

Insuring Non-Verifiable Losses¹

Neil A. Doherty², Christian Laux³, and Alexander Muermann³

January 2012

¹We thank Brett Green, John Major, Kjell Nyborg, Sandrine Spaeter, and seminar participants at Georgia State University, the European Group of Risk and Insurance Economists Meeting 2011, the European Winter Finance Summit 2011, the Financial Intermediation Research Society Meeting 2011, and the NBER Insurance Workshop 2011 for their helpful comments. We thank Robert Kremslehner for excellent research assistance.

²Department of Business and Public Policy, The Wharton School, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104-6302, USA, email: doherty@wharton.upenn.edu

³Department of Finance, Accounting, and Statistics, Vienna University of Economics and Business, Heiligenstaedter Str. 46-48, 1190 Vienna, Austria, email: christian.laux@wu.ac.at; alexander.muermann@wu.ac.at

Abstract

Insurance contracts are often complex and difficult to verify outside the insurance relation. We show that standard one-period insurance policies with an upper limit and a deductible are the optimal incentive-compatible contracts in a competitive market with repeated interaction. Optimal group insurance policies involve a joint upper limit but individual deductibles and insurance brokers can play a role implementing such contracts for the group of clients. Our model provides new insights and predictions about the determinants of insurance.

Key Words deductible insurance, upper limit, implicit insurance contracts, insurance brokers

JEL Classification D86, G22, L14

1 Introduction

Insurance contracts are often annual contracts with a deductible and an upper limit. We show that these contracts can be the optimal incentive-compatible contracts when insured losses are not verifiable. By providing a novel explanation for the types of contracts we observe in practice, we show that these contracts are robust to problems of contract enforcement. Problems of enforceability are important in insurance relations since claims often depend on some attributes that are hard to specify or identify in advance and hard to measure. Thus, insurance contracts can be complex and policyholders are concerned that insurers may defect on their promise to pay reliably and quickly. Our model provides new insights and predictions about the demand for insurance.

We consider a model with infinite periods where, in each period, risk-averse policyholders incur a random loss. Insurers are risk neutral and policyholders can switch insurers at no cost when an insurance contract expires. The insurance market is competitive. The incurred loss covered by the contract is not verifiable but can be observed by the policyholder and the insurer. Thus, claims contingent on the realized loss have to be self enforcing, and incentive compatibility requires that the insurer has an incentive to make payments as promised. The only leverage that the policyholder has over the insurer is to switch the insurer: the policyholder continues to do business with the insurer only if the insurer honors the contract. The threat to switch the insurer is credible as the policyholder can always find a new insurer in a competitive market.

However, for the same reason, the contractual promises that the policyholder can make to the insurer are limited. In a one-period insurance contract, the policyholder can only commit to a fixed payment, i.e., the insurance premium. The threat to change the insurer matters only if insurers can earn a rent from retaining policyholders. Thus, incentive compatibility requires that the policyholder pays an insurance premium in excess of the expected claims payment so that the insurer earns a rent when signing the contract. The expected future rent from continued business induces the insurer to make a payment despite the non-verifiability of losses. The maximum payment from the insurer determines the level of the required rent. Losses below the maximum payment can be insured at a fair premium, which equals the

expected payment. An upper limit reduces the required rent. At the same time, reducing coverage in high-loss states increases the marginal utility in these states and a deductible becomes optimal.

A natural question is whether long-term insurance contracts can be used to relax the incentive-compatibility constraint of the insurer and thereby improve risk sharing. Indeed, the insurer can be induced to make a higher claims payment in exchange for higher future premiums in a multi-period contract. However, in this case, the policyholder is locked into the insurance relation over multiple periods which undermines the policyholder's ability to threaten the insurer to switch to another insurer. This problem can be solved since promising higher future premiums is akin to borrowing. It is not necessary to tie borrowing with insurance, which requires multi-period insurance contracts. The policyholder is better off with single-period insurance contracts and a separate possibility to borrow against future income in the case of high losses.

More generally, borrowing and saving can be used to smooth income risk over time and can be a substitute for insurance. But insurance will continue to be used if the borrowing rate exceeds the savings rate, and the access to borrowing and saving does not change the form of the optimal one-period insurance contract.

Policyholders can benefit from pooling their threat to switch the insurer. Each individual pays a rent to the insurer, which results in a maximum willingness to pay. If the upper limit is binding for one policyholder but not binding for another policyholder, risk sharing can be improved through joint contracting. We show that the optimal contract for a group of policyholders includes an individual deductible for each policyholder's loss and a joint upper limit on total claims payments to the group of policyholders.

Implementing a contract with a joint upper limit requires that the policyholders observe each others' losses and that they collectively switch to another insurer if the incumbent insurer shirks on any single one of them. Thus, the coordination cost of directly writing an explicit joint contract with the insurer would be high and such contracts are generally not observed in practice. An insurance broker may step in as an intermediary to implicitly implement the joint contract. By bundling the policies of multiple policyholders, the broker has a higher bargaining power than individual policyholders: the broker can threaten to

leave with all clients if the insurer shirks on one. Thereby, the broker increases the expected payments to policyholders for a given level of rent paid to the insurer. Moreover, by coordinating claims settlement, the broker oversees the claims and insurance payments of its clients.

This role for the insurance broker is reflected in the contractual arrangement with insurers. It is normal for brokers to “own the renewal rights” on the book of business they place with the insurer. That is, the broker is free to recommend to its clients that they renew with the current insurer or switch to a rival. Accordingly, the insurer revokes any right to directly solicit business placed through the broker. This provision vests the broker with considerable bargaining power and enables the broker to elicit higher transfers for non-verifiable losses than would be possible (at the same rent) with individual contracting.

Our model provides new insights and predictions about determinants of insurance. First, the cost of insurance is (positively) related to the policyholder’s probability of financial distress. Second, the cost of insuring a high-severity-low-frequency event is higher than for a low-severity-high-frequency event (with the same expected loss) even absent counterparty risk of the insurer or cost of holding capital. Third, capital is important for insurers to not only assure solvency but also to hold excess capacity to underwrite business.

Without market frictions, it is optimal for a risk-averse policyholder to purchase full insurance coverage from risk-neutral insurers (Mossin, 1968). Given the prevalence of deductibles and upper limits in insurance contracts, the insurance literature has devoted considerable attention to understanding these features in insurance contracts. An important result is that, with proportional loading, the insurer optimally covers claims in excess of some threshold (deductible) (see, e.g., Arrow, 1963; Raviv, 1979; Gollier and Schlesinger, 1995). The optimality of deductible insurance is also derived by Cummins and Mahul (2004) assuming an exogenous upper policy limit. Huberman, Mayers, and Smith (1983) show that an upper limit can be optimal when policyholders are protected by limited liability (e.g., third party insurance) and the coverage schedule must be non-decreasing. In Garratt and Marshall (1996) an upper limit in property insurance is an optimal response to moral hazard when the option to rebuild a different home has a positive value for the policyholder. Al-

ternative explanations for a deductible involve moral-hazard or adverse-selection problems (see, e.g., Holmström, 1979). The only friction that we introduce is the non-verifiability of losses. Thereby we capture the idea that individuals and firms face many risks for which it is difficult to write enforceable insurance contracts. A standard insurance contract with a deductible and upper limit is the optimal self-enforcing contract where the insurer makes transfers to retain future business.

There is a large literature that analyzes situations where contracts cannot be enforced (in a court) so that parties can strategically renege on contractual promises. Applications include financial contracting when cash-flow claims cannot be enforced (Bulow and Rogoff, 1989; Bolton and Scharfstein, 1990; Hart and Moore 1998; Kehoe and Levine, 1993; Bond and Krishnamurthy, 2004; Fluck, 1998; Myers, 2000), subjective performance evaluation and incentive contracts (Lazear, 1979; Shapiro and Stiglitz, 1984; Bull, 1987; McLeod and Malcomson, 1989; Levin, 2003), and unobservable product quality (Klein and Leffler, 1981). However, when parties repeatedly interact, contractual non-compliance can be punished through termination of the business relation. In this case, contracts have to be self enforcing such that it is in the parties' best interest to comply. The self-enforcing constraint limits the types of contracts that can be implemented: in particular, the utility derived from the best alternative available outside option limits the transfers that can be self enforcing.

The setting has also been applied to risk sharing (e.g., Kimball, 1988; Thomas and Worrall, 1988; Coate and Ravallion, 1993; Allen and Gale, 1999). Kimball (1988) and Coate and Ravallion (1993) analyze self-enforcing informal risk sharing between risk-averse agents in an infinitely repeated setting. The threat of exclusion from future risk sharing provides an agent with incentives to make payments to another agent. In this setting, first best risk sharing is feasible up to the maximum transfer that can be implemented under the threat of exclusion. These authors analyze a setting of rural, less developed countries where no formal risk sharing arrangements are available, and risk-averse agents mutually insure themselves but cannot rely on enforceable contracts. Both agents are risk averse and can be excluded from future risk sharing after shirking. As we discuss in greater detail below, problems of enforceability are also important in developed countries, even with formal insurance contracts. One difference to our paper is that we assume that insurance is offered

by risk-neutral insurers. Thomas and Worrall (1988) consider a related setting where a risk-neutral firm provides a risk-averse employee with insurance against volatile spot market wages. Either agent can move to the spot market immediately after having observed the spot market wage. Wages have to be self enforcing. Again, both agents (the employee and the firm) are excluded from any form of future risk sharing arrangement after shirking. Our paper differs from this strand of literature by assuming that policyholders cannot be excluded from access to other insurers. Excluding an agent from purchasing insurance is difficult in an anonymous insurance market where insurers compete for business. Thus, we assume that a policyholder can always switch to another insurer. As a consequence, self-enforcing transfers from the policyholder to the insurer are not possible. Instead, contracts have to be enforceable. However, enforceable contractual promise to make future payments are akin to borrowing and should not be linked to insurance contracts since it reduces the insurer's willingness to pay.

