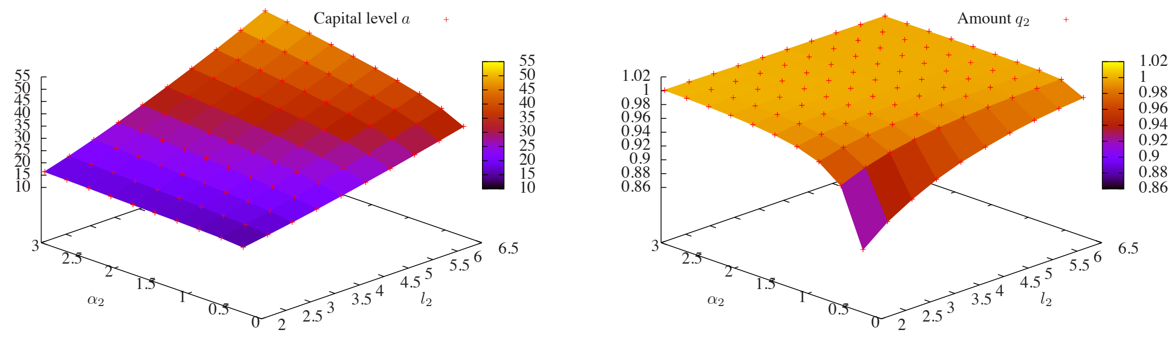
$$\begin{aligned} \frac{q_i\, \frac{\partial \mathrm{ES}_{\alpha^*}(I)}{\partial q_i}}{\mathrm{ES}_{\alpha^*}(I)} &= \text{const} \times \sum_{k_1=0}^{N_1} \ldots \sum_{k_i=1}^{N_i} \ldots \sum_{k_m=0}^{N_m} \binom{N_1}{k_1} \ldots \binom{N_i-1}{k_i-1} \ldots \binom{N_m}{k_m} \\ &\quad \times \pi_1^{k_1} \ldots \pi_i^{k_i} \ldots \pi_m^{k_m} (1-\pi_1)^{N_1-k_1} \ldots (1-\pi_i)^{N_i-k_i} \ldots (1-\pi_m)^{N_m-k_m} \\ &\quad \mathbf{1}_{\{k_1 q_1 l_1 + \ldots + k_m q_m l_m \geq a\}}\, q_i\, l_i, \end{aligned} \tag{41}$$

i.e. it is of a similar form as (40) but now 1) does not contain the adjustment based on the marginal utilities $\frac{\partial \mathbb{P}^*}{\partial \mathbb{P}}$, and 2) the state-specific loss for a consumer in group i, $(q_i\, l_i)$, is not scaled by the aggregate loss—which, in the counterparty-driven allocation, is a consequence of the proportional partitioning of the recoveries in states of default.

To assess the consequences of these adjustments, we consider the case of two different types of consumers with differing loss exposure and risk aversion levels. Specifically, we assume consumers in the first group have wealth level $\mathrm{w}_1 = 7$, 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 $\mathrm{w}_2 = 7$, CARA preferences, and face a loss with an occurrence probability of $\pi_2 = 10\%$, but we vary their risk aversion level α_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 α_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.[22]

[22]For the participation amount of consumers in group 1, q_1, we observe the opposite relationship, i.e. it is decreasing



(a) Capital levels a

(b) Amounts q_2

Figure 2: Optimal capitalization a and optimal coverage level q_2 for group 2 loss levels l_2 between between 2 and 6.5 and group 2 risk aversion levels α_2 between 0.25 and 2.95.

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 and the ES allocation rule as a function of l_2 and α_2. From the figure, the functional relationship appears to be non-smooth, which is due to genuine discontinuities emerging from the discreteness of the resulting aggregate loss distribution. To better illustrate the relationship, we therefore include a 5×5-polynomial fit as well as contour lines based on the smoothed surface.

To interpret the results, it is important to note 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.

The difference is negative (about -1%) for relatively low loss and risk aversion levels and positive (about 2%) for relatively large loss and risk aversion levels—resulting in an overall range of about 3% across all combinations. Hence, as in the previous example, we find that 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, ES-based allocations are not conservative enough—i.e. the counterparty-based principle allocates more to the high risk group than ES.

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.

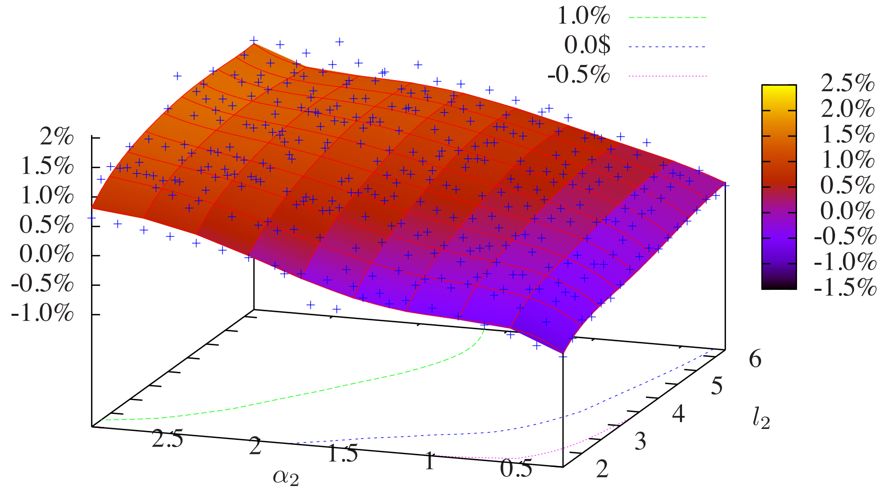


Figure 3: Difference in allocation of capital to group 2 between the counterparty-based allocation rule and the ES allocation rule for group 2 loss levels l_2 between between 2 and 6.5 and group 2 risk aversion levels α_2 between 0.25 and 2.95.

