Multiline Insurance with Costly Capital and Limited Liability

March 14, 2008

Abstract

We study a competitive multiline insurance industry, in which insurance companies with limited liability choose which insurance lines to cover and the amount of capital to hold. Premiums are determined by no-arbitrage option pricing methods. The results are developed under the realistic assumptions that insurers face friction costs in holding capital and that the losses created by insurer default are shared among policyholders following an ex post, pro rata, sharing rule. In general, the equilibrium ratios of premiums to expected claims and of default costs to expected claims will vary across insurance lines. We characterize the situations in which monoline and multiline insurance offerings will be optimal. Insurance lines characterized by a large number of essentially independent risks will be offered by very large multiline firms. Insurance lines for which the risks are asymmetric or correlated may be offered by monoline insurers. The results are illustrated with examples.

1 Background

The optimal allocation of risk in an insurance market was studied in the seminal work of Borch (1962), who showed that without frictions a Pareto efficient outcome can be reached. Furthermore, in a friction-free setting, insurers can hold sufficient capital to guarantee they will pay all claims. Two frictions, however, appear to be important in practice: 1) excess costs to holding capital, leading insurers to conserve the amount of capital they hold, and 2) limited liability, creating conditions under which insurers may fail to make payments to policyholders. When markets are incomplete, in the sense that policyholders face a counterparty risk that cannot be independently hedged, the existence of the two frictions can have a significant impact on the industry equilibrium, including the amount of capital held, the premiums set across insurance lines, and the industry structure regarding which insurance lines are associated with monoline versus multiline insurers.

In special cases, the impact of such frictions may be negligible and a no-frication — *perfect market* — approach approach is warranted. Such an approach is for example taken in Phillips, Cummins, and Allen (1998) (henceforth denoted PCA), in which no costs are associated with collateral, and markets are assumed to be complete. However, in many cases, the no friction assumption may be too simplistic. It is generally more realistic to presume that frictions do exist and that they may have an important impact on the equilibrium. For example, Froot, Scharsfstein, and Stein (1993) emphasize the importance of capital market imperfections for understanding a variety of corporate risk management decisions, with the tax disadvantages to holding capital within a firm an especially common and important factor. For insurance firms, Merton and Perold (1993), Jaffee and Russell (1997), Cummins (1993), 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.

The risk of insurer default in paying policyholder claims has lead to the imposition of strong regulatory constraints on the insurance industry in most countries. Capital requirements are one common form of regulation, although there is no systematic framework for determining the appropriate levels. As Cummins (1993) and Myers and Read (2001) point out, it is likely that the capital requirements are being set too high in some jurisdictions and too low in others, and similarly for the various lines of insurance risk, in both cases leading to inefficiency. It is thus important to have an objective framework for identifying the appropriate level of capital based on each insurer's particular book of business.

Insurance regulation also focuses on the industry structure, requiring certain high-risk insurance lines to be provided on a monoline basis. Monoline restrictions require that each insurer dedicate its capital to pay claims on its monoline of business, thus eliminating the diversification benefit in which a multiline firm can apply its capital to pay claims on any and all of its insurance lines.[1] Jaffee (2006), for example, describes the monoline restrictions

[1] Monoline restrictions do not preclude an insurance holding company from owning an amalgam of both monoline and multiline subsidiaries. The intent of monoline restrictions is that the capital of a monoline division must be dedicated to paying claims from that division alone.

imposed on the mortgage insurance industry, an industry, as it happens, currently at significant risk to default on its obligations as a result of the subprime mortgage crisis. Jaffee conjectures that the monoline restrictions were imposed as consumer legislation to protect the policyholders on relatively safe lines from an insurer default that would be created from large losses on a line with more catastrophic risks. It is thus valuable to have a framework in which the optimality of monoline versus multiline formats can be determined.

Although this paper is developed in the context of an insurance market, we believe the framework will be applicable to the issues of counterparty risk and monoline structures that are pervasive throughout the financial services industry. For example, the 1933 Glass Steagall Act forced US commercial banks to divest their investment bank divisions, while the 1956 Bank Holding Company Act of 1956 similarly separated commercial banking from the insurance industry; both Acts, repealed in 1999, were in effect monoline restrictions. In a similar fashion, Leland (2007) develops a model in which single-activity 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.

This paper provides a detailed analysis of the structure of an insurance market under the assumptions of costly capital, limited liability, incomplete markets and perfect competition between insurance companies. We specifically focus on the following questions:

1. Premia: For an insurance company offering insurance in multiple insurance lines, what will be the price structure across lines?

2. Cost allocation: For an insurance company offering insurance in multiple insurance lines, how should the firm allocate costs between these lines?

3. Choice of insurance lines: How will firms choose the basket of insurance lines to offer to their customers?

4. Choice of capital: Given a choice of insurance lines, what level of capital will an insurance company choose?

We introduce a consistent model to analyze each of these questions. Our results significantly extend and generalize the analyses in earlier papers, e.g., in Phillips, Cummins, and Allen (1998) and Myers and Read (2001). Three factors lead to these differences. First, we consider a competitive market, in which insurance companies (insurers) compete to attract risk averse agents who wish to insure risks (insurees). This competition severely restricts the monoline and multiline structures that may exist in equilibrium.

Second, we rely on the existence of a pricing kernel to price any risk,[2] but we make the additional assumption that insurees cannot replicate the insurance payoffs by trading in the

[2] In most parts of the paper, the pricing kernel can be quite general, although in some parts we will make the additional assumption that individual agents risks are idiosyncratic, so that the pricing operator coincides with the discounted objective expectations operator.

market.[3] This restriction implies that the insurance company provides value by tailoring insurance products that are optimal for its customers. Without the restriction, the insurance company is redundant — since any payoff can be replicated by trading in the asset market — and therefore the structure of the industry would be indeterminate.

Third, in the case of insurer bankruptcy, we assume that the insurer's available assets are distributed to the policyholders following what we call the *ex post pro rata rule*. Under this rule, the available assets are allocated to policyholder claims based on each claimant's share of the total claims. This rule has sensible properties and generally reflects the actual practice.[4]

The paper is organized as follows: In section 2 we give a review of related literature. In section 3, we introduce the basic framework for our analysis. In section 4, we analyze the first two questions — the pricing and cost allocation across lines in the case when the amount of capital and the choice of insurance lines is given. In section 5, we analyze the monoline versus multiline choice and the implications for industry structure. We analyze the insurance line choices in a competitive market from two angles: In what we denote the *traditional case*, there are many, essentially independent, risks available: In this case, insurance companies will be massively multiline oriented. In contrast, in what we call the *nontraditional case*, the market may be best served by monoline insurance companies. This may occur if there are a few lines, if losses between lines are highly correlated, or if loss distributions between lines are asymmetric. Finally, section 6 makes some concluding remarks.

2 Literature review

Financial models of insurance pricing and capital allocation were first developed by applying the principles of the capital asset pricing model (CAPM, see, for example, Fairley (1979) and Hill (1979)), or a discounted cash flow model (see, for example, Myers and Cohn (1987)). Both models, however, have significant drawbacks. The CAPM applications have the basic problem that they fail to incorporate the default risk faced by policyholders as a result of the insurance firms limited liability. The CAPM is also not well suited to pricing risks with heavy tails, as would be plausible for various lines of catastrophic disasters and terrorist attacks. The discount cash flow models must apply a risk-adjusted discount rate, but the derivations of this rate incorporate neither the frictional costs of holding capital nor the default risk for policyholders.

A major advance occurred by applying option valuation methods to the questions of insurance pricing and capital allocation, starting with the monoline models of Doherty and Garven (1986) and Cummins (1988). These papers specify the default risk faced by policyholders as a put option held by the equity owners of the insurer. The value of the option

[3]This could be because firms aggregate several sources of risks, and the market thereby is incomplete. It could also be that insurees do not have access to the stock market, or even that they are not sophisticated enough construct replicating portfolios in an asset market.

[4]This assumption seems to reflect the contracts offered to insurees in practice, as discussed below.

depends on the range of possible outcomes for policy claims and the amount of capital held by the insurer (which is the strike price of the option). The premium is then determined as the expected losses on the policy line minus the value of the default option. In other words, the ratio of the premium to the expected claim is less than 1, since the claims are not always paid in full.

The extension of option pricing methods to a multiline insurer was first provided by Phillips, Cummins, and Allen (1998) (PCA). Their analysis embeds the simplifying assumption that claims for all the lines are realized at the same date, which has the very useful implication that insurer default is simultaneously determined for all insurance lines. If the insurers assets equal or exceed the actual policyholder claims, then all claims are paid in full. Whereas, if the actual claims exceed the available assets, the insurer defaults, and pro rated payments are made to policyholders following a loss allocation rule. The specific rule used by PCA is that each policyholder is allocated a share of the shortfall based on the amount of her initially expected claims relative to the total of all initially expected claims. Since the shortfall shares are based on the expected claims as of the initial date, we will refer to this as the *ex ante* rule.

The PCA ex ante allocation rule implies that the premium to expected claims ratio will be constant across lines, since the default cost relative to expected claims is constant across lines. While this result is very powerful, the ex ante rule is a very special case, with the undesirable feature that the allocation of the shortfall in case of default is allocated to policyholders on the basis of only the aggregate shortfall and the initially expected loss on each insurance policy. In other words, the amount of the shortfall allocated to a policyholder depends only on her expected loss and is independent of that policyholders actual claims. Since the expected claims are not observable in the market, policyholders would have no basis to validate the share of the shortfall imposed on them. Moreover, policyholders with small expected losses would have to make payments to other policyholders with larger expected losses.

Mahul and Wright (2004) note that while an *ex ante* allocation rule may lead to optimal risk sharing among policyholders, in practice *ex post* payments from policyholders to other policyholders will be difficult to enforce.[5] Instead, in our model as explained below, we apply an *ex post* pro rata rule in which policyholders share the default shortfall in proportion to their actual claims. The result is that claimants always receive some net payment from the insurer, albeit less than their total claim when the insurer defaults.

How to allocate capital within a multiline insurer is another important question. PCA take the position that the allocation of capital is not needed for price determination when the insurance is sold in informationally efficient, competitive insurance markets. In our case, however, the frictional costs of holding capital make it imperative to allocate capital in order to determine the appropriate insurance premiums across lines. Two studies have been undertaken to work on this problem, Merton and Perold (1993) and Myers and Read (2001).

[5]The proper rule for allocating the default shortfall is related to the question of optimal contract design when insurer default may occur. Mahul and Wright (2004) is the most recent contribution to the small literature that discusses this issue, which also includes Schlesinger (2000) and Doherty and Schlesinger (1990).

In both cases, the amount of capital to be allocated to an insurance line is determined by an experiment in which the size of each line is changed and a computation is made of the resulting change in capital if the firms overall risk is to be constant.

The Merton and Perold (MP) experiment, more precisely, entirely removes an insurance line from the multiline firm and then computes the reduction in the insurers total capital requirements. The resulting reduction in capital is interpreted as the capital amount to be allocated when that line is part of the multiline firm. This procedure is then repeated for each line that the insurer covers. The MP method has the attribute that the sum of the capital allocations across all the lines will be less than the total capital required of the firm when it offers all the lines. The reason is that the overall firm receives a benefit of diversification that cannot be allocated to any of the individual lines.

The Myers and Read (MR) model also uses a marginal method to compute the capital allocation, but instead of removing each entire line from the total, they change the coverage amount of each line only be small incremental amounts. MR demonstrate that the capital allocations determined by their incremental technique satisfy the adding constraint whereby the sum of the capital allocations exactly equals the total amount of capital to be allocated. Our model, as developed below, applies the same concept for determining capital allocations and takes advantage of the same adding up condition. Our results differ from MR, however, because the MR computations are based on the PCA model's *ex ante* allocation rule for insurer shortfalls, whereas our results are developed on the basis of our *ex post* allocation rule.[6]

The last major topic considered in this paper is the determination of the industry structure in terms of which insurance lines are efficiently provided by monoline versus multiline insurers. We know of no papers that have considered this question within an analytic framework.