We also analyze how risk sharing can be improved through joint contracts and discuss the role of brokers. Brokers act as agents who match trading partners and generate information (e.g., Rubinstein and Wolinsky, 1987; Biglaiser, 1993; Cummins and Doherty, 2006). In our model information is also important, but in addition, the broker improves efficiency through the implementation of joint contracts.

The paper proceeds as follows. In the next section we discuss non-verifiability in insurance relations. We introduce the model in Section 3 and derive the optimal incentive-compatible contract for a single policyholder in Section 4. In Section 5, we examine the optimal joint contract for a group of policyholders and discuss the role of brokers. We conclude in Section 6 by discussing the implications of our results. All proofs are in the appendix.

2 Non-verifiability in insurance relations

As discussed above, there is a large literature analyzing situations where contracts are not available or cannot be enforced. In much of this literature, the contracting parties have the relevant information (about cash flows, performance, loss, etc.) but this information is not verifiable by third parties. Thus, it is difficult to enforce contracts in courts. Enforceability

is also an important issue in insurance relations; policyholders are concerned about the possibility that the insurer reneges on promised claims payments. The policyholder's possibility to switch to another insurer, and the insurer's concern for its reputation, provide incentives to honor the claims payments as intended in the original contract.

Insurance contracts cover losses, costs, and required investments triggered by specific events and circumstances. Households and firms have to decide which parts of these losses, costs or investments to insure and which to retain (i.e., not insure). For example, corporations continuously invest in repair and maintenance of buildings, have to bear losses from failed product placements, and finance marketing campaigns to counter competitive threats. It would not be efficient to insure the major part of these losses, costs, and investments. Thus, an insurance contract has to distinguish between those losses, costs, and investments that should be insured and those that are associated with normal business operations and should not be covered. However, it is often difficult to disentangle the two. The specific losses and costs to be covered differ for different corporations and circumstances and depend on the comparative advantage of bearing losses, information and incentive problems, taxes, etc. Nevertheless, given their expertise, the contracting parties may have a mutual conceptual understanding *ex ante* and/or *ex post* of what it is they would like to insure. But, when it comes to putting this conceptual understanding *ex ante* into contracts, there are so many situation-specific variables, that it would be a formidable task to write an appropriate contract. For third parties it is then often impossible to verify *ex post* which part of the losses the insurance parties wished to insure and which part they did not want to insure. Unless the insurance contract is incentive compatible, the insurance parties have no incentives *ex post* to truthfully reveal the original intentions of the contract or any specific circumstances that affect the level of claims payments. In the following we provide examples for problems of determining the level of claims payments that are underlying the non-verifiability of insured losses.

The insurance of an office building or factory is generally limited to the costs of replacement with something of similar condition, potentially including current environmental and safety standards. However, these costs are not well defined unless incurred, i.e., unless the office or factory is "replaced". But policyholders often do not replace the old building by

something of similar condition. Instead, they may choose to build a very different building in which case it is difficult to separate the “replacement” cost from the investment cost.

Another example is related to a firm’s reputation. Reputation is important for corporations, as AIG recently experienced. Its subsidiary Chartis now offers the insurance policy *ReputationGuard*,¹ which covers “proactive costs and response costs resulting from reputation threats and reputation attack.” The policyholder’s payoff is not all that clear. The policy covers the costs of engaging consultants to defend against a reputational incident and possibly the cost of implementing their recommendations, but there is no explicit criteria to determine when such an event has occurred, how much consulting service will be provided and there is only a vague reference to covering the costs of defensive actions recommended by consultants. However, the contracting parties presumably have a clear mutual understanding of what it is in principle that they want to insure. But putting this into effect after the fact calls for situation-specific judgment calls and thus verification is difficult. Instead, the contracting parties might rely on their understanding and the insurer’s concern for its reputation.

Some insurers, such as Chubb Insurance Company, have made and protected a reputation for going the extra mile to ensure that policyholders are adequately compensated and happy with their claims settlements. The strategy is to choose a policy language that is somewhat restrictive in coverage with an understanding that ambiguity over the amount of coverage is resolved in the policyholder’s favor. However, should there be suspicions about the nature of the claim or the policyholder’s conduct, then the insurer could seek protection behind the policy language. This flexibility cannot be written in an enforceable contract. Chubb does business mainly through independent agents and brokers who can help policyholders to receive the advertised favorable treatment.

A crucial issue for policyholders is the speed with which the insurer settles claims. This aspect of quality is also difficult to enforce, in particular, if the insurer can haggle over the terms of the contract. Such haggling and delay in claims payments can be very costly for a corporate policyholder who might have purchased the policy to protect against financial

¹See press release “Chartis Launches ReputationGuard®, Reputational Risk Insurance,” 10/11/2011, http://www.chartisinsurance.com/nglobalweb/internet/US/en/files/ReputationGuard%2010%207%2011_tcm295-377116.pdf

distress. Reputation plays an important role in the insurer's willingness to settle claims quickly and reliably.

To deal with problems of adverse selection and moral hazard, contracts often exclude specific events from coverage or require precautionary actions. An example is product recall, which is difficult to insure because of adverse selection and moral hazard. The contracting parties have to tackle the problems through contract adjustments and specific clauses. But these adjustments and clauses make contracts more complex and enforcement as intended more difficult. Appropriate clauses might be numerous and vague. It might then be optimal and necessary for policyholders to rely on incentive compatible rather than on enforceable contract clauses.

The relation between an insurer and a reinsurer is particularly interesting as reinsurance contracts are very simple despite the complexity of the underlying risks. They rarely specify the underwriting and claims settlement practices to be adopted by the primary insurer and often are not specific in defining coverage. Instead of writing complex contracts that try to foresee all contingencies, reinsurance contracts rely to a large extent on continued business relations, trust, and reputation rather than on the enforceability of specific clauses. The mechanism works because the contracting parties know very well the intention of the contract that they signed, even if it may be incomplete from the perspective of an outsider.

In Europe, in 2007, a major cyclonic weather system spawned a series of storms with wind, hail, and flood damages causing high reinsurance claims. For the specific level of claims payments it was crucial whether damages were attributed to a single event or to multiple events. Since the reinsurance contract did not provide clear guidance, the parties could not rely on enforcement of the contract. Instead, the threat of the insurer to terminate the business relation with reinsurers, the intervention of a broker, as well as the promise of additional (compensating) business played an important role in "determining" the level of claims payment.²

Finally, non-contractible claims can also arise because of unanticipated events, which cannot be specified in a formal contract. However, generally excluding or including unanticipated events may not be optimal. For example, consider toxic mold, which burst onto

²The example stems from private conversations with insurers.

the insurance scene as an unanticipated loss. Although it may be optimal to insure against toxic mold, its coverage carries significant moral hazard since insurance may be seen as a substitute for proper repair and maintenance of property. Thus, the issue of insurability is more complex than merely including (or excluding) unanticipated perils since coverage should not only depend on the peril, but also on the moral hazard it might engender. When contracting parties know that losses may arise in the future, but the specific nature is not known *ex ante*, a self-enforcing insurance relation might still provide a significant degree of insurance.³

3 The Model

We consider an infinite-period economy with risk-averse policyholders (households or firms) and risk-neutral insurers. All policyholders are infinitely lived and identical with strictly increasing and concave utility u from consumption in each period.⁴ In each period t , each policyholder j receives a fixed income w_0 and faces a loss of random size L_{jt} . Losses are identically and continuously distributed on the interval $[0, \bar{l}]$. Policyholders use the discount factor δ . Thus, given a random consumption stream $\{c_{jt}\}_{t=1, \dots}$, a policyholder's present value of expected utility is $V_j(\cdot) = \sum_{t=1}^{\infty} \delta^t E[u(c_{jt})]$.

The insurance market is competitive and insurers simultaneously offer insurance contracts. Each policyholder j and the insurer observe the realized loss l_{jt} . However, this loss is not verifiable so that a promised claims payment for a given loss realization cannot be enforced in a court. Instead, insurers can choose the level of transfer to their policyholders after the loss has been realized. In particular, they may choose to deviate from the coverage schedule which they initially offered. Each insurer has sufficient wealth to honor all claims if it wishes to do so. Insurers use the risk-free rate r to discount periods; there is no discounting within periods.

³Allen and Gale (1999) also discuss the role of self-enforcing (implicit) contracts to insure against unforeseen contingencies and nonspecific risks. They focus on the difference between financial transactions that are carried out through intermediaries and direct market transactions.