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 in an incomplete market setting with frictional capital costs. 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 of the measures, the confidence level in VaR, or 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 relying on risk measure-based solvency assessment to one relying on market (counterparty) discipline will not necessarily mitigate the oft-lamented failure of financial institutions to penalize "tail" risk.

Numerous extensions are possible. Our setting is but one possible specification of the profit maximization problem. Others could include informational frictions with regards to consumer endowments and preferences, explicit consideration of managerial incentives in a decentralized organization of the multi-line business, a more detailed modeling of the firm's capital raising and capital structure decision, or any number of other complications. All of these these would of course lead to different risk measures, underscoring the point that risk measures chosen for their technical properties generally fail to yield correct pricing and efficient allocation of capital.

One can also contemplate changes of perspective: The calculations in this paper are done from the perspective of a profit-maximizing firm, but one could also consider 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 their value 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. One such exercise is performed by Acharya et al. (2010) who study a particular environment which leads to a new risk measure, dubbed *systemic expected shortfall*.

In general, the particular specifications of the economic environment will lead to different risk measures. Going forward, the challenge for companies and regulators will be to choose risk measures consistent with their economic objectives and constraints. These issues are intriguing ones 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

$$\tilde{v}_i\left(\mathrm{w}_i - p_i^*(q_i), q_i; \tilde{a}, \tilde{q}_1, \ldots, \tilde{q}_N\right) = \mathbb{E}\left[U_i\left(\mathrm{w}_i - p_i^*(q_i) - L_i + \tilde{R}_i\right)\right], \tag{42}$$

where:

$$\tilde{R}_i = \tilde{R}_i\left(q_i; \tilde{a}, \tilde{q}_1, \ldots, \tilde{q}_N\right) = \min\left\{I_i(L_i, q_i), \frac{\tilde{a}}{\sum_{j=1}^N I_j\left(L_j, \tilde{q}_j\right)} I_i(L_i, q_i)\right\}.$$

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.[23]

The firm's objective function is identical to the previous one, except that 1) the firm now specifies a price function rather than a single price point, and 2) the firm fixes the recovery rates for purposes of consumer incentive compatibility by choosing $\tilde{a}$ and $\{\tilde{q}_i\}$ instead of the "true" levels of a and $\{q_i\}$:

$$\max_{\tilde{a}, \{p_i^*(\cdot)\}, \{\tilde{q}_i\}} \left\{\sum p_i^*(\tilde{q}_i) - \sum e_i - \tau\tilde{a}\right\} \tag{43}$$

The firm still faces the previous constraints (7) and (8),

$$\begin{aligned} v_i\left(\tilde{a}, \mathrm{w}_i - p_i^*(\tilde{q}_i), \tilde{q}_1, \ldots, \tilde{q}_N\right) &\geq \gamma_i, \\ s\left(\tilde{q}_1, \ldots, \tilde{q}_N\right) &\leq a, \end{aligned}$$

[23] Alternatively, we could also specify

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

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:

$$\tilde{q}_i \in \arg\max_{q_i} \tilde{v}_i(\mathrm{w}_i - p_i^*(q_i), q_i\,; \tilde{a}, \tilde{q}_1, ..., \tilde{q}_N), \ \ \forall i. \tag{44}$$

Equation (44) 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 $\hat{a}$, $\{\hat{q}_i\}$, 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:

$$\hat{q}_i \in \arg\max_{q_i} \tilde{v}_i(\mathrm{w}_i - p_i^*(q_i), q_i\,; \hat{a}, \hat{q}_1, ..., \hat{q}_N), \ \ \forall i.$$

The following lemma shows that this function exists.

Lemma A.1. *Suppose $\hat{a}$, $\{\hat{q}_i\}$, 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:

$$g(x) = \tilde{v}_i(\mathrm{w}_i - x, 0\,; \hat{a}, \hat{q}_1, ..., \hat{q}_N)$$

is monotonically decreasing and, hence, invertible, so that we may uniquely define:

$$p_i^*(0) = g^{-1}(\gamma_i), \tag{45}$$

which obviously satisfies:

$$\tilde{v}_i(\mathrm{w}_i - p_i^*(0), 0\,; \hat{a}, \hat{q}_1, ..., \hat{q}_N) = \gamma_i.$$

Furthermore, let $p_i^*(\cdot)$ be a solution to the differential equation (initial value problem):[24]

$$\frac{\partial p_i^*(x)}{\partial x} = \frac{\frac{\partial}{\partial x}\tilde{v}_i(\mathrm{w}_i - p_i^*(x), x\,; \hat{a}, \hat{q}_1, ..., \hat{q}_N)}{\frac{\partial}{\partial \mathrm{w}}\tilde{v}_i(\mathrm{w}_i - p_i^*(x), x\,; \hat{a}, \hat{q}_1, ..., \hat{q}_N)}, \ p_i^*(0) = g^{-1}(\gamma_i), \tag{46}$$

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 $\frac{\partial \tilde{v}_i}{\partial \mathrm{w}}$, $\frac{\partial \tilde{v}_i}{\partial q_i} > 0$, we know that the function is monotonically increasing.

