Equilibrium with Monoline and Multiline Structures*

Rustam Ibragimov† Dwight Jaffee‡ Johan Walden§

Abstract

We study a competitive market for risk sharing, in which risk tolerant providers of risk protection, who face frictional costs in holding capital, offer coverage over a range of risk classes to risk averse agents. We distinguish monoline and multiline industry structures and characterize when each structure is optimal. Markets for which the risks are limited in number, asymmetric or correlated will be served by monoline structures, whereas markets characterized by a large number of essentially independent risks will be served by many multiline firms. Our results are consistent with observed structures within insurance, and also have general implications for the financial services industry.

JEL: G2, G22.

Keywords: Insurance, financial services, risk sharing, diversification.

*In loving memory of Dwight Jaffee, 1943-2016. We thank seminar participants at the University of Tokyo, the Symposium on Non-Bayesian Decision Making and Rare, Extreme Events, Bergen, Norway, the 2008 American Risk and Insurance Association meetings, Portland, Oregon, and the 2013 NBER Insurance Working Group Meeting. We also thank two anonymous referees and the co-editor, Andrew Ellul, for helpful comments and suggestions. Walden thanks the NUS Risk Management Institute for support. Ibragimov gratefully acknowledges support provided by the Russian Science Foundation, Project No. 16-18-10432.

†Imperial College Business School, U.K., and Innopolis University, Russia. Tanaka Building, South Kensington Campus, London, SW7 2AZ, U.K. *E-mail*: `irustam@imperial.ac.uk`.

‡Haas School of Business, University of California at Berkeley, 545 Student Services Building #1900, CA 94720-1900.

§Corresponding author. Haas School of Business, University of California at Berkeley, 545 Student Services Building #1900, CA 94720-1900. *E-mail*: `walden@haas.berkeley.edu`.

1 Introduction

The equilibrium outcome in a friction-free Walrasian market for risk sharing is well-understood: Idiosyncratic risks will be diversified away so that no agent is exposed to them, and there is no risk premium associated with such risks. The remaining common risk will be divided so that risk averse agents take on less risk than risk tolerant agents, who in turn are compensated via higher expected returns (Borch (1962), Arrow (1964), Sharpe (1964), Lintner (1965), Ross (1976)). All agents thus share the common risk but their exposures differ. In the Capital Asset Pricing Model, for example, the common risk is represented by the market portfolio.

In some real-world markets the observed outcome is reasonably close to these predictions. In stock markets, for example, index funds and ETFs allow investors to diversify idiosyncratic risk while loading up on systematic risk. In other markets, however, the outcome is significantly different, as usually explained by the frictions present in practice. We provide several examples of such frictions. First, because of the vast number of risks that agents are exposed to in the real world, extremely complex and likely infeasible multilateral contracts, contingent on the joint realization of all risks, would be needed to reach the optimal risk sharing outcome. Second, various costs are present in practice, e.g., transaction costs and taxes, that effectively limit the potential for perfect risk-sharing. Third, there are exogenous constraints on the contracts that may be written. For example, firms have limited liability and therefore cannot be forced to make arbitrary large payments, their employees cannot be forced to stay at the firm, etc. A fourth friction, which will not be our focus, is moral hazard that makes it infeasible to write contracts contingent on agents' private information and actions.

Another common market outcome is one of one-sided risk protection, in which risk tolerant agents—insurers—provide coverage to risk averse agents—insurees. Indeed, a sizable portion of risk in the economy is covered in traditional insurance markets: Gross insurance premiums in the U.S. in 2013 were almost USD 1.3 Trillion,[1] and the *values* insured were orders of magnitude higher, with insured building values alone exceeding USD 40 Trillion.[2] Insurance-like contracts also arise in financial markets, e.g., through credit default swaps and catastrophe bonds. More broadly, a multitude of products implicitly provide one party with insurance against bad outcomes through a default option, the most common such product being standard corporate debt.

[1] "Annual Report on the Insurance Industry," Federal Insurance Office, 2014.
[2] "Catastrophe risk and the cost of real estate insurance," Zurich American Insurance Corporation, 2013.

The aforementioned frictions have been proposed to play a role in explaining the relatively simple contract structures observed in these markets (see, e.g., Arrow (1963), Doherty, Laux, and Muermann (2014), and Jaffee and Walden (2014)).

It is an open question what will be the industry equilibrium structure in a market with one-sided risk protection. Indeed, to the best of our knowledge, even fundamental questions like what will be the number of firms, and how these firms will organize to serve the market remain unanswered. In this paper, we address these questions in the context of an insurance industry, while stressing that our framework is applicable more broadly to markets characterized by one-sided risk-protection contracts.

Our starting point is a market in which risk neutral insurers offer protection against the risks faced by risk averse insurees. In a friction-free setting, the optimal outcome is for the insurer to write standard contracts that protect against all risk, and the industry structure is irrelevant since all policy holders are fully protected regardless of the structure. In practice, frictions make such an outcome infeasible. First, virtually all insurers are now limited liability corporations, which eliminates the unlimited recourse to partners' external (private) assets that was once common.[3] To avoid counterparty risk, a large amount of capital therefore needs to be held within the firm. Second, the excess costs of holding capital, such as corporate taxes, asymmetric information, and agency costs, create deadweight costs to holding internal (on balance sheet) capital, providing a strong incentive for the insurers to limit the amount of internal capital they hold.[4] In practice, policyholders therefore face counterparty risk. Third, although it is no longer clear whether standard insurance contracts are optimal with these frictions, such contracts are still observed in practice, possibly because of their simplicity as previously discussed. Moreover, in case of insurer default, the shortfall is typically allocated according to a simple so-called *ex post* payout rule in proportion to actual claims (see Ibragimov, Jaffee, and Walden (2010)). In our analysis, we take these frictions as given.

We introduce a parsimonious model with insurance firms with limited liability and costly

[3] The insurer Lloyds of London once provided a credible guarantee to pay all claims, based on the private wealth of its "names" partners. In the aftermath of large asbestos claims, however, Lloyds now operates primarily as a standard "reserve" insurer with balance sheet capital and limited liability.

[4] This second factor was first emphasized with respect to general corporate finance by Froot, Scharsfstein, and Stein (1993), and Froot and Stein (1998) later extended that analysis to financial intermediaries. For insurance firms, Cummins (1993), Merton and Perold (1993), Jaffee and Russell (1997), Myers and Read (2001), and Froot (2007) all emphasize the importance of various accounting, agency, informational, regulatory, and tax factors in raising the cost of internally held capital.

internal capital. Each firm has the option to offer coverage against one or more of the existing insurance lines to risk-averse policy holders, i.e., whether to organize as a monoline or a multiline company. The contract space is restricted to include one-sided risk protection, and in case of default the ex post payout rule is used. The choice of monoline versus multiline structure is important because it determines in which states default occurs, and thereby counterparty exposures to such default risk.

We consider a competitive market, in which firms compete, which severely restricts the feasible equilibrium industry structures. We relate our equilibrium concept to so-called core stability in hedonic coalition games. We know of no other paper that provides an analytic framework for determining the industry structure that will prevail for an industry that may contain both monoline and multiline structures.

Our model has several important industry structure implications. A priori, one may expect the benefit of diversification to cause the equilibrium outcome to have one gigantic massively multiline insurer that takes full advantage of diversification and serves all the insurees in the market. Such a fully diversified outcome is typically *not* the outcome that will prevail, however. Instead, the industry will typically be served by several multiline companies, each holding a different amount of capital, and there may also even be a role for some companies to choose a monoline structure.

There are two forces that counteract the diversification benefits of having all risks covered by the same company. First, different levels of internal capital may be optimal for different types of risks. Specifically, an increase in internal capital decreases the risk of insurer default, but also increases the total cost of holding such capital. By choosing a multiline structure, a firm is forced to choose a single "compromise" capital level. In contrast, with a monoline structure each line can be served by a firm with an amount of internal capital tailored for that specific line. We show that when there is a sufficient number of insurance lines in an economy with risk distributions that are sufficiently similar, the vast majority of policy holders will be served by multiline structures, but there will be many such firms and thus the outcome is still far away from the fully diversified case. Some risks that would be uninsurable under a monoline structure, in that the policy holders would not demand insurance at the price at which a monoline insurer would be willing to sell it, may be insurable by a multiline company. Moreover, special cases aside, there may always be a role for some monoline structures, for which the optimal level of capital is significantly different from all other lines.

Second, significantly asymmetric risk distributions across lines may also work against multiline outcomes in equilibrium. Specifically, when risks are asymmetric it is often the case that one line subsidizes another under a multiline structure. In a competitive market, the agents in the subsidizing line may then rather be served by a monoline structure, which is therefore the outcome that prevails. In addition, when risks are positively correlated across lines, monoline structures are all else equal more favored. This is unsurprising, because the benefit of diversification is known to decrease when correlation increases. To support our intuition, we provide a detailed example of an economy with two insurance lines, in which we can completely characterize when monoline and multiline structures are optimal.

Our results have several potential policy implications. Briefly, each equilibrium outcome is constrained Pareto efficient with respect to the levels of internal capital chosen by firms, but there may be room for Pareto improvements by changing the industry structure in terms of which types of structures (monoline or multiline) serve which lines. Also, some efficient allocations may not be sustainable as equilibrium outcomes. To respond, a policy maker may wish to introduce monoline regulation, and in some cases also multiline regulation.

We stress that if any of the three aforementioned frictions (costs of holding internal capital, limited liability, and standard insurance contracts) is removed, the tension between monoline and multiline outcomes disappears. Without costs of holding internal capital, it is cost-free for the insurer to choose a very high capital level, in which case monoline and multiline outcomes are equivalent. This is also the case without limited liability, in which case the insurer can commit to cover all losses even without holding internal capital. Finally, under complete contracting freedom, a multiline insurer may offer monoline-like contracts, so the multiline outcome can never be dominated.

Our results fit broadly with several observations. For one thing, insurance companies operate with different ratings—AAA, AA, A, etc.—reflecting different amounts of capital and safety. For another thing, catastrophe lines of insurance illustrate one type of risks in which the benefits of diversification may be low enough to make a monoline structure preferable. The catastrophe lines create a potential bankruptcy risk if a "big one" should create claims that exceed the insurer's capital resources. Thus, an insurer that offers coverage against both traditional diversifiable risks and catastrophe risks may impose a counterparty risk on the policyholders of its traditional lines, a counterparty risk that would not exist if the insurer did not offer coverage on the catastrophe line. This negative externality for the policyholders on the diversifiable lines can be avoided if catastrophe insurers operate on a monoline basis, each one offering coverage against just one risk class and holding just the amount of capital that its policyholders deem optimal.

The paper is organized as follows: In the next section we discuss related literature, and in Section 3 we provide the framework for our analysis. In Section 4, we show in an example with two risks that little can be said about the industry structure in the general case. In Section 5, we then study markets with many independent risks. We first focus on the outcomes that are feasible in markets with many risk classes, and then on the actual outcomes that will prevail in equilibrium in such markets. A general implication of our analysis for such markets is that the multiline outcome will be predominant but that there may still be room for some monoline companies, even as the number of insurance lines becomes very large. We also discuss potential policy implications. Finally, Section 6 concludes.

2 Literature Review

The distinction between monoline and multiline structures is important within the financial services industry, and specifically within insurance. United States' insurance regulations require that the major financial guarantee insurance lines—mortgage insurance, municipal bond insurance, and various credit default swaps—be provided only on a monoline basis in order to protect the policyholders in multiline firms from an insurer default that could result from losses on these high-risk lines.[5] These regulations have been successful in the sense that while a number of large monoline insurers failed due to the financial crisis, there was no contagion to policyholders on other insurance lines.[6]

Within the financial services industry, the Glass Steagall Act imposed a monoline restriction by forcing U.S. commercial banks to divest their investment bank divisions. Winton (1995) develops a banking industry equilibrium in which larger banks have the advantage of greater

[5] As developed in Jaffee (2006), there is some variation in the state-based monoline restrictions, but most states adhere closely to the "model law" created by the National Association of Insurance Commissioners.

[6] Most of the major monoline insurers, including ACA, AMBAC, CIFG, FGIC, MBIA, and Radian, have faced serious financial stress and major down ratings, ending in a number of cases with complete failure and liquidation. Details are available from the Association of Financial Guaranty Insurers at: `http://www.afgi.org/`. AIG stands alone as a large, failed, multiline insurer, due to its credit default swap (CDS) positions. It is noteworthy, moreover, that these CDS risks were held in an operating division of the holding company, and not in a monoline insurance subsidiary.

diversification, but the drawback of a lower capital ratio. Kahn and Winton (2004) study the optimal subsidiary structure for a financial institution that chooses between projects that differ with respect to quality and riskiness, in the presence of moral hazard. Their focus is on an institution with two "lines," for which they distinguish between a bipartite structure—corresponding to our monoline structure—and a unitary structure—corresponding to our multiline structure. They derive conditions under which one structure or the other is optimal. Their financial institution is risk neutral, so diversification plays no role in their model. Rather, it is moral hazard in the form of an incentive for the institution to risk shift that drives their results. Their main conclusion is that a unitary structure allows the financial institution to extract value from lenders in both "lines," by risk shifting, whereas the bipartite structure insulates the lenders of safer loans. As a result, an efficient outcome is typically easier to reach under the bipartite structure. Our model, in contrast, has no moral hazard but rather focuses on the efficiency of different company structures in achieving risk diversification, and the related optimal choices of internal capital. Another difference between Kahn and Winton (2004) and our study is that they focus on an individual institution, whereas our analysis concerns a whole industry.

Bolton and Oehmke (2015), discuss the optimal resolution mechanism for the orderly liquidation of global financial institutions called for in the Dodd-Frank Act, and specifically whether a so-called single-point-of-entry (SPOE) resolution, in which the loss-absorbing capital is shared across multiple jurisdictions, or a multiple-point-of-entry (MPOE) resolution, in which separate capital is held under each jurisdiction, is optimal. The two variations may be viewed as multiline and monoline structures, respectively, separated by location in this case.

Ibragimov, Jaffee, and Walden (2011) analyze diversification and systemic risk in a model where financial institutions can choose different structures that can be interpreted as monoline and massively multiline. Briefly, in their model the institutions either completely share risks —corresponding to a massively multiline structure—or act in isolation—corresponding to an industry with only monoline structures. In contrast to our one-sided risk-protection framework, Ibragimov, Jaffee, and Walden (2011) focus on symmetric risk-sharing agreements. Also, their model does not allow for endogenously determined levels of internal capital, which is a key component for our results.

Within the theory of conglomerates in corporate finance, Leland (2007) develops a model in which single-activity operating corporations can choose the optimal debt to equity ratio, whereas multiline conglomerates obtain a diversification benefit but can only choose an average debt to equity ratio for the overall firm. Thus, here too there is a tension between the diversification benefit associated with a multiline structure and the benefit of separation allowed by a monoline structure.[7]

Within insurance, Lakdawalla and Zanjani (2012) studies the equilibrium in an insurance market in which catastrophe bonds and diversified reinsurers compete as risk transfer mechanisms. The catastrophe bonds in Lakdawalla and Zanjani (2012) are similar to our monoline insurers in that each mechanism covers only a single risk. Their catastrophe bonds, however, are assumed to be fully collateralized, whereas we allow our monoline insurers to choose the optimal degree of collateralization. Their reinsurers are also similar to multiline insurers in that each institution provides benefits of diversification. Our analysis, however, is more general because we allow multiline insurers to cover any optimal number of lines. The result is that our set of potential equilibrium outcomes is substantially broader than those considered by Lakdawalla and Zanjani (2012).

A key assumption of our model is that of an excess cost for an insurer to hold capital, as would arise from corporate taxes or other frictions. This factor causes insurers to conserve on capital, leading to the possibility of insurer default. Zanjani (2002) also studies the effect of costly internal capital for insurers. His focus, however, is not on the industry structure but on the pricing effects of variations in the capital ratios, a topic also studied in Ibragimov, Jaffee, and Walden (2010). Moreover, Zanjani makes several simplifying assumptions, including normal risk distributions and exogenous demand functions, features that we substantially generalize in this paper.

Several papers in the insurance literature have extended the classical analysis of the optimal insurance contract in Arrow (1963) to include possible insurer default, e.g., because of limited liability, and/or costly internal capital. Doherty and Schlesinger (1990) analyze optimal insurance in case of possible insurer default, but default is modeled as an exogenous event in their paper, as is the case in Cummins and Mahul (2003), Mahul and Wright (2004), and Mahul and Wright (2007). Cummins and Mahul (2004) allow for an upper bound on compensation, e.g., because of limited liability, but treat this bound as exogenous. Kaluszka and Okolewski (2008)

[7] Diamond (1984) also notes the comparable role that diversification may play for banks, insurers, and operating conglomerates. However, his analysis does not consider the possible advantage of the monoline structure when the diversification benefits are limited.

allow for a general cost function that depends both on expected losses and maximal claims, whereas Jaffee and Walden (2014) make assumptions of costly internal capital in line with those in our study. Doherty, Laux, and Muermann (2014) show that when losses are non-verifiable, upper bounds on compensation naturally arise. In contrast to our paper, all of these studies focus on the case with a single insurer and policy holder.