⁴In the context of corporations, concave utility functions are used to model the rationale for corporate risk management, e.g., Froot, Scharfstein, and Stein (1993).

4 Individual contracting

In this section we analyze the optimal insurance contract when each policyholder chooses a coverage schedule that depends on the individual loss only. To focus our discussion, we first assume that there is no intertemporal borrowing and saving and that insurers offer only one-period contracts. Later, we discuss the role of borrowing and saving and show that one-period insurance contracts are indeed optimal. Without borrowing and saving, policyholders consume their end of period net income in each period and we can drop the time index t since all time periods are identical (steady state). Because all policyholders are identical, we can also suppress the index j .

The sequence of events in each period is as follows. First, insurers simultaneously offer insurance contracts by quoting an insurance premium P and a non-negative coverage schedule specifying claims payments $I(l)$ for all loss realizations $l \in [0, \bar{l}]$. Second, losses are realized and each policyholder and the insurer observe the realized loss. Third, insurers choose the transfer to their policyholders and all policyholders consume their end of period net income, $w(l) = w_0 - l - P + I(l)$.

When losses are verifiable and insurance contracts can be enforced, insurers compete away all rents and offer insurance at a fair premium $P = E[I(L)]$. It is well known that in this case full insurance is optimal and $I(l) = l$ for all $l \in [0, \bar{l}]$.

In our setting the loss is not verifiable and it must be in the insurer's self interest to make the promised payment. That is, the coverage schedule has to be incentive compatible. The only mechanism that can provide the insurer with incentives to honor the promise is the policyholder's threat to choose a different insurer in the future if the current insurer shirks on $I(l)$. In our setting the threat to terminate the business relation is indeed credible since a policyholder is at least as well off with another insurer. Thus, given that the policyholder terminates the business relation if the insurer shirks, a contract is incentive compatible if, for all $l \in [0, \bar{l}]$, the present value of the future rent from continued business is at least as large as the required claims payment:

$$I(l) \leq \frac{P - E[I(L)]}{r}. \quad (1)$$

(1) is satisfied for all payments if it is satisfied for the maximum claims payment. Thus, (1) can be replaced by $P \geq E[I(L)] + rI^{\max}$, where $I^{\max} = \max_{l \in [0, \bar{l}]} I(l)$. In a competitive market, (1) will be binding for I^{\max} and $P = E[I(L)] + rI^{\max}$. Therefore, in addition to the expected claims payment, $E[I(L)]$, the premium includes a rent, rI^{\max} , even in a competitive insurance market. A rent is necessary to assure incentive-compatibility of the coverage schedule.

Maximizing $V(\cdot)$ is equivalent to maximizing policyholder's expected utility from end of period consumption, $E[u(\cdot)]$. With a competitive insurance market, the incentive-compatible coverage schedule is determined by the following optimization problem:

$$\max_{(P, I(\cdot))} E[u(w(L))] \quad (2)$$

$$\text{s.t. } P = E[I(L)] + rI^{\max}, \quad (3)$$

$$I^{\max} = \max_{l \in [0, \bar{l}]} I(l),$$

$$0 \leq I(l) \text{ for all } l \in [0, \bar{l}].$$

We state the solution to the optimization problem in the following proposition.

Proposition 1 *The optimal contract includes a strictly positive deductible, $D^* > 0$, an upper limit, $I^{\max*} < \bar{l} - D^*$, and full compensation of losses in excess of the deductible until the upper limit is reached: $I^*(l) = \min\{(l - D^*)^+, I^{\max*}\}$.*

The contract is piece-wise linear and resembles a standard insurance contract with a deductible and an upper limit. The novel feature is that we derive this contract for non-verifiable losses and that we obtain an endogenous upper limit. Arrow (1963), Raviv (1979), and Gollier and Schlesinger (1995) show that with verifiable losses a straight deductible insurance policy is optimal for risk-averse policyholders if insurance involves a frictional cost that is proportional to each claims payment. In this case, the marginal cost of providing an additional dollar of insurance coverage is constant. With non-verifiable losses, the friction stems from the incentive-compatibility constraint (3), which is binding and requires that the insurer earns a rent that is proportional to the maximum claims payment. In this case,

the marginal frictional cost of providing an additional dollar of insurance coverage is zero below the maximum coverage. An upper limit thus reduces the cost of providing incentive-compatible insurance and full insurance of losses below the upper limit is possible at a fair premium. However, a straight upper limit policy implies that the policyholder's marginal utility for losses above the upper limit is higher than for losses below the upper limit. A deductible reduces the premium level and thereby allows the policyholder to transfer income from those states with low marginal utility to those with high marginal utility. This motive for a deductible is also discussed by Cummins and Mahul (2004) for verifiable losses and an exogenous upper limit; a related argument is made by Doherty and Schlesinger (1990) who show that counterparty risk of the insurer makes it optimal for policyholders to choose partial insurance.

Financial distress and the level of insurance The interest rate r plays a critical role in the level of rent that is required for an incentive-compatible maximum claims payment. Alternatively, we could have assumed that there is an exogenous probability that the policyholder (firm) will stop purchasing insurance, e.g., because the firm goes bankrupt. A model with zero discounting but an exogenous probability of termination yields equivalent results as our model. A higher probability of termination is equivalent to a higher interest rate.

Proposition 2 *There exists a level of interest rate \bar{r} such that no insurance is optimal for all $r \geq \bar{r}$, i.e., $D^* = \bar{l}$ and thus $I^*(l) = 0$ for all $l \in [0, \bar{l}]$. If preferences exhibit constant absolute risk aversion (CARA), D^* is strictly increasing and $D^* + I^{\max*}$ is strictly decreasing in r for all $r < \bar{r}$.*

For CARA preferences, the optimal level of insurance decreases in r , as the optimal deductible D^* increases and the upper limit $I^{\max*}$ decreases. This observation is particularly interesting for an exogenous probability of termination. It implies that firms with a higher probability of distress, which have a higher probability of not buying insurance in the future, might optimally reduce the insurance coverage. Thus, we provide an additional justification for why highly levered or distressed firms choose a lower level of insurance that does not

rely on risk shifting or wealth effects. It is well known that with CARA, insurance demand decreases when the cost of insurance (loading) increases. However, the interesting effect here is that the cause for the higher cost stems from the policyholder's higher probability of financial distress. Although the probability of distress is unrelated to the insured risk, it increases the cost of insurance since the required rent to assure incentive compatibility increases in the probability of financial distress. One reason for a policy limit is that it reduces the insurer's risk of financial distress. In contrast, in our model, the policy limit is related to the policyholder's risk of financial distress. Because of the higher required rent, it may be optimal for a policyholder with a higher probability of financial distress to choose a lower policy limit.

Blackmailing If, for a given loss l , the required claims payment $I(l)$ is lower than the maximum claims payment I^{\max} , the insurer would be willing to pay more for the continued renewal of the contract than required by the contract. Thus, the policyholder may be tempted to blackmail the insurer and require a total payment up to I^{\max} . However, blackmailing will not occur in equilibrium if, after having been blackmailed once, the insurer expects a level of blackmailing of $E[x] > 0$ in every future period. Even if $E[x]$ is very small, the insurer will not yield to blackmailing. The reason is that the expectation of future blackmailing reduces the maximum claims payment that the insurer is willing to pay, which in turn reduces the policyholder's willingness to purchase insurance from this insurer in the future.

After blackmailing, the insurer's incentive-compatibility constraint becomes

$$I(l) \leq \frac{P - E[I(L)] - E[x]}{r}.$$

Given that the incentive constraint was binding for I^{\max} without blackmailing, the insurer no longer has an incentive to pay I^{\max} when the policyholder incurs high losses. Thus, blackmailing today reduces the insurer's maximum willingness to pay in the future. Given this adverse effect of blackmailing, it is not optimal for the policyholder to renew the insurance contract after blackmailing. Instead, it is optimal for the policyholder to switch the insurer after the incumbent insurer made the payment. Anticipating these non-renewal incentives,

the insurer will not make any payment today. We note that the policyholder cannot commit to purchasing insurance in the future because with such a commitment the insurer's incentive-compatibility constraint will also be violated.

Blackmailing is payoff-equivalent to requiring the insurer to reduce the price for insurance right before renewal of a contract. However, this cannot be optimal if such a request is likely to come again in the future. The price was chosen to maximize the policyholder's expected utility subject to the insurer's incentive-compatibility constraint. Therefore, reducing the price cannot be in the interest of the policyholder, who will choose a different insurer if the insurer is willing to reduce the price.

Borrowing and saving We now allow the policyholder to borrow at an interest rate r_b to cover realized losses or to save at a savings rate r_s . We assume that $r_s \leq (1 - \delta)/\delta \leq r_b$ and that there is no default (risk-free borrowing and saving). The policyholder will not borrow or save if there is no income risk or if losses are fully insured at a fair premium. However, if losses are not fully insured, borrowing and saving can be used to smooth income risk across periods which is akin to self-insurance. The policyholder will transfer income to future periods if current losses are low and borrow against future income if current losses are high. The benefit of self-insurance is that the policyholder can choose the level of borrowing and saving after observing the loss realization and the (fixed) repayment associated with borrowing and saving is enforceable.