[24]Here, $\frac{\partial \tilde{v}_i}{\partial \mathrm{w}}$ 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:

$$\begin{aligned}
& \tilde{v}_i(\mathrm{w}_i - p_i^*(q_i), q_i\,; \hat{a}, \hat{q}_1, ..., \hat{q}_N) \\
= \;& \gamma_i + \int_0^{q_i} \left[\frac{\partial}{\partial q_i} \tilde{v}_i(\mathrm{w}_i - p_i^*(x), x\,; \hat{a}, \hat{q}_1, ..., \hat{q}_N) - \frac{\partial}{\partial \mathrm{w}} \tilde{v}_i(\mathrm{w}_i - p_i^*(x), x\,; \hat{a}, \hat{q}_1, ..., \hat{q}_N) \times \frac{\partial p_i^*(x)}{\partial x} \right] dx \\
= \;& \gamma_i + 0, \; q_i > 0. \qquad (47)
\end{aligned}$$

In particular,

$$\tilde{v}_i(\mathrm{w}_i - p_i^*(\hat{q}_i), \hat{q}_i\,; \hat{a}, \hat{q}_1, ..., \hat{q}_N) = \gamma_i,$$

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

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

proving the first part of the lemma. Moreover, (47) directly implies that:

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

proving the second part. □

B Identities in Section 2

B.1 Derivation of Equation (10)

Let λ_k be the Lagrange multiplier associated with the participation constraint (7) for consumer k, and let ξ the multiplier associated with (8). Then, the first order conditions for an interior solution of Problem 6 are:[25]

$$[q_i] \qquad -\sum_k \frac{\partial e_k}{\partial q_i} + \sum_k \lambda_k \frac{\partial v_k}{\partial q_i} - \frac{\partial s}{\partial q_i}\xi = 0, \qquad (48)$$

$$[a] \qquad -\sum_k \frac{\partial e_k}{\partial a} - \tau + \sum_k \lambda_k \frac{\partial v_k}{\partial a} + \xi = 0, \qquad (49)$$

$$[p_i] \qquad 1 - \lambda_i \frac{\partial v_i}{\partial \mathrm{w}} = 0. \qquad (50)$$

Using (48) and (50), we see that (9) may be rewritten as:

$$\begin{aligned}
\frac{\partial p_i^*}{\partial q_i} &\overset{(50)}{=} \frac{1}{v_i'} q \left[\frac{\partial v_i}{\partial q_i} + \mathbb{E}\left[\mathbf{1}_{\{I \geq a\}}\, U_i' \,\frac{a}{I^2}\, I_i \,\frac{\partial I_i}{\partial q_i} \right] \right] \\
&\overset{(48)}{=} \sum_k \frac{\partial e_k}{\partial q_i} + \frac{\partial s}{\partial q_i}\,\xi - \sum_{k \neq i} \frac{\frac{\partial v_k}{\partial q_i}}{v_k'} + \frac{\mathbb{E}\left[\mathbf{1}_{\{I \geq a\}}\, U_i' \,\frac{a}{I^2}\, I_i \,\frac{\partial I_i}{\partial q_i} \right]}{v_i'}
\end{aligned}$$

[25] 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 (49):

$$\frac{\partial p_i^*}{\partial q_i} = \frac{\partial e_i^Z}{\partial q_i} + \frac{\partial s}{\partial q_i}\left[\sum_k \frac{\partial e_k}{\partial a} + \tau - \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}\right] + \mathbb{E}\left[\mathbf{1}_{\{I\geq a\}} \sum_k \frac{U_k'}{v_k'}\frac{a}{I^2} I_k \frac{\partial I_i}{\partial q_i}\right].$$

Moving on:

$$\frac{\partial p_i^*}{\partial q_i} = \frac{\partial e_i^Z}{\partial q_i} + \frac{\partial s}{\partial q_i}\left[\mathbb{P}(I \geq a) + \tau - \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}\right] + \frac{\mathbb{E}\left[\mathbf{1}_{\{I\geq a\}} \sum_k \frac{U_k'}{v_k'}\frac{1}{I^2} I_k \frac{\partial I_i}{\partial q_i}\right]}{\sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}} \times \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'} \times a$$

and, therefore, (10).

B.2 Derivation of Equation (18)

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

$$\max_{a,\{q_i\},\{p_i\},\{K_j\},\{\mathrm{w}_{ij}\}} \sum p_i - \sum e_i - \tau\, a,$$

subject to:

$$\begin{aligned} v_i &\geq \gamma_i, \\ s(q_1,\ldots,q_N) &\leq a, \\ \sum_j \pi_j\, K_j &= 1, \\ \sum_j \pi_j\, \mathrm{w}_{ij} &= \mathrm{w}_i. \end{aligned}$$

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

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

The first order condition for $\{\mathrm{w}_{ij}\}$ is:

$$\begin{aligned} [\mathrm{w}_{ij}] \qquad & \lambda_i \frac{\partial v_i}{\partial \mathrm{w}_{ij}} - \eta_i\, \pi_j = 0 \qquad (51) \\ & \Leftrightarrow \lambda_i\, p_j^{(S)}\, \mathbb{E}^{\mathbb{P}}\left[U_i'(W_i - p_i - L_i + R_i)|\omega_j^{(S)}\right] - \eta_i\, \pi_j = 0, \end{aligned}$$

where $\{\eta_i\}$ are the Lagrange multipliers for the individual wealth constraints. Since

$$\begin{aligned} 0 &= \sum_j \left(\lambda_i \, p_j^{(S)} \, \mathbb{E}^{\mathbb{P}} \left[U_i'(W_i - p_i - L_i + R_i) | \omega_j^{(S)} \right] - \eta_i \, \pi_j \right) \\ &= \lambda_i \, \frac{\partial v_i}{\partial \mathrm{w}} - \eta_i, \end{aligned}$$