3 A competitive multiline insurance market

We first study the case of only one insured risk class. Consider the following one-period model of a competitive insurance market. At $t = 0$, an *insurer* (i.e., an insurance company) in a competitive insurance market sells insurance against a risk, $\tilde{L} \geq 0$,[7] to an *insuree*.[8] The expected loss of the risk is $\mu_L = E[\tilde{L}]$, $\mu_L < \infty$.

The insurer has limited liability and reserves capital within the company, so that A is available at $t = 1$, at which point losses are realized and the insurer satisfies all 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

[6] Zanjani provides are alternative approach to premium setting and capital allocation when there is default risk by specifying that the demand for insurance depends in part on the quality of the insurer. Although this approach can provide general conclusions comparable to the PCA and MR models, it lacks the quantitative precision that is provided by the default option approach. For our paper, this attribute of the option approach is critical in determining the optimal industry structure between monoline and multiline insurers.

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

[8] It is natural to think of each risk as an insurance line.

the additional amount that is due. 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.[9] The price for the insurance is P. Throughout the paper, the risk-free discount rate is normalized to zero.[10]

We assume that there are friction costs when holding capital within an insurer, including both taxation and liquidity costs.[11] The cost is δ per unit of capital. This means that to ensure that A is available at $t = 1$, $(1 + \delta) \times A$ needs to be reserved at $t = 0$.

There is a friction-free complete market for risk, admitting no arbitrage. The price for $\tilde{L}$ risk in the market is

$$P_L = Price\left(\tilde{L}\right) = E^*[\tilde{L}] = E[\tilde{m} \times \tilde{L}],$$

where $Price$ is a linear pricing function, which can be represented by the risk-neutral expectations operator $E^*[\cdot]$, or with the state-price kernel, $\tilde{m}$, in the objective probability measure,[12] and we assume that $E^*[\tilde{L}] < \infty$. Similarly, the price of the option to default is

$$P_Q = Price\left(\tilde{Q}(A)\right) = E^*[\tilde{Q}(A)] = E[\tilde{m} \times \tilde{L}].$$

Since the market is competitive and the cost of holding capital is δA, the price charged for the insurance is

$$P = P_L - P_Q + \delta A. \quad (1)$$

To ensure that A is available at $t = 1$, the additional amount of $A - P_L + P_Q$ needs to be reserved by the insurer. The total market structure is summarized in Figure 1, which also shows how noarbitrage pricing in the market for risk determines the price for insurance in the competitive insurance market.

It is natural to ask why insurees would impose the costs of holding capital by buying from an insurer instead of going directly to the market for risk. We make the assumption that insurees do not have direct access to the market for risk and that they can only insure through the insurers.[13]

The generalization to the case when there are multiple risk classes is straightforward. If

[9] When obvious, we suppress the A dependence, e.g., writing $\tilde{Q}$ instead of $\tilde{Q}(A)$.

[10] The results are qualitatively the same with a non-zero risk-free rate.

[11] Jaffee and Russell (1997) discuss a variety of costs that arise regarding insurer capital, including a risk of firm takeover as well as liquidity and tax costs. Myers and Read (2001) and and Cummins (1993) make similar assumptions.

[12] See, e.g., Duffie (2001) for standard results on existence and uniqueness of pricing functions under these completeness and noarbitrage assumptions.

[13] For example, if we think of the market for risk as a reinsurance market, this may be a natural constraint. A similar assumption, which would lead to identical results, is if the insuree faces costs that are equal to or higher than the costs faced by the insurer, in which case it will be optimal to buy from the insurer. Finally, if the market is incomplete and the insurer is risk-neutral, there may be no way to replicate the payoffs in the market for risk, leaving the insurance market as the sole market available for the insuree.

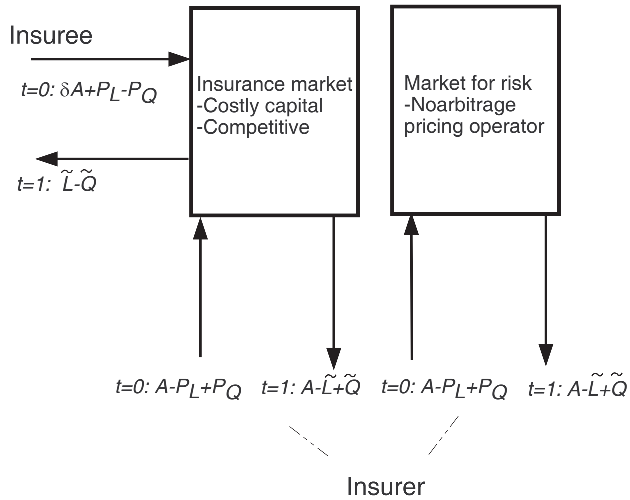

Figure 1: *Structure of model. Insurers can invest in market for risk or in a competitive insurance market. There is costly capital, so to ensure that A is available at $t = 0$, $(1 + \delta)A$ needs to be reserved at $t = 1$. The premium, $\delta A + P_L - P_Q$, is contributed by the insuree and $A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

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

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

where $\tilde{L} = \sum_i \tilde{L}_i$ and $\tilde{Q}(A) = \max(\tilde{L} - A, 0)$. The total price for the risks is, $P \stackrel{\text{def}}{=} \sum_i P_i$, where P_i is the price for insurance against risk i, is once again on the form (1).

Now consider an insurance market, in which M insurers sell insurance against $N \geq M$ risks. We partition the total set of N risks into $\mathcal{X} = \{X_1, X_2, \ldots, X_M\}$, where $\cup_i X_i = \{1, \ldots, N\}$, $X_i \cap X_j = \emptyset, i \neq j$, $X_i \neq \emptyset$. The partition represents how the risks are insured by M monoline or multiline insurance businesses. The total industry structure is then characterized by the duple, $\mathcal{S} = (\mathcal{X}, \mathbf{A})$, where $\mathbf{A} \in \mathbb{R}_+^M$ is a vector with i:th element representing the capital available in the multiline business that insures the risks for agents in X_i.[14] The number of sets in the partition is denoted by $M(\mathcal{X})$. Two two polar cases are the fully multiline industry structure, with $\mathcal{X} = \{\{0, 1, \ldots, N\}\}$ and the monoline industry structure, with $\mathcal{X} = \{\{0\}, \{1\}, \ldots, \{N\}\}$

Our analysis so far is thus based on the following assumptions:

1. *Market completeness*: The market for risk is arbitrage-free and complete, such that

[14]We use the notation $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$ and $\mathbb{R}_- = \{x \in \mathbb{R} : x \leq 0\}$.

there is a linear pricing operator.

2. *Limited liability*: Insurers have limited liability.

3. *Costly capital*: There is a cost for insurers to hold capital.

4. *Competitive insurance markets*: Prices for insurance are set competitively.

5. *Access to markets*: Insurees do not have direct access to the market for risk.

These assumptions completely determine the pricing of risk and cost allocation between different lines, as shown in section 4.

To understand the prevailing market structure, $\mathcal{S} = (\mathcal{X}, \mathbf{A})$, in an economy – the objective of section 5 – we also need assumptions about insurees. For simplicity, we assume that there are N distinct insurees. Each risk is insured by one insuree with expected utility utility function u, where u is a strictly concave, increasing function defined on the whole of $\mathbb{R}_-$. For some of the results we need to make stronger conditions on u.[15] The risk can not be divided between multiple insurers.[16] Finally, we assume that expected utility, U, is finite, $U = Eu(-\tilde{L}) > -\infty$.

We will make extensive use of the certainty equivalent as the measure of the size of a risk. For a specific utility function, u, the certainty equivalent of risk $\tilde{L}$, $CE_u(-\tilde{L}) \in \mathbb{R}$ is defined such that $u(CE_u(-\tilde{L})) = E[u(-\tilde{L})]$, where $E[\cdot]$ is the (objective) expectations operator.

Finally, we will assume that the risks are idiosyncratic, i.e., that risk-neutral expectations coincide with objective expectations, $E^*[\cdot] \equiv E[\cdot]$.

To summarize, in section 5, the following additional assumptions are made:

6. *Risk-averse insurees*: Insurees are risk averse.

7. *Nondivisibility*: Risks are nondivisible.

8. *Idiosyncratic risks*: The insurance risks are idiosyncratic, i.e., the risk-neutral expectations operator coincides with the objective expectations operator.

For many types of individual and natural disaster risks, such as auto and earthquake insurance, etc., this seems a reasonable assumption, although, of course, there will be some mega-disasters and corporate risks for which it is not true.

[15] We do not distinguish between lines of risks and individual risks, for simplicity assuming that there is one insuree within each line. Obviously, so far this is no restriction, since, in principle, N can be very large. If we wish to study a case with a "small" N, for the special case when there are several identical agents with perfectly correlated risks, we can treat such a situation as there being one representative insuree facing one risk, collapsing many risks into one line. In case of many i.i.d. risks a similar argument may be made.

[16] Such insurances are seldom seen in practice; maybe because of the agency problems that would prevail between insurers when handling split insurance claims.

4 Pricing and cost allocation

We now turn to the insurer's main question, how costs are allocated and premiums are set across insurance lines (questions 1 and 2 in the introduction). For the time being, we assume that the cost of holding capital is zero, $\delta = 0$, and focus on premium setting, i.e., how the total premium, P should be split between the insurees.

What is missing is a rule for how payments are shared between claim-holders in case of default. There are obviously many such rules, but a minimal set of consistency requirements is

Condition 1 Consistency requirement:

1. *In case of no default: The payment to each insuree is exactly the amount claimed:* $Payment_i = \tilde{L}_i$.
2. *In case of default: The payment to each insuree is bounded above by the claim, and below by zero:* $0 \leq Payment_i \leq \tilde{L}_i$.
3. *No-claim policy: Insurees with no claims do not receive payments:* $\tilde{L}_i = 0 \Rightarrow Payment_i = 0$.
4. *Linearity: If* $\tilde{L}_1, \ldots, \tilde{L}_N$ *is insured, with capital* A*, leading to payments* $(Payment_1, \ldots, Payment_N)$*, then insurance against risks* $c\tilde{L}_1, \ldots, c\tilde{L}_N$*, with capital* $c \times A$ *leads to payments* $(c \times Payment_1, \ldots, c \times Payment_N)$ *for all* $c > 0$.

We deviate from the optimal contracting setting here, in that we focus on contracts that are present in practice. If more general contracts are possible, it may for example be optimal to have ex post transfers from claimants who did not impose any losses to claimants who did; such contracts, which would effectively turn the insuree into an insurer in some states of the world, do not seem to exist in practice. It is out of the scope of this paper to analyze why such contracts are rarely seen — we refer to existing literature, e.g., Mahul and Wright (2004).

We first focus on an *ex post pro rata* sharing rule that specifies that:

$$Payment_i = \frac{\tilde{L}_i}{\tilde{L}} \times Total\ Payment = \tilde{L}_i - \tilde{L}_i \times \frac{\tilde{Q}(A)}{\tilde{L}}, \tag{2}$$

i.e., in case of default the insurees share the total payments according to their fraction of total losses. In this case, the market price for insurance in line i is

$$P_i = P_{L_i} - Price\left(\tilde{L}_i \times \frac{\tilde{Q}(A)}{\tilde{L}}\right). \tag{3}$$

It is easy to check that the *ex post* sharing rule satisfies the consistency condition. Moreover, it seems to correspond well to the rules used in practice.[17] An important variable is the *default value fraction*, r_i, i.e., the fraction of the default option value that is allocated to line i:

$$r_i \stackrel{\text{def}}{=} Price\left(\frac{\tilde{L}_i}{\tilde{L}} \times \frac{\tilde{Q}(A)}{P_Q}\right). \tag{4}$$

Per definition, we have $\sum_i r_i = 1$, and the *default value per unit of risk* for line i, z_i, is then

$$z_i \stackrel{\text{def}}{=} \frac{r_i P_Q}{P_{L_i}} = \frac{1}{P_{L_i}} \times Price\left(\frac{\tilde{L}_i}{\tilde{L}_i} \times \tilde{Q}(A)\right). \tag{5}$$

Equation (3) can then be written

$$P_i = P_{L_i} - r_i P_Q, \tag{6}$$

and we have

$$\frac{P_i}{P_{L_i}} = 1 - z_i, \tag{7}$$

where $\frac{P_i}{P_{L_i}}$ is the *premium-to-liability* ratio as defined in Phillips, Cummins, and Allen (1998). Clearly, the premium-to-liability ratio may vary with i under the *ex post* sharing rule (verified in an example in the appendix). Equation (6) is the fundamental equation for how premiums will be set.