As we discuss in the appendix, borrowing and saving does not affect the optimal structure of the insurance contract derived in Proposition 1. If $r_s = r_b = (1 - \delta)/\delta$, the policyholder can borrow and save to consume $w_0 - E[L]$ in each period. Thus, perfect self-insurance is possible with infinite periods and no insurance coverage is purchased. However, for $r_s < (1 - \delta)/\delta < r_b$, it is not possible to implement a constant consumption stream with purely borrowing and saving and the policyholder has to bear consumption risk. Insurance can now be optimal.

Multi-period contracting Although losses are not verifiable, claims payments are verifiable. A natural question to ask is whether multi-period insurance contracts can be used to relax the incentive compatibility constraint by increasing future premiums or reducing

expected future claims payments in exchange for a higher current payment. To implement such contracts, the policyholder must not be able to cancel the policy early to avoid higher premiums or lower future claims payments. Thus, the policyholder is locked into the insurance relation. Commitment to renewal limits the policyholder's ability to switch the insurer and, all else equal, the insurer's willingness to pay. Any potential benefit of conditioning the premium of future claims payments on the history of claims payments in a multi-period contract must therefore overcompensate the loss in flexibility to terminate the relation.

Both, an increase in future premiums and a reduction in expected future claims payments is akin to a transfer from the policyholder to the insurer. Given risk aversion, the policyholder prefers to make a fixed transfer in the future, which can be implemented through a higher premium, which the policyholder has to pay (since the contract must not be canceled early).⁵ To increase the current claims payment by a given amount, the present value of the future premium has to increase by the same amount. Thus, the mechanism through which the current claims payment can be increased is equivalent to borrowing. "Borrowing" relaxes the incentive-compatibility constraint because the amount that is borrowed is repaid in the future; it does not provide additional insurance coverage. Instead, it can be interpreted as self-insurance. But no insurance contract is needed to borrow or self-insure.

As we discuss more formally in the appendix, we can therefore conclude that the policyholder is strictly better off with one-period insurance contracts and a separate possibility to save or borrow against future income. Thus, in a competitive market where policyholder's expected utility is maximized, one-period insurance contracts will be offered. The freedom to switch insurers easily assures the highest level of insurance coverage for a given level of rent.

⁵Note that it is not possible to choose a higher transfer in states where losses are lower than the deductible, which the policyholder would prefer, since losses are not verifiable. The policyholder has no incentive to pay voluntarily since the insurer cannot enforce it and termination of the contract through the insurer has no bite since the policyholder can go to another insurer, saving the payment. Reducing future claims payments (i.e., the transfer from the insurer to the policyholder) is not optimal since in this case the transfer in states where losses are lower than the deductible would be zero.

5 Joint contracting

We now consider a group of n policyholders with individual loss exposures L_1, \dots, L_n . The sequence of events is equivalent to the one with individual insurance. The notable difference is that a joint insurance contract is possible where each policyholder's claims payment is now given by $I_j(l^n)$ for $l^n = (l_1, \dots, l_n) \in [0, \bar{l}]^n$ and $j = 1, \dots, n$. Each policyholder in the group observes the entire realization of losses $l^n = (l_1, \dots, l_n)$ and claims payments to all group members and coordinates his or her action with the other group members in the following way. If the insurer pays $I_j = I_j(l^n)$ to each policyholder $j = 1, \dots, n$, then each policyholder purchases insurance from the insurer again in the future. However, if the insurer shirks on one of the policyholders, then all policyholders switch to a rival insurer.

The incentive-compatibility constraints ensuring that the insurer honors the claims payments are then

$$\sum_{j=1}^n I_j(l^n) \leq \frac{nP - E \left[\sum_{j=1}^n I_j(L^n) \right]}{r}$$

for all $l^n \in [0, \bar{l}]^n$. The necessary and sufficient constraint is determined by the maximum aggregate claims payment to the group, i.e., $nP \geq E \left[\sum_{j=1}^n I_j(L^n) \right] + rI_n^{\max}$, where $I_n^{\max} = \max_{l^n \in [0, \bar{l}]^n} \sum_{j=1}^n I_j(l^n)$. The end of period consumption of policyholder j is $w(l_j, l^n) = w_0 - l_j - P + I_j(l^n)$, for all $l^n \in [0, \bar{l}]^n$. The optimal premium P and incentive-compatible coverage schedules $(I_j(\cdot))_{j=1, \dots, n}$ that maximize the policyholders' expected utilities are given by the solution to the following optimization problem

$$\begin{aligned} & \max_{(P, (I_j(\cdot))_{j=1, \dots, n})} \sum_{j=1}^n E[u(w(L_j, L^n))] & (4) \\ \text{s.t. } & nP \geq E \left[\sum_{j=1}^n I_j(L^n) \right] + rI_n^{\max}, \\ & I_n^{\max} = \max_{l^n \in [0, \bar{l}]^n} \sum_{j=1}^n I_j(l^n), \\ & 0 \leq I_j(l^n) \text{ for all } l^n \in [0, \bar{l}]^n \text{ and } j = 1, \dots, n. \end{aligned}$$

In the following proposition we state the structure of the optimal insurance contracts.

Proposition 3 *The optimal joint contract that maximizes the policyholders' expected utili-*

ties includes a strictly positive individual deductible $D^* > 0$ for each policyholder's loss and a joint upper limit $I_n^{\max*} < n(\bar{l} - D^*)$ on total claims payments to the group of policyholders. The optimal individual allocation (claims payment) to each policyholder can be implemented by setting

$$I_j^*(l^n) = \begin{cases} (l_j - D^*)^+ & \text{if } \sum_{j=1}^n (l_j - D^*)^+ \leq I_n^{\max*} \\ (l_j - D(l^n))^+ & \text{otherwise} \end{cases}$$

where $D(l^n)$ is given by $\sum_{j=1}^n (l_j - D(l^n))^+ = I_n^{\max*}$.

If the joint upper limit is binding ($\sum_{j=1}^n (l_j - D^*)^+ > I_n^{\max*}$), joint contracting allows for an improved allocation of the total transfer from the insurer to the policyholders: capital is allocated so that policyholders' marginal utilities are equal (subject to the constraint that claims payments are not negative). Thus, the claims payment of each policyholder depends on the loss realization of all other policyholders and can be lower than the claims payment with individual contracting. However, commitment to such an allocation is ex ante optimal. The optimal claims payments can be implemented by increasing each policyholder's deductible from D^* to $D(l^n)$, where $D(l^n)$ is chosen such that the sum over all claim payments $I_j^*(l^n) = (l_j - D(l^n))^+$ equals the maximum willingness to pay, i.e., $\sum_{j=1}^n (l_j - D(l^n))^+ = I_n^{\max*}$. The optimal contract for a binding upper limit resembles the contract proposed by Mahul and Wright (2004) for catastrophic risk sharing arrangements within a pool of policyholders when financial resources are limited.

However, in our setting, a joint contract is beneficial even when the joint upper limit is not binding and each policyholder receives the claims payment $I_j^*(l^n) = (l_j - D^*)^+$, which does not depend on the other policyholders' losses. The reason is that a joint upper limit allows for a more efficient use of the rent that the policyholders have to pay to the insurer. As an example, assume that there are two policyholders ($n = 2$) and that the joint upper limit equals the sum of the individual upper limits ($2I^{\max*} = I_2^{\max*}$). Thus, the total rent under the two contracts is equal. Now suppose that the realized losses are such that the individual upper limit is binding for policyholder j but not for policyholder i ($l_i < D + I^{\max*} < l_j, i \neq j$). Thus, policyholder j 's maximum payment is $I^{\max*}$ with separate contracting but can be increased to $\min\{l_j - D, 2I^{\max*} - (l_i - D)^+\}$ under joint contracting without changing the transfer to policyholder i .

Our analysis highlights a novel benefit of joint contracting: increased bargaining power ex post. This contrasts with higher bargaining power ex ante (to reduce the premium), which is of no value here given a competitive insurance market. It is interesting to note that the benefit of increased bargaining power can be realized even when policyholders do not accept a lower transfer ex post when the joint upper limit is binding. That is, with joint contracting it is possible to implement a contract where the policyholder always receives at least as much as with the individual contract, but possibly more when the own loss exceeds the individual upper limit which the policyholder would have with an individual contract. The claims payment can be increased through the joint threat of termination of the business relation if the other members of the group have a low loss realization.

The role of brokers Implementing a joint contract requires that policyholders observe each others' losses and the insurer's payments. Moreover, they jointly have to leave the insurer and switch to another insurer if the insurer shirks on any one of them. The cost of joint contracting between an insurer and n policyholders may therefore be very high and, indeed, we generally do not observe such contracts formally in practice.

However, implementing such contracts may be an important role of intermediaries such as brokers. A broker monitors losses and transfers to its clients and thereby has the required information readily available. If the insurer shirks on one of its clients, the broker can recommend all clients to switch to another insurer.

The brokers' ability to recommend to clients to switch insurers is a central element of brokerage. Typically, brokers "own the renewal rights" on the book of business they place with the insurer. That is, the broker is free to recommend to its clients that they renew with the current insurer or switch to a rival. Correspondingly, the insurer revokes any right to direct solicitation of business placed through the broker.