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

$$\frac{\pi_j}{p_j^{(S)}} = \frac{\mathbb{E}^{\mathbb{P}} \left[U_i'(W_i - p_i - L_i + R_i) | \omega_j^{(S)} \right]}{\frac{\partial v_i}{\partial \mathrm{w}}}. \tag{52}$$

As before, we seek a pricing function satisfying:

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

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

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

Simplifying, we obtain:

$$\begin{aligned} \frac{\partial p_i^*}{\partial q_i} &= \frac{\partial e_i^Z}{\partial q_i} + \frac{\partial s}{\partial q_i} \left[\sum_k \frac{\partial e_k}{\partial a} + \tau - \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'} \right] + \mathbb{E}^{\mathbb{P}} \left[\mathbf{1}_{\{I \geq K\, a\}} \sum_k \frac{U_k'}{v_k'} \, \frac{K\, a}{I^2} \, I_k \, \frac{\partial I_i}{\partial q_i} \right] \\ &= \frac{\partial e_i^Z}{\partial q_i} + \frac{\partial s}{\partial q_i} \left[\mathbb{E}^{\mathbb{Q}} \left[K \, \mathbf{1}_{\{I \geq K\, a\}} \right] + \tau - \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'} \right] + \frac{\mathbb{E}^{\mathbb{P}} \left[\mathbf{1}_{\{I \geq K\, a\}} \sum_k \frac{U_k'}{v_k'} \, \frac{K}{I^2} \, I_k \, \frac{\partial I_i}{\partial q_i} \right]}{\sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'}} \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'} \, a \\ &= \frac{\partial e_i^Z}{\partial q_i} + \frac{\partial s}{\partial q_i} \left[\mathbb{E}^{\mathbb{Q}} \left[K \, \mathbf{1}_{\{I \geq K\, a\}} \right] + \tau - \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'} \right] + \tilde{\phi}_i \times a \times \left[\sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'} \right], \end{aligned}$$

i.e. (17), where:

$$\tilde{\phi}_i = \frac{\mathbb{E}^{\mathbb{P}} \left[\mathbf{1}_{\{I \geq K\, a\}} \sum_k \frac{U_k'}{v_k'} \, \frac{K}{I^2} \, I_k \, \frac{\partial I_i}{\partial q_i} \right]}{\mathbb{E}^{\mathbb{P}} \left[\mathbf{1}_{\{I \geq K\, a\}} \sum_k \frac{U_k'}{v_k'} \, K \, \frac{I_k}{I} \right]} \stackrel{(52)}{=} \frac{\mathbb{E}^{\mathbb{Q}} \left[\mathbf{1}_{\{I \geq K\, a\}} \sum_k \frac{U_k'}{\mathbb{E}^{\mathbb{P}}[U_k' | \omega^{(S)}]} \, \frac{K}{I^2} \, I_k \, \frac{\partial I_i}{\partial q_i} \right]}{\mathbb{E}^{\mathbb{Q}} \left[\mathbf{1}_{\{I \geq K\, a\}} \sum_k \frac{U_k'}{\mathbb{E}^{\mathbb{P}}[U_k' | \omega^{(S)}]} \, K \, \frac{I_k}{I} \right]},$$

i.e. (18). In the limiting case of a complete market, clearly $U_i'(\mathrm{w}_{ij} - p_i - l_{ij} + r_{ij})$ 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:[26]

$$[q_i] \qquad -\sum_k \frac{\partial e_k}{\partial q_i} - V\mathbb{E}\left[\frac{\partial I_i(L_i,q_i)}{\partial q_i}\,\frac{\mathbf{1}_{\{I=a\}}}{dt}\right] + \sum_k \lambda_k \frac{\partial v_k}{\partial q_i} - \xi\frac{\partial s}{\partial q_i} = 0, \tag{53}$$

$$[a] \qquad -\sum_k \frac{\partial e_k}{\partial a} - \tau + \sum_k \lambda_k\,\frac{\partial v_k}{\partial a} + \xi + V\,f_I(a) = 0, \tag{54}$$

$$[p_i] \qquad 1 - \lambda_i \frac{\partial v_i}{\partial \mathrm{w}} = 0, \tag{55}$$

where—as before—λ_i and ξ 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:

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

Therefore:

$$\begin{aligned}
\frac{\partial p_i^*}{\partial q_i} &= \frac{1}{\frac{\partial v_i}{\partial \mathrm{w}}}\left[\frac{\partial v_i}{\partial q_i} + \mathbb{E}\left[1_{\{I>a\}}\,\frac{a}{(I)^2}\,I_i\,\frac{\partial I_i}{\partial q_i}\,U_i'\right]\right] \\
&= \sum_k \frac{\partial e_k}{\partial q_i} + \xi\,\frac{\partial s}{\partial q_i} + V\,\mathbb{E}\left[\frac{\partial I_i(L_i,q_i)}{\partial q_i}\,\frac{\mathbf{1}_{\{I=a\}}}{dt}\right] - \sum_{k\neq i}\frac{\frac{\partial v_k}{\partial q_i}}{v_k'} + \frac{\mathbb{E}\left[1_{\{I>a\}}\,\frac{a}{(I)^2}\,I_i\,\frac{\partial I_i}{\partial q_i}\,U_i'\right]}{v_i'} \\
&= \frac{\partial e_k^Z}{\partial q_i} + \left[\mathbb{P}(I\geq a) + \tau - \sum_k \frac{\frac{\partial v_k}{\partial a}}{v_k'} - V\,f_I(a)\right]\,\frac{\partial s}{\partial q_i} + V\,f_I(a)\,\mathbb{E}\left[\frac{\partial I_i}{\partial q_i}\middle|\,I=a\right] \\
&\quad + \mathbb{E}\left[1_{\{I>a\}}\,\sum_k \frac{U_k'}{v_k'}\frac{a}{(I)^2}\,I_k\,\frac{\partial I_i}{\partial q_i}\right]
\end{aligned}$$

and since:

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

we obtain (24).

