

**ON THE DESIGN OF OPTIMAL  
INSURANCE POLICIES UNDER  
MANIPULATION OF AUDIT COST**

by

Pierre PICARD\*

N° 9605

June 1996

---

\*I am grateful to the participants in the Franco-American Seminar on the Economics of Insurance (Bordeaux, June 23-25, 1995), and in seminars at ENPC and THEMA (Paris), and more particularly to Bernard Caillaud and Keith Crocker, for helpful comments and suggestions.

## ON THE DESIGN OF OPTIMAL INSURANCE POLICIES UNDER MANIPULATION OF AUDIT COST

**Abstract** : This paper characterizes optimal insurance policies under deterministic auditing, in a situation where the policyholders can misrepresent their losses. First, it is shown that a pure deductible contract (i.e. a 100 per cent coverage above a deductible) is optimal when the policyholders are in a position allowing them to inflate their claims by intentionally increasing the damages. The paper then focuses on the situation where policyholders can manipulate the audit cost and the insurer is unable to observe the cost incurred by his auditor, which leads to a moral hazard problem and to an auditor-insurer cost-sharing rule. When the auditor is risk-averse, the optimal insurance policy involves some degree of coinsurance. An upper limit on coverage is optimal when the auditor is infinitely risk-averse.

Keywords : insurance, audit, coinsurance.

JEL classification number : D8

## CONTRATS D'ASSURANCE OPTIMAUX AVEC MANIPULATION DES COÛTS D'AUDIT

**Résumé** : L'objet de cet article est de caractériser les polices d'assurance optimales dans un contexte où les assurés sont en mesure de masquer le montant exact de leurs pertes et où les assureurs mettent en place des procédures d'audit déterministe. On montre tout d'abord qu'il est optimal d'offrir un contrat de franchise pure (c'est-à-dire une couverture à 100 % au-delà d'une franchise) lorsque les assurés peuvent accroître leurs demandes d'indemnité en accroissant délibérément le montant de leurs pertes. On considère ensuite la situation où les assurés peuvent manipuler le coût d'audit et où l'assureur ne peut observer le coût supporté par son auditeur, ce qui conduit à un problème de risque moral et à une règle de partage des coûts entre l'auditeur et l'assureur. Lorsque l'auditeur a de l'aversion pour le risque, la police d'assurance optimale inclut un certain degré de coassurance. De plus, il est optimal de fixer une limite supérieure à la couverture lorsque l'auditeur a une aversion au risque infinie.

Mots clefs : assurance, audit, coassurance.

JEL numéro de classification : D8

# 1 INTRODUCTION

In a situation where insurance purchasers can misrepresent their losses, and where verifying the actual extent of damages is costly, an optimal insurance policy should reach a compromise between two conflicting objectives : sharing the risk between the insured and the insurer optimally and minimizing the expected verification cost. This trade-off leads to an optimal second-best policy that departs from the perfect information full insurance solution. In particular, a deterministic audit policy specifies whether there is verification or not, as a function of the magnitude of damages. The purpose of this paper is to characterize optimal insurance contracts, under such a deterministic auditing policy. We shall more particularly focus our attention on the possibility for policyholders to inflate claims by manipulating audit costs.

The analysis of optimal contracting under exogenous audit cost will provide a useful benchmark for our analysis of audit cost manipulation. Under deterministic auditing<sup>1</sup> and exogenous audit cost as shown by Townsend (1979), a risk-neutral insurance company should offer policies that involve positive coverage only if the loss exceeds a minimum level. More specifically, we shall provide conditions under which the optimal policy actually involves 100 percent coverage above a deductible (what we shall call a *pure deductible contract*). As a matter of fact, Townsend's solution corresponds to a discontinuous indemnity schedule that may encourage the policyholder to create further damage when the loss is less than the deductible. The optimal indemnity schedule should trade-off risk-sharing and audit cost minimization, but it should also urge the insured not to inflate the size of the claim by intentionally increasing the damages. As a matter of fact, we will show that the optimal policy is a pure deductible contract. This proposition extends a result due to Huberman, Mayers and Smith (1983) to a costly state verification<sup>2</sup>

---

<sup>1</sup>For simplicity's sake in this paper, we shall limit our attention to deterministic audit procedures in which claims are either monitored with certainty or not monitored at all, depending on the size of damages. This simplification of the auditing procedure will allow us to focus on the effect of the auditor-insurer agency relationship on the shape of the indemnity schedule in a tractable way. On random auditing, see Townsend (1979), Baron and Besanko (1984) and Mookherjee and Png (1989).

<sup>2</sup>Huberman's, Mayers' and Smith's proof (1983) is rather informal, while we prove the

and it provides another motive for offering pure deductible contracts, besides the wellknown interpretations in terms of transaction costs (Arrow, 1971) or moral hazard (Holmström, 1979).

In many cases, policyholders are also in a position allowing them to inflate claims by manipulating the audit costs so as to make the monitoring activity more difficult<sup>3</sup>. This manipulation is costly to the policyholder since it typically involves some kind of collusion with middlemen (such as healthcare providers, garage owners or lawyers) who are in position to falsify the damages and to make it more costly to prove that the claim has been built up. In such a situation the insurer may not be able to observe the cost incurred by his auditor (be he a consulting physician, an expert, a claim handler, an investigator or a lawyer) which leads to a moral hazard problem. The payment from the insurer to the auditor should then be contingent on the amount of losses so as to induce efficient auditing in order to dissuade the policyholder from manipulating the audit cost. This will result in larger expected cost if the auditor is risk-averse. We will show that some degree of *coinsurance* above a deductible is then optimal in order to decrease the risk premium paid to the auditor. In particular, when the auditor is infinitely risk-averse, the optimal policy entails an *upper limit* on coverage. In some cases, the optimal policy entails coinsurance of large losses or an upper limit on coverage without any deductible for small losses, although there are fixed costs per claim.

These results may be compared to other cases for coinsurance that have been put forward in the literature. Raviv (1979) shows that coinsurance is optimal if the insurer is risk-averse or if insurance costs are a convex function of the coverage. He also justifies ceilings on coverage through regulatory constraints. Likewise, Huberman, Mayers and Smith (1983) argue that bankruptcy rules may lead insurers to offer undercoverage of high losses. We here provide a complementary explanation for coinsurance and for an upper limit on coverage, which results from imperfect monitoring of the auditor's activity by the insurer.

---

optimality of a standard deductible contract by using an optimal control argument.

<sup>3</sup>Our description of build up as pertaining to the possibility for the insureds to manipulate audit costs follows Bond and Crocker (1995). However, Bond and Crocker assume that the auditor's activity (and particularly the cost of monitoring claims) is perfectly observed by the insurer, while the present paper focuses on the agency relationship between the insurer and the auditor, in a context where audit costs are not observed by the insurer.

Our results are particularly relevant to markets covering medical and disability insurance, fire, theft or liability insurance for consumers and small businesses where coinsurance clauses or ceilings on coverage are commonly observed. For instance, in the case of health care, co-payment rates on prescription charges for drugs are frequently combined with a deductible system that makes the consumer pay 100 % up to a certain level above which the expenditures are subsidized. In the above mentioned markets, individual losses are small compared to the aggregate payments made by the insurer, therefore assuming that insurers are risk neutral seems to be quite acceptable. Furthermore, there is no obvious reasons for assuming that the costs associated with such insurance payments are a strictly convex function of the coverage. On the contrary, assuming fixed costs of processing claims and a nonincreasing marginal cost associated with insurance payments seems much more realistic (see Gollier, 1987). Under such assumptions, Raviv's (1979) analysis predicts that pure deductible contracts should be observed, contrary to many insurers' actual behavior. However, the above mentioned insurance markets involve significant auditing costs, and the policyholders may make it more costly to detect build up by colluding for instance with health care providers or by appealing to lawyers for help<sup>4</sup>. Furthermore, the consulting physicians, the experts or the lawyers hired as auditors by the insurers may conceivably be risk-averse, or their behavior may be affected by limited liability constraints. In such cases, our model predicts that a coinsurance clause or an upper limit on coverage should be combined with a deductible system and also that the optimal auditor-insurer contract corresponds to a simple cost-sharing formula.

