January 11, 2010

Insurance Contract Design when the Insurer has Private Information on Loss Size*

Qin Lian
Harris Schlesinger

Abstract

This paper examines the optimal indemnity contract in an insurance market, when the insurer has private information about the size of an insurable loss. Both parties know whether or not a loss occurred, but only the insurer knows the true value of the loss and/or to what extent the losses are covered under the policy. The insured may verify the insurer's loss estimate for a fixed auditing cost. The optimal contract reimburses the auditing costs in addition to full insurance for losses less than some endogenous limit. For losses exceeding this limit, the contract pays a fixed indemnity and requires no monitoring. The optimal contract is compared with the contracts obtained in cases where it is only the insured who can observe the loss size.

Keywords: Contract design, Fraud, Insurance, Monitoring, Private information

JEL classification: D52, D82, G22

Contact author:
Harris Schlesinger, 200 Alston Hall, University of Alabama, Tuscaloosa, AL 35487-0224 (USA)
E-mail: hschlesi@cba.ua.edu

---

* Qin Lian is Assistant Professor of Finance at Louisiana Tech University. Harris Schlesinger is Professor of Finance at the University of Alabama. The authors thank Keith Crocker, Georges Dionne (the editor), Pierre Picard and two anonymous referees for helpful comments on an earlier version of this paper.

1. Introduction

Following a loss, the insurer and the insured generally have opposing interests with regards to the amount of an insurance claim. For example, Doherty and Muermann (2005) note that "Insurers are now more likely to dispute large claims, to offer less than 100 cents on the dollar, or to try to get away without paying." Indeed, larger claims tend to be underpaid, as documented by Crocker and Tennyson (2002). On the other side, the insured wants to obtain as high an indemnity as possible.

To date, the theoretical literature on insurance fraud only has considered private loss information held by the insured, who can then inflate loss claims. The insurer can verify the claim, but only at a cost.[1] However, insurers often have as much, if not more, information about the true loss size as the insured. Moreover, any indemnity payment for a loss depends not only upon the economic value of damages, but also upon contract terms and stipulations about what losses are covered and to what extent.

This paper analyzes a model of costly state verification in an insurance market where the insurer knows the magnitude of the covered damages but the insured does not. This assumption regarding information asymmetry is the reverse of that usually assumed in the fraud literature. We derive the optimal indemnity schedule and compare it to those in the extant fraud literature. Assuming that there are no non-monitoring transactions costs, this schedule is almost a "mirror image" of that obtained by Bond and Crocker (1997) when the insured has private information about the size of the loss.

[1] See Picard (2000) for a review of the literature.

Although a more realistic model would incorporate private information by both the insurer and the insured, aspects of our optimal schedule can be seen in real-world contracting. For example, automobile collision damages are typically deemed a "total loss" once repair costs reach a high-enough percent of the car's actual cash value.[2] The same rationale applies to marine insurance, where a "constructive total loss" applies to situations in which the projected losses are high enough not to assess the damages further and simply "write them off" as a total loss.

As an example of the type of asymmetry that we have in mind, consider an automobile collision accident that had minimal visible damage to the car, but may have weakened the frame of the car. The structural damage might not be noticeable immediately, but it can subject the car to further breakdowns and costly repairs in the future. Oftentimes the insurer will have its own mechanics (or subcontracted mechanics) assess the damage. Should the car's owner accept the insurer's assessment of the damage and the corresponding indemnity payment?

Likewise, consider an accident that involves a physical injury. The insured may feel much pain, but she might not know whether the pain is just a minor sprain or something more serious. Some minor treatment might ease the pain and correct the condition temporarily. At a later date, it may be impossible to prove that the original accident was a proximate cause for any ensuing untoward health effects. Suppose that the insurer has its own medical expert access the extent of the injury and then offers to pay for some specific treatment as the sole indemnification. Should the insured

[2] According to an on-line report by Steve Siler (2008), some insurers will declare a total loss once the repair costs reach as little as 51% of the car's actual cash value, though most will not declare a total loss until the bill is around 80% of the actual cash value. If monitoring the actual damages and/or monitoring the car's actual cash value involve costs to the insured, she might be willing to accept such an offer. Other transactions costs of the insurer also might help to explain the existence of such caps.

accept the payment as a remedy, or should she seek a second opinion on the extent of the damages, even if the second opinion is costly to the insured?

In other settings, such as with homeowners insurance, the contract itself is often puzzling to the insured. For example, it might not be clear exactly which losses are covered. If multiple perils compete as the proximate cause of the damages, the insurer will have an incentive to attribute most of the damage to whichever peril has the lowest level of coverage.[3] Claims adjusters typically are either employees of the insurance company, or else they are so-called "independent adjusters," who also work as hired agents of the insurance company. Such adjusters obviously have more experience with the claims process as well as more knowledge of the insurance contract provisions than does the consumer.