An alternative expression for the premium-to-liability ratio, using the state price kernel kernel, $\tilde{m}$, is

$$\frac{P_i}{P_{L_i}} = 1 - z_i = 1 - \frac{1}{P_{L_i}} \times E\left[\tilde{m} \times \frac{\tilde{L}_i}{\tilde{L}} \times \tilde{Q}(A)\right], \tag{8}$$

from which we see that z_i will be larger for risks that tend to make up a large fraction of total losses in the states of the world when a company defaults, and for risks that are positively correlated with the market. This leads to:

Implication 1

- *For risks that are not related to market risk, with losses that tend to be large in states of the world in which the insurer defaults, premium-to-liability ratios will be low.*
- *For risks that are not related to market risk, with losses that tend to small in states of the world in which the insurer defaults, premium-to-liability ratios will be high.*

[17] For example, see National Association of Insurance Commissioners, Insurer Receivership Model Act, October 2007. An ex-post payout rule is also specified in the contracts offered by the California Earthquake Authority for circumstances in which the Authority has insufficient resources all of its claims. See also (Mahul and Wright 2004) for a discussion of alternative payoff rules and the complications created by rules other than ex-post prorated payments.

- *For risks that have high correlation with market risk, premium-to-liability ratios will be low.*

- *For risks that have low correlation with market risk, premium-to-liability ratios will be low.*

These implications are different than the predictions in Phillips, Cummins, and Allen (1998). According to PCA (Hypothesis 2, page 609), the default premium should be constant across lines.[18]

The source behind these differences is that PCA's results are based on an *ex ante* sharing rule, in which the default option value is allocated between lines according to the value of the default-free claim at $t = 0$ (instead of according to the realized values at $t = 1$).The payment under the *ex ante* sharing rule is

$$Payment_i = \tilde{L}_i - \frac{P_{L_i}}{P_L} \times \tilde{Q}(A), \tag{9}$$

and it is easy to check that in this case

$$z_i = \frac{P_Q}{P_L}, \qquad \text{and} \qquad \frac{P_i}{P_{L_i}} = 1 - \frac{P_Q}{P_L}, \tag{10}$$

which do not vary with i. This is equation (18) in PCA (assuming that discount rate and inflation are 0). We also note that the ex ante sharing rule through (10) implies that

$$r_i = \frac{P_{L_i}}{P_L}, \tag{11}$$

which, in general, is different from (4). As we have argued, we believe that, in practice, ex post sharing rules are more common. Moreover, the ex ante rule will in general not satisfy the consistency condition:

Implication 2 *The* ex post *sharing rule satisfies the consistency condition, whereas the* ex ante *sharing rule, in general, does not.*

In the appendix, we introduce a simple example that emphasizes the difference between the two sharing rules, and that violates consistency for the *ex ante* sharing rule.

We now turn to question 2, how capital, A, and costs, δA, should be allocated between business lines. Our analysis is along the lines of Myers and Read (2001), which is also based

[18]PCA assume that $\delta = 0$ — an assumption that we will relax in subsequent analysis. However, their model is also more general: It is dynamic and includes an inflation premium, as well as risky processes for the returns of A over time. Our model could be generalized along such lines, but the sources to the differences between our and PCA's approach are not related to these factors, which is why we can keep our analysis simple.

on an *ex ante* sharing rule, but since we assume an *ex post* sharing rule, our results are different. For the time being, we keep assuming that $\delta = 0$.

In general, we wish to allocate the assets A among lines, $A = \sum_i A_i = \sum_i v_i A$, where v_i is the *relative asset allocation.* As showed in Myers and Read (2001), any allocation rule, v_i, ($\sum_{i=1}^N v_i = 1$) implies a "summing up" relationship, i.e., if we study insurance portfolios

$$\sum_i q_i \tilde{L}_i,$$

then the payout of the default option is

$$\tilde{Q}_q = \max\left(\sum_i q_i(-\tilde{L}_i - v_i A_i), 0\right),$$

and the summing up rule states that

$$\sum_i q_i \frac{\partial \tilde{Q}_q}{\partial q_i} \equiv \tilde{Q}_q$$

in all states of the world, for all $q_1, \ldots, q_N > 0$ (we show how the rule arises in the appendix; the argument is similar to Myers and Read (2001)).

The rule immediately implies that

$$\sum_i q_i \frac{\partial P_{Q_q}}{\partial q_i} = P_{Q_q}.$$

However, this in turn implies that the marginal price of buying one extra unit of $\tilde{L}_i$ risk, at $q_1 = q_2 = \cdots = 1$, is

$$P_i = P_{L_i} - \frac{\partial P_{Q_q}}{\partial q_i}. \tag{12}$$

Therefore, using the definition of r_i, we can conclude that any cost allocation rule leads to the allocation of relative default option value

$$r_i = \frac{\partial P_{Q_q}/\partial q_i}{P_Q}. \tag{13}$$

So far, we have made no restrictions on the v_i's except for that they should sum to one. However, as shown in the appendix, consistency requires that there is bijection, $R : \{r_i\} \leftrightarrow \{v_i\}$, between relative asset allocations, v_i, (satisfying $\sum_i v_i = 1$) and allocation of relative default option values, r_i's, (also satisfying $\sum_i r_i = 1$), so that given r_i's, the v_i's are uniquely defined. Any other formula will lead to mispricing in the existing lines. We note that v_i also determines the surplus allocations: $S_i = v_i A - P_{L_i}$ and thereby the relative surplus allocations, $s_i \stackrel{\text{def}}{=} S_i/P_{L_i} = A\frac{v_i}{P_{L_i}} - 1$.

It is shown in the appendix that the correct choice under the *ex post* sharing rule takes the particularly simple form $v_i = r_i, i = 1, \ldots, N$. This is our fundamental formula for how capital and costs should be allocated to different lines, which we formulate in the following

Implication 3 *For $\delta = 0$, if an* ex post *sharing rule is used, then the only asset allocation rule that does not lead to redistribution between new and old insurees in case of a marginal expansion in a specific insurance line is $v_i = r_i$, $i = 1, \ldots, N$.*

This is not the same allocation as suggested in Myers and Read (2001). On page 554 it is suggested that the marginal contribution to default value, d_i — in our notation, defined as $d_i \stackrel{\text{def}}{=} \frac{\partial P_Q}{\partial P_{L_i}}$, i.e., the increase in default option value for a one dollar increase in liability — should be chosen such that d_i is the same across lines. Via (13), and the relationship $\frac{\partial P_Q}{\partial P_{L_i}} = (\partial P_Q/\partial q_i)/P_{L_i}$, we have

$$d_i = P_Q \times \frac{r_i}{P_{L_i}}, \tag{14}$$

which, under the *ex post* sharing rule, is (in general) not constant. However, under the *ex ante* sharing rule, we get $d_i = P_Q/P_L$ (since $r_i = P_{L_i}/P_L$), so the equal allocation rule proposed in Myers and Read (2001) is consistent with the *ex ante* sharing rule. In the appendix we show in an example that the cost allocation rule used in Myers and Read (2001) leads to redistribution between old and new insurees under the *ex post* sharing rule. This means that the original prices — which were calculated without taking the possibility of redistribution into account — are incorrect. Thus, the equal allocation rule, as expected, is not consistent with the *ex post* sharing rule.

We now generalize the analysis to the case when $\delta > 0$. Following rule (4), the corresponding v_i should be used to allocate assets: $\{v_i\} = R(\{r_i\})$. Therefore, in the case of costly capital, $\delta > 0$, the pricing formula becomes:

$$P_i = P_{L_i} + r_i \delta A - r_i P_Q. \tag{15}$$

Thus, we have

Implication 4 *For all $\delta \geq 0$, if an* ex post *sharing rule is used, then assets $A_i = v_i A$ should be allocated to line i, where $v_i = r_i$ and r_i is defined as in (4). This implies the pricing formula (15).*

We summarize the differences between the *ex post* and *ex ante* in Table 1.

5 Industry structure

We now study how the industry structure – monoline versus multiline – and the related capital allocations are determined (questions 3 and 4 of the introduction). To analyze these

Variable	**Definition**	***ex post*** **sharing rule**	***ex ante*** **sharing rule**
Default option payment, $\tilde{Q}(A)$	$\tilde{Q}(A) = \min(\tilde{L}, A)$		
Payment to insuree		$\tilde{L}_i - \frac{\tilde{L}_i}{\tilde{L}} \times \tilde{Q}(A)$	$\tilde{L}_i - \frac{P_{L_i}}{P_L} \times \tilde{Q}(A)$
Price of default-free liability, P_{L_i}			
Price of defaultable insurance, P_i			
Allocation of default option	$r_i = \frac{P_{L_i} - P_i}{P_Q}$	$r_i = Price\left[\frac{\tilde{L}_i \times \tilde{Q}(A)}{\tilde{L} \times P_Q}\right]$	$r_i = \frac{P_{L_i}}{P_L}$
Default option per unit liability	$z_i = \frac{r_i}{P_{L_i}} \times P_Q$	$r_i = Price\left[\frac{\tilde{L}_i \times \tilde{Q}(A)}{\tilde{L} \times P_{L_i}}\right]$	$z_i = \frac{P_Q}{P_{L_i}}$
Relative asset allocation,	$v_i = \frac{A_i}{A}$	$v_i = r_i$	$\{v_i\} = R(\{r_i\})$
Premium-to-liability ratio	$\frac{P_i}{P_{L_i}}$	$1 - z_i$	$1 - \frac{P_Q}{P_L}$
Marginal contribution to default value	$d_i = \frac{r_i}{P_{L_i}} \times P_Q$	$Price\left[\frac{\tilde{L}_i \times \tilde{Q}(A)}{\tilde{L} \times P_{L_i}}\right]$	$\frac{P_Q}{P_L}$
Surplus allocation	$S_i = v_i A - P_{L_i}$	$S_i = r_i A - P_{L_i}$	$S_i = v_i A - P_{L_i}$
Relative surplus allocation	$s_i = \frac{v_i}{P_{L_i}} \times A - 1$	$s_i = \frac{r_i}{P_{L_i}} \times A - 1$	$s_i = \frac{v_i}{P_{L_i}} \times A - 1$

Table 1: Relationship between variables under *ex post* and *ex ante* sharing rules.

questions given a fixed level of capital and prices — although quite straightforward — may give quite misleading results. For example, an insurance company choosing to be massively multiline may wish to have a lower level of capital than the total capital of monoline business insuring the same risks. Moreover, the risk structure of insurance in a multiline business may be quite different than in a monoline business.

In a competitive market, we would expect such differences to have pricing implications, since insurees have propensities to pay different amounts for different risk structures. Therefore, before answering the questions of industry structure, we must first study how capital, A, and price, P, are *endogenously* determined in a competitive market:

- Monoline pricing: For an insurance company offering insurance in a single insurance line in a competitive market, what price will be charged for insurance as a function of the level of capital?

- Monoline level of capital: For a monoline insurance company, what level of capital will be chosen?

In section 5.1, we analyze these questions in a competitive market setting with costly capital, $\delta > 0$. Then in sections 5.2 and 5.3, we extend this analysis to the multiline setting. Only then can we analyze questions 3 and 4 of the introduction.