The broker can monitor the execution of a joint upper limit. Even if the joint upper limit is not explicitly stated in insurance contracts, by coordinating his clients' choice of insurer, the broker has bargaining power that allows him to implement the same allocation. In this case, the broker bargains on behalf of clients for a more generous transfer. However, when total losses are very high, there are also policyholders who will receive less than what they

would receive under individual contracting given their loss realizations. Thus, it is important that policyholders trust that their broker will act in their best interest. Trust is particularly important since the insurer (policyholders) may be tempted to collude with the broker to reduce (increase) transfers in its (their) interest. Of course, such collusion with the insurer or with individual policyholders is not allowed in practice and therefore difficult to arrange. If the broker receives side payments from the insurer or other parties, the broker's clients may not trust the broker to improve risk sharing compared to an individual direct contract.

In practice, where insurance contracts are incomplete and difficult to enforce, the relationship often is brokered. For example, Chubb uses a set of independent agents and brokers. These intermediaries "own" the renewal rights and can advise clients to move business if they become unhappy with Chubb's claims performance. Chubb entrusts its reputation to these agents and brokers to commit to its promised claim settlement strategy. As another example, a feature of the reinsurance market is the ubiquity of brokers. If a reinsurer shirks, the broker will know of it. The consequence for the reinsurer might be not only a loss of that contract, but a diversion of other business from the broker to other reinsurers. This leveraging of reputation enhances the bargaining power of the primary insurer.

6 Discussion

We derive the structure of the optimal insurance contract when losses are non-verifiable. The optimal contract to insure individual losses resembles a standard one-period insurance contract with a deductible and an upper limit on the maximum claims payment. The optimal joint contract to insure the losses of a group of policyholders involves a deductible on individual losses and a joint upper limit on the total maximum claims payment to the group of policyholders.

Insurers honor policyholders' claims because they earn a rent from future business, which they lose if they fail to pay those claims. The loss of future business is less critical for the insurer if the insurer's capacity to underwrite business is constrained, e.g., due to limited capital. The threat to terminate business is then diluted and the insurer's maximum willingness to pay is correspondingly reduced. Many insurers hold capital well in excess of

regulatory capital constraints. Reasons are that excess capital is held as buffer to deal with business cycles (Gron, 1994; Winter, 1994) or that policyholders require lower default rates than implied by regulatory standards (Epermanis and Harrington, 2006). Both reasons are related to counterparty risk of the insurer. Our analysis provides a novel justification even absent any effect on default risk: well capitalized insurers value their business relations more highly and have a higher willingness to pay for a given level of rent that they can earn from continuing their business with a client in the future.

To limit the rent, it is optimal for policyholders to introduce an upper limit. With proportional loading, the frictional cost of insurance depends only on the expected claims payment and not whether the underlying risk is a high-severity-low-frequency event or a low-severity-high-frequency event. In our setting, where losses are non-verifiable, the rent and thus the frictional cost is proportional to the maximum claims payment. For high-severity-low-frequency events, the maximum loss relative to the expected loss is particularly large and insuring non-verifiable high-severity-low-frequency risk is particularly costly relative to the expected coverage. This is consistent with high prices in the reinsurance market for catastrophe risk. Froot (2001) and Froot and O’Connell (2008) show that reinsurance premiums are a multiple of expected losses; during some periods up to seven times. This multiple is particularly high for low-frequency layers. Froot (2001) reviews several explanations and argues that shortage of risk-taking capital in the reinsurance market due to capital market imperfections is the most convincing reason. In our context, if insured losses due to catastrophic events are difficult to verify, high premiums in the reinsurance market for catastrophe risks result from the high rent which the insurer has to pay for providing the reinsurer with sufficient incentives not to dispute the validity of claims or to delay payments. Whereas Froot (2001) argues that it is costly for a reinsurer to hold sufficient capital to cover claims (ability to pay), we suggest that it is also costly to induce a reinsurer to honor claims (willingness to pay).

The required rent is also increasing if there is a positive probability that the insurance relation terminates for exogenous reasons, for example, because of the client’s financial distress. If the client’s probability of financial distress increases, covering non-verifiable losses will become more costly. Large catastrophic losses in the reinsurance market tend to increase

reinsurance premiums and reduce quantities through a reduction of reinsurance capacity and increase in counterparty risk (Gron, 1994; Winter, 1994; Cagle and Harrington, 1995; Cummins and Danzon, 1997; Froot, 2001; Froot and O'Connell, 2008). Both reasons originate from the protection seller. Our paper suggests another reason related to the protection buyer. An increase in the probability of financial distress of the protection buyer due to a large catastrophic loss reduces the likelihood that a business relation is continued. Thus, premiums for the reinsurance of losses increase and demand for reinsurance decreases if some of the losses are non-verifiable.

A Appendix: Proofs

A.1 Proof of Proposition 1

The solution to the optimization problem (2) is identical to the solution of the following problem

$$\max_{(P, I(\cdot), I^{\max})} E[u(w(L))] \quad (5)$$

$$\text{s.t. } P \geq E[I(L)] + rI^{\max}, \quad (6)$$

$$0 \leq I(l) \leq I^{\max} \quad (7)$$

with $w(L) = w_0 - L - P + I(L)$. Let f denote the probability density function of the loss L . The Lagrangian function to this optimization problem is then given by

$$\begin{aligned} \mathcal{L} &= u(w(l)) f(l) + \lambda (P - I(l) - rI^{\max}) f(l) \\ &+ \xi(l) I(l) + \zeta(l) (I^{\max} - I(l)) \end{aligned}$$

where λ , $\xi(l)$, and $\zeta(l)$ are the Lagrange multipliers to the constraints (6) and (7), respectively, with

$$\xi(l) \begin{cases} = 0 & \text{if } I(l) > 0 \\ \geq 0 & \text{if } I(l) = 0 \end{cases}$$

and

$$\zeta(l) \begin{cases} = 0 & \text{if } I(l) < I^{\max} \\ \geq 0 & \text{if } I(l) = I^{\max} \end{cases}.$$

Variational calculus implies

$$\int_0^{\bar{l}} \left(\frac{\partial \mathcal{L}}{\partial I} \delta I(l) + \frac{\partial \mathcal{L}}{\partial P} \delta P + \frac{\partial \mathcal{L}}{\partial I^{\max}} \delta I^{\max} \right) dl = 0$$

The necessary conditions for optimality are

$$u'(w(l)) f(l) - \lambda f(l) + \xi(l) - \zeta(l) = 0 \quad (8)$$

$$\int_0^{\bar{l}} (-u'(w(l)) f(l) + \lambda f(l)) dl = 0 \quad (9)$$

$$\int_0^{\bar{l}} (-\lambda r f(l) + \zeta(l)) dl = 0 \quad (10)$$

The sufficient condition is satisfied since $u(w)$ is jointly concave in $(P, I(\cdot), I^{\max})$ and the constraints (6) and (7) are linear in $(P, I(\cdot), I^{\max})$. Condition (9) implies

$$\int_0^{\bar{l}} u'(w(l)) f(l) dl = \lambda \quad (11)$$

and condition (10) yields

$$\int_0^{\bar{l}} \zeta(l) dl = \lambda r. \quad (12)$$

Combining conditions (8), (11), and (12) implies

$$\int_0^{\bar{l}} \xi(l) dl = \lambda r. \quad (13)$$

For inner solutions $0 < I^*(l) < I^{\max}$ we have $\xi(l) = \zeta(l) = 0$ and condition (8) simplifies to $u'(w(l)) = \lambda$. Marginal utility is thus constant which implies $I^{*'}(l) = 1$ for all l such that $0 < I^*(l) < I^{\max}$. Furthermore, because u is strictly increasing, condition (11) implies $\lambda > 0$. Since $r > 0$, condition (13) implies that there exists a subset in $[0, \bar{l}]$ for which $\xi(l) > 0$, i.e., for which $I^*(l) = 0$. Since $I^*(l)$ is strictly increasing for inner solutions, this implies that the subset is of the form $[0, D]$ for some deductible level $D > 0$. Analogously, condition (12) implies that there exists a subset in $[0, \bar{l}]$ for which $\zeta(l) > 0$, i.e., for which $I^*(l) = I^{\max*}$. Again, since $I^*(l)$ is strictly increasing for inner solutions, this subset is of the form $[U, \bar{l}]$ for some $U < \bar{l}$. Combining these three conditions implies that $U = D + I^{\max*}$ and $I^*(l) = \min\{(l - D)^+, I^{\max*}\}$ with $D > 0$ and $D + I^{\max*} < \bar{l}$.