If the insurer offers a questionable settlement, the insured might initiate a costly verification process, to determine whether or not the insurer's loss estimate was justified. For example, public claims adjusters can be hired by the insured, and thus represent the insured's interest. If the verification procedure can be used by the courts, the insurer will have an incentive to accurately settle any claims, lest it be held liable for punitive damages in addition to any costs of indemnification.[4]

The paper follows the costly-state-verification literature, where the uninformed party has the opportunity to verify claims. Our approach is most similar to that of Bond and Crocker (1997). This approach previously had been used by Townsend (1979), Kaplow (1994) and Picard (1996) in studying the impact of potential fraud by the insured on the design of insurance contracts.

---

[3] This argument was brought to the forefront with the recent disputes about damage caused by flooding versus damage caused by wind following Hurricane Katrina.

[4] As a practical matter, it may be hard for the courts to award policyholders punitive damages for hard-to-verify injuries, since discrepancies in the loss valuation might be portrayed as differences of opinion, rather than bad faith by the insurer. This may explain the large discrepancy between claimed losses and settled losses for hard-to-verify injuries, as documented by Crocker and Tennyson (2002).

In contrast to this literature, we assume that the insurer is the party who has private information about loss size. The optimal insurance contract stipulates a flat indemnity payment for losses exceeding an endogenously determined loss amount. There is no verification for losses exceeding this critical level. The insured verifies any loss below this critical level and receives an indemnity payment based on the verified value of the loss. The optimal indemnity consists of ex-post reimbursement of incurred verification costs, plus full insurance for damages below this critical level. This contract induces a truthful revelation of the damage amount by the insurer, i.e. it is a "revelation mechanism."

One issue that presents itself quite differently from existing literature is the commitment to monitoring. Ex ante, the insured needs to commit to monitor losses below a critical level; but if the potential *ex post* gain to the insured is less than the cost of monitoring, this commitment breaks down. We deal with this problem directly by having the insurer pay only a reimbursement for *incurred* monitoring cost. This stipulation is not moot. A cash indemnity for the insured damages might or might not be used to replace or repair the damaged object. However, the optimal insurance arrangement will pay (or reimburse) only for verification costs that are actually incurred. The consumer cannot receive cash for the expected monitoring cost, and then spend the money elsewhere. Thus, the monitoring cost to the consumer *ex post* is essentially zero and the consumer will not renege on the commitment to monitor. Of course, this cost is not zero ex ante, since it will be reflected in the insurance premium. But by accepting this contract and paying the up-front premium, the insured essentially commits to monitoring.

Although our model is normative in nature, there are some types of monitoring cost that might be covered in existing real-world contracts. For example, even if the insurer provides an initial medical diagnosis, the insured might be covered for the cost of obtaining a second opinion.

The next section begins with a model of the insurer's information regarding losses. We also show how an optimal contract handles loss-verification by the insured and how the indemnity behaves within the non-monitoring region. Section 3 develops the optimal contractual form for losses within the monitoring region. We also explain how the optimal indemnity structure is second-best Pareto efficient. We conclude with some limitations of the model and directions for future extension.

2. A model with insurer information about losses

We consider a risk-averse economic agent (the "insured"), who faces a loss exposure. The insured's preferences are represented within an expected-utility framework by the twice differentiable utility function U. There also is a risk-neutral insurer that offers insurance in exchange for an up-front premium payment.

The loss distribution is common knowledge to both the insured and to the insurers. In particular, we assume that a loss occurs with probability p, $0 < p < 1$. The loss size is denoted by x, where x is the realized value of a random loss-size variable. So as to avoid bankruptcy considerations, the maximum loss is assumed to be less than W, where W represents the initial wealth of the insured. The conditional distribution of loss size, given that a loss has occurred, is assumed to be continuous over the interval $(0,\bar{x}]$ with the density function $g(x)$. The insured chooses an insurance contract to maximize the expected utility of her final wealth. Her utility function of wealth, U, is assumed to be twice differentiable with U strictly increasing and strictly concave, connoting risk aversion.

When a loss occurs, both the insured and insurer are aware of the loss but only the insurer observes the exact loss amount, x. The insurer announces a loss size $\hat{x}$, which obliges the insurer to pay the contracted indemnity $I(\hat{x})$ to the insured. Depending on the message $\hat{x}$ from the insurer, the insured can decide whether to obtain an independent audit at a fixed cost of $c > 0$. If performed, the audit is assumed to be perfectly informative in that it reveals the true size of the loss, x. In that case, the insurer pays the indemnity based on the true loss size.[5]

We consider only insurance contracts that are characterized via deterministic audits. As we show below, the insured pre-commits to verify loss announcements within a particular range and part of the indemnity payment is conditional on performing such an audit.[6] Of course, one of the obstacles in designing the optimal contract is to make sure that the insured does not renege on this pre-commitment to monitor *ex post*.