5.1 Capital and price in the monoline case

So far, we have relied on noarbitrage, ensuring the existence of a risk-neutral expectations operator. We continue with this general set-up, assuming that the risk-neutral measure, is equivalent to the objective probability measure.[19]

In the rest of the section, we focus on the monoline case, $N = 1$. We assume that $\tilde{L}$ has an absolutely continuous, strictly positive, p.d.f with support on the whole of $\mathbb{R}_+$. We define the default option's "Eta", i.e., $\eta(A) = \frac{\partial E^*[\tilde{Q}(A)]}{\partial A}$. Since $\tilde{L}$ has an absolutely continuous, strictly positive, distribution, $\eta(A)$ is a continuous strictly negative function on $(0, \infty)$ regardless of the distribution of $\tilde{L}$, and the risk-neutral measure (Ingersoll 1987). We study the price of insurance, P as a function of capital, A. It is straightforward to show that

Lemma 1 *The price of insurance as a function of capital, A, satisfies the following conditions*

i) $P(0) = 0$,
ii) $P'(A) = \delta - \eta(A) > 0$,
iii) $P(A) = P_L + \delta A + o(1)$, *for large* A,
iv) $P''(A) < 0$.

Thus, regardless of the distributional form, $P(A)$ will be a strictly increasing, strictly concave function with known asymptotics.

[19] That is, $E[\tilde{L}] = 0 \Leftrightarrow E^*[\tilde{L}] = 0$ for all risks, $\tilde{L}$

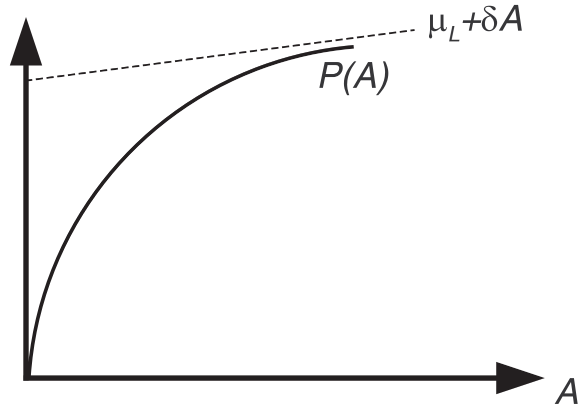

Figure 2: *Insurance premium as a function of capital.*

The conditions in Lemma 1 are natural. The first condition states that if the insurer does not put aside any capital, it may charge no premium (anything else would be an arbitrage opportunity). The second condition relates a small change in capital, A with the change in premium P. In light of equation (1), it is natural that the change in option value will enter into such a relation. The third condition shows that as A becomes large, the premium approach the price of insurance with unlimited liability, P_L (as the option value of defaulting disappears) and costs of keeping capital within the firm (which will become large as they are proportional to capital). The fourth condition, which follows as a direct consequence of the convexity of an option's value as a function of strike price (see Ingersoll 1987), states that P is concave.

The optimal (A, P) pair will depend on the preferences of the insuree. We therefore turn to the insuree's problem. In a general model with other sources of risk, we would relate the risk-neutral measure to the insuree's expected utility function. However, to keep things simple, we make additional assumptions about the risks. Specifically, from here on, we rely on assumptions 6-8 in the introduction, i.e., that insurees are risk-averse, that risks are nondivisible and idiosyncratic. We also make the fairly standard partial equilibrium assumption, that the $\tilde{L}$-risk is the only source of risk the insuree faces.[20]

The pricing relation (1) can then be written:

$$P(A) = \mu_L + \delta A - \mu_Q, \tag{16}$$

where we have defined $\mu_Q = \mu_{Q(A)} = E[\tilde{Q}(A)]$.

Given the competitiveness in the insurance market, the insurer will choose capital, A that maximizes the expected utility of the insuree, i.e., since the total payoff to the insuree

[20]Or, in the special case of CARA utility, that any other source of risk is independent of $\tilde{L}$-risk.

is $-P(A) - \tilde{L} + (\tilde{L} - \tilde{Q}(A)) = -P(A) - \tilde{Q}(A)$,

$$A^* = \underset{0 \le A < \infty}{\arg\max}\, Eu[-P(A) - \tilde{Q}(A)]. \tag{17}$$

For example, any lower value of A than A^* would allow a competitor to take over the whole market by offering a contract with a preferable, that is to say higher, value of A. In general, A^* is a set, i.e., there can be multiple solutions to (17).

If $\delta = 0$, it is easy to show that the company will reserve an arbitrary large large amount of capital. Formally, the solution is $A^* = \{\infty\}$ and the price is $P = \mu_L$. We call this the *friction-free outcome*, since the insurer never defaults and all risk is transferred from the insuree to the insurer in an optimal manner. In this case the expected utility of the insuree is $U = u(-\mu_L)$ and the certainty equivalent of his utility decrease is the same as if he were risk-neutral, $CE_u(\tilde{L}) = -\mu_L$.

When capital is costly, $\delta > 0$, it is not possible to obtain the friction-free outcome. We assume that the cost of holding capital is small compared with expected losses. Specifically, we assume that

Condition 2 *$CE_u(-P(A) - \tilde{Q}(A)) < -\mu_L(1 + \delta)$ for all $A \in [0, \mu_L]$.*

This implies that each risk is potentially insurable in that if an insurer could guarantee default-free insurance against a risk by capitalizing the *expected* losses, the agent would be willing to buy such insurance, paying the cost of holding capital for the expected losses. In reality, the insurance company would keep a higher level of capital and would still risk default. In the case of costly capital, the best we can therefore hope for, is for the insuree to reach a certainty equivalent of $-\mu_L(1 + \delta)$. We therefore call an outcome in which an agent obtains $CE_u = -\mu_L(1 + \delta)$ the *ideal outcome with frictions*.

It is easy to show that the set of solutions to (17) is compact and nonempty. However, it may be that $0 \in A^*$, i.e., it is optimal not to offer insurance. In fact, for insurees that are close to risk neutral, we would expect no insurance to be optimal, since the loss of reserving capital would always be greater than the gain from reduced risk. We first wish to understand in which situations it is potential for insurance to exist, i.e., when there exists a utility function such that $0 \notin A^*$. We have

Proposition 1

For a risk $\tilde{L}$ and cost of holding capital $\delta > 0$, there exists a strictly concave utility function, u, such that $0 \notin A^*$ for an insuree with utility function u, if and only if there is a level of capital, A, such that the price, P, satisfies $P < A$, where P is defined in (1).

The "only if"-part of the proposition is immediate, since if it fails, it would be less expensive for the insuree to reserve the capital than to buy the insurance. Clearly, we would only expect this to be the case in cases of very large δ. The "if"-part is proved in the appendix.

We also wish to be able to rank risks when the insurance market is present. Without an insurance market, stochastic dominance can be used. Given two risks, with payoff $-\tilde{L}_1$ and $-\tilde{L}_2$, with $\mu_{L_1} = \mu_{L_2} = \mu_L$, $Eu(-\tilde{L}_1) \geq Eu(-\tilde{L}_2)$ for all utility functions, if and only if $-\tilde{L}_1$ second order stochastically dominates $-\tilde{L}_2$,

$$-\tilde{L}_1 \succeq -\tilde{L}_2. \tag{18}$$

If F_1 and F_2 are the c.d.f.'s of $-\tilde{L}_1$ and $-\tilde{L}_2$ respectively (with range in $\mathbb{R}_-$), we know from Rothschild and Stiglitz (1973) that second order stochastic dominance is equivalent to the so-called integral condition:

$$\int_{-\infty}^{t} F_1(x)dx \leq \int_{-\infty}^{t} F_2(x)dx,$$

for all $t < 0$.

Is there a similar ranking when the insurance market is present? To analyze this question, we define $\tilde{Q}_1$ and $\tilde{Q}_2$ as the option payoffs from default, for risk 1 and 2 respectively. In what follows, we restrict our attention to cases in which it is optimal for an insurer to buy insurance and the optimal capital is greater than the expected loss, $A^* > \mu_L$. This is obviously a situation which that we expect to have in a standard insurance setting.

We recall that

$$P = \delta A + \mu_L - \mu_Q, \tag{19}$$

which we use to rewrite

$$U = Eu\left[-P - \tilde{Q}\right] = Eu\left[-\mu_L - \delta A + (\mu_Q - \tilde{Q})\right] \tag{20}$$

For a given A, (20) implies that regardless of utility function, an investor will be better off facing risk $\tilde{L}_1$, than $\tilde{L}_2$ if and only if

$$-(\tilde{Q}_1 - \mu_{Q_1}) \succeq -(\tilde{Q}_2 - \mu_{Q_2}). \tag{21}$$

Clearly, (21) is not the same as (18), so we can not expect second order stochastic dominance to allow us to rank risks in the presence of an insurance market. Instead, we have the stronger condition

Proposition 2 *Given an insurer with capital A, if for all $t < -\mu_L$,*

$$\int_{-\infty}^{t+\mu_L} F_1(x)dx \leq \int_{-\infty}^{t} F_2(x)dx, \tag{22}$$

then any insuree with a strictly concave utility function will prefer to insure risk $\tilde{L}_1$ over risk $\tilde{L}_2$ in a competitive monoline insurance market.

We study the differences in the following example:

Example 1 *Consider the risks* $\tilde{L}_\beta$, $\beta \geq 1$, *where the c.d.f. of* $-\tilde{L}_\beta$ *is* $F_\beta(x) = e^{\beta(x+1)-1}$, $x < 1/\beta - 1$ *that are shifted, reflected, exponential distributions.*[21] *It is clear that* $\mu_L = E[\tilde{L}_\beta] = 1$ *and, furthermore, it is easy to check that for* $\beta_1 > \beta_2$, $F_{\beta_1}(x) - F_{\beta_2}(x-1) < 0$ *for* $x < -\beta_2/(\beta_1 - \beta_2)$, *and* $F_{\beta_1}(x) - F_{\beta_2}(x-1) \geq 0$ *otherwise. Therefore* $\int_{-\infty}^{t}(F_{\beta_1}(x) - F_{\beta_2}(x-1))dx$ *realizes a maximal value at* $t = \mu_L = -1$, *and it is straightforward to check that*

$$\beta_1 \geq \beta_2 e^{\beta_2}$$

is a necessary and sufficient condition for the conditions in Proposition 2 to be satisfied. This is obviously a stronger condition than $\beta_1 \geq \beta_2$, *which is what is needed for second order stochastic dominance.*

5.2 The monoline versus multiline business choices

We now have almost all of the machinery to study questions 3 and 4 under the competitive market assumption. What we still need is a notion of competitive markets in the multiline setup. We have already used the assumption of competitive markets to completely understand the pricing and choice of level of capital in the monoline case. Specifically, we used the argument to only study outcomes that satisfied equation (17). In the multiline case, however, the analysis is slightly more complex, since several possible industry structures, $\mathcal{S}$, may be possible, and since there is now a trade-off between providing utility to multiple agents. Our restriction will therefore be to require Pareto efficiency.

We first note that for N risks, $\tilde{L}_1, \dots, \tilde{L}_N$, 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, $i \in X_j$, is

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

His expected utility is therefore $Eu_j(-P_i(A_i) + \tilde{K}_i(\mathcal{S}))$. Moreover, for a set of agents, $u_1, \dots, u_N$, each wishing to insure risk $\tilde{L}_i$, an industry structure, $\mathcal{S}$, where $\mathcal{X} = \{X_1, \dots, X_M\}$ and $\mathbf{A} = (A_1, \dots, A_M)^T$, is Pareto efficient, if there is no industry structure $\mathcal{S}'$ such that $E[u_i(-P_i(A_i) + \tilde{K}_i(\mathcal{S}))] \leq E[u_i(-P_i(A_i) + \tilde{K}_i(\mathcal{S}'))]$ for all i and $E[u_i(-P_i(A_i) + \tilde{K}_i(\mathcal{S}))] < E[u_i(-P_i(A_i) + \tilde{K}_i(\mathcal{S}'))]$ for some i.[22]

In a Pareto dominated industry structure, we would expect insurers to enter the market with improved offerings, thereby outcompeting existing insurers. In fact, we make a somewhat stronger requirement, that there should be no way to way to improve the situation for a single insuree by offering that insuree a monoline insurance — even if that makes other agents worse off.