The paper is organized as follows. Section 2 presents the model when audit cost cannot be manipulated. In this framework, we establish the optimality of pure deductible contracts when the policyholders can create extra damages in order to inflate their claims. Such policies with straight deductibles are benchmarks for the subsequent analysis of audit cost manipulation. In section 3, we characterize the insurer-auditor relationship and the

---

<sup>4</sup>An interesting example is provided by Dionne and St-Michel (1991) in their analyses of workers' compensation. They describe the behavior of a worker who tries to find and convince a physician to write a medical report permitting more days of absence from work than the level that would be decided upon under perfect information. They show that this build-up mechanism is particularly significant for injuries (such as lower back pain or spinal disorder) that are difficult to diagnose.

optimal insurance contract when the policyholders can manipulate the cost born by the auditor. In particular, we establish the optimality of coinsurance or of ceilings on coverage. Section 4 concludes. All proofs are in appendix.

## 2 OPTIMAL INSURANCE CONTRACTS WHEN AUDIT COSTS CANNOT BE MANIPULATED

Identical insurance buyers own an initial wealth  $W$  and they face a risk of loss  $x$ , where  $x$  is a random variable with a support  $[0, \bar{x}]$  and a cumulative distribution  $F(x)$ . The no-loss outcome may be reached with positive probability. Therefore, we assume that  $x$  is distributed according to a mixture of discrete and continuous distributions :  $x$  has a mass of probability  $f(0)$  at  $x = 0$  and there is a continuous probability density function  $f(x)$  over  $(0, \bar{x}]$ , with  $f(x) > 0$  for all  $x$ . In other words  $f(x)/[1 - f(0)]$  is the density of damages conditional on a loss occurring.

The insurance policy is described as a contract specifying the (non-negative) payment  $t(x)$  from the insurer to the policyholder if the loss is  $x$  as well as a premium  $P$  paid by the policyholder. The realization of  $x$  is known only to the policyholder unless there is verification. The cost of an insurance contract includes the payment  $t(x)$ , a fixed administrative cost  $c_f$  whenever a claim is filed and an audit cost  $c_a$  when a claim is verified. We note  $c = c_f + c_a$ .

The policyholder's final wealth is  $W_f = W - P - x + t(x)$ . Policyholders are risk-averse. They maximize the expected utility of final wealth  $EU(W_f)$ , where  $U(\cdot)$  is a twice differentiable von Neumann-Morgenstern utility function, with  $U' > 0$ ,  $U'' < 0$ .

We define a deterministic audit policy as a verification region  $M \subset (0, \bar{x}]$ , with complement  $M^c$  that specifies when there is to be verification.

For simplicity's sake, we assume that at zero cost the policyholder can misrepresent the magnitude of damages downward : this assumption will straightforwardly ensure that the actual payment is a non-decreasing function of the damages and that the verification region  $M$  is an upper interval in  $[0, \bar{x}]$  in all the variants of our model -see lemma 1 below<sup>5</sup>. The implementation of

---

<sup>5</sup>Assuming that damages can be misrepresented downward is not essential to establish the main results of the paper but it simplifies the proofs significantly. See the proof of

an insurance contract is as follows. A policyholder who experiences a loss  $x$  may choose to file a claim  $\hat{x}$  and he can misrepresent the damages at any level  $\bar{x}$  such that  $\bar{x} \leq x$ . If  $\hat{x} \in M$ , the claim is audited, the loss  $\bar{x}$  is observed and the payment is  $t(\bar{x})$ . If  $\hat{x} \in M^c$ , the claim is not audited and the payment to the policyholder is  $t(\hat{x})$ .

A contract  $\delta = \{t(\cdot), M, P\}$  is said to be *incentive compatible* if the policyholder truthfully reveals the actual loss, i.e. if  $\hat{x} = \bar{x} = x$  is an optimal strategy for the policyholder. Lemma 1 establishes that any contract is weakly dominated<sup>6</sup> by a non-decreasing incentive compatible contract, in which the claims are audited when they exceed a threshold  $m \in [0, \bar{x}]$ .

**Lemma 1** *Any contract  $\delta = \{t(\cdot), M, P\}$  is weakly dominated by an incentive compatible contract  $\tilde{\delta} = \{\tilde{t}(\cdot), \tilde{M}, \tilde{P}\}$  such that:*

- (i)  $\tilde{M} = (m, \bar{x}]$  with  $m \in [0, \bar{x}]$
- (ii)  $\tilde{t}(x)$  is non-decreasing with  $\tilde{t}(x) = t_0 \geq 0$  if  $x \leq m$  and  $\tilde{t}(x) > t_0$  if  $x > m$

Hence the optimal contract may be described by  $P, m \geq 0, t_0 \geq 0$  and  $t(\cdot) : (m, \bar{x}] \rightarrow \mathbb{R}_+$  such that  $t(\cdot)$  is non-decreasing and  $t(m_+) \geq t_0$ <sup>7</sup>. The optimal contract maximizes the policyholder's expected utility

$$EU = \int_0^m U(W - x - P + t_0) dF(x) + \int_{m_+}^{\bar{x}} U(W - x - P + t(x)) dF(x)$$

subject to the constraint which assures that the expected profit of the insurer is non-negative, as well as to monotonicity and sign constraints.

When  $x \in [0, m]$ , a claim is filed only if  $t_0 > 0$ . The non-negative profit constraint may then be written :

$$P \geq (t_0 + c_f)F(m) + \int_{m_+}^{\bar{x}} [t(x) + c]dF(x) \quad \text{if } t_0 > 0$$

and

---

proposition 3.2 in Townsend's paper when such an assumption is not made.

<sup>6</sup>Dominance is here defined with respect to the expected utility of the policyholder and to the expected profit of the insurer.

<sup>7</sup>For any function  $\varphi(\cdot) : [0, \bar{x}] \rightarrow \mathbb{R}$ , we note  $\varphi(x_+) = \lim_{\substack{y \rightarrow x \\ y > x}} \varphi(y)$ .

$$P \geq \int_{m_+}^{\bar{x}} [t(x) + c]dF(x) \quad \text{if } t_0 = 0$$

Consider a contract  $\delta$  characterized by  $t_0 = \bar{t}_0 > 0$ ,  $t(\cdot) = \bar{t}(\cdot)$ ,  $m = \bar{m}$  and  $P = \bar{P}$ . Then define the contract  $\delta'$  such that  $t_0 = 0$ ,  $t(\cdot) = \bar{t}(\cdot) - t_0$ ,  $P = \bar{P} - t_0 - c_f F(\bar{m})$  and  $m = \bar{m}$ .  $\delta$  and  $\delta'$  yield an identical expected profit and  $\delta'$  gives a larger expected utility than  $\delta$ . Hence  $t_0 = 0$  is optimal, which is quite intuitive since a zero payment in the no-verification region saves fixed administrative costs<sup>8</sup>.