If the verification cost (or "monitoring cost") c is too high, it will preclude insurance from being offered at all, except for the possibility of a fixed payment for any loss occurrence, with no monitoring. If the monitoring cost c is negligible, there is no effect of the asymmetric information and the market effectively is fully informed. We assume that the monitoring cost c lies somewhere between these two extremes.

The optimal insurance indemnity schedule, I, consists of two regions that are contingent on the insurer's announcement of the loss size $\hat{x}$. We define the monitoring region M as the set of announced losses for which verification of the loss size occurs. Ex ante, the insured must commit to monitor the loss in order to enforce

---

[5] Since an audit will perfectly reveal an underpayment by the insurer, the insurer will have no reason to understate a loss, if it knows its announcement of $\hat{x}$ will lead to an audit by the consumer. Alternatively, we could assume a small penalty that is assessed to the insurer for lying.

[6] Mookherjee and P'ng (1989) have shown that a stochastic design typically dominates a deterministic scheme. Fagart and Picard (1999) examined optimal insurance under random audits, for the case with private information by the insured. However, we restrict our attention to deterministic contracts as a first step in characterizing the optimal contract in the presence of the insurer's private information.

the indemnity payment $I = I(x)$ if $\hat{x} \in M$; otherwise the insurer will always offer the smallest possible compensation for losses. If the announced loss size is in the complement of *M*, denote M^c, there is no verification and the insured receives the indemnity $I(\hat{x})$.

The indemnity payment, *I*, is assumed to be non-decreasing in the loss size. We also restrict the insurance indemnity to be non-negative $I(x) \geq 0$ and we assume that the indemnity is zero when no loss occurs.[7] We allow for the insurance indemnity to include a reimbursement for *incurred* monitoring costs, and we show below that it is optimal to do so. We further assume that insured will not monitor if there is nothing to be gained.[8] The fact that monitoring costs are only covered via reimbursement for incurred costs is not innocuous. If the insurer announces a loss size $\hat{x}$ for which a higher indemnity is possible, the insured will monitor the loss size with no cost of her own. On the other hand, if the insurer announces a loss size $\hat{x}$ for which some maximal indemnity is paid, there are no grounds for monitoring, even though it would not be costly to the insured. With nothing to gain *ex post* from monitoring, the insurance premium will be lower *ex ante*, since the insurer will save on reimbursing monitoring costs. *Thus, we have established a method to pre-commit the insured to monitoring all loss claims made by the insurer, except for those paying the maximum indemnity.*

In the absence of asymmetric information or verification costs, there would be no obstacle in achieving a first-best contract. However, since observation of the true loss is costly to the insured, we can only consider a second-best solution for insurance contracting. The implementation of such an optimal contract will cause the insurer to

[7] We need to distinguish here between no loss occurring and loss occurring with $x = 0$. Although the latter is a zero-probability event, we do allow for the limit of $I(x)$ to be positive as x approaches zero.

[8] In other words, if the insured is indifferent (*ex post*) to monitoring and not monitoring, we assume that no monitoring occurs.

reveal the true loss amount to the insured. We make the common assumption that the insurer will announce the loss truthfully, when faced with the same payoff for several different announced loss sizes.

The Revelation Principle allows us to restrict attention to implementation through a so-called direct mechanism.[9] Since no verification occurs in the non-monitoring region, we obtain the following result, which we label as Lemma 1.

Lemma 1: *For the optimal contract, there exists a constant indemnity payment,* $I(\hat{x}) = I_0$*, for any announced loss in the non-monitoring region,* $\hat{x} \in M^c$.

The proof is simple. If $I(\hat{x})$ is not constant across the non-monitored states, the insurer always has the incentive to announce the non-monitored state with the lowest indemnity payment.

Clearly, any indemnity payment in the monitoring region must be no larger than payments in the non-monitoring region. If this were not the case, the insurer would misrepresent the actual loss to be in the non-monitoring region, to seek a more favorable settlement. If $I(\hat{x}) = I_0$ for some $\hat{x} \in M$, there is no gain to monitoring, which by assumption contradicts $\hat{x} \in M$. This result is summarized as Lemma 2.

Lemma 2: *For all settlement announcements in the monitoring region,* $\hat{x} \in M$*, the indemnity payment is strictly lower than the constant payment in the non-monitoring region,* $I(\hat{x}) < I_0$.

---

[9] This means that the set of contracts under consideration can be limited to ones for which the insurer announces the true loss size *x*. See Meyerson (1982).

Given the features of indemnity payments over the two regions (monitoring and non-monitoring), Lemma 3 characterizes the monitoring and non-monitoring regions as being convex sets, with only the smaller loss announcements monitored.

Lemma 3: *There exists a critical loss announcement level m, such that* $M = (0, m)$ *and* $M^c = [m, \overline{x}]$.