[26] For simplicity, we omit the "t" super- and subscripts in case no ambiguity arises.

C Identities in Section 4

C.1 Derivation of Equation (36)

For consumer N, $L_N \sim Exp(\nu)$ 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

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

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

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

For part $i.$, we have

$$\begin{aligned}
& q\, \mathbb{E}\left[L_N\, \mathbf{1}_{\{q(L_{-N}+L_N)<a\}}\right] \\
= \quad & q \int_0^\infty \int_0^\infty \mathbf{1}_{\{i+l<a/q\}}\, l\, \nu\, e^{-\nu\, l}\, \frac{\nu^{N-1}}{(N-2)!}\, i^{N-2}\, e^{-\nu\, i}\, dl\, di \\
= \quad & q \frac{\nu^N}{(N-2)!} \int_0^{a/q} \int_0^{a/q-i} l e^{-\nu\, l}\, dl\; i^{N-2}\, e^{-\nu\, i}\, di \\
= \quad & q \frac{\nu^N}{(N-2)!} \int_0^{a/q} \left[\frac{1}{\nu^2} - \frac{1}{\nu}\left(\frac{a}{q}+\frac{1}{\nu}\right) e^{-\nu\, a/q}\, e^{\nu\, i} + i \frac{1}{\nu}\, e^{-\nu\, a/q}\, e^{\nu\, i}\right]\, i^{N-2}\, e^{-\nu\, i}\, di \\
= \quad & q \frac{\nu^{N-2}}{(N-2)!} \int_0^{a/q} i^{N-2}\, e^{-\nu\, i}\, di - q \frac{\nu^{N-1}}{(N-2)!}\, e^{-\nu\, a/q}\, \left[\left(\frac{a}{q}+\frac{1}{\nu}\right) \int_0^{a/q} i^{N-2}\, di - \int_0^{a/q} i^{N-1}\, di\right] \\
= \quad & \frac{q}{\nu} \Gamma_{N-1,\nu}(a/q) - q \frac{\nu^{N-1}}{(N-1)!} e^{-\nu\, a/q} \left(\frac{a}{q}\right)^{N-1} \left[\frac{1}{\nu} + \frac{1}{N}\frac{a}{q}\right].
\end{aligned}$$

Therefore, since all consumers are identical, the objective function (6) takes the form displayed in (36). For condition (7), on the other hand, we have

$$\begin{aligned}
V \quad = \quad & V_N = \mathbb{E}\left[U\left(\mathrm{w} - p - L_N + R_N\right)\right] \\
= \quad & \underbrace{\mathbb{E}\left[U\left(\mathrm{w} - p - (1-q)\, L_N\right)\, \mathbf{1}_{\{q(L_{-N}+L_N)<a\}}\right]}_{\text{part } i.} \\
& + \underbrace{\mathbb{E}\left[U\left(\mathrm{w} - p - L_N + a \frac{L_N}{L_{-N}+L_N}\right)\, \mathbf{1}_{\{q(L_{-N}+L_N)\geq a\}}\right]}_{\text{part } ii.}
\end{aligned}$$

For part i, we obtain

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

For part $ii.$, note that

$$\begin{aligned}
&\mathbb{E}\left[U\left(\mathrm{w}-p-((L_{-N}+L_N)-a)\,\frac{L_N}{L_{-N}+L_N}\right)\,\mathbf{1}_{\{q(L_{-N}+L_N)\geq a\}}\right] \\
=&\ \int_0^1\int_0^\infty e^{-\alpha(\mathrm{w}-p-(l-a)y)}\mathbf{1}_{\{l\geq a/q\}}\,\frac{\nu^N}{(N-1)!}\, l^{N-1}\, e^{-\nu\, l}\,(N-1)\,(1-y)^{N-2}\, dl\, dy \\
=&\ -e^{-\alpha(\mathrm{w}-p)}\int_{a/q}^\infty\frac{\nu^N}{(N-1)!}e^{-\nu\, l}\, l^{N-1}\underbrace{\int_0^1 e^{(\alpha(l-a))y}(N-1)\,(1-y)^{N-2}\, dy}_{\mathrm{mgf}_{\mathrm{Beta}(1,N-1)}(\alpha(l-a))}\, dl \\
=&\ -e^{-\alpha(\mathrm{w}-p)}\int_{a/q}^\infty\frac{\nu^N}{(N-1)!}e^{-\nu\, l}\, l^{N-1}\sum_{k=0}^\infty\frac{(\alpha(l-a))^k}{(N-1+k)!}\, dl \\
=&\ -e^{-\alpha(\mathrm{w}-p)}\sum_{k=0}^\infty\sum_{j=0}^{N-1}\binom{N-1}{j}\, a^j\, e^{-\nu\, a}\frac{\alpha^k}{(N-1+k)!}\,\frac{(N+k-j-1)!}{\nu^{-j+k}} \\
&\qquad\times\int_{a/q}^\infty\frac{\nu^{N+k-j}}{(N+k-j-1)!}e^{-\nu\,(l-a)}\,(l-a)^{N+k-j-1}\, dl \\
=&\ -e^{-\alpha(\mathrm{w}-p)}\sum_{k=0}^\infty\sum_{j=0}^{N-1}\frac{(N-1)!}{(N-1+k)!}\frac{(a\nu)^j}{j!}\, e^{-\nu\, a}\left(\frac{\alpha}{\nu}\right)^k\frac{(N+k-j-1)!}{(N-1-j)!}\,\bar{\Gamma}_{N-j+k,\nu}\left(a\left(\frac{1-q}{q}\right)\right).
\end{aligned}$$