A.2 Proof of Proposition 2

From Proposition 1 we have $I^*(l) = \min\{(l - D^*)^+, I^{\max*}\}$ with $D^* > 0$ and $D + I^{\max*} < \bar{l}$; D^* and $I^{\max*}$ maximize the expected utility, that is

$$\begin{aligned} & \max_{(D, I^{\max})} E[u(w_0 - L - P^* + I^*(L))] \\ &= \max_{(D, I^{\max})} \int_0^D u(w_0 - l - P) f(l) dl + u(w_0 - D - P) (F(D + I^{\max}) - F(D)) \\ & \quad + \int_{D+I^{\max}}^{\bar{l}} u(w_0 - l - P + I^{\max}) f(l) dl \end{aligned}$$

with $P^* = E[I^*(L)] + rI^{\max*}$. For arbitrarily small $r > 0$, the rent included in the premium is close to zero and some strictly positive level of insurance is optimal. For $r = 100\%$, the premium is greater than the maximum payment of the insurance policy and it is thus optimal not to buy insurance, i.e. $I^*(l) = 0$ for all $l \in [0, \bar{l}]$. Since the solutions $D^* = D^*(r)$ and $I^{\max*} = I^{\max*}(r)$ to the above maximization problem are continuous in r (see the first-order conditions below), there exists a level \bar{r} such that, for all $r \geq \bar{r}$, $I^*(l) = 0$ for all $l \in [0, \bar{l}]$. Note that

$$P = \int_D^{D+I^{\max}} (l - D) f(l) dl + I^{\max} (1 - F(D + I^{\max}) + r)$$

and thus $\partial P / \partial D = -(F(D + I^{\max}) - F(D))$ and $\partial P / \partial I^{\max} = 1 - F(D + I^{\max}) + r$. The first-order conditions for inner solutions $D^* = D^*(r)$ and $I^{\max*} = I^{\max*}(r)$ of the maximization problem are

$$\begin{aligned} \frac{\partial E[u(w(L))]}{\partial D} &= (F(D^*(r) + I^{\max*}(r)) - F(D^*(r))) \\ &\cdot (E[u'(w_0 - L - P^* + I^*(L))] - u'(w_0 - D^*(r) - P^*)) = 0 \end{aligned}$$

and

$$\begin{aligned} \frac{\partial E[u(w(L))]}{\partial I^{\max}} &= -(1 - F(D^*(r) + I^{\max*}(r)) + r) E[u'(w_0 - L - P^* + I(L))] \\ &+ \int_{D^*(r)+I^{\max*}(r)}^{\bar{l}} u'(w_0 - l - P^* + I^{\max*}(r)) f(l) dl = 0. \end{aligned}$$

These two equations can be simplified to

$$E[u'(w_0 - L - P^* + I^*(L))] - u'(w_0 - D^*(r) - P^*) = 0 \quad (14)$$

and

$$\begin{aligned} (1 - F(D^*(r) + I^{\max*}(r)) + r) u'(w_0 - D^*(r) - P^*) \\ - \int_{D^*(r)+I^{\max*}(r)}^{\bar{l}} u'(w_0 - l - P^* + I^{\max*}(r)) f(l) dl = 0. \end{aligned} \quad (15)$$

Implicitly differentiating condition (14) with respect to r yields

$$\begin{aligned}
& \frac{d}{dr} (E [u' (w_0 - L - P^* + I^* (L))] - u' (w_0 - D^* (r) - P^*)) \\
&= -P^{*'} (r) E [u'' (w_0 - L - P^* + I^* (L))] \\
&\quad - D^{*'} (r) u'' (w_0 - D^* (r) - P^*) (F (D^* (r) + I^{\max * *} (r)) - F (D^* (r)) - 1) \\
&\quad + I^{\max * *} (r) \int_{D^*(r)+I^{\max *}(r)}^{\bar{l}} u'' (w_0 - l - P^* + I^{\max * *} (r)) f (l) dl \\
&= D^{*'} (r) (F (D^* (r) + I^{\max * *} (r)) - F (D^* (r))) (E [u'' (w_0 - L - P^* + I^* (L))] - u'' (w_0 - D^* (r) - P^*)) \\
&\quad - I^{\max * *} (r) \left((1 - F (D^* (r) + I^{\max * *} (r)) + r) E [u'' (w_0 - L - P^* + I^* (L))] \right. \\
&\quad \quad \left. - \int_{D^*(r)+I^{\max *}(r)}^{\bar{l}} u'' (w_0 - l - P^* + I^{\max * *} (r)) f (l) dl \right) \\
&\quad - I^{\max * *} (r) E [u'' (w_0 - L - P^* + I^* (L))] + D^{*'} (r) u'' (w_0 - D^* (r) - P^*) = 0
\end{aligned}$$

Note that

$$\begin{aligned}
P^{*'} (r) &= -D^{*'} (r) (F (D^* (r) + I^{\max * *} (r)) - F (D^* (r))) \\
&\quad + I^{\max * *} (r) (1 - F (D^* (r) + I^{\max * *} (r)) + r) + I^{\max * *} (r).
\end{aligned}$$

CARA implies $u'' (w) = -\alpha u' (w)$ where α is the coefficient of absolute risk aversion. The two conditions (14) and (15) then imply

$$E [u'' (w_0 - L - P^* + I^* (L))] - u'' (w_0 - D^* (r) - P^*) = 0$$

and

$$\begin{aligned}
& (1 - F (D^* (r) + I^{\max * *} (r)) + r) E [u'' (w_0 - L - P^* + I^* (L))] \\
&\quad - \int_{D^*(r)+I^{\max *}(r)}^{\bar{l}} u'' (w_0 - l - P^* + I^{\max * *} (r)) f (l) dl = 0.
\end{aligned}$$

Therefore

$$\begin{aligned}
& \frac{d}{dr} (E [u' (w_0 - L - P^* + I^* (L))] - u' (w_0 - D^* (r) - P^*)) \\
&= (-I^{\max * *} (r) + D^{*'} (r)) u'' (w_0 - D^* (r) - P^*) = 0
\end{aligned}$$

and thus

$$D^{*'}(r) = I^{\max*}(r) > 0.$$

Implicitly differentiating the second condition (15) with respect to r yields

$$\begin{aligned} & \frac{d}{dr} \left((1 - F(D^*(r) + I^{\max*}(r)) + r) u'(w_0 - D^*(r) - P^*) \right. \\ & \quad \left. - \int_{D^*(r)+I^{\max*}(r)}^{\bar{l}} u'(w_0 - l - P^* + I^{\max*}(r)) f(l) dl \right) \\ &= -P'(r) \left((1 - F(D^*(r) + I^{\max*}(r)) + r) u''(w_0 - D^*(r) - P^*) \right. \\ & \quad \left. - \int_{D^*(r)+I^{\max*}(r)}^{\bar{l}} u''(w_0 - l - P^* + I^{\max*}(r)) f(l) dl \right) \\ & \quad - I^{\max*'}(r) \int_{D^*(r)+I^{\max*}(r)}^{\bar{l}} u''(w_0 - l - P^* + I^{\max*}(r)) f(l) dl \\ & \quad - D^{*'}(r) (1 - F(D^*(r) + I^{\max*}(r)) + r) u''(w_0 - D^*(r) - P^*) + u'(w_0 - D^*(r) - P^*) \\ &= 0 \end{aligned}$$

CARA and condition (15) imply

$$\begin{aligned} & \frac{d}{dr} \left((1 - F(D^*(r) + I^{\max*}(r)) + r) u'(w_0 - D^*(r) - P^*) \right. \\ & \quad \left. - \int_{D^*(r)+I^{\max*}(r)}^{\bar{l}} u'(w_0 - l - P^* + I^{\max*}(r)) f(l) dl \right) \\ &= -I^{\max*'}(r) (1 - F(D^*(r) + I^{\max*}(r)) + r) u''(w_0 - D^*(r) - P^*) \\ & \quad - D^{*'}(r) (1 - F(D^*(r) + I^{\max*}(r)) + r) u''(w_0 - D^*(r) - P^*) + u'(w_0 - D^*(r) - P^*) \\ &= 0 \end{aligned}$$

and thus

$$\begin{aligned} I^{\max*'}(r) &= -D^{*'}(r) + \frac{u'(w_0 - D^*(r) - P^*)}{u''(w_0 - D^*(r) - P^*)} \cdot \frac{1}{1 - F(D^*(r) + I^{\max*}(r)) + r} \\ &= -D^{*'}(r) - \frac{1}{\alpha} \cdot \frac{1}{1 - F(D^*(r) + I^{\max*}(r)) + r} < 0. \end{aligned}$$

Therefore

$$\frac{d}{dr} (I^{\max*}(r) + D^*(r)) = -\frac{1}{\alpha} \cdot \frac{1}{1 - F(D^*(r) + I^{\max*}(r)) + r} < 0.$$

A.3 Borrowing and saving

In each period, the policyholder can borrow an amount $x_t > 0$ to cover losses at a borrowing rate r_b or save $x_t < 0$ at a savings rate r_s . The policyholders' present value of expected utility is

$$\sum_{t=1}^{\infty} \delta^t E[u(w_t(x_{t-1}, x_t))]$$

$$\text{s.t. } w_t(x_{t-1}, x_t) = w_0 - l_t - (1 + r_{t-1})x_{t-1} + x_t$$

$$r_t = \begin{cases} r_b & \text{if } x_t > 0 \\ r_s & \text{if } x_t < 0 \end{cases}$$

$$x_0 = 0$$

If $r_s = \frac{1-\delta}{\delta} = r_b$, the policyholder by setting $x_t = l_t - E[L] + \frac{x_{t-1}}{\delta}$ consumes $w_t(x_{t-1}, x_t) = w_0 - E[L]$ in each period. This borrowing and savings plan achieves the same level of expected utility as full insurance at a fair premium. Since it is costly to insure non-verifiable losses, no insurance is purchased.