Before proving Lemma 3, we note that monitoring is always conditional on the occurrence of a loss, which is observable. Thus, an announcement of $\hat{x} = 0$ would be taken to imply that a loss has occurred, but the insurer announces that there were zero damages. Since this is a zero probability event under our assumptions we assume without any loss of generality that $\hat{x} > 0$.

<u>Proof of Lemma 3</u>: The proof is by contradiction. Suppose that there exists $\hat{x}_1 \in M^c$ and $\hat{x}_2 \in M$ with $\hat{x}_1 < \hat{x}_2$. From Lemma 1, the indemnity payment I_0 is constant over the non-monitoring region. The non-decreasing property of the indemnity payment implies that $I(\hat{x}_1) \le I(\hat{x}_2)$, since $\hat{x}_1 < \hat{x}_2$. However, it then follows from Lemma 2 that $I(\hat{x}_1) = I_0 > I(\hat{x}_2)$, which is a contradiction. To see why, at the boundary, $m \in M^c$ and not $m \in M$, note that $x = m$ is a zero-probability event, so that the indemnity payment $I(m)$ does not affect the insurer's expected profit. Since the insured strictly prefers the higher indemnity payment, $I_0 > I(\hat{x})$ for all $\hat{x} \in M$, the optimal contract will stipulate $m \in M^c$. ■

The lemmata above assert that the insured's monitoring decision depends on the insurer's announcement offer, $\hat{x}$, and the critical level, *m*. If a small loss amount is observed, the insurer will announce the small loss $\hat{x}$, even though this implies that this announcement will entail verification by the insured. Thus, there is no reason for the insurer to falsely deflate minor loss amounts. On the other hand, if a large loss is observed, the insurer must announce the loss in the non-monitoring set. If the insurer does not announce such a loss, the loss will be monitored and the insurer will ultimately pay I_0.

For the insured, if the announced loss $\hat{x}$ is high enough to be in the non-monitoring region, there is no benefit to verification, even if costless. This follows since the indemnity will be the same for any verified loss size in the non-monitoring region, $I(x) = I_0$. Therefore, by our assumption, verification of the true loss amount will not take place. If the loss is in the monitoring region, we show below that insured will indeed monitor the loss *ex post*.

3. The optimal contract

In the extant literature, such as Kaplow (1994), Bond and Crocker (1997), and Picard (1996), the insured is the party assumed to have private information regarding the actual loss amount, while in this paper it is the insurer who knows the true loss amounts. This change in the direction of the information symmetrically modifies the structure of implementable contracts. However, this systematic change can be interpreted in light of the results in the above-mentioned papers.

An insurance contract is said to be incentive compatible, if the insurer truthfully reveals the magnitude of the loss. The implementable contract, as characterized by the lemmata can be expressed as follows.

$$(1) \qquad I = \begin{cases} I_0 & for \quad x \in M^c = [m, \overline{x}] \\ I(x) < I_0 & for \quad x \in M = (0, m). \end{cases}$$

Given such an implementable contract, the insured's expected utility is given by

$$(2) \qquad \begin{aligned} V \equiv (1-p)U(W-\pi) &+ \int_0^m U\left(W-\pi-x+I(x)-c\right) f(x)dx \\ &+ \int_m^{\overline{x}} U\left(W-\pi-x+I_0\right) f(x)dx, \end{aligned}$$

where π denotes the insurance premium and where for notational convenience we let $f(x) = pg(x)$.

The expected profit of a risk-neutral insurer with no costs other than indemnification costs may be written as,

$$(3) \qquad \textit{Expected Profit} = \pi - \int_0^m I(x) f(x)dx - \int_m^{\overline{x}} I_0 f(x)dx.$$

Individual rationality (i.e. the participation constraint of the insurer) requires that the expected profit (3) be non-negative.

If the insurer has no costs besides indemnification costs, the optimal contract will maximize the insured's expected utility, subject to the insurer's participation. We assume for now that this requires an expected profit of zero for the insurer, although as we explain below, the analysis is easily extended to any contracts in which the insurer receives any fixed level of profit. The objective is to choose the insurance indemnity schedule $I(x)$, together with the choice of insurance premium π and monitoring region, defined via the choice of m, such that (2) is maximized subject to (1) and subject to the expected profit in (3) equaling zero.

The optimal contract is characterized in the following Proposition. A formal proof of this Proposition appears in the Appendix.

Proposition: *The optimal insurance contract in a market where the insurer possesses private information on loss size has the following properties*:

(a) for $x \in M = (0, m)$, $I(x) = x + c$;

(b) for $x \in M^c = [m, \overline{x}]$, $I(x) = I_0$, *where* $I_0 > m + c$.

An efficient contract is illustrated in Figure 1. When the insurer's assessment is less than the threshold level *m*, the insured always verifies the loss amount. The optimal contract in this region entails verification costs plus full compensation for losses. In the non-monitoring region, a constant payment is offered, which saves on monitoring cost and helps to lower the premium.