3 The Model

We first study the case of only one insured risk to introduce the basic concepts and notation, and then proceed to the main focus of our study, which contains multiple risks.

3.1 One Risk

Consider the following one-period model. At $t = 0$, an *insurer*, representing an insurance company, or more generally a financial institution offering one-sided risk-protection, in a competitive market sells insurance against an idiosyncratic risk, $\tilde{L} \geq 0$ to an *insuree*.[8] The expected loss of the risk is $\mu_L = E[\tilde{L}]$, $\mu_L < \infty$.

The insuree is risk averse, with expected utility function u, where u is an increasing strictly concave function that satisfies some technical conditions.[9] We also assume that the risk cannot be divided between multiple insurers. We note that sharing risks is uncommon in practice, reflecting the fixed costs of evaluating risks and selling policies, as well as the agency problems between insurers when handling split insurance claims. Finally, we assume that expected utility, U, is finite, $U = Eu(-\tilde{L}) > -\infty$.

For many types of individual and natural disaster risks, such as auto and earthquake insurance, etc., it seems reasonable to assume that risks are idiosyncratic, i.e., that they do not carry a premium above the risk-free rate in a competitive market for risk. In the analysis here, we assume that the risks are idiosyncratic in this sense.

At $t = 0$, the insurer takes on risks, receives premium payments, and contributes its own equity capital. The premiums and contributed capital are invested in risk-free assets, so that

[8] Throughout the paper we use the convention that losses take on positive values.

[9] Specifically, u is twice continuously differentiable function defined on the whole of $\mathbb{R}_- = (-\infty, 0]$, and $u'(0) \geq C_1 > 0$, $u''(x) \leq C_2 < 0$, for constants C_1, C_2, for all $x \leq 0$. For some of the results stronger conditions on u will be introduced when needed.

the assets A are available at $t = 1$, at which point losses are realized. Without loss of generality, we normalize the risk-free discount rate to zero.

The insurer has limited liability, and satisfies claims by paying $\tilde{L}$ to the insuree, as long as $\tilde{L} \leq A$. But, if $\tilde{L} > A$, the insurer pays A and defaults on the additional amount that is due.[10] Thus, the payment is

$$Payment = \min(\tilde{L}, A) = \tilde{L} - \max(\tilde{L} - A, 0) = \tilde{L} - \tilde{Q}(A),$$

where $\tilde{Q}(A) = \max(\tilde{L} - A, 0)$, i.e., $\tilde{Q}(A)$ is the payoff to the option the insurer has to default. As shown in Cummins and Mahul (2004), see also Jaffee and Walden (2014), the optimal contract with limited liability and capital has this form, but includes a deductible. In the analysis here, we assume that the deductible is zero since this property of the insurance contract is of second order importance to our analysis, and since this assumption simplifies the analysis substantially. The premium for the insurance is P.

The price for $\tilde{L}$ risk in a competitive market is $P_L = \mu_L$, because the risk is idiosyncratic. More generally, given that there is a market for risk that admits no arbitrage, there is a risk-neutral expectations operator that determines the premium in a competitive insurance market. Since the risk is idiosyncratic, the risk-neutral expectation coincides with the true expectation for the risks that we consider. Similarly, the value of the option to default is $P_Q = E[\tilde{Q}(A)] = \mu_Q$.

There are deadweight frictional costs that apply when an insurer holds internal capital; we refer to these as the excess costs of internal capital. The most obvious source is the taxation of corporate income, although asymmetric information, agency issues and bankruptcy costs may create similar costs. We specify the excess cost of internal capital as δ per unit of capital, i.e., δ provides a reduced form summary of the total excess cost per unit risk.[11] This assumption is comparable to the standard corporate finance assumption of a tax shield provided by corporate debt. The only difference is an algebraic sign: insurers hold net positive positions in financial

[10] In the context of insurance, for some lines of consumer insurance (e.g. auto and homeowner) there exist state guaranty funds through which the insurees of a defaulting insurer are supposed to be paid by the surviving firms for that line. In practice, delays and uncertainty in payments by state guaranty funds leave insurees still facing a significant cost when an insurer defaults; see Cummins (1988). More generally, our analysis applies to all the commercial insurance lines and catastrophe lines for which no state guaranty funds exist.

[11] This is the assumption used in a series of papers by Froot, Scharsfstein, and Stein (1993), Froot and Stein (1998), and Froot (2007). It also implies that an additional dollar of equity capital raises the firm's market value by less than a dollar. Since we assume a competitive industry, this excess cost is recovered through the higher premiums charged to policyholders. It is also for this reason that the amount of capital is chosen to maximize policyholder utility.

instruments as capital assets, while most operating corporations are net debt issuers. Our model shows that, even with the deadweight cost of internal capital, insurers maintain net positive positions in financial assets precisely because it reduces the counterparty risk faced by their policyholders. The result is that to ensure that a capital amount A is available at $t = 1$, $(1 + \delta)A$ needs to be reserved at $t = 0$. Since the market is competitive and the cost of internal capital is δA, the premium charged for the insurance is[12]

$$P = P_L - P_Q + \delta A = \mu_L - \mu_Q + \delta A. \tag{1}$$

We assume, in line with practice, that premiums are paid upfront. To ensure that A is available at $t = 1$, the additional amount of $A + \delta A - P = A - P_L + P_Q$ needs to be contributed by the insurer. Through the remainder of the paper, we refer to A as the insurer's assets or capital, it being understood that the amount $P_L - P_Q + \delta A$ is paid by the insurees as the premium, and the amount $A - P_L + P_Q$ is contributed by the insurer's shareholders. The total market structure is summarized in Figure 1.[13]

3.2 Multiple Risks

We follow Phillips, Cummins, and Allen (1998) by assuming that claims on all the insured risks are realized at the same time, $t = 1$.[14] At $t = 1$ the insurer either pays all claims in full (when assets exceed total claims) or defaults (when total claims exceed the assets).

If coverage against N risks is provided by one insurer, the total payment made to all policyholders with claims, taking into account that the insurer may default, is

$$Total\ Payment = \tilde{L} - \max(\tilde{L} - A, 0) = \tilde{L} - \tilde{Q}(A),$$

where $\tilde{L} = \sum_i \tilde{L}_i$ and $\tilde{Q}(A) = \max(\tilde{L} - A, 0)$. The shortfall in total assets for a defaulting

[12] The premium setting and capital allocations build on the no-arbitrage, option-based, technique, introduced to insurance models by Doherty and Garven (1986), then extended to multiline insurers by Phillips, Cummins, and Allen (1998) and Myers and Read (2001), and further developed in Ibragimov, Jaffee, and Walden (2010).

[13] Note that we implicitly assume that insurees do not have direct access to the market for risk or for other reasons prefer to go through the financial intermediary, a common assumption in the literature. Otherwise they would go directly to the market and avoid the friction cost of internal capital.

[14] This assumption is invaluable in that it allows tractability in computing the risk sharing and risk transfer attributes of the equilibrium outcomes of the model. It does, however, also mean that we are unable to study a variety of explicitly dynamic questions. Although the study of such dynamic factors would certainly provide additional insight, we believe that they would not change the basic results that are emphasized in this paper.

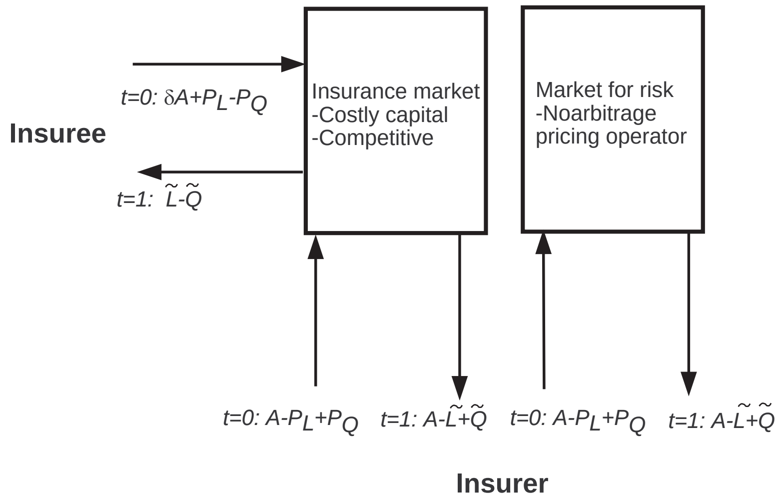

Figure 1: *Structure of model. Insurers can invest in market for risk and in a competitive insurance market. There is costly capital, so to ensure that A is available at $t = 0$, $(1 + \delta)A$ needs to be reserved at $t = 1$. The premium, $\delta A + P_L - P_Q$, is contributed by the insuree and $A - P_L + P_Q$ by the insurer. The discount rate is normalized to zero. Competitive market conditions imply that the premium for insurance is $P = P_L - P_Q + \delta A$.*

insurer is allocated across claimants in proportion to their actual claims. This is the so-called *ex post* sharing rule, which is extensively discussed in Ibragimov, Jaffee, and Walden (2010). With this rule, the payments made to insuree i is then

$$Payment_i = \frac{\tilde{L}_i}{\tilde{L}} A = \tilde{L}_i - \frac{\tilde{L}_i}{\tilde{L}} \tilde{Q}(A). \tag{2}$$

We thus use this rule in line with our previous discussion about contract simplicity, although with unconstrained contract structures there may be a very complex multilateral contracts that specify how the shortfall should be shared for all joint risk realizations, and that provides improved risk-sharing.

Following Ibragimov, Jaffee, and Walden (2010), we define the binary default option

$$\tilde{V}(A) = \begin{cases} 0 & \tilde{L} \leq A, \\ 1 & \tilde{L} > A, \end{cases} \tag{3}$$

and the price for such an option in the competitive friction free market, $P_V = E[\tilde{V}]$. The total price for the risks is, $P \stackrel{\text{def}}{=} \sum_i P_i$, where P_i is the premium for insurance against risk i. It follows that

$$P_i = P_{L_i} - r_i P_Q + v_i \delta A, \tag{4}$$

where

$$r_i = E\left[\frac{\tilde{L}_i}{\tilde{L}} \times \frac{\tilde{Q}}{P_Q}\right], \qquad v_i = E\left[\frac{\tilde{L}_i}{\tilde{L}} \times \frac{\tilde{V}}{P_V}\right]. \tag{5}$$

Thus, Equations 2-5 completely characterize payments and prices for all policyholders in the general case with multiple risks.

Our model does not include reinsurance, but our "insurers" could in principle also include reinsurance companies, so one could argue that such companies do not need to be explicitly modeled. With reinsurance, however, an additional dimension is added in that contracts are written also between the insurers in our model, complicating the risk-sharing structure.[15] Also, reinsurance contracts may be more sophisticated than standard insurance contracts, adding further complexity. We view the incorporation of reinsurance to the model as an interesting future extension.

4 Equilibrium Market Structure With Two Risks

It is helpful to begin with a market in which there are two risks. For simplicity, we assume that the two insurees have expected utility functions defined by $u(x) = -(-x+t)^\beta$, $\beta > 1$, $t \geq 0$, $x < 0$, and that the risks, $\tilde{L}_1$ and $\tilde{L}_2$, have (scaled) Bernoulli distributions: $\mathbb{P}(\tilde{L}_1 = 1) = p$, $\mathbb{P}(\tilde{L}_1 = 0) = 1-p$, $\mathbb{P}(\tilde{L}_2 = 2) = q$, $\mathbb{P}(\tilde{L}_2 = 0) = 1-q$, $\mathrm{corr}(\tilde{L}_1, \tilde{L}_2) = \rho$. Depending on $0 < p < 1$ and $0 < q < 1$, there are restrictions on the correlation, ρ. For example, ρ can only be equal to 1 if $p = q$.

There are two possible market structures in this case. The two risks may be insured by two separate "mono-risk" insurers, or alternatively jointly by one "multi-risk" insurer. We note that if any of the three frictions in our model (costs of holding internal capital, limited liability, and standard insurance contracts) is removed, the tension between mono-risk and multi-risk insurer outcomes disappears. Without costs of holding internal capital, the insurer will choose a high capital level, insure all risk, and never default. The insuree is therefore indifferent between a

[15] We thank a referee for pointing this out.

mono-risk and multi-risk insurance outcome in this case. Similarly, without limited liability, the insurer can commit to cover all losses even without holding internal capital, and again the insuree is indifferent between a mono-risk and a multi-risk insurance outcome. Finally, under complete contracting freedom, it is feasible for a multi-risk insurer to offer contracts with identical payments in all states as with mono-risk insurance contracts, so the multi-risk insurance outcome can never be dominated. Thus, all the three frictions are necessary for a constructive analysis of the mono-risk versus multi-risk insurance tradeoff.

To analyze the market structure given a fixed level of capital and premiums, although quite straightforward, may give misleading results because the capital held and the structure chosen are jointly determined. For example, an insurance company choosing to insure many risks may choose to hold a lower level of capital than the total capital of a set of mono-risk firms jointly insuring the same risks. We therefore need to allow the level of capital to vary. Specifically, we will compare a multi-risk insurance structure where one insurer sells insurance against both risks when reserving capital A, with a mono-risk insurance structure where two companies insure the two risks separately, reserving capital A_1 and A_2, respectively.

Given our assumptions about competitive markets, we would expect a multi-risk insurer to dominate if it can choose a level of capital that makes both insurees better off than what they can get from a mono-risk insurer. On the other hand, if there is a way for a mono-risk insurer to choose a level of capital that improves the situation for the first insuree, then we would expect this insuree to go with the mono-risk insurer, and the mono-risk outcome will dominate. A similar argument can be made if a mono-risk offering dominates the multi-risk insurance outcome for the second risk.

To formalize this intuition, we let $U_1^{MONO}(A_1)$ and $U_2^{MONO}(A_2)$ denote the expected utilities of the first and second insuree with mono-risk insurance, when capital A_1 and A_2 is reserved, respectively. Similarly, $U_1^{MULTI}(A)$ and $U_2^{MULTI}(A)$ denote the expected utilities of the first and second insuree when insured by a multi-risk insurer with capital A. The multi-risk insurance structure is said to dominate if there is a level of internal capital, A, such that for $U_1^{MULTI}(A) > U_1^{MONO}(A_1)$ for all A_1, and $U_2^{MULTI}(A) > U_2^{MONO}(A_2)$, for all A_2. Otherwise, the mono-risk insurance structure is said to dominate.

In case of a mono-risk market structure, we would expect the competitiveness between insurers to lead to an outcome where the level of capital is chosen to maximize the insuree's expected utility. From the analysis in Jaffee and Walden (2014) we know that under general conditions there is a unique level of capital, $A^* \geq 0$ that maximizes the insuree's expected utility under a mono-risk insurance structure. Thus, the multi-risk structure dominates if there is a level of capital, A, such that $U_1^{MULTI}(A) > U_1^{MONO}(A_1^*)$ and $U_2^{MULTI}(A) > U_2^{MONO}(A_2^*)$.

We study the case when $\beta = 7$, $t = 1$, $\delta = 0.2$, $p = 0.25$, and $q = 0.65$. When the industry is structured as two mono-risk insurers, the optimal capital levels are $A_1^* = 0.78$ and $A_2^* = 1.53$ respectively, i.e., these are the levels of capital that maximize the respective expected utilities of the two insurees. The expected utilities for insuree 1 and 2 for these choices of mono-risk capital are $U_1^{MONO}(A_1^*) = -12.06$, and $U_2^{MONO}(A_2^*) = -933.7$. These utility levels are represented by the solid straight horizontal and vertical lines in Figure 2.

The curves in Figure 2 show the multi-risk insurance outcome as a function of capital, A, for correlations $\rho \in \{0.1, 0.2, 0.3, 0.4\}$. When correlations are low, the outcome can be improved for both insurees by moving to a multi-risk insurance structure, reaching an outcome somewhere on the efficient frontier of the multi-risk utility possibility curve. For $\rho = 0.4$, however, insuree 2 will not participate in the multi-risk insurance solution, regardless of capital, since the utility is always lower than what he achieves in a mono-risk offering. The mono-risk insurance outcome will therefore prevail.

Similar results hold if we allow for deductibles. Recall from Section 3.1 that we exclude deductibles for tractability in the general analysis. In this simplified setting, however, they are straightforward to include. Including risk-specific deductibles in the multi-risk insurance setting does introduce extra degrees of freedom in how risks may be shared in different states, whereas in the mono-risk insurance setting with Bernoulli distributed risks, the presence of deductibles is irrelevant (see Jaffee and Walden (2014)). The simplest possible extension of the model just analyzed, however, with one three-point risk distribution and one two-point distribution, leads to similar results when deductibles are included as those we just derived. We provide an example in the appendix.