If $r_s < \frac{1-\delta}{\delta} < r_b$, it is not possible to obtain a constant consumption stream by implementing the same borrowing and savings plan. The policyholder will bear some risk and insurance may be optimal. If it is optimal to purchase insurance, the optimal one-period contract resembles the one derived in Proposition 1.

First, the optimal contract involves an upper limit I^{\max} since the cost of insurance is proportional to the maximum claims payment. Given I^{\max} , full insurance can be purchased at a fair premium for losses below I^{\max} . Thus, it cannot be optimal to borrow if $l \leq I^{\max}$ since $\frac{1-\delta}{\delta} < r_b$. However, borrowing can be optimal if $l > I^{\max}$, which reduces the optimal I^{\max} . Second, assume full insurance of losses below I^{\max} . If saving is optimal, then the same level of savings is optimal for all $l \leq I^{\max}$. Since $r_s < r_b$, the marginal utility is identical for all $l \leq I^{\max}$ but higher for $l > I^{\max}$. Thus, by the same line of reasoning as in Proposition 1 it is optimal to choose a strictly positive deductible.

A.4 Multi-period contracts

Multi-period insurance contracts are dominated by a combination of one-period insurance contracts and borrowing. Without loss of generality, we consider a two-period insurance contract with pre-

miums $\{P_1, P_2(I_1)\}$ and coverage schedules $\{I_1(l_1), I_2(l_2, I_1)\}$. The premium and claims payment in the second period can depend on the claims payment in the first period.⁶ Incentive compatibility is required to hold in both the second and first period. Thus,

$$I_2 \leq \frac{1+r}{2r+r^2} \left[P_1 - E[I_1] + \frac{E[P_2(I_1) - I_2(I_1)]}{1+r} \right] \quad (\text{IC2})$$

$$I_1 - I_1^s \leq \frac{P_2(I_1) - P_2(I_1^s) - E[I_2(I_1)]}{1+r} + \frac{1}{2r+r^2} \left[P_1 - E[I_1] + \frac{E[P_2(I_1) - I_2(I_1)]}{1+r} \right], \quad (\text{IC1})$$

where $I_1^s \in \arg \max_{\hat{I}_1} I_1 - \hat{I}_1 + \frac{P_2(\hat{I}_1)}{1+r}$ maximizes the insurer's present value when shirking in the first period.

The right hand side of (IC2) is the present value of the rent from future business when the two-period contract is renewed after the second period. The right hand side of (IC1) also includes this present value (discounted one more period since the contract is renewed only after the second period); in addition, the insurer will not make a payment to the policyholder in the second period when shirking in the first period.

Let $I_1^s = 0$, $P_2(I_1) - P_2(I_1^s) = 0$ and $E[I_2(I_1)] = E[I]$ so that the two-period contract replicates the one-period contract with the only difference that the policyholder cannot change the insurer after the first period. In this case, the right hand side of (IC2) equals $\frac{P-E[I]}{r}$ and (IC2) is equivalent to (1). (IC1) can be rewritten as $I_1 \leq \frac{P-E[I]}{r} - \frac{P}{(1+r)}$. Thus, the maximum willingness to pay in the first period is $\hat{I}_1^{\max} = I^{\max} - \frac{E[I]+rI^{\max}}{(1+r)}$ and strictly lower than with one-period contracts.

Commitment to a multi-period contract can only be beneficial if (IC2) can be relaxed by increasing the future premium $P_2(I_1)$ or reducing the expected future claims payment $E[I_2(I_1)]$ for an increase in I_1 . Given risk aversion, the policyholder prefers to shift the transfer to states where there is no loss. Since losses are non-verifiable, the best alternative is to increase the premium and choose $E[I_2(I_1)] = E[I]$. (IC2) can be rewritten as

$$I_1 \leq \frac{P-E[I]}{r} - \frac{P}{(1+r)} + \frac{P_2(I_1) - P_2(I_1^s)}{(1+r)} + I_1^s$$

For $I_1 \leq \hat{I}_1^{\max}$, (IC2) is satisfied for a constant P_2 , which is optimal given risk aversion, and $I_1^s = 0$ and $P_2(0) = P_2(\hat{I}_1)$. To increase I_1 above \hat{I}_1^{\max} , the premium must increase by

⁶Net transfers from the policyholder to the insurer cannot depend on the realized loss since conditional transfers cannot be enforced (given non-verifiability of losses) and are not incentive-compatible (given that the policyholder can go to another insurer if the contract is terminated by the insurer).

$P_2(I_1) - P_2(\hat{I}_1^{\max}) = (1+r)(I_1 - \hat{I}_1^{\max})$. The mechanism by which (IC2) can be relaxed is akin to borrowing from the insurer: the policyholder receives $I_1 - \hat{I}_1^{\max}$ today (in the form of a higher claims payment) and repays $(1+r)(I_1 - \hat{I}_1^{\max})$ in the next period (in the form of an increased premium). Thus, the relaxation of the incentive constraint stems from the possibility to borrow. Borrowing is contractible: the policyholder can promise to repay $(1+r)X$ tomorrow for X today (there is not default).

A two-period insurance contract is not necessary to borrow. The policyholder is better off if the insurer separates the borrowing component from insurance and offers one-period insurance contracts and a separate possibility to borrow. The one-period contract has the benefit of giving more bargaining power to the policyholder.

A.5 Proof of Proposition 3

The proof is analogous to the proof of Proposition 1. The solution to the optimization problem (4) is identical to the solution of the following problem

$$\max_{(P, (I_j(\cdot))_{j=1, \dots, n}, I_n^{\max})} \sum_{j=1}^n E[u(w(L_j, L^n))] \quad (16)$$

$$\text{s.t. } nP \geq E\left[\sum_{j=1}^n I_j(L^n)\right] + rI_n^{\max}, \quad (17)$$

$$0 \leq I_j(l^n) \text{ and } \sum_{j=1}^n I_j(l^n) \leq I_n^{\max} \text{ for all } j = 1, \dots, n. \quad (18)$$

with $w(L_j, L^n) = w_0 - L_j - P + I_j(L^n)$. Let f^n denote the joint probability density function of $L^n = (L_1, \dots, L_n)$ with support $[0, \bar{l}]^n$. The Lagrangian function to this optimization problem is then given by

$$\begin{aligned} \mathcal{L} = & \sum_{j=1}^n u(w(l_j, l^n)) f(l^n) + \lambda \left(nP - \sum_{j=1}^n I_j(l^n) - rI_n^{\max} \right) f(l^n) \\ & + \sum_{j=1}^n \xi_j(l^n) I_j(l^n) + \zeta(l^n) \left(I_n^{\max} - \sum_{j=1}^n I_j(l^n) \right) \end{aligned}$$

where λ , $\xi_j(l^n)$, and $\zeta(l^n)$ are the Lagrange multipliers to the constraints (17) and (18), respectively, with

$$\xi_j(l^n) \begin{cases} = 0 & \text{if } I_j(l^n) > 0 \\ \geq 0 & \text{if } I_j(l^n) = 0 \end{cases}$$

and

$$\zeta(l^n) \begin{cases} = 0 & \text{if } \sum_{j=1}^n I_j(l^n) < I_n^{\max} \\ \geq 0 & \text{if } \sum_{j=1}^n I_j(l^n) = I_n^{\max} \end{cases}.$$

Variational calculus implies

$$\int_0^{\bar{l}} \dots \int_0^{\bar{l}} \left(\sum_{j=1}^n \frac{\partial \mathcal{L}}{\partial I_j} \delta I_j(l^n) + \frac{\partial \mathcal{L}}{\partial P} \delta P + \frac{\partial \mathcal{L}}{\partial I_n^{\max}} \delta I_n^{\max} \right) dl^n = 0.$$

The necessary conditions for optimality are

$$u'(w(l_j, l^n)) f(l^n) - \lambda f(l^n) + \xi_j(l^n) - \zeta(l^n) = 0 \text{ for all } j = 1, \dots, n \quad (19)$$

$$\int_0^{\bar{l}} \dots \int_0^{\bar{l}} \sum_{j=1}^n (-u'(w(l_j, l^n)) f(l^n) + n\lambda f(l^n)) dl^n = 0 \quad (20)$$

$$\int_0^{\bar{l}} \dots \int_0^{\bar{l}} (-\lambda r f(l^n) + \zeta(l^n)) dl^n = 0 \quad (21)$$

The sufficient condition is satisfied since $u(w)$ is jointly concave in $(P, (I_j(\cdot))_{j=1, \dots, n}, I_n^{\max})$ and the constraints (17) and (18) are linear in $(P, (I_j(\cdot))_{j=1, \dots, n}, I_n^{\max})$. Condition (20) implies

$$\int_0^{\bar{l}} \dots \int_0^{\bar{l}} u'(w(l_j, l^n)) f(l^n) dl^n = \lambda \quad (22)$$

and condition (21) yields

$$\int_0^{\bar{l}} \dots \int_0^{\bar{l}} \zeta(l^n) dl^n = \lambda r. \quad (23)$$

Combining conditions (19), (22), and (23) implies

$$\int_0^{\bar{l}} \dots \int_0^{\bar{l}} \xi_j(l^n) dl^n = \lambda r \text{ for all } j = 1, \dots, n. \quad (24)$$

We first show that $I_j^*(l^n) = (l_j - D)^+$ for some $D > 0$ and all l^n if $\sum_{j=1}^n I_j(l^n) < I_n^{\max}$. For inner solutions $I_j(l^n) > 0$ we have $\xi_j(l^n) = \zeta(l^n) = 0$ for all $j = 1, \dots, n$ and condition (19) simplifies to $u'(w_0 - l_j - P + I_j(l^n)) = \lambda$. This implies $I_j(l^n) = l_j - D$ for some $D \geq 0$. Next we show that $D > 0$. Since u is strictly increasing, condition (22) implies $\lambda > 0$ and condition (24) thus yields that there exists a subset in $[0, \bar{l}]^n$ for which $\xi_j(l^n) > 0$ for all $j = 1, \dots, n$, i.e., for which $I_j(l^n) = 0$. Since $I_j(l^n)$ is strictly increasing in l_j for inner solutions and does not depend on other policyholders' losses, this implies $I_j(l^n) = 0$ for $l_j \leq D$ and $D > 0$.