--- INSERT FIGURE 1 ABOUT HERE ---

Note that for $\hat{x} = m$, on the boundary of the monitoring region, there is no monitoring and $I(m) = I_0 > m + c$, as explained in the proof of Lemma 3. In this setting with no transactions costs, the insurer charges a fair premium, which equals the expected loss reimbursement plus the monitoring cost.

Over the verification region, the insured receives full insurance together with full reimbursement of incurred monitoring expenses, while she receives a constant payment in absence of verification. This is similar to the models in which the insured possesses the private information, except that the monitoring region in those models is for large losses and the monitoring is done by the insurer. Indeed, our contract is almost a mirror image of the contracts in Kaplow (1994), Picard (1996) and Bond and Crocker (1997).

In both settings, the asymmetry of information makes first-best efficient risk-sharing (i.e. full insurance) suboptimal. Since we assume that an indemnity payment depends on whether the insured audits a loss or not, full insurance plus auditing costs are reimbursed only when the insured audits the loss. Thus, the insured will consistently commit to auditing losses in the monitoring region *ex ante* and will indeed monitor them *ex post*, since they are effectively costless to the insured *ex post*.[10] This last point is crucial in coping with the pre-commitment problem.

A more general objective would to analyze all (second-best) Pareto-efficient contracts, in this setting of asymmetric information. This is a fairly trivial extension, accomplished by restricting the insurer's profit to be some positive constant. That is, we can use equation (3), but set the expected profit of the insurer equal to some arbitrary constant $k \geq 0$. Thus, we maximize the expected utility of the insured for a given level of insurer profit. The analysis is then the same as in the Proposition.

Such a setting might be within a competitive insurance market, if we assume that the competitive market allows for some fixed level of "normal profit" in the insurance business. But we do not need to restrict ourselves to competitive markets. For example, one special case is a monopoly market, in which case the insurer would simply extract all consumer surplus. The insurer will set an indemnity schedule that yields only a negligible gain over the insured's reservation level of utility, which is the no-insurance level.

4. Conclusion

We examined the optimal contract in the presence of private information by the insurer regarding the size of the loss. The contract specifies two regions: (1) a non-

[10] Since the insurer incurs the monitoring costs in models with private information by the insured, the optimal indemnity schedule in these models does not include monitoring costs; see Bond and Crocker (1997) and Kaplow (1994).

monitoring region in which losses exceed a critical value and (2) a monitoring region in which losses are less than this critical level. Over the non-monitoring region, a fixed indemnity payment is scheduled. Over the monitoring region, a full-insurance policy that also includes a full reimbursement of any incurred monitoring costs is optimal. The inclusion of monitoring costs, but only as a reimbursement, ensures that the ex ante monitoring commitment of the insured is met ex post. Such a contract is second-best Pareto efficient and it induces the insurer to truthfully reveal information.

Our model is one of costly state verification. We do not consider the case of costly state falsification. It certainly might be the case that insurers engage in costly activities with the sole purpose of distorting the true size of the loss. For example, an insurer might pay for some repair that disguises more serious damage, before offering to pay the insured for some more obvious small damage amounts.

In a more realistic setting, both the insurer and the insured are likely to possess some private information about the loss damages. Moreover, it is likely that each party has some signal about the information possessed by the other. Although these extensions go beyond our scope in the present paper, we believe that some of the characteristics of our optimal contract already show up in extant contracts. We mentioned the notion of constructive total loss as one example. Caps on certain types damages, is another.

Obviously, in more robust settings, certain contract features might serve several purposes. For instance, caps on damages, in our setting, allow for the insurer to assess the damage to be high enough to pay a fixed payment and be done with the claim. But such caps also serve to prevent fraud by the insured, who may otherwise tend to inflate her claims. The issue then is what is the appropriate cap level?[11] Hopefully, the results in our paper can contribute to the understanding of such contract features and their purposes.

[11] For example, the provincial governments of Alberta and New Brunswick in Canada recently capped insurance payments for "minor" physical injuries in auto accidents, although Alberta's courts failed to uphold the law as constitutional, and the law is being reconsidered on appeal.

APPENDIX

Proof of the Proposition:

Consider a fixed value of *m*. We first show that $I(x) \geq c$ for an optimal contract. Suppose that $I(x) < c$ for some $x < m$. Since insured wealth is non-increasing in the loss size over the monitoring region, it follows that $W - \pi > W - \pi - x - c + I(x)$ for all $x \in (0, m)$.