The optimal contract  $\delta = \{t(\cdot), m, P\}$  is characterized by maximizing

$$EU = \int_0^m U(W - x - P)dF(x) + \int_{m_+}^{\bar{x}} U(W - x - P + t(x))dF(x) \quad (1)$$

subject to

$$P \geq \int_{m_+}^{\bar{x}} [t(x) + c]dF(x) \quad (2)$$

$t(x)$  being non-negative and non-decreasing over  $(m, \bar{x}]$  with  $m \geq 0$ .

Proposition 1 characterizes the optimal contract.

**Proposition 1** *An optimal contract satisfies the following conditions:*

$$\begin{aligned} t(x) &= 0 \text{ if } x \leq m \\ t(x) &= x + z - m \text{ if } x > m \\ \text{with } m &> 0, 0 < z < m \text{ and } M = (m, \bar{x}]. \end{aligned}$$

The optimal contract is described in figure 1. No claim is filed when the loss is less than a threshold  $m > 0$ . When the damages are larger than  $m$ , the optimal contract entails verification of the claims and partial insurance. Furthermore marginal damages are fully insured, i.e.  $t'(x) = 1$  for  $x > m$ .

---

<sup>8</sup>Note that this result would not hold any more if the fact that the policyholder had suffered some loss were publicly observable, while the size of the damages is private information known only to the policyholder. Paying a positive indemnity in the no-verification region would then be optimal (see Bond and Crocker, 1995).



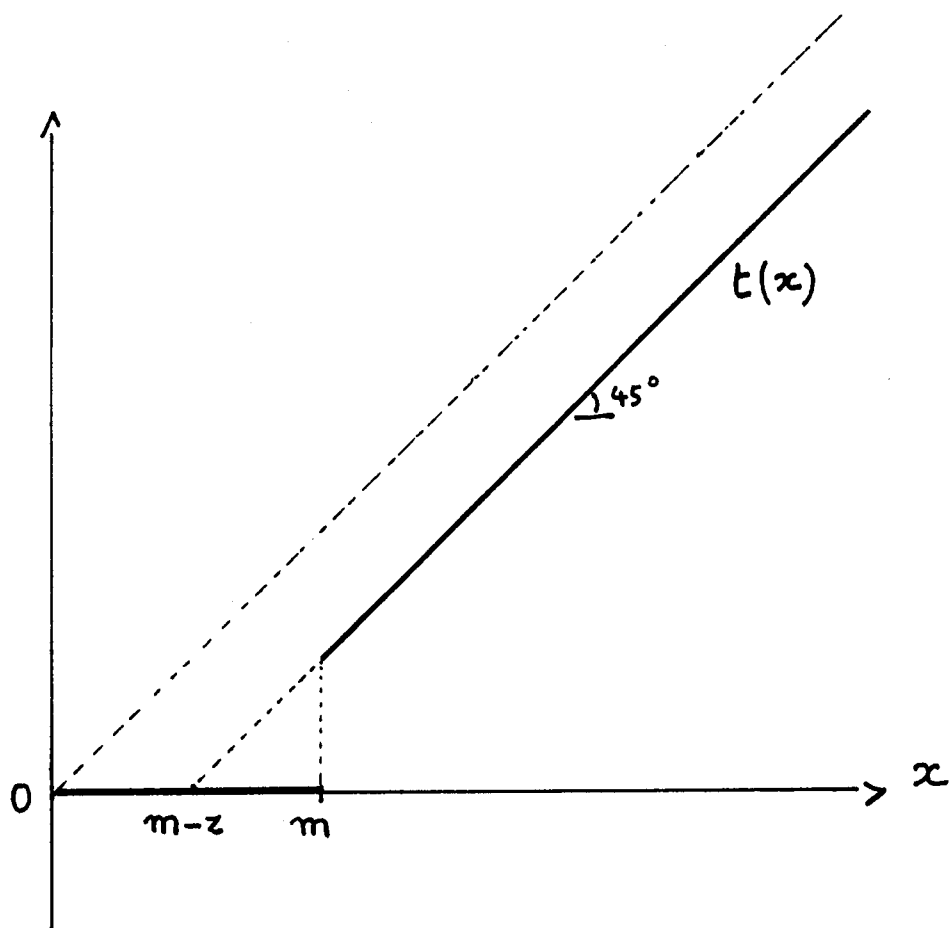


Figure 1

The shape of the optimal indemnity schedule  $t(x)$  may be interpreted as follows. For cost minimization reasons, it is optimal to choose a positive threshold  $m$  so as to decrease administrative and audit costs. Then,  $t(\cdot)$  should equalize the marginal utility of final wealth in each state where there is verification with the expected marginal utility of final wealth. This is an obvious consequence of the fact that any increase in the insurance payment in an interval  $[a, b] \subset (m, \bar{x}]$  should be compensated by an increase in the insurance premium which is paid in all states of nature. Because no claim is filed under the threshold, the expected marginal utility of final wealth is larger than the marginal utility in the no-loss state. Consequently, only partial insurance is optimal when the threshold is crossed.

The optimal contract that we have just described does not look like the policies that are most frequently offered by insurers and the explanation to this discrepancy is probably the following : because of the upward discontinuity of the optimal indemnity schedule at  $x = m$ , once some damage has occurred, the insured may be incited to create further damage and to inflate the size of the claim. More precisely, a policyholder who suffers a loss  $x$  less than  $m$  but greater than  $m - z$  will profit by increasing the damage up to  $x = m$  insofar as the insurer is not able to distinguish the initial damage and the extra damage. A policyholder would never increase the damage if and only if  $t(x) - x$  were non-increasing over  $[0, \bar{x}]$ . When  $t(x)$  is non-decreasing, this no-manipulability condition implies that  $t(x)$  should be a continuous function.

In what follows, we will refer to A1 as to the assumption according to which the policyholders can create extra damage so as to inflate their claims<sup>9</sup>. Under A1, an insurance contract is said to be incentive compatible when policyholders do not create further damage and they truthfully announce their loss.

---

<sup>9</sup>Extra damages may be deliberately brought about by the insured, as for instance when a firm lets stocks in warehouse burn so as to pocket the insurance compensation (In the US, the Insurance Committee for Arson Control (ICAC) estimates that arson costs insurance companies nearly 30 % of the total dollars spent on fire losses !). However, in many cases, a loss increase is made thanks to the help of a middleman : think for instance of a physician who increases the prescription drugs or the hospital charges, so as to exceed the threshold of the insurance policy of his patient. Garage owners or lawyers may also be in a position allowing them to inflate the cost of road accidents or casualties. In such cases, A1 means that eliciting verifiable information about overpayment is too costly or too time consuming to the insurer.

It is straightforward to check that under A1, the contract defined in Proposition 1 is dominated by a pure deductible contract defined by  $t(x) = \text{Sup}\{0, x - m + z\}$  and  $M = (m - z, \bar{x}]$ . To establish the optimality of pure deductibles, we proceed in two steps. Lemma 2 establishes that under A1, any contract is weakly dominated by an incentive compatible contract in which the largest claims are verified. Proposition 2 then says that pure deductible contracts are optimal.

**Lemma 2 :** *Under A1, any contract  $\delta = \{t(\cdot), M, P\}$  is weakly dominated by an incentive compatible contract  $\tilde{\delta} = \{\tilde{t}(\cdot), \tilde{M}, \tilde{P}\}$  such that:*

- (i)  $\tilde{M} = (m, \bar{x}]$  with  $m \in (0, \bar{x}]$
- (ii)  $\tilde{t}(x)$  is continuous and non-decreasing and  $\tilde{t}(x) - x$  is non-increasing with  $\tilde{t}(x) = t_0 \geq 0$  if  $x \leq m$  and  $\tilde{t}(x) > t_0$  if  $x > m$ .