[21] The restriction $\beta \geq 1$ can be extended to $\beta > 0$ at the cost of allowing for $\tilde{L}$ to be less than zero. All derivations go through in this case too.

[22] Here, Pareto efficiency is defined given the (restricted) set of limited liability contracts available.

Definition 1

- *A Pareto efficient outcome, $\mathcal{S}$, is said to be* robust to monoline blocking*, if there is no insuree, $i \in \{1,\ldots,N\}$ such that $E[u_i(\tilde{K}_i(\mathcal{S}))] < E[u_i(-P(A) - \tilde{Q}(A)]$ for some $A \geq 0$.*
- *The set of Pareto efficient outcomes robust to monoline blocking is denoted by $\mathcal{O}$.*

Remark 1 *The concept of robustness to monoline blocking is similar to the* core *concept used in coalition games, although, in general, $\mathcal{O}$ is neither a subset, nor a superset of the core.*[23]

We are interested in $\mathcal{O}$, since we believe that it may be easier for a competitor to compete for customers within one line of business, than in for a customers in multiple lines simultaneously. Technically, the monoline blocking condition allows us to show that $\mathcal{O}$ is always nonempty in (as opposed to the core in our setting).

The following existence and compactness results are straightforward to derive

Lemma 2

- *$\mathcal{O}$ is nonempty.*
- *The set of A's such that $A = \mathbf{A}_i$ for some $(\mathcal{X}, \mathbf{A}) \in \mathcal{O}$ is compact.*

What can we say about industry structure when there are many risks available? Intuitively, when capital is costly and there are many risks available, we would expect that 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 than the monoline business.[24] The argument is very general, as long as there are enough risks to pool, that are not too correlated. For example, in our model, under general conditions, the multiline business can reach an outcome arbitrary close to the ideal outcome with frictions. We have:

[23] See, e.g., Osborne and Rubinstein (1984). In our model, monoline structures may dominate multilines, leading to non-cohesiveness, which means that the core may contain Pareto-dominated outcomes. Therefore, there may be outcomes in the core that are not in $\mathcal{O}$. On the other hand, the core is robust to blocking/competition by any insurance company (monoline or multiline) which is stricter than the mono-blocking condition for $\mathcal{O}$, and moreover, $\mathcal{O}$ may contain other structures than the partition into one massively multiline business, so $\mathcal{O}$ may contain elements that are not in the core. In the case of cohesive games, the core is a subset of $\mathcal{O}$, since any element in the core will be Pareto efficient. If, in addition, there are only two lines, the core is the same as $\mathcal{O}$, since only mono-blocking is possible.

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

Proposition 3 *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 three times continuously differentiable, that u''' is bounded by a polynomial of degree q, and that the risks L_i are such that $\left|\sum_{i=1}^{N} \mu_i\right| \geq CN$ for some $C > 0$ and $E|L_i|^p \leq C$ for $p = 4 + 2q$ and some $C > 0$.*

Then, regardless of the per unit cost of holding capital, δ, as N grows, a fully multiline industry, $\mathcal{X} = \{\{1, \ldots, N\}\}$ can choose capital A, to reach an outcome that converges to the ideal outcome with frictions as N grows, i.e.,

$$\min_i CE_u(\tilde{K}_i((\mathcal{X}, A))) = \mu_{Li}(1 + \delta) + o(1).$$

Remark 2 *Proposition 3 can be generalized in several directions, e.g., to allow for dependence. As follows from the proof of the proposition, it also holds for all (possibly dependent) risks L_i with $E|L_i|^p < C$ that satisfy Rosenthal inequality (see Rosenthal (1970)). Rosenthal inequality 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)). Using Phillips-Solo device (see Phillips and Solo (1992)) similar to the proof of Lemma 12.12 in Ibragimov and Phillips (2004), one can show that it is also satisfied 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 obtain extensions of Proposition 3 to the case of losses that satisfy nonlinear moment assumptions.*

The Proposition provides an upper bound of the number of risks that need to be pooled to get close to the friction-free outcome. For a lower bound on the number of risks needed, we have the following proposition

Proposition 4 *If, in additions to the assumptions of proposition 3, the risks are uniformly bounded: $\tilde{L}_i \leq C$ (a.s.) for all i, and Condition 2 is satisfied, then for every $\epsilon > 0$, there is an N such that $\lim_{\epsilon \searrow 0} N(\epsilon) = \infty$ and such that any partition that has*

$$CE_u(\tilde{K}_i((\mathcal{X}, \mathbf{A}))) \geq \mu_{L_i}(1 + \delta_i) - \epsilon, \qquad \text{for } i \in X, \tag{23}$$

has $|X| \geq N$ (i.e., X contains at least N elements).

Remark 3 *Proposition 4 also holds if the condition of bounded risks is replaced by utility functions having deceasing absolute risk aversion, and the* expectation *of the risks being uniformly bounded ($E[\tilde{L}_i] < C$ for all i) as shown in the proof in the appendix.*

We formally define what an industry structure to be massively multiline as N grows to mean that the average number of lines for insurers grows without bonds, i.e., $\lim_{N\to\infty} N/M(\mathcal{P}) = \infty$. With this definition, we have

Proposition 5 *Under the conditions of propositions 3 and 4, any sequence of Pareto efficient industry structures will be massively multiline as N approaches infinity.*

These asymptotic results suggest that when there is a large number of essentially independent risks that are "small," the multiline insurance structure is optimal. For standard risks — like auto and life insurance — it can be argued that these conditions are reasonable. However, the results also provide an indication of when a multiline structure may not be optimal:

Implication 5 *A multiline structure may be suboptimal*

- *If there is a limited number of risks.*
- *If risks are asymmetric, for example, when some risks are heavy-tailed and others are not.*
- *If risks are dependent.*

One type of risks, that seem to satisfy all these sources of multiline failure is catastrophic risks. Consider, for example, residential insurance against earthquake risk in the bay area in California.[25] The outcome for different households within this area will obviously be heavily dependent, in case of an earthquake, making the pool of risks essentially behave as one large risk, without diversification benefits. Moreover, many catastrophic risks are known to have heavy tails. This further reduces the diversification benefits, even when risks are independent. Thus, even though an earthquake in California and a hurricane in Florida may be considered independent events, the gains from diversification of such risks may be limited due to their heavy-tailedness.

We now show in an example that the previous intuition indeed holds. Specifically, we show that asymmetry between risks and dependence of risks makes the monoline outcome more likely.

[25] See, e.g., Ibragimov, Jaffee, and Walden (2008).

5.3 One versus two lines - An example

The case with general N and risk distributions is complicated. We therefore focus on a special case: We look at a situation with two insurance line and compare the two industry structures $\mathcal{X}^{SL} = \{\{1\},\{2\}\}$ (monoline) with $\mathcal{X}^{ML} = \{\{1,2\}\}$ (multiline). In the first partition, we know how $\mathbf{A}^1 = (A_1, A_2)^T$ should be chosen from our previous analysis, leading to industry structure $\mathcal{S}^{SL} = (\mathcal{X}^{SL}, \mathbf{A}^1)$. In the second partition, there is typically a whole range of capital, $A \in [\underline{A}, \overline{A}]$, leading to competitive outcomes, $\mathcal{S}^{ML} = (\mathcal{X}^{ML}, A)$. The condition for the multiline business to be optimal is now that there is an $A \in [\underline{A}, \overline{A}]$, such that $\mathcal{S}^{ML}$ offers an improvement for both agents, i.e., $Eu[\tilde{K}_i(\mathcal{S}^{SL})] \leq Eu[\tilde{K}_i(\mathcal{S}^{ML})]$, $i = 1, 2$. We study the conditions under which this is satisfied.

For simplicity, we assume that insurees have expected utility functions defined by $u(x) = -(-x+t)^\beta$, $\beta > 1$, $x < 0$, and that $\tilde{L}_1$ and $\tilde{L}_2$ have Bernoulli distributions: $\mathbb{P}(\tilde{L}_1 = 1) = p$, $\mathbb{P}(\tilde{L}_2 = 1) = q$, $\text{corr}(\tilde{L}_1, \tilde{L}_2) = \rho$.[26] Specifically, we study the case $\beta = 1.5$, $t = 1$, $\delta = 0.008$ and $p = 0.1$. We first choose $q = 0.3$ and compare the monoline outcome with the multiline outcome for $\rho \in \{-0.1, 0, 0.1\}$ in Figure 3. The solid lines show optimal expected utility for insuree 1 and 2 respectively in the monoline case (which occurs at capital levels $A_1 = 0.7965$ and $A_2 = 0.8997$). For the case of negative and zero correlation, the situation can be improved for both insurees by moving to a duo-line solution, reaching an outcome somewhere on the efficiency frontier of the duo-line utility possibility curve. For the case of $\rho = 0.1$, insuree 1 will not participate in a duo-line solution, and the monoline outcome will therefore prevail.

In Figure 4 we plot the regions in which monoline duo-line solutions will occur respectively, as a function of q and ρ, given other parameter values given above ($p = 0.1$, $\beta = 1.5$, $\delta = 0.008$, $t = 1$). In line with our previous discussion, summarized in Implication 5, it is clear that, all else equal, increasing correlation decreases the prospects for a duo-line solution. Also, increasing the asymmetry ($|p - q|$) between risks decreases the prospects for a duo-line outcome.

Thus, in line with Implication 5, we find that multiline insurers choose lines in which

- Losses are uncorrelated/have low correlation.
- Loss distributions are similar/not too asymmetric.

6 Concluding remarks

This paper developed a model of the insurance market under the assumptions of costly capital, limited liability, incomplete markets and perfect competition. We focus on the determination of 4 key variables for the general case of a multiline insurer: (1) premiums, (2) capital allocations across the insurance lines, (3) the aggregate amount of capital, and (4) the choice between monoline and multiline firm structures.

[26]Depending on $0 < p < 1$ and $0 < q < 1$, there are restrictions on the correlation, ρ. Only for $p = q$ can ρ take on any value between -1 and 1.