Now increase the premium π by some small amount $\Delta\pi$, $\Delta\pi < c - I(x)$. Simultaneously increase I_0 by this same amount, leaving *m* unchanged. Thus, wealth (and, hence, expected utility) conditional on a loss $x \geq m$ would be unchanged. The increase to $I_0 + \Delta\pi$ also is self-financing over the region $x \in \left[m, \overline{x}\right]$. We can increase $I(x)$ for all $x \in (0, m)$ by the same $\Delta\pi$ as well, but this still leaves the extra premium collected in the no-loss state, $(1 - p)\Delta\pi$, to be disbursed, if we wish to maintain the same zero expected profit for the insurer.

Before dispersing this extra premium, note that the wealth in the no-loss state is $w - \pi - \Delta\pi$, which is higher than the wealth for loss states within the monitoring region, $W - \pi - \Delta\pi - x - c + I(x) + \Delta\pi = W - \pi - x - c + I(x)$. Now we can use the additional premium proceeds to increase $I(x)$ a bit more within the monitoring region. Thus, the total effect of the higher premium, conditional on not having a loss exceeding *m* (i.e. conditional on either having no loss or having a loss that falls within the monitoring region), would be a mean-preserving improvement via second-order stochastic dominance: less wealth in the no-loss state and more wealth in the monitoring-region loss states. Since the consumer is risk averse, the effect would be to increase her expected utility of wealth conditional on $x \in [0, m)$, while leaving the expected utility conditional on $x \in [m, \overline{x}]$ unaffected. Obviously, such a change cannot be possible for any optimal contract. As a consequence, we must have $I(x) \geq c$ for every $x \in (0, m)$ in any optimal contract.

We next turn our attention to the optimal indemnity payment, *I*(*x*), in the monitoring region. The indemnity maximizes the insured's expected utility (2) subject to insurer's incentive constraints (1) and non-negative profit constraint for the insurer (3) ≥ 0. For any fixed values of *m*, I_0 and *π*, the Hamiltonian for $x \in (0, m)$ can be written as:

$$(4) \qquad H = U(W - \pi - x - c + I(x)) f(x)$$
$$+\theta_1[\pi - (I(x) f(x) + \int_m^{\bar{x}} I_0 f(x) dx)] + \theta_2 (I_0 - I(x)).$$

If $\theta_2 > 0$, then $I(x) = I_0$ for all *x*, which we are assuming is not optimal. Thus, $\theta_2 = 0$. Differentiating *H* with respect to $I(x)$, for $x \in (0, m)$ yields

$$(5) \qquad \frac{\partial H}{\partial I(x)} = U'(W - \pi - x + I(x) - c) f(x) - \theta_1 f(x).$$

Since we have just established that $I(x) \geq c$, it follows from (5) that

$$(6a) \qquad I(x) = c \qquad \text{if } \; U'(W - \pi - x) f(x) - \theta_1 f(x) < 0$$

and

$$(6b) \qquad I(x) > c \quad \text{for } x > 0 \text{ if } \; U'(W - \pi - x + I(x) - c) f(x) - \theta_1 f(x) = 0 \qquad .$$

Raviv (1979) uses a formulation similar to (6a) and (6b) above, and argues for the existence of a deductible.[12]

If (6b) holds for all $x \in (0, m)$, the optimal coverage entails paying full insurance for the damage plus fully reimbursing the verification costs, $I(x) = x + c$. To see this, note that (6b) implies that $I(x) = x + c + \psi$ for all $x \in (0, m)$, where ψ is a constant.

[12] In the case where $c = 0$, so that the loss size is essentially observable, the result above is precisely Raviv's result for the case with zero transactions costs.

Letting $x \to 0^+$, we must have $\psi \geq 0$, since $I(x) \geq c$. Suppose $\psi > 0$. For notational convenience, denote the following probabilities

$$q_0 \equiv 1 - p, \quad q \equiv \int_0^m f(x)dx \quad \text{and} \quad q_3 \equiv \int_m^{\bar{x}} f(x)dx \qquad ,$$

which obviously depend upon m.

By our assumption, $W - \pi < W - \pi - x - c + I(x) = W - \pi + \psi$. Now reduce the indemnity for $x \in (0, m)$ by some small amount $\Delta\psi$, $0 < \Delta\psi < \psi$. At the same time, reduce the indemnity for $x \in [m, \bar{x}]$ by an amount $\Delta\pi$ and simultaneously reduce the insurance premium by the same $\Delta\pi$. Thus, for $x \in [m, \bar{x}]$ the insured's wealth is unchanged. To keep the insurer's expected profit unchanged at zero, we require $\Delta\pi = q\Delta\psi + q_3\Delta\pi$, or equivalently $\Delta\pi = [q/(q_0 + q)]\Delta\psi < \Delta\psi$. This assumption also keeps the insured's expected wealth unchanged.