So far we have interpreted $\tilde{L}$ as the risk faced by an individual insuree, but in many cases we may alternatively think of $\tilde{L}$ as a whole insurance line, or risk class, representing total risk faced by a large number of individual insurees with similar individual exposures. With this interpretation, the mono-risk and multi-risk market outcomes then represent a market with two monoline companies and a market with one multiline company, respectively.

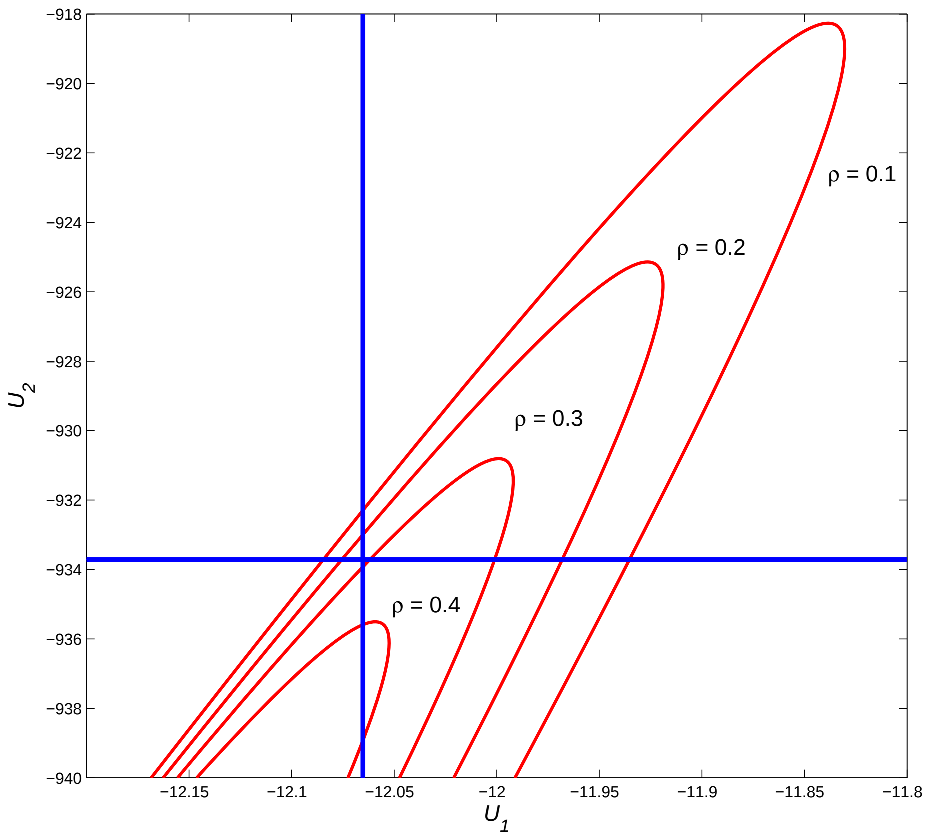

Figure 2: *The solid vertical and horizontal lines show optimal expected utility for insuree 1 and 2 respectively when the industry is structured as two mono-risk insurers. The curved lines show the utility combinations for a multi-risk insurer, based on 4 different correlations between risks 1 and 2, and for all possible capital levels, A. The mono-risk outcome dominates when* $\rho = 0.4$*, because the multi-risk structure is suboptimal for insuree 2. For* $\rho = 0.3$*,* $\rho = 0.2$ *and* $\rho = 0.1$*, the multi-risk insurance structure dominates since it is possible to improve expected utility for insuree 2, as well as for insuree 1. Parameters:* $p = 0.25$*,* $q = 0.65$*,* $\delta = 0.2$*,* $\beta = 7$*.*

Two potential issues arise with such an interpretation. The first is that of indivisibility: whereas it is natural to assume that the insurance of an individual risk cannot be split between multiple insurers, it may be possible for a whole line of risks to be split, i.e., a multiline insurer may choose to offer insurance against fractions of risk classes, e.g., selling insurance to half of the agents facing risks in one class, and against all the risks in the other class. In practice, such splitting may not be feasible, since insurance lines may be effectively indivisible because of high fixed costs of underwriting line risk.[16]

Even when splitting is feasible, the outcomes may not be affected. The results in our current example actually remain very similar even if line splitting is permitted. It is straightforward to show that when $\rho = 0.4$ in the previous case, *any* multiline insurance structure covering a fraction $0 < a < 1$ of the insurees in the first risk class and $0 < b < 1$ of those in the second will make either insurees in the first or in the second risk class worse off. Specifically, a multiline insurer offering insurance against $a\tilde{L}_1$ in line 1 and $b\tilde{L}_2$ in line 2 will be dominated by a monoline offering in either line 1 or line 2 (or both). So *no* multiline offerings dominate the monoline outcomes, even when such fractional offerings are possible (this argument also holds for the residual fraction of $1 - a$ policyholders in line 1 and $1 - b$ in line 2, of course).

The second potential issue with a line interpretation of the risks is whether a representative agent utility specification for the insurees in the line is feasible. If all agents within a line have identical utility functions and individual risks within the line are perfectly correlated, the representative agent interpretation is straightforward, as it is in the more general case in which only some agents have realized losses, but those that do have perfectly correlated loss realizations. In an even more general setting, Jaffee and Walden (2014) show that optimal insurance contracts are symmetric when many identical agents in a line face identical and symmetrically distributed risks. A consequence of this symmetry is that the ex ante expected utilities of all agents in such a line is the same, in turn allowing the construction of an (ex ante) representative agent utility function. In what follows, we take as given such a representative agent utility specification and proceed with line interpretations of the $\tilde{L}_i$ risks, stressing that the results, while always valid under an individual risk interpretation, rely on further assumptions under a line interpretation.

In Figure 3, we plot the regions in which the monoline and multiline structures occur, as a function of q and ρ, using the parameter values $p = 0.1$, $\beta = 1.2$, $\delta = 0.01$ and $t = 1$. The

[16] We thank a referee for making this point.

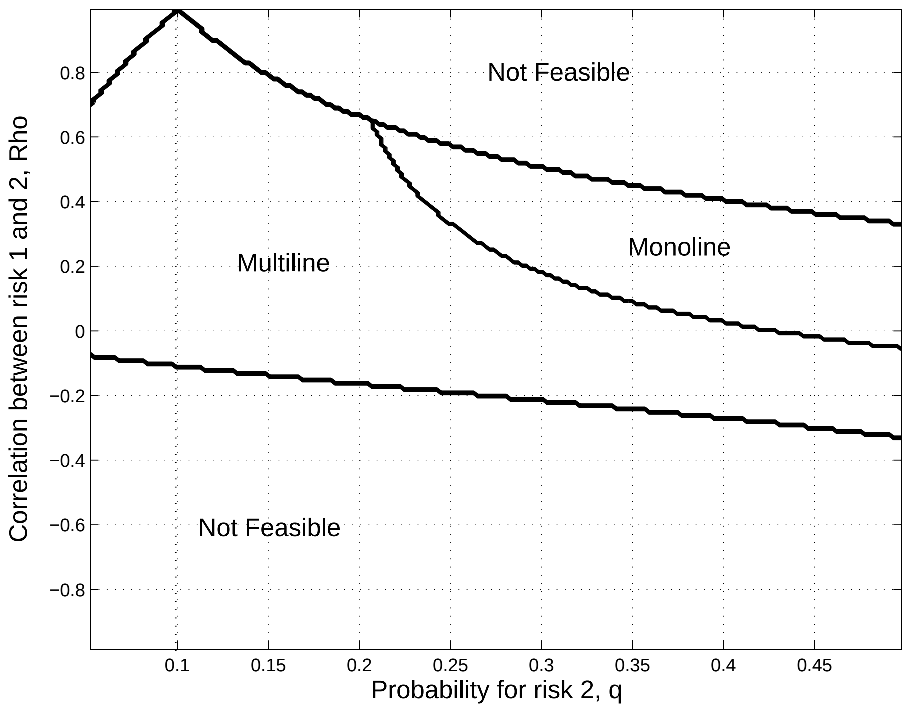

Figure 3: *Regions of q and ρ, in which monoline and multiline structure is optimal. All else equal: Increasing ρ (correlation), given q makes monoline structure more likely. Increasing asymmetry of risks $(q-p)$ also makes monoline structure more likely. Correlations can not be arbitrary for the two (Bernoulli) risks, so there are combinations of q and ρ that are not feasible. Parameters: $p = 0.1$, $\delta = 0.01$, $\beta = 1.2$.*

figure shows that, all else equal, increasing the correlation decreases the prospects for a multiline structure, as does increasing the asymmetry $(q-p)$ between risks. The two regions labeled "not feasible" arise because certain combinations of q and p are not possible for the assume parameter values.

The intuition behind these results is quite straightforward, once we understand the multiple forces at play. First, diversification — a major rationale for insurance in the first place — benefits a multiline structure. It allows the insurer to decrease the risk of default, given a constant level of capital per unit of risk insured. Alternatively, it allows the insurer to decrease the amount of capital reserved for a constant level of risk, decreasing the total cost of internal capital. So, from a diversification perspective, multiline structures should be more efficient than monoline structures.

But a multiline structure also forces insurees in the two lines to compromise on the level of capital. This effect benefits a monoline structure, which allows line-specific capital choices.

Also, when risk distributions are more asymmetric (here represented by an increasing difference between, q and p), insurees in one line may increasingly subsidize those in another. For example, insurees in low-risk property and casualty insurance lines may not be willing to group together with insurees in high risk, heavy-tailed, catastrophe insurance line, since their capital requirements may be quite different and, moreover, in case of insurer default, chances are high that losses in the catastrophe line will be very large, dwarfing any claims in the other line.

Finally, when correlation (ρ) between lines increases, a monoline structure becomes relatively more preferable. This is in line with the intuition that the diversification benefits of a multiline structure decreases when risks are positively correlated and, all else equal, the multiline structure therefore becomes relatively less beneficial.

Clearly, the tradeoff between these forces is nontrivial, and we do not expect a complete equilibrium characterization to be feasible in the general case. In the next section, we analyze the case with many risk classes and show that in this case, when risks have limited asymmetry and dependence, the typical outcome is one with many multiline structures, although there may still be a role for a few monoline structures.

5 Equilibrium Market Structure With Many Risk Classes

We extend the concepts from the previous section to markets with multiple risks classes.

5.1 Definitions and Efficiency

Consider a market, in which $M \geq 1$ insurers provide insurance against $N \geq M$ risk classes, each risk class held by a representative insuree. The set of risk classes is $X = \{1, \ldots, N\}$, which is partitioned into $\mathcal{X} = \{X_1, X_2, \ldots, X_M\}$, where $\cup_i X_i = X$, $X_i \cap X_j = \emptyset, i \neq j$, $X_i \neq \emptyset$. The partition represents how the risks are insured by M monoline or multiline insurers. The total industry structure is characterized by the duple, $\mathcal{S} = (\mathcal{X}, \mathbf{A})$, where $\mathbf{A} \in \mathbb{R}_+^M$ is a vector with ith element representing the capital available in the firm that insures the risks for agents in X_i. We call $\mathbf{A}$ the capital allocation and $\mathcal{X}$ the industry partition. If we want to stress the set of insurees included in a market, we write $\mathcal{S}_X$.

The number of sets in the industry partition is denoted by $M(\mathcal{X})$. Two polar cases are the fully multiline industry partition, $\mathcal{X}^{MULTI} = \{\{1, \ldots, N\}\}$ and the monoline industry partition,

$\mathcal{X}^{MONO} = \{\{1\}, \{2\}, \ldots, \{N\}\}$. Of course, for a fully multiline industry structure, $M = 1$ and $\mathbf{A} = A$. Given an industry structure, $\mathcal{S}$, the premium in each line is uniquely defined through Equation 4, and we write $P_i = P_i(\mathcal{S})$.

Note that when a monoline insurer chooses capital $A = 0$, this is equivalent to the insuree not purchasing insurance ($P_i = 0$, $\tilde{Q}_i \equiv \tilde{L}_i$). We identify such an outcome in which there is no market for the ith insurance line with a monoline insurer that chooses $A = 0$.

For N risks, $\tilde{L}_1, \ldots, \tilde{L}_N$, and a general industry structure, $\mathcal{S} = (\mathcal{X}, \mathbf{A})$, when the ex post sharing rule is used, the residual risk for an insuree (i.e., the net risk after claims are received), $i \in X_j$, is then

$$\tilde{K}_i(\mathcal{S}) = \frac{\tilde{L}_i}{\sum_{i' \in X_j} \tilde{L}_{i'}} \min\left(A_i - \sum_{i' \in X_j} \tilde{L}_{i'}, 0\right), \tag{6}$$

and his expected utility is $U_i(\mathcal{S}) = Eu_i(-P_i(\mathcal{S}) + \tilde{K}_i(\mathcal{S}))$.

We use Pareto dominance to rank different industry structures. For N agents with utility functions, u_i, $1 \leq i \leq N$, where each agent wishes to insure risk $\tilde{L}_i$, an industry structure, $\mathcal{S}'$, Pareto dominates another structure, $\mathcal{S}$, if $E[u_i(-P_i(\mathcal{S}) + \tilde{K}_i(\mathcal{S}))] \leq E[u_i(-P_i(\mathcal{S}') + \tilde{K}_i(\mathcal{S}'))]$ for all i and $E[u_i(-P_i(\mathcal{S}) + \tilde{K}_i(\mathcal{S}))] < E[u_i(-P_i(\mathcal{S}) + \tilde{K}_i(\mathcal{S}'))]$ for at least one i. We also say that $\mathcal{S}'$ is a Pareto improvement of $\mathcal{S}$. An industry structure, $\mathcal{S}$, for which there is no Pareto improvement is said to be Pareto efficient. An industry structure, $(\mathcal{X}, \mathbf{A})$, is said to be constrained Pareto efficient (given $\mathcal{X}$), if there is no $\mathbf{A}'$ such that $(\mathcal{X}, \mathbf{A}')$ is a Pareto improvement of $(\mathcal{X}, \mathbf{A})$. An industry partition $\mathcal{X}$ is said to be Pareto efficient if there is a capital allocation, $\mathbf{A}$, such that $(\mathcal{X}, \mathbf{A})$ is Pareto efficient. Given the finite expected utility for all insurees, the continuity of expected utility as a function of capital for any given multiline structure that follows from Equation 6, the linear cost of capital, and the finite number of possible industry structure partitions, it follows that there is a constrained Pareto efficient outcome for any given industry partition, $\mathcal{X}$, and that the set of Pareto efficient industry structures, which we denote by $\mathcal{P}$, is nonempty.

5.2 Feasible Outcomes

When there are numerous independent risk classes in the economy, asymptotic analysis becomes feasible. Our first objective is to understand how powerful diversification is in providing value to the insurees in this case, which we analyze in the next section. We then analyze what the implications are for equilibrium outcomes in the following two sections. We use the certainty equivalent as a measure of the size of a risk facing an insuree. For a specific utility function, u, the certainty equivalent of risk $\tilde{L}$, $CE_u(-\tilde{L}) \in \mathbb{R}_-$ is defined such that $u(CE_u(-\tilde{L})) = E[u(-\tilde{L})]$.

When capital is costly, $\delta > 0$, it is not possible to obtain the friction-free outcome, in which full insurance is offered at the price of expected losses, μ_L. To ensure that the friction cost is not so high as to rule out insurance coverage, we assume that the cost of holding capital for each agent i is sufficiently small compared with expected losses, such that

Condition 1 $CE_{u_i}(-P(A) - \tilde{Q}(A)) < -\mu_{L_i}(1+\delta)$ *for all* $A \in [0, \mu_{L_i}]$.

Condition 1 implies that the risk is potentially insurable. Specifically, the right-hand-side of Condition 1 represents the *ideal risk-free outcome* with costly internal capital, in which the ith insuree would pay $\mu_{L_i}(1+\delta)$ to completely offload the risk to the insurer. In practice, an insurer holding capital equal to the expected loss would still create a counterparty risk, so this is indeed an ideal risk-free outcome. The condition implies that both no insurance (corresponding to $A = 0$ on the left-hand-side) and insurance under a monoline structure (corresponding to $A > 0$) are dominated by the ideal risk-free outcome.

Under Condition 1, some risk classes may be uninsurable when there is a limited number of insurance lines, e.g., if the cost of internal capital is high. Such outcomes are avoided by imposing the following conditions:

Condition 2 *For all i in the monoline industry partition, optimal capital is strictly positive, i.e., for all i, $A_i^* > 0$.*

In our subsequent analysis we will sometimes impose a slightly stronger condition:

Condition 3 *For all i in the monoline industry partition, optimal capital is uniformly strictly positive, i.e., there is a constant, $\epsilon > 0$, such that for all i, $A_i^* > \epsilon$.*

Note that Condition 3 implies Condition 2, but that Condition 1 neither implies, nor is implied, by any of Conditions 2 and 3.