Next, if $\sum_{j=1}^n (l_j - D)^+ \geq I_n^{\max}$ we show that $I_j^*(l^n) = (l_j - D(l^n))^+$ for some $D(l^n) > 0$ where $D(l^n)$ is defined by $\sum_{j=1}^n (l_j - D(l^n))^+ = I_n^{\max}$. For solutions $I_j(l^n) > 0$ we have $\xi_j(l^n) = 0$ and condition (19) simplifies to $u'(w_0 - l_j - P + I_j(l^n)) = \lambda + \zeta(l^n)/f(l^n)$. For reasons of symmetry this implies that $l_j - I_j(l^n)$ is identical for all j with $I_j(l^n) > 0$, and thus $I_j(l^n) = l_j - D(l^n)$. As shown above, condition (24) implies $I_j(l^n) = 0$ for $l_j \leq D(l^n)$.

Finally, condition (23) implies that there exists a subset in $[0, \bar{l}]^n$ for which $\zeta(l^n) > 0$, i.e., for which $\sum_{j=1}^n I_j^*(l^n) = I_n^{\max*}$. Again, since $\sum_{j=1}^n I_j^*(l^n)$ is strictly increasing in l_j for inner solutions, this subset is of the form $\{l^n \in [0, \bar{l}]^n \mid \sum_{j=1}^n (\bar{l} - l_j) \leq \bar{l} - U\}$ for some $U < \bar{l}$. Since $I_j^*(l^n) = (l_j - D)^+$, $I_n^{\max*} = U - nD < n(\bar{l} - D)$.

References

- [1] Allen, F., and D. Gale, 1999, “Innovation in Financial Services, Relationships, and Risk Sharing”, *Management Science* 45(9), 1239-1253.
- [2] Arrow, K.J., 1963, “Uncertainty and the Welfare Economics of Medical Care”, *American Economic Review* 53(5), 941-973.
- [3] Biglaiser, G., 1993, “Middlemen as Experts”, *RAND Journal of Economics* 24(2), 212-223.
- [4] Bolton, P., and D.S. Scharfstein, 1990, “A Theory of Predation Based on Agency Problems in Financial Contracting”, *American Economic Review* 80(1), 93-106.
- [5] Bond, P., and A. Krishnamurthy, 2004, “Regulating Exclusion from Financial Markets”, *Review of Economic Studies* 71(3), 681-707.
- [6] Bull, C., 1987, “The Existence of Self-Enforcing Implicit Contracts”, *Quarterly Journal of Economics* 102(1), 147-159.
- [7] Bulow, J., and K. Rogoff, 1989, “Sovereign Debt: Is to Forgive to Forget?”, *American Economic Review* 79(1), 43-50.
- [8] Cagle, J.A.B., and S.E. Harrington, 1995, “Insurance Supply with Capacity Constraints and Endogenous Insolvency Risk”, *Journal of Risk and Uncertainty* 11(3), 219-232.
- [9] Coate, S., and M. Ravallion, 1993, “Reciprocity Without Commitment: Characterization and Performance of Informal Insurance Arrangements”, *Journal of Development Economics* 40(1), 1-24.
- [10] Cummins, J.D., and P.M. Danzon, 1997, “Price, Financial Quality, and Capital Flows in Insurance Markets”, *Journal of Financial Intermediation* 6(1), 3-38.
- [11] Cummins, J.D., and N.A. Doherty, 2006, “The Economics of Insurance Intermediaries”, *Journal of Risk and Insurance* 73(3), 359-396.
- [12] Cummins, J.D., and O. Mahul, 2004, “The Demand for Insurance With an Upper Limit on Coverage”, *Journal of Risk and Insurance* 71(2), 253-264.
- [13] Doherty, N.A., and H. Schlesinger, 1990, “Rational Insurance Purchasing: Consideration of Contract Nonperformance”, *Quarterly Journal of Economics* 105(1), 243-253.
- [14] Epermanis, K., and S.E. Harrington, 2006, “Market Discipline in Property/Casualty Insurance: Evidence from Premium Growth Surrounding Changes in Financial Strength Ratings”,

- Journal of Money, Credit, and Banking* 38(6), 1515-1544.
- [15] Fluck, Z., 1998, "Optimal Financial Contracting: Debt versus Outside Equity, Review of Financial Studies 11(2), 383-418.
- [16] Froot, K.A., 2001, "The Market for Catastrophe Risk: A Clinical Examination", *Journal of Financial Economics* 60(2-3), 529-571.
- [17] Froot, K.A., and P.G.J. O'Connell, 2008, "On the Pricing of Intermediated Risks: Theory and Application to Catastrophe Reinsurance", *Journal of Banking and Finance* 32(1), 69-85.
- [18] Froot, K.A., D.S. Scharfstein, and J.C. Stein, 1993, "Risk Management: Coordinating Corporate Investment and Financing Policies", *Journal of Finance* 48(5), 1629-1658.
- [19] Garratt, R., and J.M. Marshall, 1996, "Insurable Interest, Options to Convert, and Demand for Upper Limits in Optimum Property Insurance", *Journal of Risk and Insurance* 63(2), 185-206.
- [20] Gollier, C., and H. Schlesinger, 1995, "Second-Best Insurance Contract Design in an Incomplete Market", *Scandinavian Journal of Economics* 97(1), 123-135.
- [21] Gron, A., 1994, "Capacity Constraints and Cycles in Property-Casualty Insurance Markets", *RAND Journal of Economics* 25(1), 110-127.
- [22] Hart, O., and J. Moore, 1998, "Default And Renegotiation: A Dynamic Model Of Debt", *Quarterly Journal of Economics* 113(1), 1-41.
- [23] Holmström, B., 1979, "Moral Hazard and Observability", *Bell Journal of Economics* 10(1), 74-91.
- [24] Huberman, G., D. Mayers, and C.W. Smith, Jr., 1983, "Optimal Insurance Policy Indemnity Schedules", *Bell Journal of Economics* 14, 415-426.
- [25] Kehoe, T.J., and D.K. Levine, 1993, "Debt-Constrained Asset Markets", *Review of Economic Studies* 60(4), 865-888.
- [26] Kimball, M.S., 1988, "Farmers' Cooperatives as Behavior Toward Risk", *American Economic Review* 78(1), 224-232.
- [27] Klein, B., and K.B. Leffler, 1981, "The Role of Market Forces in Assuring Contractual Performance", *Journal of Political Economy* 89(4), 615-641.

- [28] Lazear, E., 1979, "Why is There Mandatory Retirement?", *Journal of Political Economy* 87(6), 1261-1284.
- [29] Levin, J., 2003, "Relational Incentive Contracts", *American Economic Review* 93(3), 835-857.
- [30] Mahul, O., and B.D. Wright, 2004, "Efficient Risk Sharing Within a Catastrophe Insurance Pool", unpublished manuscript.
- [31] McLeod, W.B., and J.M. Malcomson, 1989, "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment", *Econometrica* 57(2), 447-480.
- [32] Mossin, J., 1968, "Aspects of Rational Insurance Purchasing", *Journal of Political Economy* 76(4), 533-568.
- [33] Myers, S., 2000, "Outside Equity", *Journal of Finance* 55(3), 1005-1037.
- [34] Raviv, A., 1979, "The Design of an Optimal Insurance Policy", *American Economic Review* 69(1), 84-96.
- [35] Rubinstein, A., and A. Wolinsky, 1987, "Middlemen", *Quarterly Journal of Economics* 102(3), 581-593.
- [36] Shapiro, C., and J.E. Stiglitz, 1984, "Equilibrium Unemployment as a Worker Discipline Device", *American Economic Review* 74(3), 433-444.
- [37] Thomas, J.P., and T. Worrall, 1988, "Self-Enforcing Wage Contracts", *Review of Economic Studies* 55(4), 541-554.
- [38] Winter, R.A., 1994, "The Dynamics of Competitive Insurance Markets", *Journal of Financial Intermediation* 3(4), 379-415.