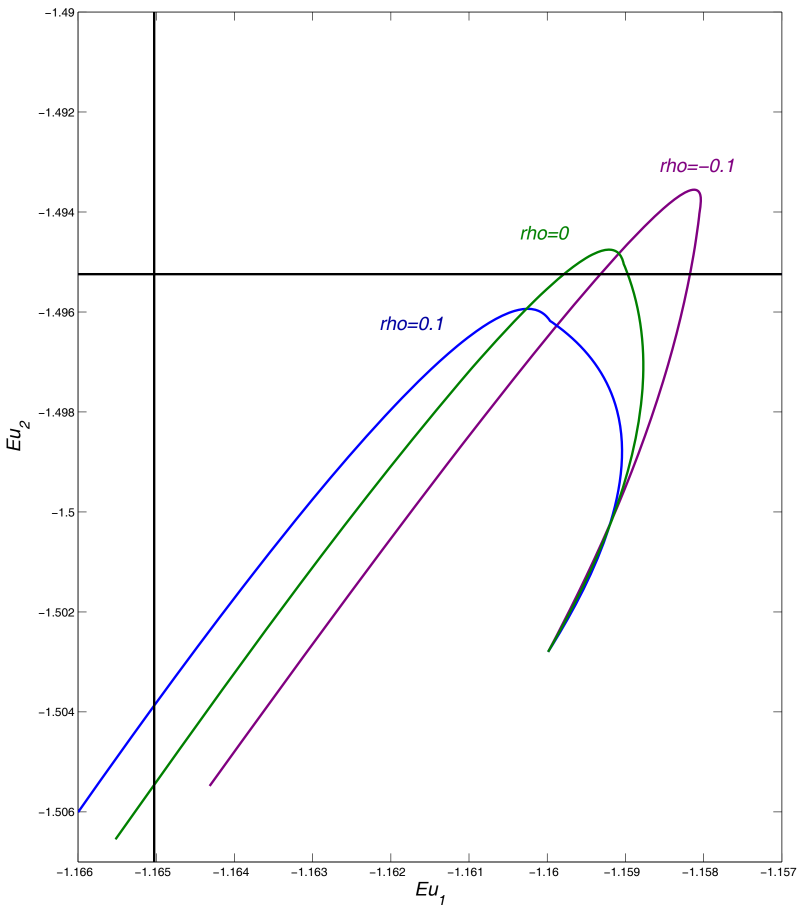

Figure 3: *Monoline versus multiline industry structure. Monoline outcome will occur when* $\rho = 0.1$*, because duo line structure is suboptimal for insuree 2. For* $\rho = 0$ *and* $\rho = -0.1$*, multiline structure occurs since it is possible to improve expected utility for insuree 2, as well as for insuree 1.*

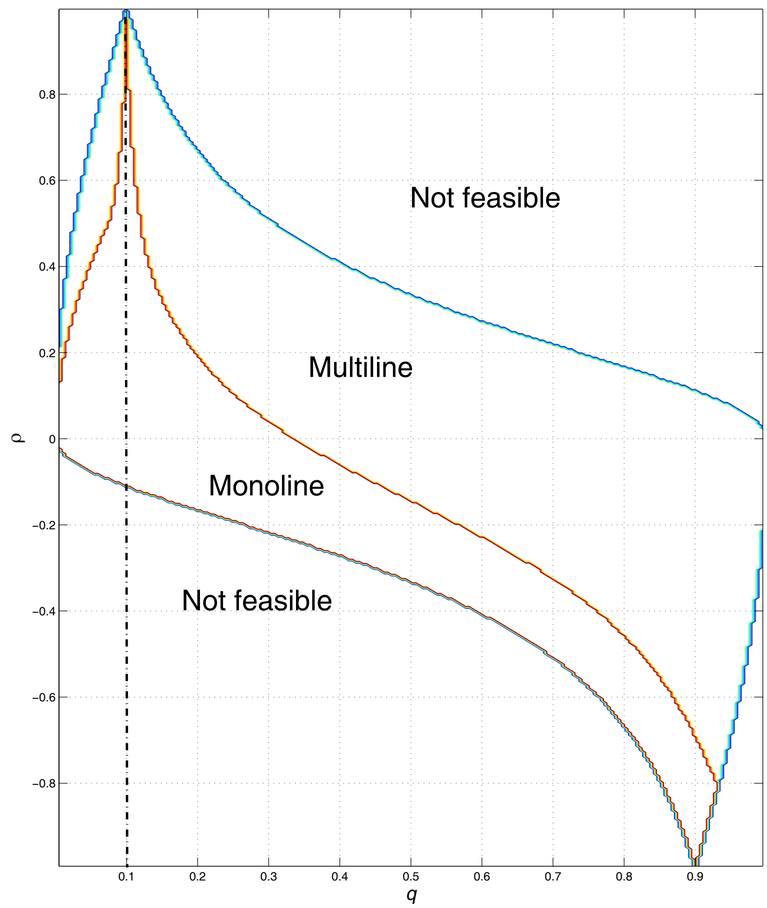

Figure 4: *Regions of q and ρ, in which monoline and multiline structure is optimal. All else equal: Increasing ρ (correlation), or $|p-q|$ (asymmetry of risks) makes monoline structure more likely.*

The premium setting and capital allocations are based on the no-arbitrage, option-based, technique, first developed in the papers by Phillips, Cummins, and Allen (1998) and Myers and Read (2001). These papers, however, apply an ex ante rule for allocating the shortfall in claim paying capacity when the insurer defaults, which has the undesirable features that (i) it is based on the unobservable initially expected loss and that (ii) it will often require that policyholders with small expected claims make payments to the policyholders facing large expected claims. Instead, this paper develops the solution when the shortfall created by an insurer default is shared among the claimants under an ex post, pro rata, rule based on the actual realized claims.

The unique contribution of this paper is that it develops a framework to determine the industry structure in terms of which insurance lines are provided by monoline versus multiline insurers. We employ an equilibrium concept based on a criterion of Pareto efficiency within a competitive industry. Pareto dominated structures are eliminated by new entrants that offer a preferred structure. The resulting equilibrium is robust to the entry of any new monoline provider.

We derive quite strong properties for this equilibrium. First, we find that the multiline structure dominates when the benefits of diversification are achieved because the underlying risks are numerous and relatively uncorrelated. We would expect this condition to hold for consumer lines such as homeowners and auto insurance. On the other hand, when the risks are difficult to diversify because they are limited in number and heavy tailed, the monoline structure may be the efficient form. This condition may hold for the various catastrophe lines, including natural disasters, security insurance, and terrorism.

Appendix

An example

Consider two independent risks, $\tilde{L}_1$ with 50% chance of being -40 and 50% chance of being zero, and $\tilde{L}_2$, with 50% chance of being -10 and 50% chance of being zero. The four states of the world are thus, $\{(0,0), (-40,0), (0,-10), (-40,-10)\}$. Further, assume that the total liability is $A = 20$, that $r = 0$, and risk-neutral propabilities coincide risk-neutral probabilities. Then, the price of $\tilde{L}_1$ is $P_{L_1} = 20$, $P_{L_2} = 5$, and $P_Q = 25\% \times 20 + 25\% \times 30 = 12.5$. Now, using (2), the payout in the four states of the world to the two insurees are $\{(0,0), (20,0), (0,10), (16,4)\}$. The sharing rule for the case when both insurees have claims stems from the fraction of losses being $\tilde{L}_1/(\tilde{L}_1 + \tilde{L}_2) = 40/(40+10) = 80\%)$, so 80% of the total capital of 20, i.e., 16 goes to insuree 1, and the remaining 4 to insuree 2. Thus, the value of the limited liability insurance against $\tilde{L}_1$ is $25\% \times 20 + 25\% \times 16 = 9$, and the value of the limited liability insurance against $\tilde{L}_2$ is $25\% \times 10 + 25\% \times 4 = 3.5$. This in turn implies that the premium-to-liability ratios are

$$\frac{P_1^*}{P_{L_1}} = \frac{9}{20} = 0.45, \qquad \frac{P_2^*}{P_{L_2}} = \frac{3.75}{5} = 0.75,$$

which are not equal. Intuitively it is clear that the premium-to-liability ratio is lower for the first risk, as it is more likely to realize losses in the states of the world when the firm defaults, and therefore does not pay back the full losses.

We now study the same example in a setup similar to PCA. The key behind the differences is that they assume an *ex ante* sharing rule, whereas we assume an *ex post* pro rata sharing rule.

We look at the payouts under (9). Since $P_{L_1}/(P_{L_1} + P_{L_2}) = 80\%$, (9) implies that the payouts in the four states of the word in this case are $\{(0,0), (24,-4), (0,10), (16,4)\}$, i.e., the rule indicates that insuree 2 should pay insuree 1 20% of the option payout of $20 = 4$, in the case where only insuree 1 realizes losses. This obviously violates both assumptions 2 and 3 of the consistency condition. Under the ex ante sharing rule, the premia are $P_1^* = 20 - 0.8 \times 12.5 = 10$ and $P_2^* = 5 - 0.2 \times 12.5 = 2.5$, so the premium-to-liability ratios are equal:

$$\frac{P_1^*}{P_{L_1}} = \frac{10}{20} = 0.5, \qquad \frac{P_2^*}{P_{L_2}} = \frac{2.5}{5} = 0.5,$$

in line with PCA's argument.

Cost allocation

The marginal interpretation is that, with a specific choice of A_i's (and thereby v_i's), the marginal default value of increasing exposure to $\tilde{L}_i$ risk is

$$\left.\frac{\partial P_{Q_q}}{\partial q_i}\right|_{q_1,\ldots,q_N=1} = r_i P_Q.$$

We are interested in the bijection between r_i's and v_i's: $R : \{r_i\} \leftrightarrow \{v_i\}$. We have, for a specific choice of A_i,

$$\tilde{Q}_q = \max\left(\sum_i q_i(\tilde{L}_i - A_i), 0\right),$$

so

$$\left.\frac{\tilde{Q}_q}{\partial q_i}\right|_{q_1,\ldots,q_N=1} = (\tilde{L}_i - v_i A_i) I_{\{-\tilde{L}-A>0\}},$$

where I is the indicator function. This immediately implies that

$$r_i P_Q = \left.\frac{\partial P_{Q_q}}{\partial q_i}\right|_{q_1,\ldots,q_N=1} = Price\left((\tilde{L}_i - A_i) I_{\{\tilde{L}-A>0\}}\right). \tag{24}$$

Given an r_i, and a pricing rule, A_i can now be chosen to satisfy (24), which defines the bijection. This immediately gives us

$$v_i = \frac{A_i}{A},$$

and in our terminology, the s_i's in Myers and Read (2001) satisfy $A_i = (1 + s_i)P_{L_i}$, so

$$s_i = \frac{A_i}{P_{L_i}} - 1.$$

The total surplus, as defined in Myers and Read (2001), is now $S = A - \sum_i P_{L_i}$, so the surplus allocated to line i is $S_i = s_i P_{L_i}$.

Given any choice of r_i's, such that $\sum_i r_i = 1$, and the corresponding A_i's, we have

$$P_Q = \sum_i r_i P_Q = Price\left(\sum_i (\tilde{L}_i - A_i) I_{\{\tilde{L}-A>0\}}\right) = Price\left(\left(\tilde{L} - \sum_i A_i\right) I_{\{\tilde{L}-A>0\}}\right),$$

but since $Price\left(\left(\tilde{L} - A\right) I_{\{\tilde{L}-A>0\}}\right) = Price(max(\tilde{L} - A, 0)) = P_Q$, this equation will be satisfied iff $\sum_i A_i = A$. Thus, $\sum_i A_i = A$, i.e., $\sum_i v_i = 1$ iff $\sum r_i = 1$.

An interpretation of the relationship between the $\{v_i\}$'s and the $\{r_i\}$'s is that the $\{v_i\}$'s tell us how much the insurer must increase assets, if there is a marginal increase of insurance in one line. If the insurer currently sells insurance $\tilde{L}_i$, and then increases risk exposure in line

i to $\tilde{L}_i(1+\Delta q)$, assets will change from A to $A(1+v_i\Delta q)$. With such a change, the value of the option to default increases (to a first order approximation) by $P_Q r_i \times \Delta q$. Any other choice will lead to a redistribution between old and new insurees and will therefore not be consistent with noarbitrage, as shown in the redistribution example below.

The rule proposed in Myers and Read (2001) is to choose

$$d_i \stackrel{\text{def}}{=} \left.\frac{\partial P_{Q_q}}{\partial q_i}\right|_{q_1,\ldots,q_N=1} = P_Q \frac{r_i}{P_{L_i}},$$

such that the d_i's are constant across i's (see page 559). Clearly, since $\sum_i r_i = 1$, the only way of doing this is to choose $r_i = \frac{P_{L_i}}{P_L}$, which is the r_i's suggested by the *ex ante* sharing rule.