What can we say about the efficiency of different industry structures? Intuitively, when capital is costly and there are many risks available, we would expect an insurer to be able to diversify by pooling many risks and, through the law of large numbers, choose an efficient A^*

per unit of risk. Therefore, the multiline structure should be more efficient in mitigating risk than the monoline structure.[17] The argument is very general, as long as there are enough risks to pool, these risks are independent, and some technical conditions are satisfied. We have:

Theorem 1 *Consider a sequence of insurees, $i = 1, 2, \ldots$, with expected utility functions, $u_i \equiv u$, holding independent risks $\tilde{L}_i$. Suppose that u'' is bounded by a polynomial of degree q, that $E[\tilde{L}_i^p] \leq C$ for $p = 2 + q + \epsilon$ and some $C, \epsilon > 0$, and $E(\tilde{L}_i) \geq C'$, for some $C' > 0$. Then, regardless of the cost of internal capital, $0 < \delta < 1$, as the number of risks in the economy, N, grows, a fully multiline industry, $\mathcal{X}^{MULTI} = \{\{1, \ldots, N\}\}$ with capital $A = \sum_{i=1}^{N} \mu_{L_i}$, reaches an outcome that converges to the ideal risk-free outcome with costly internal capital, i.e.,*

$$\min_{1 \leq i \leq N} CE_u(-P_i((\mathcal{X}^{MULTI}, A)) + \tilde{K}_i((\mathcal{X}^{MULTI}, A))) = -\mu_{Li}(1 + \delta) + o(1).$$

It follows that in large economies that satisfy the assumptions of Theorem 1, and in which Condition 1 is satisfied, multiline structures will asymptotically dominate both monoline structures and no insurance. A consequence of the theorem is that multiline structures may have the potential to serve insurance lines in a market that would fail with monoline structures.

Theorem 1 can be further generalized in several directions, e.g., to allow for dependence. For example, as follows from the proof of the Theorem, it also holds for all (possibly dependent) risks $\tilde{L}_i$ with $E|\tilde{L}_i|^p < C$ that satisfy the Rosenthal inequality.[18]

Theorem 1 shows that with enough risks, a solution can be obtained arbitrarily close to the ideal risk-free outcome with costly internal capital. In the Appendix, we also show the opposite result, namely that with too few risks it is not possible to get arbitrarily close to the ideal

[17] This type of diversification argument is, for example, underlying the analysis and results in both Jaffee (2006) and Lakdawalla and Zanjani (2012).

[18] The Rosenthal inequality (see Rosenthal (1970)) and its analogues are satisfied for many classes of dependent random variables, including martingale-difference sequences (see Burkholder (1973) and de la Peña, Ibragimov, and Sharakhmetov (2003) and references therein), many weakly dependent models, including mixing processes (see the review in Nze and Doukhan (2004)), and negatively associated random variables (see Shao (2000) and Nze and Doukhan (2004)). Furthermore, using the Phillips-Solo device (see Phillips and Solo (1992)) in a similar fashion of the proof of Lemma 12.12 in Ibragimov and Phillips (2008), one can show that Theorem 1 also holds for correlated linear processes $\tilde{L}_i = \sum_{j=0}^{\infty} c_j \epsilon_{i-j}$, where (ϵ_t) is a sequence of i.i.d. random variables with zero mean and finite variance and c_j is a sequence of coefficients that satisfy general summability assumptions. Several works have focused on the analysis of limit theorems for sums of random variables that satisfy dependence assumptions that imply Rosenthal-type inequalities or similar bounds (see Serfling (1970), Móricz, Serfling, and Stout (1982) and references therein). Using general Burkholder-Rosenthal-type inequalities for nonlinear functions of sums of (possibly dependent) random variables (see de la Peña, Ibragimov, and Sharakhmetov (2003) and references therein), one can also obtain extensions of Theorem 1 to the case of losses that satisfy nonlinear moment assumptions.

risk-free outcome with costly internal capital. The two results together illuminate the power of diversification in eliminating counterparty risk, as long as there are a sufficient number of risk classes and the loss distributions are sufficiently well behaved. They also address the issue of risk of default driven by other risk classes that we saw in the previous section, since no such risk is present after full diversification.

The results do not, however, address the question of optimal industry structure when capital levels are endogenous. In fact, it is straightforward to show that the ideal risk-free outcome with costly internal capital leads to a capital level that is "too high" in the sense that insurees would always prefer a lower level of capital than the risk-free level. The intuition behind this result is simple. The marginal utility benefit of decreasing capital, A, slightly below the ideal risk-free outcome, in terms of reduced cost of capital, will always outweigh the marginal cost of introducing some risk, since the former factor is a first order effect, whereas the latter factor is of second order close to the ideal risk-free outcome. With endogenous capital levels, the situation becomes more complicated, because we expect insurees to have different views on how large the capital reductions from the asymptotically risk-free outcome should be. It is a priori unclear which industry structures will prevail in this case, a question we focus on henceforth.

5.3 A Strategic Game

In the multiline case with $N > 2$, there are many possible industry structures. We wish to extend the concept of dominated structures that we used in the two risk-class example to such a general multiline setting. We note that the arguments made in the Section 4 are quite similar to those in coalition games without transferable payoffs, and especially to those in the literature on hedonic coalition games, see Dreze and Greenberg (1982), Osborne and Rubinstein (1984), Greenberg and Weber (1993), Demage (1994), Bogomolnaia and Jackson (2002), and Banerjee, Konishi, and Sönmez (2001). Specifically, an insuree-centered interpretation of the comparison between monoline and multiline outcomes in Section 4 is as a decision between the two insurees about whether to form a "coalition" or not. Such a coalition dominates if both insurees are better off than they would be alone — that is, than what they would be under the monoline structure. In other words, if there are choices of capital for which the multiline outcome Pareto dominates the monoline outcome, this outcome will prevail because neither insuree has an incentive to block the multiline outcome by leaving the multiline coalition. The multiline outcome is therefore *core* *stable*, see Greenberg and Weber (1993), Demage (1994), and Bogomolnaia and Jackson (2002).

In the general case with multiple insurance lines ($N > 2$), the core stability concept requires that there is no alternative coalition that blocks a stable outcome.

Definition 1 *An industry structure $\mathcal{S}_X$ is said to be* robust to all blocking *if there is no set of insurees, $Y \subset X$, such that an insurer can make a fully multiline offering to all insurees in Y by choosing some capital level A, so that all these insurees are at least as well of as they were before, and at least one insuree is strictly better off, i.e., $U_i(\mathcal{S}_Y) \geq U_i(\mathcal{S}_X)$ for all $i \in Y$ and the inequality is strict for at least one $i \in Y$. Here, $\mathcal{S}_Y = (\{Y\}, A)$.*

Definition 2 *The* core stable set *of industry structures, $\mathcal{C}$, is defined as the set of industry structures that are robust to all blocking.*

In general, strong assumptions are needed for a characterization of $\mathcal{C}$ (e.g., to guarantee nonemptiness) in hedonic coalitions games. In our setting, however, we argue that arbitrary blocking "coalitions" of an industry structure would be difficult to achieve. Specifically, it would be challenging for a new entrant to capture specific insurance lines from several different insurers in a coordinated effort. We therefore impose weaker robustness to blocking requirements than in Definition 1.

We focus on two specific types of robustness which we believe are especially important in an equilibrium industry structure: robustness to aggregation and to monoline offerings. Aggregation occurs by merging all the lines of two or more insurers. This outcome could, e.g., be achieved if one insurance firm acquired the other. Indeed, if a more efficient offering could be made to all insurees of the merged firm, we would expect competition to lead to such mergers. Similarly, we assume robustness to monoline offerings, based on the argument that a competing monoline insurer would attract the customers in a line with an improved offering, if such an offer existed. By focusing on these two types of robustness, we are implicitly assuming that more complex competitive strategies by competitors are difficult to implement.

We formalize these concepts with the following definitions:

Definition 3 *An industry structure, $\mathcal{S}_X$, is said to be* robust to aggregation*, if there is no set of insurers in $\mathcal{X}$, $Y = \cup_{i=1}^{n} X_{k_i}$, $X_{k_i} \in \mathcal{X}$, $k_1 < k_2 \cdots < k_n$, $n \geq 1$, such that a fully* *multiline offering to all insurees in Y with some capital level A can be made that makes all these insurees at least as well of as they were before, and at least one insuree strictly better off, i.e., $U_i(\mathcal{S}_Y) \geq U_i(\mathcal{S}_X)$ for all $i \in Y$ and the inequality is strict for at least one $i \in Y$. Here, $\mathcal{S}_Y = (\{Y\}, A)$.*

Definition 4 *The set of industry structures that are robust to aggregation is denoted by $\bar{\mathcal{O}}$.*

Definition 5 *An industry structure, $\mathcal{S}$, is said to be* robust to monoline blocking, *if there is no insuree, $i \in \{1, \ldots, N\}$ such that $U_i(\mathcal{S}) < U_i^{MONO}(A)$ for some $A \geq 0$, where $U_i^{MONO}(A)$ is the expected utility insuree i achieves under a monoline offering with capital A.*

Definition 6 *The* equilibrium set, *$\mathcal{O}$, is defined as the set of industry structures that are robust to both monoline blocking and aggregation.*

Note that to dismiss a candidate equilibrium outcome, it is sufficient for one monoline structure to dominate that outcome. So, for example, in an economy with three lines, if a candidate equilibrium is $\mathcal{X} = \{\{1, 2\}, \{3\}\}$, and insurance line 1 is made better off by a monoline offering, then $\mathcal{X}$ cannot be an equilibrium outcome, even if line 2 is worse off under either of the alternative structures $\{\{1\}, \{2, 3\}\}$ or $\{\{1\}, \{2\}, \{3\}\}$.

Several implications of these definitions follow immediately. First, robustness to aggregation implies constrained Pareto efficiency, since an insurer covering risk classes X_i, who chose a Pareto inefficient level of capital could otherwise be improved upon by choosing $Y = X_i$ with a superior capital level. Second, Pareto efficient outcomes are robust to aggregation, $\mathcal{P} \subset \bar{\mathcal{O}}$, since otherwise they could be improved upon by such aggregation. Third, it follows from the definitions that $\mathcal{C} \subset \mathcal{O} \subset \bar{\mathcal{O}}$. Naturally, robustness to all blocking is a stronger condition than joint robustness to monoline blocking and aggregation, since it implies that there is no structure that dominates for *some* (not all) the insurees of one or more insurers.

The restrictions on the equilibrium set are quite weak, which implies that it is never empty:[19]

Proposition 1 *The equilibrium set is nonempty, $\mathcal{O} \neq \emptyset$.*

[19] This result is easily seen. Given a constrained Pareto efficient industry structure, $\mathcal{S}^0$, which is robust to monoline blocking, it is either the case that $\mathcal{S}^0 \in \mathcal{O}$, or that a Pareto improvement can be achieved by aggregation, leading to a new constrained Pareto efficient industry structure, $\mathcal{S}^1$. This new structure is obviously also robust to monoline blocking. The argument can be repeated and since the number of risk classes is finite, it must terminate for some $\mathcal{S}^n \in \mathcal{O}$. $\mathcal{S}^0$ can now be chosen to be a constrained Pareto efficient *monoline* structure, which per definition is robust to monoline blocking, and the result thus follows.

The weak assumptions are a major strength of our results. One may of course wish to impose additional constraints to further restrict the possible industry outcomes. Our main results on industry structure will hold for all elements in $\mathcal{O}$, so they will of course then also hold for any subset of $\mathcal{O}$. For example, since $\mathcal{C} \subset \mathcal{O}$, all results that hold for industry structures in the equilibrium set will also hold for industry structures in the core stable set. There is no guarantee that the core stable set is nonempty though. In fact, we may easily imagine a situation with three risk classes, in which it is optimal for two of the insurees to be insured by a multiline company. This situation is similar to a majority game, since any structure with two insurees may be blocked by one in which one of the two insurees is replaced by the third. The core stable set is therefore empty in this case. The analysis of the conditions under which the core stable set is nonempty in our setting provides an interesting avenue for future research.

We stress that although we use terminology from coalition games in our analysis, our market mechanism is *not* based on coalitions. It is the competitiveness of firms that ensure outcomes that can not be improved by other industry structures.

5.4 Equilibrium Outcome

We introduce the following definitions: An industry partition is said to be massively multiline if, as the number of lines in the industry, N, grows, the average number of lines per insurer grows without bounds, i.e., $\lim_{N\to\infty} N/M(\mathcal{X}) = \infty$.[20] Here, $M(\mathcal{X})$, defined in Section 5.1, is the number of insurers in the industry. We let $\mathcal{X}^{MASS}$ (or $\mathcal{X}_N^{MASS}$, if we want to stress the number of risks) represent a massively multiline industry partition.

For some of the results, we need to ensure that the risks are not too "asymmetric," by imposing the following condition:

Condition 4 *There are positive, functions, $g : \mathbb{R}_+ \to \mathbb{R}_+$ and $h : \mathbb{R}_+ \to \mathbb{R}_+$, such that g is strictly decreasing, h is nonincreasing, $\lim_{x\to\infty} h(x) = 1$, and for all i,*

$$F_i(x) \in [g(h(x)x), g(x)].$$

Here, $F_i(x)$ is the complementary cumulative distribution function of $\tilde{L}_i$, $F_i(x) \stackrel{\text{def}}{=} \mathbb{P}(\tilde{L}_i > x)$.

[20] Strictly speaking, we are analyzing a sequence of industries, with a growing number of risk classes.

Note that in the special case when the risks are i.i.d., $g = F_1$ and $h(x) \equiv 1$ can be chosen in Condition 4.

Our first result shows that monoline structures will be rare in large economies:

Theorem 2 *Under the conditions of Theorem 1, if the risks, $\tilde{L}_i$, have absolutely continuous distributions with strictly positive probability density functions on $\mathbb{R}_+$, then for large N,*

1. *If Conditions 3 and 4 are satisfied, there is a constant $C < \infty$ that does not depend on N, such that any industry structure in $\bar{\mathcal{O}}$ has at most C monoline insurers.*

2. *If Condition 1 is satisfied, then the fully multiline industry partition, $\mathcal{X}^{MULTI}$, with capital $A = \sum_i \mu_{L_i}$ Pareto dominates the fully monoline insurance structure, $\mathcal{X}^{MONO}$.*

The first part of Theorem 2 states that in large economies any industry structure that is robust to aggregation must contain very few monoline structures. Note that since both the equilibrium set, $\mathcal{O}$, and the set of Pareto efficient outcomes, $\mathcal{P}$, are subsets of $\bar{\mathcal{O}}$, this immediately implies that the property holds for both equilibrium industry structures and Pareto efficient industry structures. The second part of Theorem 2 states that when Condition 1 is satisfied, of the two opposite extremes—a fully multiline industry structure and an industry structure with only monolines—the fully multiline structure dominates.

The previous result implies that monoline structures are rare in large economies. We are also interested in the degree to which equilibrium structures will be massively multiline. We have:

Theorem 3 *Under the conditions of Theorem 1, if the risks, $\tilde{L}_i$, have absolutely continuous distributions with strictly positive probability density functions on $\mathbb{R}_+$, then for large N,*

1. *If Conditions 3 and 4 are satisfied, there is a massively multiline industry partition in the equilibrium set, $\mathcal{X}^{MASS} \in \mathcal{O}$.*

2. *If the risks are i.i.d. and Condition 3 is satisfied, the fully multiline industry partition is in the equilibrium set, $\mathcal{X}^{MULTI} \in \mathcal{O}$.*

Thus, the equilibrium structure that we may expect to see, driven by aggregation, is one which mainly consists of very large multiline structures.

The difference between the first and the second result in Theorem 3 is important. If risks are identically distributed, then the agents, having the same utility functions, will all agree upon the optimal level of internal capital, A^*. They may therefore agree to insure in one fully multiline company. If, on the other hand, the risks have different distributions (or equivalently, if the utility functions are different), then the insurees will typically disagree about what is the optimal level of internal capital. Recall that increasing capital has two offsetting effects. It decreases the risk of insurer default and thereby increases the expected utility of the insurees, but it also increases the total cost of internal capital, which decreases their expected utility. With many different types of risks, it will therefore be optimal to have several massively multiline structures that choose different levels of internal capital, instead of one fully multiline structure.