The net effect of the above changes is to increase wealth in the no loss states to $W - \pi + \Delta\pi$ and reduce it in states for $x \in (0, m)$ to $W - \pi + \psi + \Delta\pi - \Delta\psi$. But, under our assumptions, this would be a mean-preserving improvement in the insured's wealth via second-order stochastic dominance. Obviously, such a change cannot be possible for an optimal contract. As a consequence, we cannot have $\psi > 0$, so that $\psi = 0$. In other words, if (6b) holds for all $x \in (0, m)$, then $I(x) = x + c$ for these losses *x*.

Now suppose that (6a) is strictly negative for some values of *x*. Then, since $U'' < 0$, we can define a deductible level $\delta \in (0, m)$ such that $I(x) = c$ for $x \in (0, \delta)$.[13] It follows from (6b) that $I(x) = c + x + \psi - \delta$ for $x \in [\delta, m)$, for some constant ψ. Letting $x \to \delta^+$, we must have $\psi \geq 0$, since $I(x) \geq c$. Suppose that $\psi > 0$. Choose a

[13] We are assuming that (6a) is not negative everywhere, as we assume that the indemnity $I(x) = c$ for all $x < m$ is not optimal.

small positive constant $\varepsilon < \psi$. Thus, the consumer's wealth for $x \in [\delta, m)$ is $W - \pi + \psi - \delta$, which is greater than $W - \pi - x$ for all $x \in [\delta - \varepsilon, \delta)$. We proceed in two steps.

Step (1): Pay a higher indemnity $I(x) = x + c + \psi - \delta$ for $x \in [\delta - \varepsilon, \delta)$.

Step (2): Reduce the indemnity by an amount $\Delta\psi < \psi$ for all $x \in [\delta - \varepsilon, m)$.

To keep expected profit at zero, we require

$$\int_{\delta-\varepsilon}^{\delta} (x + \psi - \delta) f(x)dx = \int_{\delta-\varepsilon}^{m} \Delta\psi\, f(x)dx.$$

The net result of steps (1) and (2) above is an increase in wealth for $x \in [\delta - \varepsilon, \delta)$ and a decrease in wealth for $x \in [\delta, m)$, which is a mean-preserving improvement in the insured's wealth via second-order stochastic dominance. Once again, such a change cannot be possible for an optimal contract. As a consequence, we cannot have $\psi > 0$, so that $\psi = 0$. In other words, $I(x) = x + c - \delta$ for $x \in [\delta, m)$.

We will now show that optimal deductible must equal zero. To do this, we need to allow for consumer choice of the parameters I_0, δ, π and *m*. Given the restrictions following from (6a) and (6b), as well as the implementability restriction (1), we write the Lagrangean for this problem below. The implementability restriction now takes the specific form $I_0 \geq m + c - \delta$.[14]

$$\mathcal{L} = (1-p)U(W-\pi) + \int_0^{\delta} U(W-\pi-x) f(x)dx + \int_{\delta}^{m} U(W-\pi-\delta) f(x)dx + \int_m^{\bar{x}} U(W-\pi-x+I_0) f(x)dx$$
$$+\lambda\{\pi - [\int_0^{\delta} cf(x)dx + \int_{\delta}^{m} (c+x-\delta) f(x)dx + \int_m^{\bar{x}} I_0 f(x)dx]\} + \mu[I_0 - (c+m-\delta)].$$

This leads to the following set of first-order conditions:

(7) $$\frac{\partial \mathcal{L}}{\partial \delta} = -\int_{\delta}^{m} [U'(W-\pi-\delta) - \lambda] f(x)dx + \mu = 0$$

[14] We use a non-strict inequality here to apply Kuhn-Tucker conditions and we show below that the inequality must be strict, which confirms Lemma 2.

(8) $$\frac{\partial \mathcal{L}}{\partial I_0} = \int_m^{\bar{x}} [U'(W - \pi - x + I_0) - \lambda] f(x)dx + \mu = 0$$

(9) $$\frac{\partial \mathcal{L}}{\partial \pi} = -(1-p)U'(W-\pi) - \int_0^{\delta} U'(W-\pi-x) f(x)dx - \int_{\delta}^{m} U'(W-\pi-\delta) f(x)dx$$
$$-\int_m^{\bar{x}} U'(W-\pi-x+I_0) f(x)dx + \lambda = 0$$

(10) $$\frac{\partial \mathcal{L}}{\partial m} = [U(W-\pi-\delta) - U(W-\pi-m+I_0)] f(m)$$
$$+\lambda[I_0 - (c+m-\delta)] f(m) - \mu = 0.$$

Note from (9) that $\lambda > 0$. If the implementability constraint $I_0 \geq c+m-\delta$ is satisfied with an equality, then (10) will be negative, since $\mu \geq 0$ by Kuhn-Tucker conditions. Hence, we would have $m = 0$, and no monitoring takes place, so that $I(x) = I_0 \;\; \forall x > 0$, which we are assuming is not optimal.[15] Thus, if monitoring occurs, we must have $I_0 > c+m-\delta$, which implies that $\mu = 0$.