Once again, because of the fixed cost of claims  $c_f$ , choosing  $t_0 = 0$  is optimal. Using Lemma 2, the optimal contract  $\delta = \{t(\cdot), m, P\}$  is derived by maximizing  $EU$  given by (1) subject to (2) and  $0 \leq dt/dx \leq 1$ ,  $t(m_+) = 0$ ,  $m \geq 0$ . This leads to Proposition 2.

**Proposition 2 :** *Under A1, the optimal contract is a pure deductible contract*

$$t(x) = \text{Sup}\{0, x - m\}$$

*with  $m > 0$  and  $M = (m, \bar{x}]$ .*

### 3 MANIPULATION OF AUDIT COST AND OPTIMALITY OF COINSURANCE

We now turn our attention to the contractual relationship between the insurer and the auditor. The latter incurs the cost  $c_a$  whenever an audit is carried out. Letting  $r$  denote the payment from the insurer to the auditor, the auditor's net wealth is  $\omega = r - \tilde{c}$  where  $\tilde{c}$  denotes the cost beared by the auditor in case of claim verification. Under the previous assumptions we had  $\tilde{c} = c_a$  in case of audit and  $\tilde{c} = 0$  otherwise. We will now focus our attention on situations where the policyholder may affect the audit cost  $\tilde{c}$  and where the insurer does not observe  $\tilde{c}$ : he can only verify (without cost) whether an audit has

been performed or not. The fixed cost  $c_f$  is borne by the insurer when a claim is filed. The auditor maximizes the expected utility of his net wealth  $EV(\omega)$ , where  $V(\cdot)$  is a twice-differentiable von Neumann-Morgenstern utility function, with  $V' > 0, V'' \leq 0$ . He is willing to participate if his expected utility is larger than a reservation utility  $\bar{v}$ .

The auditor sends a report  $\tilde{x} \in [0, \bar{x}]$  to the insurer which is an evaluation of the size of the loss in case of audit. When no audit is performed, we let  $\tilde{x} = \phi$ . Messages  $\tilde{x} \in [0, \bar{x}]$  can be sent only after auditing, which costs at least  $c_a$  to the auditor. However, the policyholder may engage in actions which increase the cost to the auditor of establishing the actual loss unambiguously. More precisely, the economic loss  $x$  is observed at cost  $c_a$  when an audit is carried out. However, by incurring expenditures  $e \geq 0$ , the policyholder increases the cost to the auditor of eliciting verifiable information on the loss, which reaches  $c_a + be$ , with  $b > 0$ . Hence, we here distinguish the cost  $c_a$  of observing the loss which is incurred once the claim is audited, and the cost of eliciting verifiable information, which is larger and may be manipulated by the policyholder. For instance, the latter may pay middlemen such as health-care providers, garage owners or lawyers to make it more costly or more time-consuming to the claim handler or to the insurer's lawyer to establish the level of damages unambiguously<sup>10</sup>. The insurer does not observe the auditor's cost, which lead to a moral hazard problem : the auditor-insurer contract should induce the auditor to make effort to gather verifiable information about fraudulent claims.

In this context, we will describe the auditor-policyholder relationship as a three stage game. At stage 0, the loss level  $x \in [0, \bar{x}]$  is realized. At stage 1, the policyholder announces a loss  $\hat{x} \in [0, \bar{x}]$  and he incurs the manipulation cost  $e$ . At stage 2, the claim is audited if and only if  $\hat{x} \in M$ . When  $\hat{x} \in M$ , the auditor observes  $x$  and he reports  $\tilde{x} \in \{x, \hat{x}\}$ . If  $\tilde{x} = x \neq \hat{x}$ , the auditor incurs the cost  $c_a + be$  so that his report incorporates verifiable information. If  $\tilde{x} = \hat{x}$  the auditor only incurs the audit cost  $c_a$ . If  $\hat{x} \in M^c$ , we have  $\tilde{x} = \phi$ .

---

<sup>10</sup>Evidence on the fact that middlemen have an action on claims build-up for auto bodily injury insurance is provided by Derrig, Weisberg and Chen (1994) in their study of the Massachusetts no-fault automobile insurance system. They particularly show that the presence of an attorney and the pattern of treatment, much more than the injury itself, determine whether the driver reaches the (\$ 2,000) monetary threshold that restricts the eligibility to file a tort claim.

This sequence will be called A2<sup>11</sup>.

In what follows, the indemnity paid to the policyholder will depend on the auditor's report and we will have :  $t = T(\tilde{x})$ . Likewise we will restrict attention to simple auditor-insurer contracts  $R = R(\tilde{x})$  that specify the transfer (net of the standard audit cost  $c_a$  in case of audit) paid to the auditor as a function of his report. We then have  $r = R(\tilde{x}) + c_a$  in case of audit, and  $r = R(\tilde{x}) = R(\phi)$  otherwise. These restrictions on contracts deserve some comments. Assuming that the indemnity payment only depends on the auditor's report makes sense if the insurer cannot prove in court that the insured has deliberately overestimated the damages (when  $\hat{x} > \tilde{x}$ ). Hence, we assume that more general insurance contracts  $t = T(\hat{x}, \tilde{x})$  are not feasible, either because the amount of the initial claim  $\hat{x}$  cannot be unambiguously verified by an outside party (while verification of the auditor's report  $\tilde{x}$  is possible)<sup>12</sup>, or because in case of established overestimation of losses, the policyholder may maintain that he was in good faith or because of prohibitive transaction costs arising from the implementation of more complex contracts. For similar reasons, the (net) transfer to the auditor depends on  $\tilde{x}$  and not on  $\hat{x}$ , i.e.  $R = R(\tilde{x})$ . As we shall see, such simple contracts  $R(\tilde{x})$  may be interpreted in terms of auditor-insurer cost-sharing rules, which gives them a straightforward interpretation. Finally, unlike the Bond-Crocker (1995) approach, the fact that the audit cost is not observed by the insurer prevents from conditioning  $T$  and  $R$  on  $\tilde{c}$ .

Insurance contracts are still denoted  $\delta = \{t(\cdot), M, P\}$ , with  $t(\cdot) : [0, \bar{x}] \rightarrow \mathbb{R}_+$  and  $M \in (0, \bar{x}]$ , where  $t(x)$  is the indemnity payment when the loss is  $x$  and claims are audited whenever  $x \in M$ . The contract  $\delta$  is said to be *admissible* if it satisfies the characterization of  $\tilde{\delta}$  in Lemma 2. In what follows, we restrict our attention to admissible contracts since under A1 any contract

<sup>11</sup>Very likely, eliciting verifiable information about the actual level of losses is easier for the auditor (or equivalently manipulating audit costs is harder for the insured) when claims' build up is totally untrue than when it follows from a deliberate increase in the damages. In other terms,  $b$  should be larger in the latter case than in the first one. This is captured by assumptions A1 and A2 : under A1,  $b$  is infinite as long as  $\hat{x}$  does not exceed the level of damages and the latter may be chosen in  $[x, \bar{x}]$  and, under A2,  $b$  is finite when  $\hat{x}$  exceeds the level of damages.