An Example

To see that the fully multiline outcome may indeed not be in the equilibrium set, even when very many risks are present, consider the following example in which the first but not the second part of Theorem 3 is satisfied: There are N independent Bernoulli distributed risk classes, $\mathbb{P}(\tilde{L}_i = 1) = p$, $\mathbb{P}(\tilde{L}_i = 0) = 1 - p$, $i = 1, \ldots, N$, and additionally N independent risk-classes with scaled Bernoulli distributions, $\mathbb{P}(\tilde{L}_i = 2) = q$, $\mathbb{P}(\tilde{L}_i = 0) = 1 - q$, $i = N + 1, \ldots, 2N$. All insurees have quadratic utility functions $u(x) = x - \rho x^2$, $x \leq 0$, $\rho > 0$. Here, N is a very large number — so large so that perfect diversification is basically achieved if all N risk classes are insured by the same insurer.[21] The parameters are $p = 19/25$, $q = 1/2$, $\delta = 1/8$, and $\rho = 2$.

In this case, a massively multiline structure in which there are two insurers, one covering all risk classes of the first type, and one covering all risk classes of the second, belongs to the equilibrium set, whereas the fully multiline outcome, in which one insurer covers all risks, does not belong to the equilibrium set. This can be seen in Figure 4. With the massively multiline outcome, capital of $A_1 = N \times 0.060$ is reserved by the insurer who insures the first N (Bernoulli) risks, and $A_2 = N \times 0.74$, is chosen by the insurer who insures the latter N (scaled Bernoulli) risks, achieving expected utilities of $U_1^{MASS} = -1.139$ and $U_2^{MASS} = -1.72$, for agents insuring the first and second risk classes, respectively. These utilities are shown by the two dash-dotted lines in the figure.

Both these outcomes are robust to monoline blocking, since the optimal monoline offerings achieve lower expected utilities of $U_1^{MONO} = -1.14$ and $U_2^{MONO} = -1.88$, respectively. These

[21] We show the asymptotic results, as N tends to infinity. Since all variables almost surely converge to this asymptotic result as N grows, for large enough finite N the asymptotic results hold in large but finite industries too.

utilities of the monoline outcomes, which are shown as solid lines in Figure 4, occur at capital levels of $A_1^{MONO} = 0$ and $A_2^{MONO} = 0.94$, respectively.

The massively multiline outcome is also robust to aggregation. To see this, note that the solid curve in the figure, which represents possible expected utility of the two insuree types as a function of capital level, is below the massively multiline outcome for the second type of insurees at all capital levels, except for $A = 0.49$, at which point it is lower for the first type of insuree (about -1.18, compared with -1.139 for the massively multiline outcome). Thus, the massively multiline outcome with two insurees does indeed belong to the equilibrium set, $\mathcal{O}$.

Finally, to see that the fully multiline outcome does *not* belong to the equilibrium set, note that for all choices of capital, it is either the case that the expected utility of an insuree of type 1 is lower than that obtained in a monoline offering, or the expected utility of an insuree of type 2 is lower, or both. Thus, monoline blocking is always possible, regardless of A, and the fully multiline outcome therefore does not belong to $\mathcal{O}$.

The results in Theorems 2 and 3 together suggest that when there is a large number of essentially independent risks that are thin-tailed, a monoline insurance structure is never optimal and that massively multiline industry structures may instead occur. For standard risks — like auto and life insurance — it can be argued that these conditions are reasonable. In contrast, a multiline structure is more likely to be suboptimal if risks are not numerous, of they are dependent, and/or asymmetric across risk classes. The intuition of why is clear. First, if there are few insurance lines, the diversification benefits may be limited. Second, if risks are dependent and/or asymmetric, the negative externalities faced by policy holders in one insurance line in case of insurer default are more severe. In both cases, the multiline outcome becomes relatively less advantageous.

5.5 Potential policy implications

Our results have several potential policy implications. Catastrophe risks, in particular, appear to satisfy all these conditions under which a multiline structure may be suboptimal. Consider, for example, residential insurance against earthquake risk in California.[22] The outcome for different households within this area will obviously be heavily dependent when an earthquake occurs, making the pool of risks essentially behave as one large risk, without diversification benefits.

[22]See, e.g., Ibragimov, Jaffee, and Walden (2009).

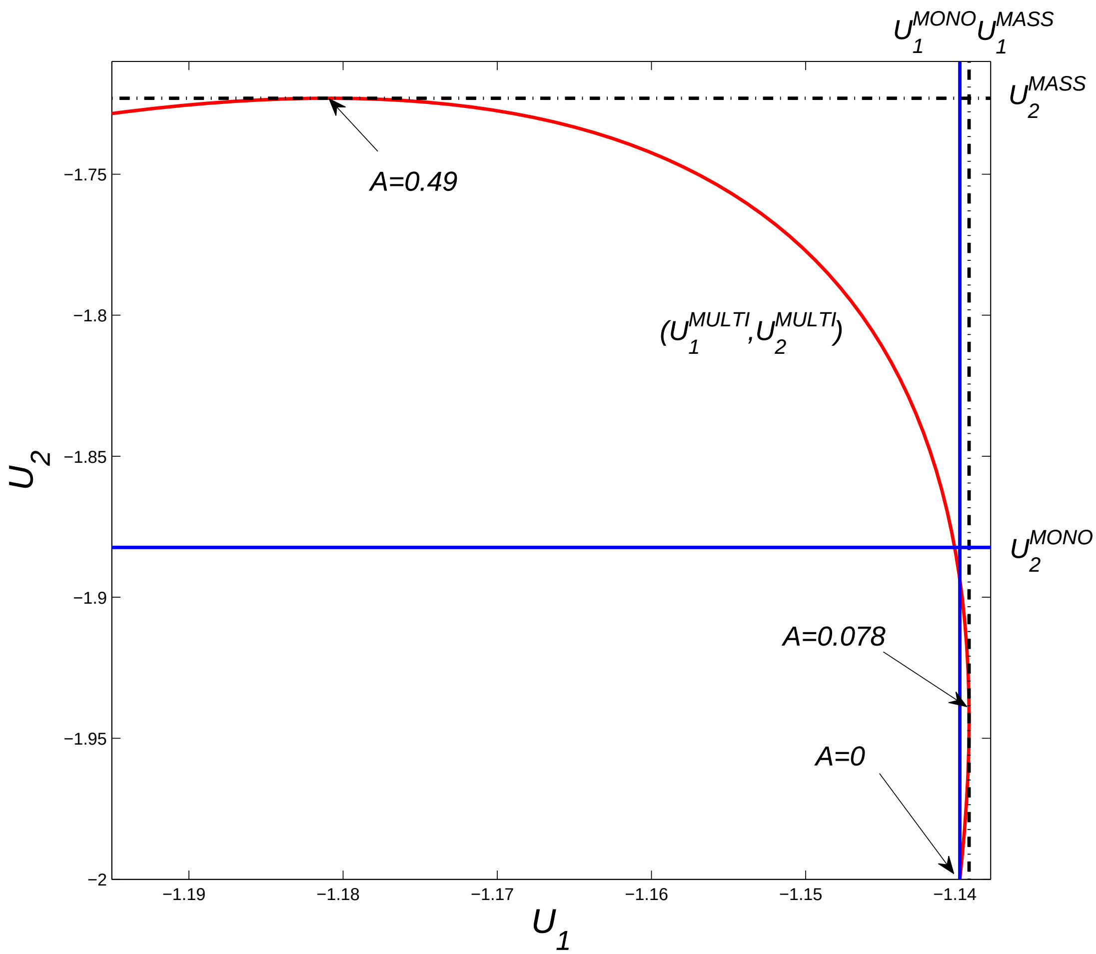

Figure 4: *Example where a massively multiline industry structure belongs to the equilibrium set, but not the fully multiline structure. The massively multiline structures (dash-dotted lines) are robust to monoline blocking (solid lines) since their utility is higher. Further, they are robust to aggregation, since the utility of one type of insurees is always dominated in the fully multiline outcome by at least one monoline offering, regardless of amount of capital (solid curve is always below one of solid lines, regardless of A). Parameters:* $p = 19/25$, $q = 1/2$, $\delta = 1/8$, $\rho = 2$.

Moreover, many other catastrophic risks are known to have heavy tails, further reducing their diversification benefits. So, even though an earthquake in California and a hurricane in Florida, and a flood on the Mississippi river may be considered independent events, the gains from diversification of such risks may be limited due to their heavy-tailedness. The risk distributions in these catastrophe insurance lines are thus quite different compared with regular lines, e.g., in property and casualty.

Interestingly, most California earthquake insurance is provided through the quasi-public California Earthquake Authority, most hurricane risk in Florida is reinsured through the state's Florida Hurricane Catastrophe Fund, and most U.S. flood insurance is provided by the National Flood Insurance Program. All three of these entities are monoline in the sense that they provide insurance or reinsurance only for the designated catastrophe. Note, however, that it is not clear that such a monoline outcome for catastrophe insurance will prevail in equilibrium. Specifically, a multiline insurance company may serve several lines, one of which is a catastrophe line that is subsidized by the other lines. The multiline structure may still be robust to monoline blocking since each insuree in the other lines may be worse off when served by a monoline structure, although excluding the catastrophe line would make all the remaining insurees better off. A policy maker wishing to avoid such cross-subsidization may choose to regulate catastrophe insurance lines so that they need to served by monoline structures.

There may also be situations in which a policy maker wishes to impose a mandatory multiline structure, e.g., to create markets for risks that would otherwise be uninsurable. Specifically, since the set of Pareto efficient structures, $\mathcal{P}$ is a subset of $\bar{\mathcal{O}}$, but in general not a subset of the equilibrium set $\mathcal{O}$, there may be efficient outcomes that cannot be sustained as equilibrium outcomes without regulation. Consider, for example, a situation where a risk class is uninsurable by a monoline insurance structure (maybe because of a too high cost of internal capital), but it could be insured by a company with two lines. If, however, the policyholders in the second line would be better served by a monoline insurer, the multiline outcome is infeasible, not being robust to monoline blocking. In such a case, regulation may be needed to sustain such an outcome. The 2002 Terrorism Risk Insurance Act (TRIA), mandating insurance companies to offer insurance against terrorist events, its extensions in 2005 and 2007, and its 2015 successor, the Terrorism Risk Insurance Program Reauthorization Act (TRIPRA), may be interpreted in light of such market failures.

Finally, our results have potentially important applications to the re-regulation of the banking industry in the aftermath of the subprime mortgage crisis. Two points, in particular, stand out from our analysis. First, while it is common to propose higher bank capital requirements, it is often not recognized that these higher requirements create larger excess costs of internal capital, some or all of which will be paid by bank depositors. Second, as an alternative regulatory device, it is useful to consider monoline restrictions for those bank lending or investment activities that pose heavy-tailed risks and thus create extraordinary counterparty risk for the bank debt holders. One form would be to require that these activities be carried out only within a separate holding-company subsidiary, so that bank debt holders would be unaffected if that monoline subsidiary were to fail. Part of the Dodd-Frank regulation is in line with this framework, e.g., the Volcker rule for the reform of U.S. prudential banking regulation, see Volcker (2010). Moreover, our analysis may be applicable to the previously discussed question of single-point-of-entry (SPOE) versus multiple-point-of-entry (MPOE) resolution for global financial institutions.

6 Concluding Remarks

This paper develops a model to analyze monoline versus multiline structures in markets for risk, under the assumptions of limited liability, costly internal capital, and a limited contracting space. We derive three important properties of such markets. First, we show that multiline structures dominate when the benefits of diversification are achieved because the underlying lines are numerous and independent. Second, even with such numerous independent lines, the industry may be served by several multiline companies, each holding a different amount of capital, as opposed to one multiline company serving all lines. There is thus a limit to the benefit of diversification even in this case with numerous nicely behaved risks. Third, the monoline structure may be efficient when risks are difficult to diversify because they are limited in number and/or heavy-tailed, as is a common characteristic of various catastrophe lines.

Our results are consistent with what is observed in practice in the insurance industry. Consumer lines such as homeowners and auto insurance are dominated by multiline insurers, while the catastrophe lines of bond and mortgage default insurance are available only on a monoline basis. Furthermore, it is a feature of our model that the default probability for a monoline bond or mortgage default insurer would likely be higher than it is for a multiline firm, offering coverage only on highly diversifiable lines.

Our results are also applicable more broadly within the financial services industry, shedding light on the original Glass Steagall Act and its rebirth within the Dodd-Frank Act as the so-called Volcker Rule. The banking application is that a multiline bank—say one that combines a traditional deposit and lending business with much riskier investment banking activities—may achieve a risk transfer from the bank's equity owners to the bank's creditors, including its depositors and/or the government's deposit insurance entity. Regulators may thus find it beneficial to ring-fence the traditional banking business from these riskier investment banking activities using what is in effect a monoline restriction.

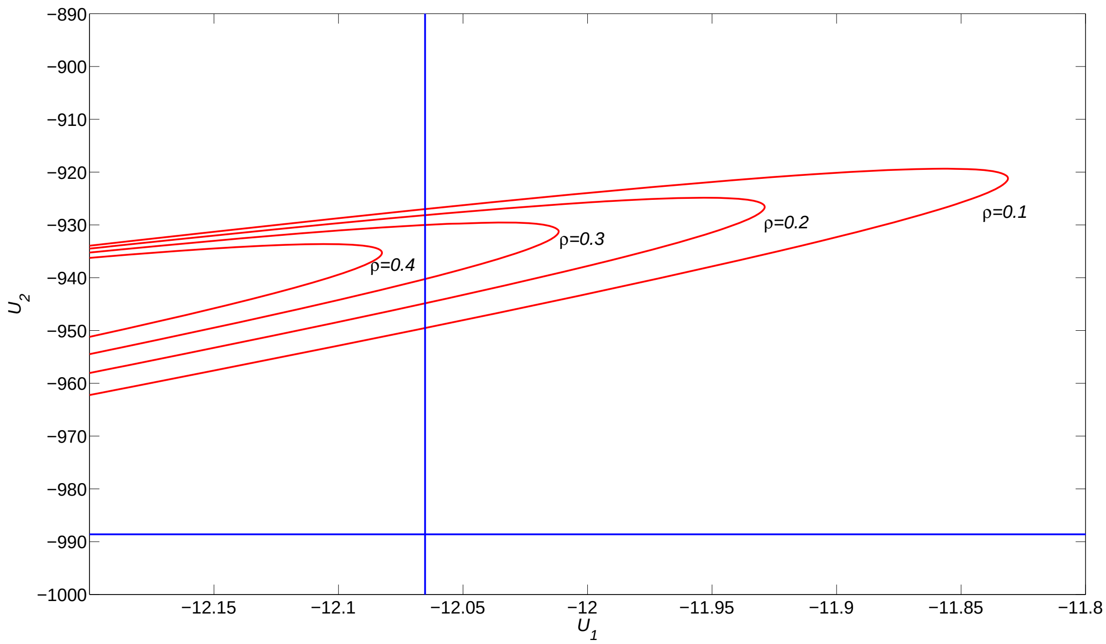

Figure 5: *The solid vertical and horizontal lines show optimal expected utility for insuree 1 and 2 respectively when the industry is structured as two mono-risk insurers. The curved lines show the utility combinations for a multi-risk insurer, based on 4 different correlations between risks 1 and 2, for all possible capital levels, A, and with optimal deductibles. The mono-risk outcome dominates when* $\rho = 0.4$, *because the multi-risk insurance structure is suboptimal for insuree 1. For* $\rho = 0.3$, $\rho = 0.2$ *and* $\rho = 0.1$, *the multi-risk structure dominates since it is possible to improve expected utility for insuree 2, as well as for insuree 1. Parameters:* $p = 0.249$, $q = 0.65$, $\delta = 0.2$, $\beta = 7$, $p_2 = 0.001$.

A Two-risk example with deductibles

We extend the example in Section 4 to include deductibles. As mentioned in the main text, since both the risks in the example of Section 4 have two-point distributions, there are only three states that need to be insured (loss for risk one, loss for risk two, and loss for both risks), and with 3 degrees of freedom in the multi-risk insurance setting (capital level, deductible in line 1, and deductible in line 2), it is possible to design an outcome that meets both insurees' demands well in all loss states for this specific example.

However, when we move marginally away from this minimal example, the situation with a tradeoff between mono-risk and multi-risk insurance structures is the same as in the case without deductibles. We extend the previous example by allowing for a third outcome for the realization of the second risk. Specifically, we assume that $\mathbb{P}(\tilde{L}_2 = 4) = p_2 = 0.001$, keeping the rest of the example the same as before (same utilities, etc.), with parameters: $\beta = 7$, $t = 1$, $\delta = 0.2$, $p = 0.249$, and $q = 0.65$.

We perform the same analysis as before, but allow for contracts with risk-specific deductibles. The curves in Figure 5 show the multi-risk insurance outcome as a function of capital, A, for correlations $\rho \in \{0.1, 0.2, 0.3, 0.4\}$, when optimal deductibles are chosen for the two risks, and also includes the optimal outcome under a mono-risk insurance structure as represented by the solid blue lines. The implication is identical as before: With high enough correlation, the mono-risk outcome dominates for one risk, whereas a multi-risk offering dominates for both risks with low correlations.