C.2 Derivation of Equation (38)

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

$$\begin{aligned}
&\mathbb{E}\left[\sum_{j=1}^N U'\left(\mathrm{w}-p-L_j+a\,\frac{L_j}{L}\right)\frac{L_j}{L}\frac{L_N}{L}\middle|\, L\right] \\
=&\ \sum_{j=1}^{N-1}\underbrace{\mathbb{E}\left[U'\left(\mathrm{w}-p-(L-a)\,\frac{L_j}{L}\right)\frac{L_j}{L}\frac{L_N}{L}\middle|\, L\right]}_{\text{part } i.}+\underbrace{\mathbb{E}\left[U'\left(\mathrm{w}-p-(L-a)\,\frac{L_N}{L}\right)\left(\frac{L_N}{L}\right)^2\middle|\, L\right]}_{\text{part } ii.}.
\end{aligned}$$

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

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

Whence, for part $i.$,

$$\begin{aligned}
&\mathbb{E}\left[U'\left(\mathrm{w}-p-(L-a)\,\frac{L_{N-1}}{L}\right)\frac{L_{N-1}}{L}\frac{L_N}{L}\,\middle|\,L\right]\\
&= \alpha\, e^{-\alpha(\mathrm{w}-p)}\int_0^1\int_0^{1-x} e^{\alpha(L-a)x}\,x\,y\,(N-1)\,(N-2)\,(1-x-y)^{N-3}\,dy\,dx\\
&= \alpha\, e^{-\alpha(\mathrm{w}-p)}\int_0^1 e^{\alpha(L-a)x}x\underbrace{\int_0^{1-x} y\,(N-1)\,(N-2)\,(1-x-y)^{N-3}\,dy}_{=(1-x)^{N-1}}\,dx\\
&= \alpha\, e^{-\alpha(\mathrm{w}-p)}\,\beta(2,N)\underbrace{\int_0^1 e^{\alpha(L-a)x}\,\frac{1}{\beta(2,N)}\,x\,(1-x)^{N-1}\,dx}_{=\mathrm{mgf}_{\mathrm{Beta}(2,N)}(\alpha(L-a))}\\
&= \alpha\, e^{-\alpha(\mathrm{w}-p)}\frac{1}{N(N+1)}(N+1)!\sum_{k=0}^{\infty}\frac{(k+1)\ (\alpha(L-a))^k}{(N+k+1)!},
\end{aligned}$$

whereas for part $ii.$,

$$\begin{aligned}
&\mathbb{E}\left[U'\left(\mathrm{w}-p-(L-a)\,\frac{L_N}{L}\right)\left(\frac{L_N}{L}\right)^2\middle|\,L\right]\\
&= \alpha e^{-\alpha(\mathrm{w}-p)}\underbrace{\mathbb{E}\left[\exp\left\{\alpha(L-a)\frac{L_N}{L}\right\}\left(\frac{L_N}{L}\right)^2\middle|\,L\right]}_{=\frac{\partial^2}{\partial t^2}\mathrm{mgf}_{\mathrm{Beta}(1,N-1)}(t)\big|_{t=\alpha(L-a)}}\\
&= \alpha\, e^{-\alpha(\mathrm{w}-p)}(N-1)!\sum_{k=0}^{\infty}\frac{(k+1)\,(k+2)\ (\alpha(L-a))^k}{(N+k+1)!},
\end{aligned}$$

so that

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

For the denominator,

$$
\begin{aligned}
& \mathbb{E}\left[\mathbf{1}_{\{q\,L\geq a\}}\sum_{j=1}^{N} U'\left(\mathrm{w}-p-L_j-a\frac{L_j}{L}\right)\frac{L_j}{L}\right] \\
= \;& \mathbb{E}\left[\mathbf{1}_{\{q\,L\geq a\}}\,N\,\alpha\,e^{-\alpha(\mathrm{w}-p)}\,(N-1)!\sum_{k=0}^{\infty}\frac{(k+1)\;(\alpha(L-a))^k}{(N+k)!}\right] \\
= \;& N!\,\alpha\,e^{-\alpha(\mathrm{w}-p)}\sum_{k=0}^{\infty}\frac{k+1}{(N+k)!}\,\alpha^k\int_{a/q}^{\infty}\frac{\nu^N}{(N-1)!}(l-a)^k\,l^{N-1}\,e^{-\nu\,l}\,dl \\
= \;& N!\,\alpha\,e^{-\alpha(\mathrm{w}-p)}\sum_{k=0}^{\infty}\left(\frac{\alpha}{\nu}\right)^k\frac{k+1}{(N+k)!}\,e^{-\nu\,a}\sum_{j=0}^{N-1}\binom{N-1}{j}(a\nu)^j\,\frac{(N+k-j-1)!}{(N-1)!} \\
& \qquad\times\int_{a/q}^{\infty}\frac{\nu^{N+k-j}}{(N+k-j-1)!}e^{-\nu(l-a)}\,(l-a)^{N+k-j-1}\,dl \\
= \;& N!\,\alpha\,e^{-\alpha(\mathrm{w}-p)}\sum_{k=0}^{\infty}\left(\frac{\alpha}{\nu}\right)^k\frac{k+1}{(N+k)!}\,e^{-\nu\,a}\sum_{j=0}^{N-1}\frac{(a\nu)^j}{j!}\frac{(N+k-j-1)!}{(N-j-1)!}\,\bar{\Gamma}_{N+k-j,\nu}\left(a\left(\frac{1-q}{q}\right)\right).
\end{aligned}
$$