<sup>12</sup>More explicitly, at stage 1, the policyholder suggests a loss level  $\hat{x}$  to the insurer (possibly by transmitting qualitative information on damages) but this message remains private : it has no legal value and cannot be put forward in court in case of litigation between the insurer and the insured or between the insurer and the auditor

is weakly dominated by an admissible contract. We assume for simplicity that  $R(x)$  is continuous over  $M = (m, \bar{x}]$

Let  $\delta = \{t(\cdot), M, P\}$  be an admissible contract and let  $\omega(x)$  specify the auditor's net gain as a function of the policyholder's loss. Let  $T(\cdot)$ ,  $R(\cdot) : [0, \bar{x}] \cup \{\phi\} \rightarrow R_+^2$  be functions that specify the indemnity payment  $T(\tilde{x})$  and the payment to the auditor  $R(\tilde{x})$  as functions of the auditor's report  $\tilde{x}$ . We will say that  $\{T(\cdot), R(\cdot)\}$  implements  $\{\delta, \omega(\cdot)\}$  under A1-A2 if

$$t(x) = T(x) \text{ and } \omega(x) = R(x) \text{ if } x \in M \quad (3)$$

$$t_0 = T(\phi) \text{ and } \omega(x) = R(\phi) \text{ if } x \in M^c \quad (4)$$

and if

$$e = 0, \hat{x} = x, \tilde{x} = \hat{x} \text{ if } \hat{x} \in M, \tilde{x} = \phi \text{ if } \hat{x} \in M^c$$

is the unique perfect equilibrium of the audit game.

In words,  $\{T(\cdot), R(\cdot)\}$  implements  $\{\delta, \omega(\cdot)\}$  if at equilibrium there is no audit cost manipulation<sup>13</sup> and the net payments are  $t(x)$  (or  $t_0$ ) and  $\omega(x)$  when the loss is equal to  $x$ . We have :

**Lemma 3** *Under A1-A2, when  $\{T(\cdot), R(\cdot)\}$  implements  $\{\delta, \omega(\cdot)\}$ , the gross payment from the insurer to the auditor satisfies*

$$r = a + \rho(x) - bt(x) + c_a \text{ if } x \in M$$

$$r = r_0 \text{ if } x \in M^c$$

where  $\rho(x) : M \rightarrow \mathbb{R}$  is a nonincreasing function such that  $\rho(m_+) = 0$  and  $a, r_0$  are constants.

Lemma 3 states that a larger level on indemnity paid to the policyholder should imply a lower net income paid to the auditor for audit cost manipulation to be deterred. The intuition behind lemma 3 is as follows. Starting from the truthful revelation of loss without manipulation of audit cost, a deviation

---

<sup>13</sup>In appendix 3, we show that allowing for cost manipulation is a weakly dominated strategy for the insurer.

at stage 1,  $\hat{x} = x' > x$ ,  $e > 0$  would be profitable to the policyholder (assuming  $t(x') > t(x)$ ) if the claim is accepted by the auditor and if  $t(x') - e > t(x)$ . Accepting the claim  $\hat{x} = x'$  is optimal for the auditor if  $R(x') \geq R(x) - be$ . Both conditions are incompatible (for all  $e$ ) if  $R(x') + bt(x') \leq R(x) + bt(x)$ , which shows that a 1\$ increase in the indemnity payment to the policyholder should lead at least to a b\$ decrease in the auditor's earnings for audit cost manipulation to be deterred.

We now are in position to characterize the optimal insurance contract. As before, it involves  $t_0 = 0$  and  $t(x) > 0$  if  $x > m$ . Then, the non-negative profit constraint and the auditor's participation constraint are respectively

$$P \geq F(m)r_0 + \int_{m_+}^{\bar{x}} [a + \rho(x) + (1 - b)t(x) + c]dF(x) \quad (5)$$

$$F(m)V(r_0) + \int_{m_+}^{\bar{x}} V(a + \rho(x) - bt(x))dF(x) \geq \bar{v} \quad (6)$$

At the optimum,  $t(\cdot) : (m, \bar{x}] \rightarrow R_+$ ,  $\rho(\cdot) : (m, \bar{x}] \rightarrow R$ ,  $m$ ,  $a$  and  $r_0$  maximize  $EU$  given by (1) subject to (5), (6),  $0 \leq dt/dx \leq 1$ ,  $t(m_+) = 0$ ,  $d\rho/dx \leq 0$ ,  $\rho(m_+) = 0$ ,  $m \geq 0$ . Characterizing the optimal solution with standard optimal control methods, leads to proposition 3. The optimal insurance policy and the payment to the auditor are illustrated in figure 2.

**Proposition 3** *Under A1-A2, when the auditor is risk-averse, the optimal insurance contract is a deductible contract with coinsurance for high levels of damages :*

$$\begin{aligned} t(x) &= 0 & \text{if } 0 \leq x \leq m \\ t(x) &= x - m & \text{if } m < x \leq x_0 \\ t'(x) &\in (0, 1) & \text{if } x_0 < x \leq \bar{x} \end{aligned}$$

with  $0 \leq m < x_0 \leq \bar{x}$  and  $M = (m, \bar{x}]$ . Furthermore, a sufficient condition for  $m > 0$  is  $c > bf(0)[1 - f(0)]/f(0_+)$ .

Proposition 3 may be interpreted as follows. As stated in lemma 3, the constraint of immunity from audit cost manipulation involves a linkage between the indemnity payment to the policyholder and the transfer to the auditor : larger indemnity payments should entail lower transfers to the auditor for the policyholder to be dissuaded from manipulating audit costs.

More precisely, lemma 3 shows that  $dr/dx$  should be less than  $-bdt/dx$ . In fact, the proof of proposition 3 shows that  $\rho(x) \equiv 0$  at the optimum, which gives

$$\begin{aligned} r &= a - bt + c_a \quad \text{if } t > 0 \\ r &= r_0 \in (a - bt(\bar{x}), a) \quad \text{if } t = 0 \end{aligned} \tag{7}$$

Equation (7) has an attractive cost-sharing interpretation: for any damage larger than  $m$ , the insurer bears the fixed cost  $c = c_a + c_f$  and the variable cost  $t$  is linearly shared between the insurer and the auditor who respectively bear  $a + (1 - b)t$  and  $bt - a$ . Hence, a greater variability in indemnity payment entails a greater variability in the auditor's net income. A risk-averse auditor should then be compensated by a higher risk-premium for his participation constraint to be satisfied. Then, some degree of coinsurance for large levels of damages decreases the risk-premium that should be paid to the auditor, which is ultimately beneficial to the policyholder<sup>14</sup>. Proposition 3 also states that a positive deductible  $m > 0$  remains optimal if the fixed cost per claim  $c$  is large enough. In appendix 1 we give a sufficient condition for  $x_0$  to be strictly lower than  $\bar{x}$ . Intuitively, this condition is satisfied when the auditor is sufficiently risk-averse.

The previous argument on the cost of providing incentives to the auditor does not hold anymore when the auditor is risk-neutral. Then the auditor becomes the insurer and a pure deductible contract remains optimal. When the auditor is infinitely risk-averse or when he is affected by a limited liability constraint, the linkage between the indemnity payments and the transfer to the auditor exerts the most extreme effects and as expressed in proposition 4 an upper bound on coverage is then optimal :

**Proposition 4** *Under A1-A2, when the auditor is infinitely risk-averse or when he is affected by a limited liability constraint, the optimal insurance contract combines a deductible with an upper bound on coverage, with full insurance of marginal damages for intermediate values :*

$$t(x) = \text{Inf}\{\text{Sup}\{0, x - m\}, x_0 - m\}$$

*with  $0 \leq m < x_0 < \bar{x}$  and  $M = (m, \bar{x}]$ . Furthermore, sufficient conditions for  $m > 0$  are  $c > bf(0)[1 - f(0)]/f(0_+)$  or  $f(\cdot)$  nonincreasing over  $(0, \bar{x}]$ .*

---

<sup>14</sup>The optimal shape of function  $t(x)$  may be further characterized under specific assumptions about preferences. For instance, the proof of proposition 3 shows that  $t(x)$  is piecewise linear when  $U(\cdot)$  and  $V(\cdot)$  are quadratic.