B Lower bound for ideal risk-free outcome

Theorem 1 shows that, with a sufficient number of risks, a solution can be obtained arbitrarily close to the ideal risk-free outcome with costly internal capital. Theorem 4 below shows the opposite, that with too few risks it is not possible to get arbitrarily close to the ideal risk-free outcome with costly internal capital:

Theorem 4 *Consider a sequence of insurees, $i = 1, 2, \ldots$. If, in addition to the assumptions of Theorem 1, the risks are uniformly bounded: $\tilde{L}_i \leq C_0 < \infty$ (a.s.) for all i, and $Var(\tilde{L}_i) \geq C_1$, for some $C_1 > 0$, for all i, then for every $\epsilon > 0$, there is an n such that $\lim_{\epsilon \searrow 0} n(\epsilon) = \infty$ and such that, as N grows, any partition with $\mathbf{A}_j = \sum_{i \in X_j} \mu_{L_i}$ for all j and*

$$\min_{1 \leq i \leq N} CE_u(-P_i((\mathcal{X}, \mathbf{A})) + \tilde{K}_i((\mathcal{X}, \mathbf{A}))) \geq \mu_{L_i}(1 + \delta_i) - \epsilon, \tag{7}$$

must have $|X_i| \geq n$ for all $X_i \in \mathcal{X}$, i.e., any $X_i \in \mathcal{X}$ must contain at least n risks.

Similarly to Theorem 1, Theorem 4 can be generalized. For example, the condition of uniformly bounded risks can be relaxed. Specifically, if the utility function, u, has deceasing absolute risk aversion, then the Theorem holds if the *expectations* of the risks are uniformly bounded ($E[\tilde{L}_i] < C$) for all i.

C Proofs

Proof of Theorem 1: To simplify the notation, in this proof we write L_i instead of $\tilde{L}_i$, L instead of $\tilde{L}$, and μ_i instead of μ_{L_i}. Moreover, C represents an arbitrary positive, finite constant, i.e., that does not depend on N. The condition that u is twice continuously differentiable with u'' bounded by a polynomial of degree q implies that

$$|u''(z)| \leq C(1 + |z|^q), \qquad z \leq 0, \tag{8}$$

for some constant, $C > 0$.

Moreover, condition $E|L_i|^p \leq C$, together with Jensen's inequality implies that $E|L_i - \mu_i|^p \leq C$, $\sigma_i^2 \leq (E|L_i - \mu_i|^p)^{2/p} \leq C$. Using the Rosenthal inequality for sums of independent mean-zero random variables, we obtain that, for some constant $C > 0$,

$$E\left|\sum_{i=1}^N (L_i - \mu_i)\right|^p \leq C \max\Big(\sum_{i=1}^N E|L_i - \mu_i|^p, \Big(\sum_{i=1}^N \sigma_i^2\Big)^{p/2}\Big), \tag{9}$$

and, thus,

$$\begin{aligned} N^{-p} E\left|\sum_{i=1}^N (L_i - \mu_i)\right|^p &\leq CN^{-p} \max\Big(\sum_{i=1}^N E|L_i - \mu_i|^p, \Big(\sum_{i=1}^N \sigma_i^2\Big)^{p/2}\Big) \\ &\leq CN^{-p/2} \max\Big(N^{-p/2} \times CN, N^{-p/2} \times (CN)^{p/2}\Big) \\ &\leq CN^{-p/2} \\ &\to 0, \end{aligned} \tag{10}$$

as $N \to \infty$, where the second to last inequality follows since $p \geq 2$.

Take $A = \sum_{i=1}^N \mu_i$. This represents a fully multiline insurer, who chooses internal capital to be equal to total expected losses. Denote

$$y_i = -\mathcal{K}_i(\mathcal{S}) = L_i \max\Big(1 - \frac{A}{\sum_{i=1}^N L_i}, 0\Big) = L_i \max\Big(1 - \frac{\sum_{i=1}^N \mu_i}{\sum_{i=1}^N L_i}, 0\Big).$$

From (4) and (6), the expected utility of insuree i is then $Eu(-\mu_i(1+\delta)+E[y_i]-y_i+\delta\mu_i(b_i-1))$. Here,

$$b_i \stackrel{\text{def}}{=} E\left[\frac{L_i}{\mu_i}\times\frac{\sum_{i=1}^N \mu_i}{\sum_{i=1}^N L_i}\times\frac{\tilde{V}}{E[\tilde{V}]}\right], \tag{11}$$

and $\tilde{V}=\left(\sum_{i=1}^N L_i-\sum_{i=1}^N \mu_{L_i}\right)_+$, in line with the definition of the binary default option in Section 3.2. Clearly, for large N, $b_i\to 1$ uniformly over the i's, i.e., $\lim_{N\to\infty}\max_{1\le i\le N}|1-b_i|=0$, so $Eu(-\mu_i(1+\delta)+E[y_i]-y_i+\delta\mu_i(b_i-1))\to Eu(-\mu_i(1+\delta)+E[y_i]-y_i)$ uniformly.

Using a Taylor expansion of order one around $-\mu_i(1+\delta)$, and the polynomial bound, (8), for u'', we get

$$u(-\mu_i(1+\delta)+E[y_i]-y_i)=u(-\mu_i(1+\delta))+u'(-\mu_i(1+\delta))(E[y_i]-y_i)+\frac{u''(\xi(y_i))}{2}(y_i-E[y_i])^2,$$
$$\xi(y_i)\in[-\mu_i(1+\delta),-\mu_i(1+\delta)+E[y_i]-y_i],\forall y_i.$$

Therefore,

$$Eu(-\mu_i(1+\delta)+E[y_i]-y_i)=u(-\mu_i(1+\delta))+\frac{1}{2}E\left(u''(\xi(y_i))(y_i-E[y_i])^2\right),$$

so

$$|Eu(-\mu_i(1+\delta)+E[y_i]-y_i)-u(-\mu_i(1+\delta))|\le C\times E\left((1+|y_i|^q)(y_i-E[y_i])^2\right)\le C'\times(E|y_i|+E|y_i|^{2+q}).$$

If the right hand side is small, then the expected utility is close to $u(-\mu_i(1+\delta))$. To complete the proof, it thus suffices to show that

$$E|y_i|^{2+q}\to 0 \tag{12}$$

as $N\to\infty$, where the speed of convergence does not depend on i, i.e., the convergence is uniform over i.

By Jensen's inequality, evidently, $E|L_i|^p\le C$ for $p=2+q+\epsilon$ with $0<\epsilon\le 2+q$. For such p and ϵ, using the obvious bound $\max\left(1-\frac{\sum_{i=1}^N\mu_i}{\sum_{i=1}^N L_i},0\right)\le 1$ and Hölder's inequality, we get, under the conditions of the Theorem,

$$\begin{aligned}E|y_i|^{2+q}=E\Big|L_i\max\Big(1-\frac{\sum_{i=1}^N\mu_i}{\sum_{i=1}^N L_i},0\Big)\Big|^{2+q}&\le E|L_i|^{2+q}\Big(\max\Big(\frac{\sum_{i=1}^N(L_i-\mu_i)}{\sum_{i=1}^N L_i},0\Big)\Big)^{\epsilon}\le\\ E|L_i|^{2+q}\Big|\frac{\sum_{i=1}^N(L_i-\mu_i)}{\sum_{i=1}^N L_i}\Big|^{\epsilon}&\le\Big(E|L_i|^p\Big)^{\frac{2+q}{p}}\Big(E\Big|\frac{\sum_{i=1}^N(L_i-\mu_i)}{\sum_{i=1}^N L_i}\Big|^p\Big)^{\epsilon/p}\le\\ C\Big(E\Big|\frac{\sum_{i=1}^N(L_i-\mu_i)}{\sum_{i=1}^N\mu_i}\Big|^p\Big)^{\epsilon/p}&\le C\Big(N^{-p}E\Big|\sum_{i=1}^N(L_i-\mu_i)\Big|^p\Big)^{\epsilon/p}.\end{aligned} \tag{13}$$

From (13) and (10) it follows that (12) indeed holds, and, since the constants do no depend on i, the convergence is uniform over i.

We now go from expected utility to certainty equivalents. Expected utility is $u(-\mu_i(1+\delta))-\epsilon=u(-\mu_i(1+\delta)-c)=u(-\mu_i(1+\delta))-cu'(\xi)$, where $\xi\in(-\mu_i(1+\delta)-c,-\mu_i(1+\delta))$, so $c=\epsilon/u'(\xi)\le\epsilon/u'(-\mu_i(1+\delta))\le\epsilon/u'(0)$, and $\epsilon\to 0$ therefore implies that $c\to 0$. The proof is complete.

∎

Proof of Theorem 2: We use the same notation as in the proof of Theorem 1.

(1.) The result follows from the compactness argument made toward the end of the proof of Theorem 3 (1.). If there would be a sequence of insurees, i_j, $j=1,\ldots,$ for which insuring with monoline insurers is robust to aggregation, with capital $A_i=\beta_i\mu_i$, there must be a limit point, i.e., a $\beta^*\in[\epsilon,1]$ for which, for each $\epsilon>0$, there is an infinite number of insurees (in the limit, as N tends to infinity) — each insuree

having optimal capital β_{i_j} — such that $|\beta_{i_j} - \beta^*| < \epsilon$. However, the argument in Theorem 3 (1.) then immediately implies that a multiline firm with capital $A = \beta^* \sum_i \mu_i$ dominates the monoline structure for these firms, contradicting the assumption that the industry partition is robust to aggregation.

(2.) From Theorem 1, it follows that the fully multiline outcome converges to the ideal risk-free outcome, which by Condition 1 in turn dominates the monoline outcome for all levels of capital $A_i \leq \mu_i$. For $A_i > \mu_i$ the ideal risk-free outcome also dominates the monoline offering, since $P(A_i) = \mu_i - \mu_{Q_i} + \delta A_i$, and thus

$$E[u(-P(A_i) - Q_i(A_i))] = E[u(-\mu_i - \delta A_i + P_{Q_i} - Q_i(A_i))] < u(-\mu_i - \delta A_i) < u(-\mu_i(1+\delta)),$$

when $A_i > \mu_i$. The result therefore follows.

∎

Proof of Theorem 3: We use the same notation as in the proof of Theorem 1, and first prove the second result since it is needed for the first result.

(2.) For the fully multiline industry structure, with capital $A = \beta \sum_i \mu_i$, the utility of agent i is $Eu(-\mu_i(1+\delta\beta) + E[y_i] - y_i)$, where

$$y_i = -\mathcal{K}_i(\mathcal{S}) = L_i \max\Big(1 - \frac{A}{\sum_{i=1}^N L_i}, 0\Big) = L_i \max\Big(1 - \beta\frac{\sum_{i=1}^N \mu_i}{\sum_{i=1}^N L_i}, 0\Big). \tag{14}$$

Here, since we we will use the results for non i.i.d. risks when proving (1.), we use the i-subscript, even though it is not needed in the i.i.d. case, in which $\mu_i \equiv \mu$.

Assume that $(\mathcal{X}^{MONO}, (A_1, A_2, A_3, \dots,))$ is a constrained Pareto efficient industry structure. Then, $\mathcal{S}^{MONO} = (\mathcal{X}^{MONO}, (A_1, A_1, A_1, \dots,))$ is obviously Pareto equivalent, i.e., it provides the same expected utility for all insurees, i, as $(\mathcal{X}^{MONO}, (A_1, A_2, A_3, \dots,))$ does. Such an industry structure is characterized by $\beta \stackrel{\text{def}}{=} A/\mu$. If we show that, regardless of β, for large enough N, and some β^{MULTI}, the industry partition, $\mathcal{X}^{MULTI}$, with capital, A, where $A = \beta^{MULTI} \times \sum_i \mu_i$, Pareto dominates S^{MONO}, then we are done, since $\mathcal{X}^{MULTI}$ is already per definition robust to aggregation.

For $\beta^{MULTI} = 1$, the argument of Theorem 1 implies that $Eu(-\mu_i(1+\delta\beta^{MULTI}) + E[y_i] - y_i) \to u(-\mu_i(1+\delta))$, i.e., the expected utility converges to the ideal risk-free outcome with costly capital, as N grows, and that the convergence is uniform in i. The utility from insuring with a monoline insurer with $\beta^{MONO} = 1$, on the other hand, is $Eu(-\mu_i(1+\delta) + E[y_i] - y_i)$, where (14) implies that $y_i = \max(L_i - \beta^{MONO}\mu_i, 0)$. Now, $E[y_i] - y_i$ is obviously second order stochastically dominated (SOSD) by 0, so for $\beta^{MONO} = 1$, the fully multiline partition with $\beta^{MULTI} = 1$, will obviously dominate the monoline offering. Similarly, the multiline partition with $\beta^{MULTI} = 1$ dominates any monoline offering with $\beta^{MONO} > 1$, since there is still residual risk for such a monoline offering *and* the total cost of internal capital is higher — effects that both make the insuree worse off.

It is also clear that for large N, β^{MULTI} for any constrained Pareto efficient outcome must lie in $[0, 1+o(1)]$, since internal capital is costly, which imposes a linear cost of increasing β, and, by the law of large numbers, all risk eventually vanishes for $\beta^{MULTI} = 1$. Thus, there is always a constrained Pareto efficient solution in $[0, 1+o(1)]$, given the fully multiline industry partition.

If we show that, for a given $0 < \beta < 1$, the fully multiline outcome will dominate monoline offerings *with the same* β, then we are done, since, regardless of a candidate β^{MONO}, choosing $\beta^{MULTI} = \beta^{MONO}$ will lead to a Pareto improvement ($\beta^{MONO} = 0$ is strictly dominated by some $\beta^{MONO} > \epsilon/\mu$ from Condition 3, so we do not need to consider $\beta^{MONO} = 0$).

For the fully multiline outcome, it is clear from (14), using an identical argument as in the proof of Theorem 1 that y_i converges uniformly in i to $(1-\beta)L_i$, and the expected utility of agent i converges uniformly to

$$Eu(-\mu_i(1+\delta\beta) + (1-\beta)(\mu_i - L_i)). \tag{15}$$

On the other hand, for the monoline offering, the expected utility is

$$Eu(-\mu_i(1+\delta\beta) + E[y_i] - y_i), \tag{16}$$

where

$$y_i = \max(L - \beta\mu_i, 0) = (1-\beta)\max\left(\mu_i - \frac{\mu_i - L}{1-\beta}, 0\right). \tag{17}$$

We define $z_i \stackrel{\text{def}}{=} \mu_i - L_i \in (-\infty, \mu_i)$ and $\alpha \stackrel{\text{def}}{=} \frac{1}{1-\beta} \in (1,\infty)$. Equations (15-17) imply that if

$$(1-\beta)z_i \succeq (1-\beta)\left(E\left[\max\left(\mu_i - \alpha z_i, 0\right)\right] - \max\left(\mu_i - \alpha z_i, 0\right)\right) \tag{18}$$

for all $\alpha \in (1,\infty)$, where $\succeq$ denotes second order stochastic dominance, then the third part of Theorem 3 is proved. However, (18) is equivalent to

$$z_i \succeq E\left[\max\left(\mu_i - \alpha z_i, 0\right)\right] - \max\left(\mu_i - \alpha z_i, 0\right),$$

and by defining $x_i \stackrel{\text{def}}{=} \max\left(\mu_i - \alpha z_i, 0\right)$, and $w_i \stackrel{\text{def}}{=} Ex_i - x_i$, to

$$z_i \succeq w_i.$$

We define $q_i(\alpha) \stackrel{\text{def}}{=} Ex_i$. From the definitions in Section 3: $\tilde{Q}(A) = \max(L - A, 0)$ and $P_Q(A) = E[\tilde{Q}(A)]$, it follows that $q_i(\alpha) = \alpha P_Q\left(\left(1 - \frac{1}{\alpha}\right)\mu_i\right)$. Therefore,

$$w_i = q_i(\alpha) - \max\left(\mu_i - \alpha z_i, 0\right). \tag{19}$$

To show that z_i second order stochastically dominates w_i, we use the following Lemma, which follows immediately from the theory in Rothschild and Stiglitz (1970):

Lemma 1 *if z and w are random variables with absolutely continuous distributions, and distribution functions $F_z(\cdot)$ and $F_w(\cdot)$ respectively, $z \in (\underline{z}, \overline{z})$, $-\infty \le \underline{z} < \overline{z} \le \infty$ (a.s.) and the following conditions are satisfied:*

1. *$E[z] = E[w]$,*
2. *$F_z(x) < F_w(x)$ for all $\underline{z} < x < x_0$, for some $x_0 > \underline{z}$,*
3. *$F_z(x) = F_w(x)$ at exactly one point, $x^* \in (\underline{z}, \overline{z})$.*

Then $z \succ w$, i.e., z strictly SOSD w.