Hence,

$$
q\,\tilde{\phi}_i \;=\; \frac{1}{N}\mathbb{E}\left[\mathbf{1}_{\{q\,L\geq a\}}\underbrace{\frac{\sum_{k=0}^{\infty}\frac{(k+1)\,(\alpha(L-a))^k}{(N+k)!}}{\sum_{k=0}^{\infty}\left(\frac{\alpha}{\nu}\right)^k\frac{k+1}{(N+k)!}\,e^{-\nu\,a}\sum_{j=0}^{N-1}\frac{(a\nu)^j}{j!}\frac{(N+k-j-1)!}{(N-j-1)!}\,\bar{\Gamma}_{N+k-j,\nu}\left(a\left(\frac{1-q}{q}\right)\right)}}_{=\hat{\psi}^*(L)}\right].
$$

For implementation purposes, the numerator can be expressed as

$$
\begin{aligned}
\sum_{k=0}^{\infty}\frac{(k+1)\,t^k}{(N+k)!}\bigg|_{t=\alpha(L-a)} &= \frac{\partial}{\partial t}\left[\sum_{k=0}^{\infty}\frac{t^{k+1}}{(N+k)!}\right]\bigg|_{t=\alpha(L-a)} = \frac{\partial}{\partial t}\left[t^{-(N-1)}\sum_{k=N}^{\infty}\frac{t^k}{k!}\right]\bigg|_{t=\alpha(L-a)} \\
&= \frac{\partial}{\partial t}\left[t^{-(N-1)}\left(e^t-\sum_{k=0}^{N-1}\frac{t^k}{k!}\right)\right]\bigg|_{t=\alpha(L-a)} \\
&= -(N-1)t^{-N}\left(e^t-\sum_{k=0}^{N-1}\frac{t^k}{k!}\right)+t^{-(N-1)}\left(e^t-\sum_{k=0}^{N-1}\frac{k\,t^{k-1}}{k!}\right)\bigg|_{t=\alpha(L-a)} \\
&= \left(e^t-\sum_{k=0}^{N-2}\frac{t^k}{k!}\right)\times\left(t^{-N+1}-t^{-N}\,(N-1)\right)+\frac{(N-1)}{(N-1)!}\,t^{-1}\bigg|_{t=\alpha(L-a)}.
\end{aligned}
$$

Finally,

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

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

References

Acerbi, C. (2002). "Spectral Measures of Risk: A Coherent Representation of Subjective Risk Aversion." *Journal of Banking and Finance* 26, 1505-1518.

Acerbi, C., and Tasche, D. (2002). "Expected Shortfall: a natural coherent alternative to Value at Risk." *Economic notes* 31, 379-388.

Acharya, V.V., Pedersen, L.H., Phillipon, T., and Richardson, M. (2010) "Measuring Systemic Risk." Working Paper, New York University.

Adrian, T., and Shin, H. (2008). "Financial Intermediary Leverage and Value-at-Risk." *Federal Reserve Bank of New York Staff Report 338*.

Artzner, P., Delbaen, F., Eber, J.M., and Heath, D. (1999). "Coherent Measures of Risk." *Mathematical Finance* 9, 203-228.

Basak, S., and Chabakauri, S. (2006). "Dynamic Mean-Variance Asset Allocation." *Review of Financial Studies* 23, 2970-3016.

Björk, T., and Slinko, I. (2006). "Towards a General Theory of Good-Deal Bounds." *Review of Finance* 10, 221-260.

Černý, A. (2003). "Generalised Sharpe Ratios and Asset Pricing in Incomplete Markets." *Review of Finance 7, 191-233*.

Cochrane, J.H., and Saá-Requejo, J. (2000). "Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets." *Journal of Political Economy* 108, 79-119.

Cummins, J.D. (1988). "Risk-based Premiums for Insurance Guaranty Funds." *Journal of Finance 43, 823-839*.

Cuoco, D., He, H., and Isaenko, S. (2008) "Optimal dynamic trading strategies with risk limits." *Operations Research* 56, 358-368.

Denault, M. (2001). "Coherent Allocation of Risk Capital." *Journal of Risk* 4, 7-21.

Dhaene, J., Tsanakas, A., Valdez, E., and Vanduffel, S. (2009). "Optimal Capital Allocation Principles." *Journal of Risk and Insurance* 79, 1-28.

Erel, I., Myers, S.C., and Read, J. (2013). "Capital Allocation." Working Paper, Ohio State University.

Föllmer, H., and Schied, A. (2002). "Convex measures of risk and trading constraints." *Finance and Stochastics* 6, 429-447.

Föllmer, H., and Schied, A. (2010). "Coherent and Convex Risk Measures." In *Encyclopedia of Quantitative Finance*, Cont, R. (Ed.), Wiley, 355-363.

Föllmer, H., and Schweizer, M. (2010). "Minimal Martingale Measure." In *Encyclopedia of Quantitative Finance*, Cont, R. (Ed.), Wiley, 1200-1204.

Frittelli, M., Gianin, E.R. (2002). "Putting Order in Risk Measures." *Journal of Banking and Finance* 26, 1473-1486.

Froot, K.A. (2007). "Risk Management, Capital Budgeting, and Capital Structure Policy for Insurers and Reinsurers." *Journal of Risk and Insurance* 74, 273-299.

Froot, K.A., Scharfstein, D.S., and Stein, J.C. (1993). "Risk Management: Coordinating Corporate Investment and Financing Policies." *Journal of Finance* 48, 1629-1658.