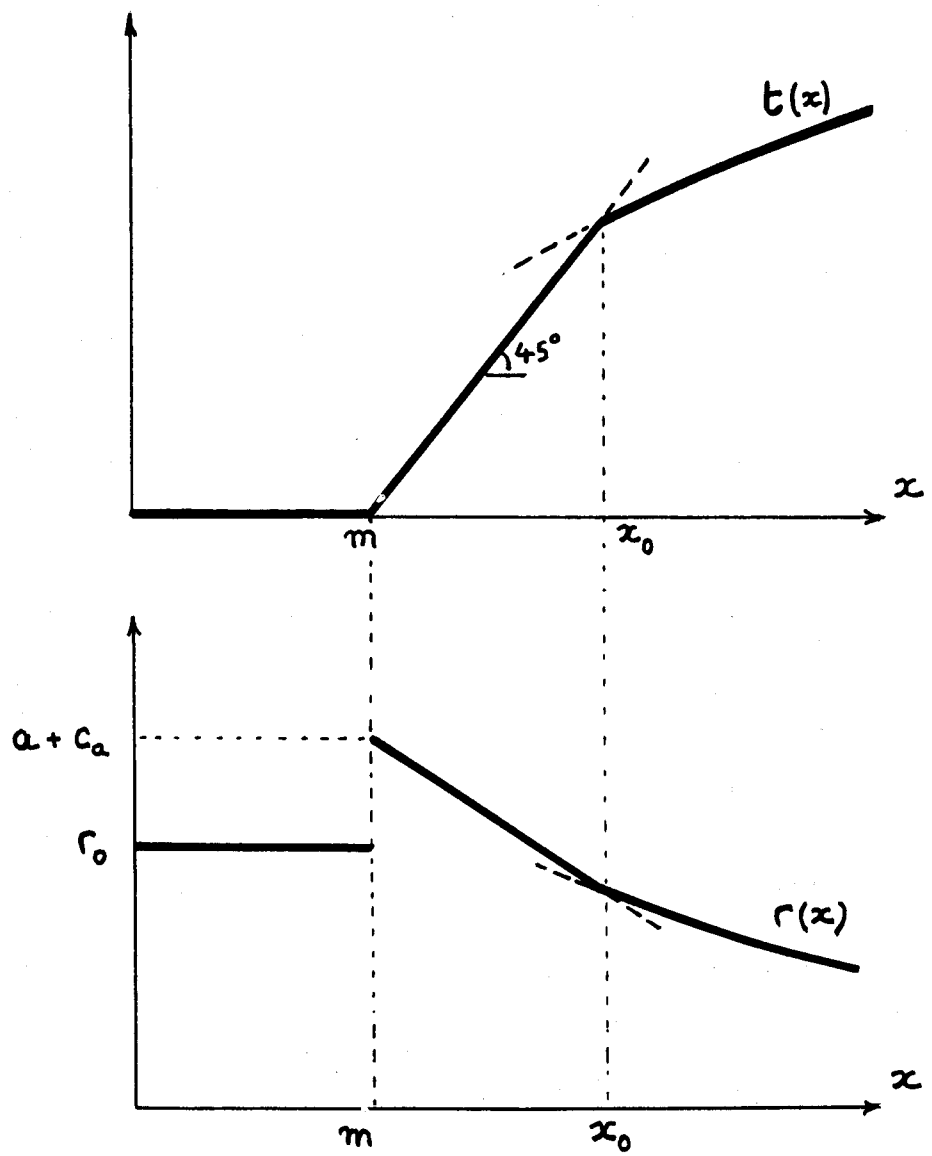


Figure 2

Propositions 3 and 4 suggest that choosing  $m = 0$  may be optimal and appendix 2 shows that this is actually the case under infinite risk aversion, when  $c/b$  is low enough and  $b$  and  $c$  are not too large. We then obtain

$$t(x) = \text{Inf}\{x, x_0\}$$

with  $0 < x_0 < \bar{x}$ , which is the opposite of a pure deductible contract : no deductible and a ceiling on coverage<sup>15</sup>. This result may be intuitively interpreted as follows. We know that full coverage above a deductible  $m$  is optimal when there is no manipulation of audit cost or when the auditor is risk-neutral. In such a case, at  $m = 0$ , the policyholder has perfect insurance and then an increase  $dm > 0$  entails no first-order risk-sharing effect. However, such an increase cuts down the expected administrative costs, which is ultimately beneficial to the policyholder. This gives  $m > 0$  at the optimum. On the contrary, when the auditor is risk-averse and audit costs can be manipulated by the policyholder, some degree of coinsurance is optimal. Then, at  $m = 0$ , increasing the threshold  $m$  for a profile  $\{dt/dx, x \in (m, \bar{x}]\}$  unchanged, entails a negative first-order risk-sharing effect because of underinsurance of large losses. This risk-sharing effect counterbalances the beneficial effect of lower administrative cost and this trade-off may tip more in favor of keeping  $m = 0$ .

## 4 CONCLUSION

Coinurance clauses such as partial insurance for large damages or ceilings on indemnity payments are frequently inserted in insurance policies. The previous literature on optimal insurance contracting has justified such clauses on the basis of assumptions on the insurer's risk-aversion, convexity of the insurance cost functions or bankruptcy rules. Our explanation does not fully depart from the first of these assumptions since it relies on the auditor's risk-aversion : the latter cannot be fully insured because of an incentive constraint. Nevertheless, by focusing on the auditor's activity, this paper has emphasized what we think is a major difficulty for the insurer who struggles

---

<sup>15</sup>Such a result strongly contrasts with the conclusions of Raviv (1979)'s model, in which upper limits on coverage are Pareto suboptimal.

against insurance fraud ; namely the fact that the verification costs can be manipulated by the policyholders, which undermines the credibility of the audit policy<sup>16</sup>. The relationship between the insurers and the agents who are in charge of proving fraud may entail agency costs and it affects the shape of the optimal indemnity schedule as well as the auditor-insurer contract. More particularly, the results of this paper suggest that straight deductibles may be optimal in markets where policyholders can easily be dissuaded from defrauding -for instance because the audit cost is low compared to the fixed administrative cost of claims- but coinsurance clauses should be included in insurance contracts when agency problems limit the efficiency of audit policy.

Of course, it should be acknowledged that these results have been obtained in the framework of a highly simplified model and considering more general settings would allow us to appraise their robustness. In particular, we may consider a generalized formulation for the audit cost activity in which the cost of eliciting verifiable information on claims is a (possibly non-linear) function of the insured's expenditures  $e$  and of the level of damages  $x$ . The issue of audit cost manipulation could also be approached in a model where some degree of damage falsification is inevitable or optimal, as in papers by Crocker and Morgan (1995) or Maggi and Rodriguez-Clare (1995). Finally, the modelling could incorporate stochastic auditing, a credibility constraint on the audit strategy or more general auditor-insurer contracts.

---

<sup>16</sup>Picard (1994) shows that credibility constraints on claims' verification may deeply affect the insurance markets and call for a cooperative action of insurers.