Clearly, $E[z_i] = E[w_i] = 0$, so the first condition of Lemma 1 is satisfied. For the second condition, it follows from (19), that for $x < q_i(\alpha)$,

$$\begin{aligned} F_{w_i}(x) &= \mathbb{P}(w_i \le x) = \mathbb{P}(q_i(\alpha) - \mu_i + \alpha z_i \le x) \\ &= \mathbb{P}(z_i \le (x + \mu_i - q_i(\alpha))/\alpha) \\ &= F_{z_i}((x + \mu_i - q_i(\alpha))/\alpha). \end{aligned} \tag{20}$$

Since $\alpha > 1$, it is clear that for small enough x, $(x + \mu_i - q_i(\alpha))/\alpha > x$, and therefore $F_{z_i}(x) < F_{w_i}(x)$, so the second condition is satisfied.

To show that the third condition is satisfied, we need the following Lemma

Lemma 2 *$q_i(\alpha) \ge \mu_i$ for all $\alpha > 1$.*

Proof: Denote by ϕ, the probability distribution function (or measure) of L_i. Thus, $\mu_i = \int_0^\infty L\phi(L)dL$. We have, for $r > 0$:

$$\int_0^r L\phi(L)dL \le r\int_0^r \phi(L)dL,$$

which leads to

$$
\begin{aligned}
\mu_i - \int_0^r L\phi(L)dL \geq \mu_i - r\int_0^r \phi(L)dL \quad &\Leftrightarrow \quad \int_r^\infty L\phi(L)dL \geq \mu_i - r\int_0^r \phi(L)dL \\
&\Leftrightarrow \quad \int_r^\infty L\phi(L)dL + \int_0^r \phi(L)dL - r \geq \mu_i - r \\
&\Leftrightarrow \quad \int_r^\infty L\phi(L)dL - r\left(1 - \int_0^r \phi(L)dL\right) \geq \mu_i - r \\
&\Leftrightarrow \quad \int_r^\infty L\phi(L)dL - r\int_r^\infty \phi(L)dL \geq \mu_i - r \\
&\Leftrightarrow \quad \int_r^\infty (L-r)\phi(L)dL \geq \mu_i - r \\
&\Leftrightarrow \quad \frac{\mu_i}{\mu_i - r}\int_r^\infty (L-r)\phi(L)dL \geq \mu_i \\
&\Leftrightarrow \quad \frac{\mu_i}{\mu_i - r}\int_0^\infty \max(L-r,0)\phi(L)dL \geq \mu_i \\
&\Leftrightarrow \quad \frac{\mu_i}{\mu_i - r}E\left[\max(L-r,0)\right] \geq \mu_i. \\
&\Leftrightarrow \quad \frac{\mu_i}{\mu_i - r}P_Q(r) \geq \mu_i. \qquad (21)
\end{aligned}
$$

Now, for $r < \mu_i$, define $\alpha = \frac{\mu_i}{\mu_i - r} \in (1, \infty)$, implying that $r = \left(1 - \frac{1}{\alpha}\right)\mu_i$. Then, the last line of (21) can be rewritten

$$q_i(\alpha) = \alpha P_Q\left(\left(1 - \frac{1}{\alpha}\right)\mu_i\right) \geq \mu_i.$$

This completes the proof of Lemma 2.

Since $z \in (\underline{z}, \overline{z}) = (-\infty, \mu_i)$, and the p.d.f. is strictly positive, $F_{z_i}(\mu_i) = 1$ and, $F_{z_i}(x) < 1$ for $x < \mu_i$. We also note that for $x < q_i(\alpha)$, $F_{w_i}(x) < 1$ is given by (20), and for $x \geq q_i(\alpha)$, $F_{w_i}(x) = 1$. Since $\overline{z} \leq q_i(\alpha)$ (by Lemma 2) the only points at which $F_{z_i}(x) = F_{w_i}(x)$ are therefore points at which $F_{z_i}(x) = F_{z_i}((x + \mu_i - q_i(\alpha))/\alpha)$, i.e., for points at which

$$x = \frac{x + \mu_i - q_i(\alpha)}{\alpha} \Rightarrow x = \frac{\mu_i - q_i(\alpha)}{\alpha - 1}. \qquad (22)$$

Now, since there is a unique solution to (22), the third condition is indeed satisfied. Thus, z_i SOSD w_i, so a monoline offering is inferior for any insuree with strictly concave utility function, and therefore (2.) of Theorem 3 holds.

(1.) It was shown above that when the risks are i.i.d., a massively multiline industry partition Pareto dominates the monoline one. There are two complications when extending the proof to nonidentical distributions. The first, main, complication is that insurees may no longer agree on the optimal level of capital, i.e., they have different β_i's. Some insurees may face "small" risks and thereby opt for limited capital, to save on costs of internal capital (a low β), whereas others may wish to have a close to full insurance (a high β). A fully multiline industry structure has only one β, and it may therefore be optimal to have several insurance companies, each with a different β. As the number of lines tends to infinity, however, within each such company the number of lines grows, making the industry structure massively — but not fully — multiline.

The second complication is technical: Since the risks are no longer i.i.d., and each insuree may therefore have a different β, most insurees will not be at their optimal β when insuring with a multiline firm, although they will be close. There is therefore a trade off between the utility loss of being at a suboptimal β, versus the utility gains of diversification, in the multiline partition. We need quantitative bounds, as opposed to the qualitative SOSD bounds above, that show that when N increases, for most insurees the second effect dominates the first. This will be ensured by Condition 4.

If Condition 4 fails, even though each insuree may eventually wish to be insured by a multiline industry, for any fixed N, it could still be that most insurees prefer a monoline solution, e.g., if insurees $i = 1, \dots, N/2$ prefer a massively multiline insurer, whereas insurees $i = N/2+1, \dots, N$ prefer to insure with a fully multiline insurance company. In this case, $M(S) = N/2 + 1$, so the average number of lines converges to 2 and the industry structure is thus not massively multiline.

The expected utility of insuree $i \in X_j$, is $Eu(-\mu_i(1+\delta\beta) + E[y_i] - y_i + \beta\delta\mu_i(b_i - 1))$, where

$$y_i = -\mathcal{K}_i(\mathcal{S}) = L_i \max\Big(1 - \frac{A}{\sum_{i\in X_j} L_i}, 0\Big) = L_i \max\Big(1 - \beta\frac{\sum_{i\in X_j}\mu_i}{\sum_{i\in X_j} L_i}, 0\Big), \tag{23}$$

and

$$b_i \stackrel{\text{def}}{=} E\left[\frac{L_i}{\mu_i} \times \frac{\sum_{i\in X_j}\mu_i}{\sum_{i\in X_j} L_i} \times \frac{\tilde{V}}{E[\tilde{V}]}\right]. \tag{24}$$

A similar argument as in Theorem 1 implies that $b_i \to 1$ uniformly, so for large $|X_j|$, we can study $Eu(-\mu_i(1+\delta\beta) + E[y_i] - y_i) \to Eu(-\mu_i(1+\delta\beta) + (1-\beta)(\mu - L_i))$.

The following two Lemmas are needed to show the result.

Lemma 3 *Under the conditions of Theorem 3.1, there is a $C > 0$, that does not depend on i, such that for all $0 \le \beta \le 1$,*

$$\left|\frac{\partial Eu(-\mu_i(1+\delta\beta) + (1-\beta)(\mu_i - L_i))}{\partial \beta}\right| \le C. \tag{25}$$

Proof Define $Z(\beta) \stackrel{\text{def}}{=} Eu(-\mu_i(1+\delta\beta) + (1-\beta)(\mu_i - L_i))$. It immediately follows that Z is concave, so $\max_{0\le\beta\le1} |Z'(\beta)|$ is realized at either $\beta = 0$ or $\beta = 1$. The derivative of Z is

$$Z'(\beta) = E[(-(1+\delta) + L_i)u'(-\mu_i(1+\delta\beta) + (1-\beta)(\mu_i - L_i))],$$

so $|Z'(0)| = |E[((-1-\delta) + L_i)u'(-L_i)]| \le C(|E[u'(0) + L_i u''(-L_i)| + |E[L_i u'(0)| + L_i^2 u''(-L_i)]| \le C(|u'(0)(\mu_i + 1) + E[(L_i + L_i^2)u''(-L_i)]|) \le C$. Here, the last inequality follows from the assumptions in Theorem 1: Since $|u''(-x)| \le c_1 + c_2 x^q$ and $|E[L_i^2 u''(-L_i)]| \le E[L_i^2(c_1 + c_2 L_i^q)] \le c_1 \times E[L_i^2] + c_2 \times E[L_i^{2+q}] \le C'$.

Thus, $|Z'(0)|$ is bounded by a constant, and moreover, since a Taylor expansion yields that $|Z'(1)| \le |Z'(0)| + \max_{0\le\beta\le1}|Z''(\beta)|$, as long as $|Z''(\beta)|$ is bounded in $0 \le \beta \le 1$ (independently of i), $|Z'(1)|$ will also be bounded by a constant. We have

$$\begin{aligned} |Z''(\beta)| &= |E[(-(1+\delta) + L_i)^2 u''(-\mu_i(1+\delta\beta) + (1-\beta)(\mu_i - L_i))]| \\ &\le |E[(-(1+\delta) + L_i)^2 u''(-\mu_i(1+\delta) - L_i))| \\ &\le E[(c_1 + c_2 L_i + c_3 L_i^2)(c_4 + c_5 L_i^q)] \le C. \end{aligned}$$

The Lemma is proved.

Lemma 3 immediately implies that for β close to the optimal β^*, the utility will be close to the utility at the optimum, since $|Z(\beta+\epsilon) - Z(\beta)| \le C\epsilon$. The second condition, that the utility provided from the asymptotic fully multiline company is bounded away from the utility provided by the monoline insurer, for a fixed β, follows from the following Lemma:

Lemma 4 *If Condition 4 is satisfied, then for any $\epsilon > 0$, there is a constant, $C > 0$, such that, for all $\beta \in [\epsilon, 1)$,*

$$Eu(-\mu_i(1+\delta\beta) + (1-\beta)(\mu_i - L_i)) - Eu(-\mu_i(1+\delta\beta) + E[y_i] - y_i) \ge C. \tag{26}$$

Here, $y_i = \max(L_i - \beta\mu_i, 0)$, is the payoff of the option to default for the monoline insurer, with capital $\beta\mu_i$.

Proof: In the notation of the proof of (2.), equation (26) can be rewritten as $Eu(-c + z_i(1-\beta)) \ge Eu(-c + w_i(1-\beta)) + C$, where $c \stackrel{\text{def}}{=} -\mu_i(1+\delta\beta)$, $z_i \stackrel{\text{def}}{=} (\mu_i - L_i)$ and $w_i \stackrel{\text{def}}{=} \frac{1}{1-\beta}(E[y_i] - y_i)$. We know from the proof that z_i SOSD w_i, so $Eu(-c+z_i(1-\beta)) > Eu(-c+w_i(1-\beta))$, but we need a bound that is independent of i.

From (20), we know that

$$F_{w_i} - F_{z_i} = F_{z_i}((x+\mu_i - q_i(\alpha))/\alpha) - F_{z_i}(x) \tag{27}$$

Here, $\alpha = \frac{1}{1-\beta} \in [\frac{1}{\epsilon}, \infty)$. From the definition of q_i it follows that $q_i(\alpha) \leq \alpha \times \mu_i$. The conditions of Theorem 1, immediately imply that $\mu_i \leq C$ for some C independent of i, so $q_i(\alpha)$ is bounded for each α, independently of i, by $\alpha \times C_0$. If we denote by F_i, the c.d.f of $-L_i$, then, since $z_i = \mu_i - L_i$, it follows that $F_{z_i}(x) = F_i(x - \mu_i)$. The risk faced by the insuree is $\frac{1}{\alpha} z_i$ in the asymptotic multiline case, and $\frac{1}{\alpha} w_i$ in the monoline case. Therefore, (27) implies that

$$e(x,\alpha) \stackrel{\text{def}}{=} F_i(x+\mu_i - q_i(\alpha)) - F_i(\alpha x)$$

is the difference between the c.d.f.'s of the monoline and fully insured offerings for largely negative x. Since $q_i(\alpha) \leq C_0\alpha$, $e(x,\alpha) \geq F_i(x - C_0\alpha) - F_i(\alpha x)$, we can choose x sufficiently negative so that $h(-x) < \frac{\alpha x - 1}{x - C_0\alpha}$. Then, for $x' < x$, from Condition 4 it follows that

$$e(x',\alpha) \geq g(-h(-x')(x' - C_0\alpha)) - g(-\alpha x') \geq g(-\alpha x' - 1) - g(-\alpha x').$$

This bound is independent of i, and since g is continuous and strictly decreasing, for each α it is the case that on $[-x-1, -x]$,

$$F_i(x+\mu_i - q_i(\alpha)) - F_i(\alpha x) > \epsilon_\alpha > 0. \tag{28}$$

We have, for a general risk, R, with c.d.f. F_R, and support on $(-\infty, r)$, where R is sufficiently thin-tailed so that expected utility is defined,

$$Eu[R] = \int_{-\infty}^{r} u(x) dF_R = u(r) - \int_{-\infty}^{r} u'(x) F_R(x) dx. \tag{29}$$

Now, since u is strictly concave, u' is decreasing, but positive, so if, $E[R_1] = E[R_2]$, $F_{R_1} > F_{R_2} + \epsilon$, on $[a,b]$, $\epsilon > 0$, and if F_{R_1} and F_{R_2} only cross at one point, then $Eu[R_1] - Eu[R_2] = \int_{-\infty}^{r} u'(x)(F_{R_2}(x) - F_{R_1}(x))dx \geq (\inf_{x \leq r} |u''(x)|) \times \epsilon \times \frac{b-a}{2}$.

Therefore, (28), implies that $Eu(-c+z_i(1-\beta)) \geq Eu(-c+w_i(1-\beta)) + C_\alpha$, and since ϵ_α is continuous in α, we can take the minimum over $\alpha \in [1, 1/\epsilon]$, to get a bound that does not depend on α. The Lemma is proved.

We have thus shown that, away from $\beta = 0$, the asymptotic fully multiline solution is uniformly better (in i) than the monoline solution, for each β (Lemma 4), and that, as long as a β close to the optimal β is offered, the decrease in utility for the insuree is small (Lemma 3). This, together with the uniform convergence toward the asymptotic risk as the number of lines covered by an insurer grows, implies that there is an $\epsilon^* > 0$ and a $C^* < \infty$, such that any firm with a β within ϵ^* distance from the optimal β for the monoline offering, and with at least C^* insurees, will make all insurees strictly better off than the monoline insurer.

Now, since $\beta \in [\epsilon, 1]$ is in a compact interval, as the number of lines grows, most insurees will be close to many insurees, making massively multiline offerings feasible for the bulk of the insurees. In fact, defining $T_N(i,\epsilon)$ to be the number of insurees, who have β's within ϵ distance from insuree i, i.e., for which $|\beta_i - \beta_j| \leq \epsilon$, in the economy with insurees $1 \leq i \leq N$, it is the case that $\forall \epsilon > 0$, and $\forall C < \infty$, all but a bounded (independently of N) number of agents belong to $\{i : T_N(i,\epsilon) \geq C\}$, due to a standard compactness argument. If this were not the case, there would be a sequence of insurees, i_j, $j = 1, \ldots$, with β's not close to other insurees, and since such a sequence must have a limit point, there must be an agent in this sequence with more than C neighbors within ϵ distance, contradicting the original assumption.

It is clear from the previous argument that for agents in $\{i : T_N(i,\epsilon) \geq C\}$, as N grows, multiline solutions will be optimal. Specifically, for $\epsilon = \epsilon^*$, a sequence of $C \to \infty$ can be chosen, and an N large enough, such that $\{i : T_N(i,\epsilon) \geq C\}/N > 1 - \delta$, for arbitrary $\delta > 0$. Then it follows immediately that all insurees can be covered by insurers with at least $C/2$ insurees, where insurer n chooses $\beta_n = n\epsilon^*/2$. This leads to a massively multiline structure, since the average number of insurees is greater than or equal to

$(1-\delta)C/2$, and, as we have shown, it dominates the monoline industry partition. We are done.