For notational convenience, denote $q_1 \equiv \int_0^{\delta} f(x)dx$ and $q_2 \equiv \int_{\delta}^{m} f(x)dx$. The first-order conditions (7) and (8) can be rewritten as follows:

(7') $$\int_{\delta}^{m} [U'(W-\pi-\delta) f(x)dx = q_2\lambda$$

and

(8') $$\int_m^{\bar{x}} [U'(W-\pi-x+I_0) f(x)dx = q_3\lambda \,.$$

Using (7') and (8'), together with (9), we obtain

(11) $$q_0 U'(W-\pi) + \int_0^{\delta} U'(W-\pi-x) f(x)dx + (q_2+q_3)\lambda = \lambda \,.$$

Now suppose that the deductible is strictly positive, $\delta > 0$. It follows from the first order condition (7') that $U'(W-\pi-\delta) = \lambda.$ From (11) it then follows that

[15] Such a contract would exchange the premium π in the no loss state for I_0-π in the loss states. For losses $x < (I_0$-$\pi)$, wealth would actually exceed the initial wealth W. Such a contract does, however, increase the greatest lower bound for possible levels of wealth.

$q_0 U'(W-\pi)+\int_0^{\delta} U'(W-\pi-x) f(x) dx=(q_0+q_1)\lambda$. But this is impossible since we have the inequality $U'(W-\pi)<U'(W-\pi-x)<U'(W-\pi-\delta)=\lambda, \quad \forall x \in (0,\delta)$. Hence, the optimal deductible level is zero so that $I(x)=x+c$ within the monitoring region. ■

REFERENCES:

Bond, E. W. and K. J. Crocker, 1997, Hardball and the soft touch: The economics of optimal insurance contracts with costly state verification and endogenous monitoring cost, *Journal of Public Economics,* 63: 239-264.

Crocker, K. J., and S. Tennyson, 2002, Insurance fraud and optimal claims settlement strategies, *Journal of Law and Economics,* 45: 469-507.

Doherty, N. A., and A. Muermann, 2005, Insuring the uninsurable: Brokers and incomplete insurance contracts, Brookings-Wharton Papers on Financial Services, ed. Robert E. Litan and Richard Herring, Brookings Institution Press, Washington, D.C., 193-211

Fagart, M.-C., and P. Picard, 1999, Optimal insurance under random auditing, *Geneva Paper on Risk and Insurance Theory*, 24: 29-54.

Kaplow, L.,1994, Optimal insurance contracts when establishing the amount of losses is costly, *Geneva Paper on Risk and Insurance Theory,* 19: 139-152.

Mossin, J., 1968, Aspects of rational insurance purchasing, *Journal of Political Economy*, 76: 533-568.

Mookherjee, D., and I. Png, 1989, Optimal auditing, insurance, and redistribution. *Quarterly Journal of Economics*, 104: 399-415.

Myerson, R., 1982, Optimal coordination mechanisms in generalized principal-agent problems, *Journal of Mathematical Economics,* 10: 67–81.

Picard, P., 1996, Auditing claims in insurance market with fraud: The credibility issue, *Journal of Public Economics*, 63: 27-56.

Picard, P., 2000, Economic analysis of insurance fraud, in: G. Dionne, ed., *Contributions to Insurance Economics* (Boston: Kluwer Academic Publishers), 315-362.

Raviv, A., 1979, The design of an optimal insurance policy, *American Economic Review* 69, 84-96.

Siler, S., 2008, A total loss, on-line at Edmunds.com, http://www.edmunds.com/advice/insurance/articles/43771/article.html .

Townsend, R. M., 1979, Optimal contracts and competitive markets with costly state verification, *Journal of Economic Theory* 21, 265-293.

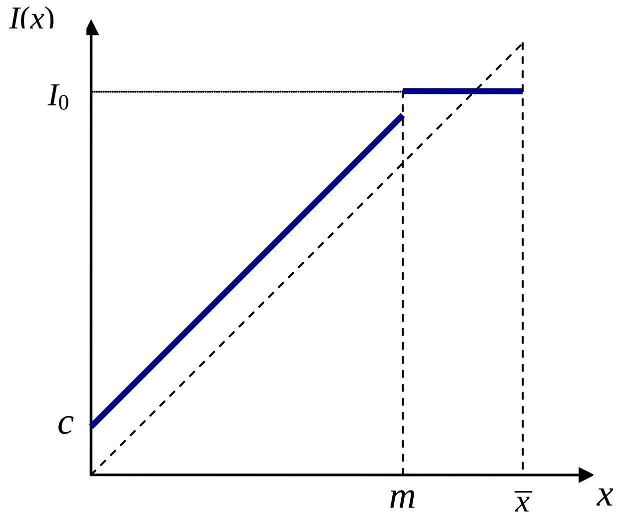

Figure 1. The Optimal Insurance Contract