## APPENDIX 1

### Proof of Lemma 1

Let  $\tilde{t}(x)$  be the maximal indemnity obtained by the policyholder when the loss is  $x$  and the contract  $\delta = \{t(\cdot), M, P\}$  is implemented.  $\tilde{t}(x)$  is non-decreasing since the policyholder can misrepresent downward the magnitude of damages.

Let  $t_0 = \tilde{t}(0)$ ,  $m = \text{Sup}\{x \mid \tilde{t}(x) = \tilde{t}(0)\}$ ,  $\tilde{M} = (m, \bar{x}]$  and  $\tilde{P} = P$ . Obviously, the contract  $\tilde{\delta} \equiv \{\tilde{t}(\cdot), \tilde{M}, \tilde{P}\}$  is incentive compatible. Hence  $\tilde{\delta}$  and  $\delta$  yield the same insurance payment.

Let  $\hat{x}(x)$  be an optimal claim of the policyholder under the contract  $\delta$  when he suffers a loss  $x$ .

Let  $x \in \tilde{M}$ . We then have  $\tilde{t}(x) > \tilde{t}(0)$  which gives  $\hat{x}(x) \in M$ , (otherwise  $\hat{x}(x)$  would be a better claim than  $\hat{x}(0)$  when  $x = 0$ ). Audit costs are thus lower under  $\tilde{\delta}$  than under  $\delta$ . ■

### Proof of Proposition 1

Let us delete the non-negativity and monotonicity constraints on  $t(x)$  over  $(m, \bar{x}]$ . We may check that the resulting optimal solution actually verifies these constraints. Assigning a multiplier  $\lambda \geq 0$  to the non-negative profit constraint, the first-order optimality conditions on  $t(\cdot)$  and  $P$  are respectively

$$U'(W - x - P + t(x)) - \lambda(x) = 0 \text{ for all } x \text{ in } (m, \bar{x}] \quad (8)$$

$$\int_0^m U'(W - x - P) dF(x) + \int_{m+}^{\bar{x}} U'(W - x - P + t(x)) dF(x) - \lambda = 0 \quad (9)$$

(8) implies that  $t(x) - x$  is constant over  $(m, \bar{x}]$ . We thus have  $t(x) = x + z - m$  for all  $x$  in  $(m, \bar{x}]$ , where  $z$  is a constant. Furthermore, (8) gives

$$\lambda = U'(W - m - P + z) \quad (10)$$

and (9) yields

$$\lambda = \int_0^m U'(W - x - P) dF(x) + [1 - F(m)] U'(W - m - P + z) \quad (11)$$

Using (10) and (11) we deduce

$$U'(W - m - P + z) = \frac{1}{F(m)} \int_0^m U'(W - x - P) dF(x) \quad (12)$$

There exists  $x_0 \in [0, m]$  such that

$$U'(W - x_0 - P) = \frac{1}{F(m)} \int_0^m U'(W - x - P) dF(x)$$

with  $0 < x_0 < m$  if  $m > 0$ . (12) gives

$$U'(W - m - P + z) = U'(W - x_0 - P)$$

which implies  $z = m - x_0 \geq 0$ ,  $0 < z < m$  if  $m > 0$  and  $z = 0$  if  $m = 0$ .

The optimality condition on  $m$  is

$$\begin{aligned} U(W - m - P)f(m_+) - U(W - m - P + z)f(m_+) + \lambda(z + c)f(m_+) &\leq 0 \\ &= 0 \text{ if } m > 0 \end{aligned} \quad (13)$$

If  $m = 0$ , we have  $z = 0$  and (13) would give  $\lambda cf(0_+) \leq 0$ , hence a contradiction. ■

### Proof of Lemma 2

For any contract  $\delta = \{t(\cdot), M, P\}$ , if the policyholder cannot inflate the size of the claim by creating further damage, we know from Lemma 1 that there exists an incentive compatible contract  $\hat{\delta} = \{\hat{t}(\cdot), \hat{M}, \hat{P}\}$  that weakly dominates  $\delta$ , with  $\hat{t}(x) = t_0$  if  $x \in \hat{M}^c$  and  $\hat{t}(x) > t_0$  if  $x \in \hat{M}$ . Furthermore  $\hat{t}(\cdot)$  is non-decreasing.

Under A1, let  $\tilde{x}(x)$ , be an optimal loss level when the actual initial loss in  $x$  and  $\hat{\delta}$  is implemented. Since  $\hat{\delta}$  is incentive compatible (for given damages),  $\tilde{x}(x)$  is an optimal solution to

$$P(x) : \text{Maximise } \hat{t}(\tilde{x}) - \tilde{x} \text{ with respect to } \tilde{x} \geq x$$

Note that  $\tilde{x}(x) = x$  if  $\tilde{x}(x) \in \hat{M}^c$ . Let  $\tilde{M} = \{x \mid \tilde{x}(x) \in \hat{M}\}$  and  $\tilde{t}(x) \equiv \hat{t}(\tilde{x}(x)) - \tilde{x}(x) + x$  and  $\tilde{P} = P$ .

If  $x \in \tilde{M}^c$ , then  $\tilde{x}(x) \in \hat{M}^c$  which implies  $\tilde{t}(x) = \hat{t}(x) = t_0$ . If  $x \in \tilde{M}$ , then  $\tilde{x}(x) \in \hat{M}$  which implies  $\tilde{t}(x) \geq \hat{t}(x) > t_0$  (by using the fact that  $\tilde{x} = x$  is a feasible solution in problem  $P(x)$ ).

1. Let us show that  $\tilde{\delta} = \{\tilde{t}(\cdot), \tilde{M}, \tilde{P}\}$  is an incentive compatible contract.

We will prove that under  $\tilde{\delta}$  the policyholder cannot profit by creating further damages from  $x$  to  $x' > x$  or by misreporting the damages downward from  $x$  to  $x' < x$ . Let  $\tilde{R}_+(x, x')$  and  $\tilde{R}_-(x, x')$  be respectively the net gains from these strategies and let

$$\tilde{R}(x, x) = \tilde{t}(x) - x = \hat{t}(\tilde{x}(x)) - \tilde{x}(x)$$

If  $x' > x$ , we have :

$$\tilde{R}_+(x, x') = \tilde{t}(x') - x' = \hat{t}(\tilde{x}(x')) - \tilde{x}(x')$$

We have  $\tilde{x}(x') \geq x' > x$ . Hence  $\tilde{x}(x')$  is a feasible solution to  $P(x)$ , which implies

$$\hat{t}(\tilde{x}(x')) - \tilde{x}(x') \leq \hat{t}(\tilde{x}(x)) - \tilde{x}(x) = \tilde{R}(x, x)$$

Hence  $\tilde{R}(x, x) \geq \tilde{R}_+(x, x')$  if  $x' > x$ .

Let  $x, x' \in [0, \bar{x}]$  with  $x < x'$ . Assume  $\tilde{t}(x') < \tilde{t}(x)$ . We then have

$$\hat{t}(\tilde{x}(x')) - \tilde{x}(x') + x' < \hat{t}(\tilde{x}(x)) - \tilde{x}(x) + x \quad (14)$$

which gives

$$\hat{t}(\tilde{x}(x')) - \tilde{x}(x') < \hat{t}(\tilde{x}(x)) - \tilde{x}(x) \quad (15)$$

(15) implies that  $\tilde{x}(x)$  is not a feasible solution to  $P(x')$ , i.e.  $\tilde{x}(x) < x'$ . Furthermore, we have

$$\hat{t}(\tilde{x}(x')) - \tilde{x}(x') + x' \geq \hat{t}(x') \quad (16)$$

since  $\tilde{x} = x'$  is a feasible solution to  $P(x')$ .