Proof of Theorem 4:

By Taylor expansion, for all x, y, $u(x+y) = u(x) + u'(x)y + u''(\zeta)\frac{y^2}{2}$, where ζ is a number between x and $x+y$. Since u'' is bounded away from zero: $-u'' \geq C > 0$, we, therefore, get

$$u(x+y) \leq u(x) + u'(x)y - C\frac{y^2}{2} \tag{30}$$

for all x, y. Using inequality (30), in the notations of the proof of Theorem 1, we obtain

$$\begin{aligned} Eu(-\mu_i(1+\delta) + E[y_i] - y_i) \leq \\ E\Big[u(-\mu_i(1+\delta)) + u'(-\mu_i(1+\delta))(E[y_i] - y_i) - C\frac{(y_i - E[y_i])^2}{2}\Big] = \\ u(-\mu_i(1+\delta)) - CVar(y_i). \end{aligned} \tag{31}$$

Consequently, if $u(-\mu_i(1+\delta)) - Eu(-\mu_i(1+\delta) + E[y_i] - y_i) < \epsilon$, then $Var(y_i) < \epsilon' = \epsilon/C$. We therefore wish to show that, for a fixed number of risks, N, $Var(y_i) \geq \epsilon_N > 0$.

We need the following two Lemmas:

Lemma 5 *For $x_2 \geq x_1$, if X is a random variable, such that $\mathbb{P}(X \leq x_1) = a$ and $\mathbb{P}(X \geq x_2) = b$, then $Var(X) \geq \frac{\min(a,b)(x_2-x_1)^2}{4}$.*

Proof: Define $e = \frac{x_2 - x_1}{2} \geq 0$. Define $M = E(X)$ and assume that $M \geq x_1 + e$. Moreover, let ϕ denote X's p.d.f. Then $Var(X) = \int (x-M)^2\phi(x)dx \geq \int_{x \leq x_1}(x-M)^2\phi(x)dx \geq \int_{x \leq x_1}(x - x_1 - e)^2\phi(x)dx \geq e^2 a = \frac{a(x_2-x_1)^2}{4}$. A similar argument shows that if $M < x_1 + e$, then $Var(X) \geq \frac{b(x_2-x_1)^2}{4}$, and since either $M \geq x_1 + e$ or $M < x_1 + e$ the Lemma follows.

Lemma 6 *If X is a random variable, such that $0 \leq X \leq C < \infty$, $E(X) = \mu$ and $Var(X) = \sigma^2 > 0$, then there are constants $d > 0$ and $\epsilon > 0$, that only depend on C, μ and σ^2, such that $\mathbb{P}(X \leq \mu - \epsilon) \geq d$ and $\mathbb{P}(X \geq \mu + \epsilon) \geq d$.*

Proof: Let ϕ be X's p.d.f. For a small $\epsilon > 0$, we have

$$\begin{aligned} \sigma^2 &= \int_0^C (x-\mu)^2\phi(x)dx \\ &= \int_{|x-\mu|<\epsilon}(x-\mu)^2\phi(x)dx + \int_{|x-\mu|\geq\epsilon}(x-\mu)^2\phi(x)dx \\ &\leq \epsilon^2 + \mu\int_{x\leq\mu-\epsilon}|x-\mu|\phi(x)dx + (C-\mu)\int_{x\geq\mu+\epsilon}|x-\mu|\phi(x)dx. \end{aligned} \tag{32}$$

Moreover, $\int_{|x-\mu|<\epsilon}(x-\mu)\phi(x)dx + \int_{|x-\mu|\geq\epsilon}(x-\mu)\phi(x)dx = 0$, and $\left|\int_{|x-\mu|<\epsilon}(x-\mu)\phi(x)dx\right| \leq \epsilon$, so

$$\int_{x\geq\mu+\epsilon}|x-\mu|\phi(x)dx + \epsilon \geq \int_{x\leq\mu-\epsilon}|x-\mu|\phi(x)dx.$$

Plugging this into (32) yields

$$C\int_{x\geq\mu+\epsilon}|x-\mu|\phi(x)dx \geq \sigma^2 - \epsilon^2 - \mu\epsilon.$$

However, since $\int_{x\geq\mu+\epsilon}|x-\mu|\phi(x)dx \leq (C-\mu)\int_{x\geq\mu+\epsilon}\phi(x)dx = (C-\mu)\mathbb{P}(X\geq\mu+\epsilon)$, we arrive at

$$\mathbb{P}(X\geq\mu+\epsilon)\geq\frac{\sigma^2-\epsilon^2-\epsilon\mu}{C(C-\mu)}.$$

A similar argument implies that

$$\int_{x\leq\mu-\epsilon}|x-\mu|\phi(x)dx+\epsilon\geq\int_{x\geq\mu+\epsilon}|x-\mu|\phi(x)dx,$$

which, when plugged into (32) yields

$$C\int_{x\leq\mu-\epsilon}|x-\mu|\phi(x)dx\geq\sigma^2-\epsilon^2-\epsilon(C-\mu),$$

and since $\int_{x\leq\mu-\epsilon}|x-\mu|\phi(x)dx\leq\mu\int_{x\geq\mu+\epsilon}\phi(x)dx=\mu\mathbb{P}(X\leq\mu-\epsilon)$, we arrive at

$$\mathbb{P}(X\leq\mu-\epsilon)\geq\frac{\sigma^2-\epsilon^2-\epsilon(C-\mu)}{C\mu}.$$

Thus, by defining

$$d\stackrel{\text{def}}{=}\min\left(\frac{\sigma^2-\epsilon^2-\epsilon\mu}{C(C-\mu)},\frac{\sigma^2-\epsilon^2-\epsilon(C-\mu)}{C\mu}\right), \tag{33}$$

the Lemma follows.

We note that the condition in Theorem 1, that $E[L_i]\geq C'$ for some $C'>0$, actually is implied by the conditions that $Var(L_i)>C_1$ and that $L_i\leq C_0$, by the following Lemma:

Lemma 7 *If X is a random variable, such that $0\leq X\leq C_0<\infty$, then $E(X)\geq\frac{Var(X)}{C_0}$.*

Proof: Define $e=E(X)$ and let ϕ denote X's p.d.f. We have $Var(X)=\int_0^C x^2\phi(x)dx-e^2\leq\int_0^C x^2\phi(x)dx\leq C\int_0^C x\phi(x)dx=C\times E[X]$.

Now, from the conditions of Theorem 4, it follows that the L_i's satisfy the conditions of Lemma 6, and that ϵ and d can be chosen not to depend on i. Moreover, $y_i=L_i\max\left(1-\frac{\sum_i\mu_i}{\sum_i L_i},0\right)$, so $\mathbb{P}(y_i\leq 0)\geq\mathbb{P}(\sum_i L_i\leq\sum_i\mu_i)\geq\prod_{i=1}^N\mathbb{P}(L_i\geq\mu_i)\geq\prod_{i=1}^N\mathbb{P}(L_i\geq\mu_i+\epsilon)\geq d^N$, where d is defined in (33).

Similarly, if $L_i\geq\mu_i+\epsilon$ for all i, then $y_i=L_i\left(1-\frac{\sum_i\mu_i}{\sum_i L_i}\right)\geq(\mu_i+\epsilon)\left(1-\frac{\sum_i\mu_i}{\sum_i\mu_i+N\epsilon}\right)=(\mu_i+\epsilon)\times\left(1-\frac{1}{1+\frac{N\epsilon}{\sum_i\mu_i}}\right)\geq\mu_i\times\left(1-\left(1-\frac{N\epsilon}{NC}\right)\right)=\mu_i\times\frac{\epsilon}{C}$, so $\mathbb{P}(y_i\geq\frac{\mu_i\epsilon}{C})\geq\prod_{i=1}^N\mathbb{P}\left(L_i\geq\mu_i+\epsilon\right)\geq d^N$.

From Lemma 7, $\mu_i\geq\frac{\sigma^2}{C}$ for all i, so we have $\mathbb{P}(y_i\geq\frac{\sigma^2\epsilon}{C^2})\geq d^N$. Therefore, y_i satisfies all the conditions of Lemma 5, with $x_1=0$, $x_2=\frac{\sigma^2\epsilon}{C^2}$, $a=b=d^N$, and therefore

$$Var(y_i)\geq C_N,\qquad\text{where } C_N=\frac{\sigma^4\epsilon^2}{4C^4}\left(\min\left(\frac{\sigma^2-\epsilon^2-\epsilon\mu}{C(C-\mu)},\frac{\sigma^2-\epsilon^2-\epsilon(C-\mu)}{C\mu}\right)\right)^N>0.$$

The constant, C_N, depends on N, σ^2 and C, but not on the specific distributions of the L_i's, so for a fixed N,

$$Eu(-\mu_i(1+\delta)+E[y_i]-y_i)\leq u(-\mu_i(1+\delta))-C\times C_N. \tag{34}$$

The argument, so far, has been for $Eu(\mu_i(1+\delta)+E[y_i]-y_i)$, whereas the utility is $Eu(\mu_i(1+\delta)+E[y_i]-y_i+\delta\mu_i(b_i-1))$, where b_i is defined in (11). However, since by definition $\sum_{i=1}^N b_i = 1$ for all N, it is clear that for all N, there must be an i, such that $b_i - 1 \leq 0$. For such an i, $Eu(\mu_i(1+\delta)+E[y_i]-y_i+\delta\mu_i(b_i-1)) < Eu(\mu_i(1+\delta)+E[y_i]-y_i)$, so $Eu(\mu_i(1+\delta)) - Eu(\mu_i(1+\delta)+E[y_i]-y_i) > \epsilon \Rightarrow Eu(\mu_i(1+\delta)) - Eu(\mu_i(1+\delta)+E[y_i]-y_i+\delta\mu_i(b_i-1)) > \epsilon$. Thus, the argument also goes through for the utility of agent.

A similar argument as in the proof of Theorem 1, takes us from utilities to certainty equivalents: $u(-\mu_i(1+\delta)-c) = u(-\mu_i(1+\delta)) - \epsilon \Rightarrow c = \epsilon/u'(-\xi)$, $\xi \leq C_0$, so $c \geq \frac{\epsilon}{u'(-C_0)}$, so if ϵ is bounded away from 0, so is c. We are done.

∎

References

Arrow, K., 1963, "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, 53, 1–42.

——— , 1964, "The Role of Securities in the Optimal Allocation of Risk Bearing," *Review of Economic Studies*, 31, 121–133.

Banerjee, S., H. Konishi, and T. Sönmez, 2001, "Core in a simple coalition formation game," *Social Choice Welfare*, 18, 135–153.

Bogomolnaia, A., and M. O. Jackson, 2002, "The stability of hedonic coalition structures," *Games and Economic Behavior*, 38, 201–230.

Bolton, P., and M. Oehmke, 2015, "Bank Resolution and the Structure of Global Banks," Working paper, Columbia University.

Borch, K., 1962, "Equilibrium in a reinsurance market," *Econometrica*, 30, 424–444.

Burkholder, D. L., 1973, "Distribution function inequalities for martingales," *Annals of Probability*, 1, 19–42.

Cummins, J. D., 1993, "Allocation of Capital in the Insurance Industry," *Risk Management and Insurance Review*, 3, 7–28.

Cummins, J. D., and O. Mahul, 2003, "Optimal insurance with divergent beliefs about insurer total default risk," *Journal of Risk and Uncertainty*, 27(2), 121–138.

——— , 2004, "The Demand for Insurance With an Upper Limit on Coverage," *Journal of Risk and Insurance*, 71(2), 253–264.

de la Peña, V. H., R. Ibragimov, and S. Sharakhmetov, 2003, "On extremal distributions and sharp L_p-bounds for sums of multilinear forms," *Annals of Probability*, 31, 630–675.

Demage, G., 1994, "Intermediate preferences and stable coalition structures," *Journal of Mathematical Economics*, 23, 45–58.

Diamond, D., 1984, "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51, 393–414.

Doherty, N., and J. Garven, 1986, "Price regulation in property-liability insurance: A contingent claims approach," *Journal of Finance*, 41, 1031–1050.

Doherty, N., and H. Schlesinger, 1990, "Rational insurance purchasing: consideration of contract nonperformance," *Quarterly Journal of Economics*, 105, 243–253.

Doherty, N. A., C. Laux, and A. Muermann, 2014, "Insuring Non-Verifiable Losses," Working paper, University of Pennsylvania.

Dreze, J., and J. Greenberg, 1982, "Hedonic coalitions: Optimality and stability," *Econometrica*, 48, 987–1003.

Froot, K. A., 2007, "Risk Management, Capital Budgeting, and Capital Structure Policy for Insurers and Reinsurers," *Journal of Risk and Insurance*, 74, 273–99.

Froot, K. A., D. Scharsfstein, and J. Stein, 1993, "Risk Management: Coordinating Corporate Investment and Financing Decisions," *Journal of Finance*, 48, 16291658.

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

Greenberg, J., and S. Weber, 1993, "Stable coalition structures with unidimensional set of alternatives," *Journal of Economic Theory*, 60, 62–82.

Ibragimov, R., D. Jaffee, and J. Walden, 2009, "Nondiversification traps in markets for catastophe insurance," *Review of Financial Studies*, 22, 959–993.

———, 2010, "Pricing and Capital Allocation for Multiline Insurance Firms," *Journal of Risk and Insurance*, 77, 551–578.

———, 2011, "Diversification Disasters," *Journal of Financial Economics*, 99(2), 333–348.

Ibragimov, R., and P. C. B. Phillips, 2008, "Regression asymptotics using martingale convergence methods," *Econometric Theory*, 28, 1–60.

Jaffee, D., 2006, "Monoline Restrictions, with Applications to Mortgage Insurance and Title Insurance," *Review of Industrial Organization*, 28, 83–108.

Jaffee, D., and T. Russell, 1997, "Catastrophe Insurance, Capital Markets, and Uninsurable Risks," *Journal of Risk and Insurance*, 64, 205–230.

Jaffee, D., and J. Walden, 2014, "Optimal Insurance with Costly Internal Capital," *Risk Management and Insurance Review*, 17, 137–161.

Kahn, C., and A. Winton, 2004, "Moral Hazard and optimal subsidiary structure for financial institutions," *Journal of Finance*, 59, 135–153.

Kaluszka, M., and A. Okolewski, 2008, "An Extension of Arrow's Result on Optimal Reinsurance Contract," *Journal of Risk and Insurance*, 75(2), 275–288.

Lakdawalla, D., and G. Zanjani, 2012, "Catastrophe bonds, reinsurance, and the optimal collateralization of risk transfer," *Journal of Risk and Insurance*, 79, 449–476.

Leland, H. E., 2007, "Financial Synergies and the Optimal Scope of the Firm: Implications for Mergers, Spinoffs, and Structured Finance," *Journal of Finance*, 62, 7650807.

Lintner, J., 1965, "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets," *Review of Economics and Statistics*, 47, 13–37.

Mahul, O., and B. D. Wright, 2004, "Implications of Incomplete Performance for Optimal Insurance," *Economica*, 71, 661–670.

——— , 2007, "Optimal Coverage for Incompletely Reliable Insurance," *Economic Letters*, 95, 456–461.

Merton, R. C., and A. F. Perold, 1993, "Theory of Risk Capital in Financial Firms," *Journal of Applied Corporate Finance*, 6, 16–32.

Móricz, F. A., R. J. Serfling, and W. F. Stout, 1982, "Moment and probability bounds with quasisuperadditive structure for the maximum partial sum," *Annals of Probability*, 10, 1032–1040.

Myers, S. C., and J. A. Read, 2001, "Capital Allocation fo Insurance Companies," *Journal of Risk and Insurance*, 68, 545–580.

Nze, P. A., and P. Doukhan, 2004, "Weak dependence: models and applications to econometrics," *Econometric Theory*, 20, 995–1045.

Osborne, M. J., and A. Rubinstein, 1984, *A Course in Game Theory*. MIT Press, Cambridge.

Phillips, P. C. B., and V. Solo, 1992, "Asymptotics for linear processes," *Annals of Statistics*, 20, 971–1001.

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

Rosenthal, H. P., 1970, "On the subspaces of L^p ($p > 2$) spanned by sequences of independent random variables," *Israel Journal of Mathematics*, 8, 273–303.

Ross, S., 1976, "The arbitrage theory of capital asset pricing," *Journal of Economic Theory*, 13, 341–360.

Rothschild, M., and J. Stiglitz, 1970, "Increasing Risk: I. A definition," *Journal of Economic Theory*, 2, 225–243.

Serfling, R. J., 1970, "Convergence properties of S_n under moment restrictions," *Annals of Mathematical Statistics*, 41, 1235–1248.

Shao, Q.-M., 2000, "A comparison theorem on moment inequalities between negatively associated and independent random variables," *Journal of Theoretical Probability*, 13, 343–356.

Sharpe, W. F., 1964, "Capital asset prices: A theory of market equilibrium under conditions of risk," *Journal of Finance*, 19, 425–442.

Volcker, P., 2010, ""Statement" before the Committee on Banking, Housing and Urban Affairs of the United State Senate," February 2, 2010.

Winton, A., 1995, "Delegated Monitoring and Bank Structure in a Finite Economy," *Journal of Financial Intermediation*, 4, 158–187.

Zanjani, G., 2002, "Pricing and capital allocation in catastrophe insurance," *Journal of Financial Economics*, 65, 283–305.