Redistribution

We show in a simple example that, under the *ex post* sharing rule, the only asset allocation rule that is consistent is choosing $A_i = r_i A$, i.e., $v_i = r_i$.[27] Assume that there are two risks, $\tilde{L}_1$ and $\tilde{L}_2$, and three states of the world. The state prices are π_j, $j = 1, 2, 3$. The losses in the different states of the world are shown in Table 2 below. The default prices for insurance

State, j	State price, π_j	$\tilde{L}_1$	$\tilde{L}_2$
1	0.5	0	0
2	0.25	10	10
3	0.25	50	30

Table 2: *Example with two risks where ex post sharing rule leads to different results than in.*

are thus, $P_{L_1} = 0.25 \times 10 + 0.25 \times 50 = 15$, and $P_{L_2} = 0.25 \times 10 + 0.25 \times 30 = 10$. We assume that the total reserves (assets) are $A = 40$. In this case, when the ex post sharing rule is used, the payout in different states of the world are summarized in Table 3. Default thus

State, j	$Payment_1$	$Payment_2$
1	0	0
2	10	10
3	25	15

Table 3: *Payments in different states of the world.*

only occurs in state 3, in which only 40 of the total 80 in liability is paid out, which in turn implies that the value of the default option is $P_Q = 0.25 \times 40 = 10$. Since the expected value of losses is $P_L = P_{L_1} + P_{L_2} = 15 + 10 = 25$, the total surplus is $S = A - P_L = 40 - 25 = 15$.

[27] The asset allocation rule is for marginal expansions in individual insurance lines. Consistent means that it does not lead to value redistribution between new and old insurees (any redistribution would be inconsistent with the method used to determine the original price).

How should the total assets, $A = 40$, be allocated between the two lines? We wish to find v_1 and v_2, such that $A = A_1 + A_2$, and $A_1 = v_1 A$, $A_2 = v_2 A$. In the case of only one state of the world in which default occurs, and the ex post payment rule, it is easy to check that our previous arguments lead to $v_1 = r_1$ and $v_2 = r_2$, which through our derived rule

$$r_i = Price\left(\frac{\tilde{L}_i}{\tilde{L}} \times \frac{\tilde{Q}(A)}{P_Q}\right), \tag{25}$$

implies that $v_1 = 0.625$, $v_2 = 0.375$, $A_1 = 25$ and $A_2 = 15$. The surplus allocation between the two lines is then, $S_1 = A_1 - P_{L_1} = 25 - 15 = 10$, $S_2 = A_2 - P_{L_2} = 15 - 10 = 5$. Finally, it is easy to check that the price of the cash-flows in Table 3 is $P_1 = 8.75$ and $P_2 = 6.25$.

The allocation implied by v_1 and v_2 is important in that it provides a rule for how assets need to change if more insurance in one line is sold. For example, assume that the insurer increases its exposure to risk 2 by $\Delta q = 10\%$. In this case, the allocation rule implies that the total assets need to increase by

$$v_2 \times \Delta q \times A = 0.375 \times 10\% \times 40 = 1.5.$$

Thus, the total assets are now $A = 41.5$, and $A_1 = 25$, $A_2 = 15 + 1.5 = 16.5$ and the expected loss of risk 2 is $P_{L_2} = 1.1 \times 10 = 11$. The new surplus allocation is therefore $S_1 = 25 - 15 = 10$, $S_2 = 16.5 - 11 = 5.5$.

Why is this the right allocation rule to use when the ex post sharing rule is used? Because, it does not change the price or risk-structure of the payout to the already insured risks, and it prices the new risk correctly. In short, there are no value transfers between insurers when the new insurance is sold. With the new risk insured, the total losses are as shown in Table 4 and the payments are as shown in Table 5. The calculations of the payments are identical to

State, j	State price, π_j	$\tilde{L}_1$	$\tilde{L}_2$
1	0.5	0	0
2	0.25	10	11
3	0.25	50	33

Table 4: *Losses when risk 2's exposure has increased by 10%.*

State, j	$Payment_1$	$Payment_2$
1	0	0
2	10	11
3	25	16.5

Table 5: *Payments when risk 2's exposure has increased by 10%, and* $A = 41.5$.

the previous ones, using the ex post sharing rule. For example, for $Payment_2$ when $j = 3$, we have $16.5 = 33/(50+33) \times 41.5$. We see in Table 5 that, in all states of the world, the payments to risk 1 are identical, and that the payments to risk 2 have increased by exactly 10%. Thus, the price per unit risk is the same as before, and 1.5 is the "correct" amount to increase assets with under the ex post sharing rule. It is easy to check that this is the only choice of v_1 (and thereby of $v_2 = 1 - v_1$) that has this property (under the ex post sharing rule). Specifically, any other choice will lead to a change the prices of the already insured risks, so that there will be "losers" and "winners" when the company scales up one insurance line.

The rule proposed in Myers and Read (2001), on the other hand, leads to redistribution under the *ex post* sharing rule. Their proposed allocation is such that the option value of default ($P_Q = 10$) is shared such that $r_1 = P_{L_1}/P_L = 15/25 = 0.6$, $r_2 = P_{L_2}/P_L = 10/25 = 0.4$, which by the relationship

$$r_i P_Q = Price\left((\tilde{L}_i - A_i) I_{\{\tilde{L}-A>0\}}\right),$$

leads to $A_1 = 26$ and $A_2 = 14$, and thereby to $v_1 = 0.65$, $v_2 = 0.35$. Therefore, if insurance 2 is scaled up by 10%, total assets should increase by

$$v_2 \times \Delta q \times A = 0.35 \times 10\% \times 40 = 1.4,$$

so assets are now $A = 41.4$. The option value of default increases with 0.4 ($r_2 \times 10\% \times P_Q$) to $P_Q = 10.4$.

The payments will be as in Table 6, using identical calculations as in Table 5, but with $A = 41.4$. However, this implies that the insurance against $\tilde{L}_1$ has become less worth by

State, j	$Payment_1$	$Payment_2$
1	0	0
2	10	11
3	24.94	16.46

Table 6: *Payments when risk 2's exposure has increased by 10%, and* $A = 41.4$.

the new investments in line 2, as has the value of insurance against $\tilde{L}_2$ risk for the original insurees. There has thus been a value transfer from the old insurees to the new ones (assuming that the price is the same), which is inconsistent.[28] Obviously, examples where the effect is larger can be constructed.

[28] If the original insurees knew that such a transfer might take place, they would not pay the premium in the first place, since the correct price for the insurance that included transfer risk would be different.

Proofs

Proof of Lemma 1: i) and ii) follow immediately from the definition of P (1). iii) is an immediate consequence of (1), and $E^*[\tilde{Q}(A)] = o(1)$ for large A, follows from $E[\tilde{L}]$ being finite together with the equivalence of the risk neutral and the objective measure. iv) follows from ii) and that $\eta' > 0$ for general distributions (see Ingersoll (1987)). ∎

Proof of Proposition 1

We first prove the "only if"-part. Assume that for all $A > 0$, $P(A) \geq A$. Let x_- denote $\min(x, 0)$. For a given A, expected utility is $Eu(-P(A) + (A - \tilde{L})_-) \leq Eu(-P(A) + A - \tilde{L}) \leq Eu(-\tilde{L}) = Eu(-P(0) + (0 - \tilde{L})_-)$, so $0 \in A^*$.

For the "if"-part: Assume that there is an A such that $P(A) < A$. Obviously, $A > 0$, since $P(0) = 0 = A$. Now, define the "utility function" $u_q(x) = (x + q)_-$. This function is concave, but only weakly so, and not twice continuously differentiable, so it is outside the class of utility functions we are studying. However, it is easy to "regularize" u_q and get an infinitely differentiable strictly increasing and concave function that are arbitrarily close u_q in any reasonable topology. We can do this by using the Gaussian test function, $\phi(x) = \frac{1}{2\sqrt{2\pi}} e^{-x^2/2}$ and define $\phi_\epsilon(x) = \phi(x/\epsilon)/\epsilon$. Finally, we define $u_{q,\epsilon}(x) = u_q * \phi_\epsilon = \int_{-\infty}^{\infty} u_q(y)\phi_\epsilon(x - y)\, dy$. Clearly, as $\epsilon \searrow 0$, $u_{q,\epsilon}$ converges to u_q. Moreover, $u_{q,\epsilon}$ is infinitely differentiable and since $u_{q,\epsilon}^{(n)} = (u_q * \phi_\epsilon)^{(n)} = u_q^{(n)} * \phi_\epsilon$, where $u_{q,\epsilon}^{(n)}$ denotes the nth derivative of $u_{q,\epsilon}$, it is easy to check that $u'_{q,\epsilon} > 0$ and $u''_{q,\epsilon} < 0$ for all q and ϵ, so $u_{q,\epsilon}$ belongs to our class of utility functions.

Now, if $A > P$, then $Eu_P(-P + (A - \tilde{L})_-) = E[(A - \tilde{L})_-] > E[(P - \tilde{L})_-] = Eu_P(-\tilde{L})$, so an insuree with "utility function" u_P is strictly better off by choosing insurance. However, since $\lim_{\epsilon \searrow 0} Eu_{P,\epsilon}(-P + (A - \tilde{L})_-) = Eu_P(-P + (A - \tilde{L})_-)$ and $\lim_{\epsilon \searrow 0} Eu_{P,\epsilon}(-\tilde{L}) = Eu_P(-\tilde{L})$, for ϵ small enough, the strict inequality also holds for a $u_{P,\epsilon}$, which belongs to our class of utility functions. Thus, insurance is optimal for an insuree with such a utility function.

We also make some straightforward observations: First, if $\tilde{L}$ has an absolutely continuous distribution in a neighborhood of 0, then $\partial Eu(A)/\partial A < 0$ at $A = 0$, i.e., insurees are always strictly worse off buying a small amount of insurance than buying no insurance at all. Second, if $\tilde{L}$ has a bounded range, with upper bound $\overline{L}$, and $\tilde{L}$ has an absolutely continuous distribution function in a neighborhood of $\overline{L}$, then $\partial Eu(A)/\partial A < 0$ at $A = \overline{L}$, i.e., insurees are always strictly worse off buying full insurance compared with buying slightly less than full insurance. These results similar to the classical results on optimal contracts having deductibles in the insurance literature. Third, if the p.d.f of $\tilde{L}$ vanishes on an interval $[a, b]$, then $Eu(A)$ is concave for $A \in [a, b]$.

∎

Proof of Proposition 2: Given A, the utility of insuring a risk is $Eu\left[-\mu_L - \delta A + (\mu_Q - \tilde{Q})\right] =$

$Eu[-\mu_L - \delta A + ((-\tilde{L} - A)_- - E[(-\tilde{L} - A)_-])]$. Here, we use the notation $x_- = \min(x, 0)$. Since $E[\tilde{L}_1] = E[\tilde{L}_2] = \mu_L$, second order stochastic dominance is therefore equivalent to $Eu((-\tilde{L}_1 - A)_- - E[(-\tilde{L}_1 - A)_-]) \geq Eu((-\tilde{L}_2 - A)_- - E[(-\tilde{L}_2 - A)_-])$, which in turn is equivalent to

$$(-\tilde{L}_1 - A)_- \succeq (-\tilde{L}_1 - A)_- + z, \tag{26}$$

where $z = E[(-\tilde{L}_1 - A)_-] - E[(-\tilde{L}_2 - A)_-]$. Now, if $z \leq 0$, (26), is implied by

$$\int_{-\infty}^{t} F_1(x)dx \leq \int_{-\infty}^{t} F_2(x)dx, \tag{27}$$

for all $t < -A$, and since $A > \mu_L$, (27) is obviously implied by (22).