Froot, K.A., and Stein, J. (1998). "Risk Management, Capital Budgeting, and Capital Structure Policy for Financial Institutions: An Integrated Approach." *Journal of Financial Economics* 47, 55-82.

Garman, M. (1997). "Taking VaR to Pieces: Presenting a method to express a portfolio's VaR as the sum of its parts." *Risk* 10, 70-71.

Goll, T., and Rüschendorf, L. (2001). "Minimax and minimal distance martingale measures and their relationship to portfolio optimization." *Finance and Stochastics* 5, 557-581.

Gourieroux, C., Laurent, J.P., and Scaillet, O. (2000). "Sensitivity analysis of Values at Risk." *Journal of Empirical Finance* 7, 225-245.

Gründl, H., and H. Schmeiser (2007). "Capital Allocation for Insurance Companies: What Good Is It?" *Journal of Risk and Insurance* 74, 301-317.

Hanif, F., Hocking, J., Jaekel, A., and Ziewer, L. (2010). "Solvency 2: Quantitative & Strategic Impact – The Tide is Going Out." Morgan Stanley and Oliver Wyman Blue Paper.

Henderson, V., Hobson, D., Howinson, S., and Kluge, T. (2005). "A Comparison of Option Prices Under Different Pricing Measures in a Stochastic Volatility Model with Correlation." *Review of Derivatives Research* 8, 5-25.

Hull, J. (2007) "VAR versus Expected Shortfall." *Risk Magazine*, March 2007.

Ibragimov, R., Jaffee, D., and Walden, J., (2010) "Pricing and Capital Allocation for Multiline Insurance Firms." *Journal of Risk and Insurance* 77, 551-578.

International Actuarial Association (2004). "A Global Framework for Insurer Solvency Assessment." IAA Insurer Solvency Assessment Working Party Research Report. Availabe at `http://www.actuaries.org/LIBRARY/Papers/Global_Framework_Insurer_Solvency_Assessment-public.pdf`.

Kalkbrener, M. (2005). "An Axiomatic Approach to Capital Allocation." *Mathematical Finance* 15, 425-37.

Keeler, T.E., and Small, K.A. (1977). "Optimal Peak Load Pricing, Investment, and Service Levels on Urban Expressways." *Journal of Political Economy* 85, 1-25.

Kou, S.G., Peng, X.H., and Heyde, C.C., (2012). "External Risk Measures and Basel Accords." *Mathematics of Operations Research*, forthcoming.

McKinsey&Company (2011). "The use of economic capital in performance management for banks: A perspective." Prepared by T. Baer, A. Mehta, and H. Samandari. *McKinsey Working Papers on Risk*, Number 24. Availabe at `http://www.mckinsey.com/clientservice/risk/pdf/risk_working_papers/24_The_Use_of_Economic_Capital.pdf`.

McNeil, A.J., Frey, R., and Embrechts, P. (2005). "Quantitative Risk Management: Concepts, Techniques, and Tools." Princeton: Princeton University Press.

Myers, S.C., and Read, J.A. (2001). "Capital Allocation for Insurance Companies." *Journal of Risk and Insurance* 68, 545-80.

National Association of Insurance Commissioners (NAIC) (2009). "Consultation Paper on Regulatory Capital Requirements and Overarching Accounting/Valuation Issues for the Solvency Modernization Initiative." Availabe at `http://www.naic.org/documents/committees_ex_isftf_iswg_capital_requirements.doc`.

Phillips, R.D., Cummins, J.D., and Allen, F. (1998). "Finanical Pricing of Insurance in the Multiple Line Insurance Company." *Journal of Risk and Insurance* 65, 597-636.

Powers, M.R. (2007). "Using Aumann-Shapley Values to Allocate Insurance Risk: The Case of Inhomogeneous Losses." *North American Actuarial Journal* 11, 113-27.

Rampini, A.A., and Viswanathan, S. (2010). "Collateral, Risk Management, and the Distribution of Debt Capacity." *Journal of Finance* 65, 2293-2321.

Rampini, A.A., and Viswanathan, S. (2012). "Financial Intermediary Capital." Working Paper, *Duke University*.

Schmock, U., and Straumann, D. (1999). "Allocation of Risk Capital and Performance Measurement." Presentation at the Risk Day 1999, Zürich. Available at `http://www.fam.tuwien.ac.at/~schmock/slides/AllocationSlidesOct1999.pdf`.

Smith, C.W., Smithson, C.W., and Wilford, S.D. (1990). "Managing Financial Risk." New York, NY, Harper Business, 355-378.

Smith, C.W., and Stulz, R.M. (1985). "The Determinants of Firms Hedging Policies." *Journal of Financial and Quantitative Analysis* 20, 391-405.

Society of Actuaries (2008). "Economic Capital for Life Insurance Companies." Research report prepared by I. Farr, H. Mueller, M. Scanlon, and S. Stronkhorst. Availabe at `http://www.soa.org/files/pdf/research-ec-report.pdf`.

Stoughton, N.M., and Zechner, J. (2007). "Optimal Capital Allocation Using RAROC and EVA." *Journal of Financial Intermediation* 16, 312-342.

Tasche, D. (2004). "Allocating Portfolio Economic Capital to Sub-portfolios." In: *Economic Capital: A Practitioner Guide*, A. Dev (ed.). Risk Books, 2004, 275-302.

Zanjani, G. (2002). "Pricing and Capital Allocation in Catastrophe Insurance." *Journal of Financial Economics* 65, 283-305.

Zanjani, G. (2010). "An Economic Approach to Capital Allocation." *Journal of Risk and Insurance* 77, 523-549.