For $z > 0$, we note that $z \leq \mu_L$, so a similar argument implies that

$$\int_{-\infty}^{t+z} F_1(x)dx \leq \int_{-\infty}^{t} F_2(x)dx, \tag{28}$$

implies the domination, which, once again is implied by (22), and we are done. ∎

Proof of Lemma 2: Second part is trivial, since for a given industry structure, we now that the set of Pareto efficient outcomes is non-empty, and that the A^*'s and thereby A's are compact. Since there are a finite number of possible partitions of N risks, the set remains compact (being a finite union of compact sets) when Pareto efficiency is taken over all Partitions. For the second part, compactness is preserved by finite intersections. Uniqueness is also straightforward: If the partition into singletons ($\mathcal{P} = \{\{1\}, \ldots, \{N\}\}$) is Pareto efficient, it is clearly robust to monoline blocking. If it is not Pareto efficient, then there is another partition that is Pareto efficient, which thereby dominates the partition into singletons. This partition is then robust to monoline blocking, since there is no way to make an agent better off by offering insurance in a monoline business. ∎

Proof of Proposition 3: The condition that u is three times continuously differentiable with u''' bounded by a polynomial of degree q implies the following uniform Lipschitz condition, with $\alpha = q$:

$$|u''(x) - u''(y)| \leq C|x - y|^{\alpha}, \qquad \text{for all } x, y.$$

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 $C > 0$,

$$E|\sum_{i=1}^{N}(L_i - \mu_i)|^p \leq C \max\Big(\sum_{i=1}^{N} E|L_i - \mu_i|^p, \Big(\sum_{i=1}^{N} \sigma_i^2\Big)^{p/2}\Big), \tag{29}$$

and, thus,

$$N^{-p}E|\sum_{i=1}^{N}(L_i-\mu_i)|^p \le CN^{-p}\max\Big(\sum_{i=1}^{N}E|L_i-\mu_i|^p, \Big(\sum_{i=1}^{N}\sigma_i^2\Big)^{p/2}\Big) \le CN^{-p/2} \to 0 \quad (30)$$

as $N \to \infty$. Take $A = \sum_{i=1}^N \mu_i$. Denote $x_i = -L_i\Big(1 - \max\Big(1 - \frac{A}{\sum_{i=1}^N L_i}, 0\Big)\Big) = -L_i\Big(1 - \max\Big(1 - \frac{\sum_{i=1}^N \mu_i}{\sum_{i=1}^N L_i}, 0\Big)\Big)$, $y_i = L_i \max\Big(1 - \frac{A}{\sum_{i=1}^N L_i}, 0\Big) = L_i \max\Big(1 - \frac{\sum_{i=1}^N \mu_i}{\sum_{i=1}^N L_i}, 0\Big)$. We have $Eu(-Ex_i(1+\delta) + L_i - x_i) = Eu(\mu_i(1+\delta) - Ey_i(1+\delta) + y_i)$. Using Taylor expansions and Lipschitz continuity of order α for u'', we get

$$|Eu(\mu_i(1+\delta) - Ey_i(1+\delta) + y_i) - u(\mu_i(1+\delta)) - \delta u'(\mu(1+\delta))Ey_i| \le C|y_i - Ey_i(1+\delta)|^{2+\alpha},$$

and, consequently,

$$|Eu(\mu_i(1+\delta) - Ey_i(1+\delta) + y_i) - u(\mu_i(1+\delta))| \le C|Ey_i| + CE|y_i|^{2+\alpha} + C|Ey_i|^{2+\alpha}. \quad (31)$$

Since, by Jensen's inequality, $|Ey_i|^{2+\alpha} \le (E|y_i|^{2+\alpha})^{1/(2+\alpha)}$, (31) implies, that, to complete the proof, it suffices to show that

$$E|y_i|^{2+\alpha} \to 0 \quad (32)$$

as $N \to \infty$.

By Jensen's inequality, we have, under the conditions of the proposition,

$$E|y_i|^{2+\alpha} = E\Big|L_i \max\Big(1 - \frac{\sum_{i=1}^N \mu_i}{\sum_{i=1}^N L_i}, 0\Big)\Big|^{2+\alpha} = E|L_i|^{2+\alpha}\Big|\max\Big(\frac{\sum_{i=1}^N (L_i - \mu_i}{\sum_{i=1}^N L_i}, 0\Big)\Big|^{2+\alpha} \le$$
$$E|L_i|^{2+\alpha}\frac{|\sum_{i=1}^N (L_i-\mu_i)|^{2+\alpha}}{|\sum_{i=1}^N \mu_i|^{2+\alpha}} \le \Big(E|L_i|^p\Big)^{1/2}\Big(E\Big|\frac{\sum_{i=1}^N (L_i-\mu_i)}{\sum_{i=1}^N \mu_i}\Big|^p\Big)^{1/2} \le$$
$$C\Big(E\Big|\frac{\sum_{i=1}^N (L_i-\mu_i)}{\sum_{i=1}^N \mu_i}\Big|^p\Big)^{1/2} \le C\Big(N^{-p}E\Big|\sum_{i=1}^N (L_i-\mu_i)\Big|^p\Big)^{1/2}. \quad (33)$$

From (33) and (30) it follows that (32) indeed holds. The proof is complete.

∎

Proof of Proposition 4:

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

$$u(x+y) \ge u(x) + u'(x)y + C\frac{y^2}{2} \quad (34)$$

for all x, y. Using inequality (34), in the notations of the proof of Proposition 3, we obtain

$$Eu(-Ex_i(1+\delta)+L_i-x_i) = Eu(\mu_i(1+\delta)-Ey_i(1+\delta)+y_i) \geq$$
$$E\Big[u(\mu_i(1+\delta))+u'(\mu_i(1+\delta))(y_i-Ey_i(1+\delta))+u''(\mu_i(1+\delta))\frac{(y_i-Ey_i(1+\delta))^2}{2}\Big] =$$
$$u(\mu_i(1+\delta))+u'(\mu_i(1+\delta))\delta Ey_i+u''(\mu_i(1+\delta))\frac{Var(y_i)+\delta^2(Ey_i)^2}{2} =$$
$$u(\mu_i(1+\delta))-u'(\mu_i(1+\delta))\delta|Ey_i|-|u''(\mu_i(1+\delta))|\frac{Var(y_i)+\delta^2(Ey_i)^2}{2}. \quad (35)$$

Consequently, if (23) is satisfied, then

$$\epsilon > u'(\mu_i(1+\delta))\delta|Ey_i|+|u''(\mu_i(1+\delta))|\frac{Var(y_i)+\delta^2(Ey_i)^2}{2} \geq$$
$$|u''(\mu_i(1+\delta))|\frac{Ey_i^2+(1-\delta^2)(Ey_i)^2}{2},$$

and, since $\delta < 1$, $Ey_i^2 < \epsilon' = 2\epsilon/|u''(\mu_i(1+\delta))|$. Take $\Delta > 0$. We have, using Jensen's inequality and the assumptions of the proposition,

$$Ey_i^2 = EL_i^2\Big[\max\Big(1-\frac{\sum_{i=1}^N \mu_i}{\sum_{i=1}^N L_i},0\Big)\Big]^2 = EL_i^2\Big[\max\Big(\frac{\sum_{i=1}^N (L_i-\mu_i)}{\sum_{i=1}^N L_i}\Big)\Big]^2 I\Big(\sum_{i=1}^N L_i > \sum_{i=1}^N \mu_i\Big) \geq$$

$$EL_i^2\Big[\max\Big(\frac{\sum_{i=1}^N (L_i-\mu_i)}{\sum_{i=1}^N L_i}\Big)\Big]^2 I\Big(\sum_{i=1}^N (L_i-\mu_i) > N\Delta\Big) \geq \mu_i^2\frac{\Delta^2}{C^2}\Big[P\Big(\sum_{i=1}^N (L_i-\mu_i) > N\Delta\Big)\Big]^2.$$

The above inequalities imply that if condition (23) is satisfied, then N must be sufficiently large so that $P\Big(\sum_{i=1}^N (L_i-\mu_i) > N\Delta\Big) < \frac{\sqrt{\epsilon'}C}{\Delta\mu} = \frac{\sqrt{2\epsilon}C}{\Delta\mu\sqrt{|u''(\mu_i(1+\delta))|}}$. ∎

Proof of Proposition 5: Proof by contradiction: Suppose that the proposition is not true. Then there is a constant integer, $C > 0$, and an increasing sequence $N_1 < N_2 < \ldots$, such that $N_i/M_i < C$ for a Pareto efficient partition containing M_i elements in the economy with N_i risks. We have $\sum_{j=1}^{M_i}|P_j| = N_i$, where we w.l.o.g. assume that the P_j's are ordered in increasing order of size. Let r denote the number of P_j's containing at most $2C$ risks. Clearly, $r \geq M_i/2$ since otherwise $\sum_{j=r}^{M_i}|P_j| > \sum_{j=M_i/2+1}^{M_i} 2C \geq 2C \times M_i/2 = N_i$, which obviously can not hold. However, from proposition 4, it is clear that for all elements in P_i, $i \leq r$, $CE_u(\tilde{K}_i)+\mu_{L_i} \geq \epsilon$, for some fixed $\epsilon > 0$, regardless of N. Now, since M_i approaches infinity as N grows, we can choose N large enough to make r arbitrary large, and the number of elements in $\cup_{i=1}^r P_i$ thereby also becomes arbitrary large. We can then use the result in proposition 3 to choose r so large, so that replacing the industry with $N_{j'}$ risks with a multiline version, $\mathcal{P} = \{\cup_{i=1}^r P_i, P_{r+1}, \ldots, P_{M_{j'}}\}$, gives for all $j \in \cup_{i=1}^r P_i$, $CE_u(\tilde{K}_j)+\mu_{L_j} < \epsilon$.

Clearly, this is a Pareto improvement, since all agents in $\{\cup_{i=1}^{r} P_i\}$ are better off, whereas all other agents are identically well off. Thus, the proposed sequence can not be Pareto optimal and we are done.

∎

References

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. (1988): "Risk-Based Premiums for Insurance Guaranty Funds," *Journal of Finance*, 43, 823–839.

——— (1993): "Allocation of Capital in the Insurance Industry," *Risk Management and Insurance Review*, 3, 7–28.

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.

DOHERTY, N., AND H. SCHLESINGER (1990): "Rational insurance purchasing: consideration of contract nonperformance," *Quarterly Journal of Economics*, 105, 243–253.

DOHERTY, N. A., AND J. R. GARVEN (1986): "Price Regulation in Property-Liability Insurance: A Contingent-Claims Approach," *Journal of Finance*, 41, 1031–1050.

DUFFIE, D. (2001): *Dynamic Asset Pricing Theory*. Princeton University Press.

FAIRLEY, W. (1979): "Investment Income and Profit Margins In Property-Liability Insurance: A Contingent Claims Approach," *Bell Journal*, 10, 192–210.

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.

HILL, R. (1979): "Profit Regulation in Property-Liability Insurance," *Bell Journal*, 10, 172–191.

IBRAGIMOV, R., D. JAFFEE, AND J. WALDEN (2008): "Nondiversification traps in markets for catastophe insurance," *The Review of Financial Studies*, Forthcoming.

IBRAGIMOV, R., AND P. C. B. PHILLIPS (2004): "Regression asymptotics using martingale convergence methods," Cowles Foundation Discussion Paper 1473. To appear in *Econometric Theory*. http://www.economics.harvard.edu/faculty/ibragimov/files/ibragimovphillipsfinal.pdf.

INGERSOLL, J. (1987): *Theory of Financial Decision Making*. Rowman & Littefield.

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.

LAKDAWALLA, D., AND G. H. ZANJANI (2006): "Catastrophe Bonds, Reinsurance, and the Optimal Collateralization of Risk-Transfer," NBER working paper No. 12742.

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.

MAHUL, O., AND B. WRIGHT (2004): "Efficient risk sharing within a catastrophe insurance pool," paper presented at the 2003 NBER Insurance Project Workshop, available at http://www.arec.umd.edu/Department/Seminars/mahulwright0204.pdf.

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 R. Cohn (1987): "Insurance Rate of Return Regulation and the Capital Asset Pricing Model," in *Fair Rate of Return in Property Liability Insurance*, ed. by J. D. Cummins, and S. Harrington.

Myers, S. C., and J. A. Read (2001): "Capital Allocation fo Insurance Companies," *The 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," *The 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.

Schlesinger, H. (2000): "The theory of insurance demand," in *Numerical Treatment of Multiscale-Problems*, ed. by G. Dionne, Handbook of Insurance. Kluwer Academic Publishers.

Serfling, R. J. (1970): "Convergence properties of S_n under moment restrictions," .

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