We also have

$$\hat{t}(x') \geq \hat{t}(\tilde{x}(x)) \quad (17)$$

since  $\hat{t}(\cdot)$  is non-decreasing. (14), (16) and (17) simultaneously give  $\tilde{x}(x) < x$ , hence a contradiction. We deduce that  $\tilde{t}(x') \geq \tilde{t}(x)$  and consequently that  $\tilde{t}(\cdot)$  is non-decreasing. This gives  $\tilde{R}(x, x) \geq \tilde{R}_-(x, x')$  if  $x' < x$ .

2. The characterization of  $\tilde{\delta}$  may then be completed as follows :

Using  $\tilde{t}(x) = t_0$  if  $x \in \tilde{M}^c$ ,  $\tilde{t}(x) > t_0$  if  $x \in \tilde{M}$  and the fact that  $\tilde{t}(\cdot)$  is non-decreasing gives  $\tilde{M} = (\tilde{m}, \bar{x}]$ . Furthermore,  $\tilde{t}(x) - x = \hat{t}(\tilde{x}(x)) - \tilde{x}(x)$  is non-increasing with respect to  $x$  (by using the definition of  $\tilde{x}(x)$ ). Lastly,

any (upward) discontinuity of  $\tilde{t}(\cdot)$  would contradict the fact that  $\tilde{t}(x) - x$  is non-increasing.

3. *Let us show that  $\tilde{\delta}$  weakly dominates  $\delta$ .*

Under A1,  $\tilde{\delta}$  and  $\hat{\delta}$  lead to the same audit costs (since  $x \in \tilde{M}$  iff  $\tilde{x}(x) \in \widehat{M}$ ) and to the same net gain  $\tilde{t}(x) - x$  for the insured. The payment from the insurer to the insured is equal to  $\tilde{t}(x)$  under  $\tilde{\delta}$  and to  $\hat{t}(\tilde{x}(x))$  under  $\hat{\delta}$ . Using the definition of  $\tilde{t}(x)$  and  $\tilde{x}(x) \geq x$  gives

$$\tilde{t}(x) = \hat{t}(\tilde{x}(x)) - \tilde{x}(x) + x \leq \hat{t}(\tilde{x}(x))$$

We deduce that  $\tilde{\delta}$  weakly dominates  $\hat{\delta}$ . Since  $\hat{\delta}$  weakly dominates  $\delta$ , the result is reached. ■

### **Proof of Proposition 2**

Let  $\mu(x) \equiv dt/dx$ . Under A1, the optimal contract maximizes  $EU$  given by (1) with respect to  $m \geq 0$ ,  $P$  and  $\{t(\cdot), \mu(\cdot)\} : (m, \bar{x}] \rightarrow \mathbb{R}^2$  subject to (2),  $0 \leq \mu(x) \leq 1$  and  $t(m_+) = 0$ .

This is an optimal control problem where  $t(x)$  is a state variable with initial point  $t(m_+) = 0$  and  $\mu(x)$  is a (piecewise continuous) control variable. Both are defined on the interval  $(m, \bar{x}]$ .

For any function  $\nu(x)$  we note  $\dot{\nu} \equiv d\nu/dx$ .

Let  $q(x)$  be the costate variable associated with  $t(x)$  and let  $\lambda$  be a Lagrange multiplier associated with the constraint (2). The Halmiltonian function is

$$H(t, \mu, q, x) = U(W - x - P + t)f(x) + q\mu - \lambda(t + c)f(x)$$

Given  $m$  and  $P$ , the necessary optimality conditions are :

$$\dot{q}(x) = -\frac{\partial H}{\partial t} = -f(x)U'(W - x - P + t(x)) + \lambda f(x) \quad (18)$$

for all  $x$  in  $(m, \bar{x}]$

$$q(\bar{x}) = 0 \quad (19)$$

$$\left. \begin{aligned} \frac{\partial H}{\partial \mu} = q(x) & \begin{cases} \geq 0 & \text{if } \mu(x) = 1 \\ = 0 & \text{if } 0 < \mu(x) < 1 \\ \leq 0 & \text{if } \mu(x) = 0 \end{cases} \end{aligned} \right\} \quad (20)$$

Furthermore optimality conditions on  $m$  and  $P$  are respectively <sup>17</sup> :

$$\left. \begin{aligned} U(W - m - P)f(m) - H(m, \mu(m), q(m), m) &= 0 \\ &\text{if } m > 0 \\ U(W - P)f(0_+) - H(0_+, \mu(0_+), q(0_+), 0_+) &\leq 0 \\ &\text{if } m = 0 \end{aligned} \right\} \quad (21)$$

$$\lambda = \int_0^m U'(W - x - P)dF(x) + \int_{m_+}^{\bar{x}} U'(W - x - P + t(x))dF(x) \quad (22)$$

Using (21) and (22) and  $t(m_+) = 0$  gives

$$q(m_+)\mu(m_+) \geq \lambda cf(m_+) > 0$$

which implies  $q(m_+) > 0$ . (18) gives

$$\dot{q}(x) = g(x)f(x) \quad \text{for all } x \text{ in } (m, \bar{x}] \quad (23)$$

with  $g(x) \equiv \lambda - U'(W - x - P + t(x))$

Let  $\Delta_0 = \{x \in (m, \bar{x}] \text{ s.t. } q(x) > 0\}$ , with  $m + \varepsilon \in \Delta_0$  for  $\varepsilon$  positive and small. (20) implies that  $\mu(x) = 1$  if  $x \in \Delta_0$ . Consequently,  $g(x)$  remains constant over  $\Delta_0$  and we may note  $g(x) = \bar{g}$  for all  $x$  in  $\Delta_0$ .

Assume  $\bar{g} \geq 0$ . Then  $\dot{q}(x) = \bar{g}f(x) \geq 0$  if  $x \in \Delta_0$ . Consequently  $[x, \bar{x}] \subset \Delta_0$  if  $x \in \Delta_0$ , which contradicts (19). We thus have  $\bar{g} < 0$ .

Let  $x_0 = \inf\{x \in (m, \bar{x}] \text{ s.t. } q(x) = 0\}$ . Assume  $x_0 < \bar{x}$ . We have  $g(x_0) = \bar{g}$  since function  $g(x)$  is continuous and  $g(x) = \bar{g}$  for  $m < x < x_0$ . (23) then gives  $\dot{q}(x_0) = f(x_0)\bar{g} < 0$ , which implies that  $q(x_0 + \varepsilon) < 0$  for  $\varepsilon > 0$ .

Let  $\Delta_1 = \{x \in (m, \bar{x}] \text{ s.t. } q(x) < 0\}$ . (20) implies that  $\mu(x) = 0$  if  $x \in \Delta_1$ . Hence

$$\dot{q}(x) = U''(W - x - P + t(x)) < 0 \text{ if } x \in \Delta_1$$

We deduce that  $(x_0 + \varepsilon, \bar{x}] \subset \Delta_1$  which contradicts (19). We thus have  $\Delta_0 = (m, \bar{x}]$ . Using  $t(m_+) = 0$  then gives  $t(x) = x - m$  if  $x \geq m$ .

Finally, assume  $m = 0$ . We then have  $t(x) = x$  for all  $x$  in  $[0, \bar{x}]$ . Using (22) then gives  $\lambda = U'(W - P)$  and  $g(x) = 0$  for all  $x$  in  $[0, \bar{x}]$ . Thus  $q(m_+) = q(\bar{x}) = 0$ , which contradicts  $q(m_+) > 0$ . ■

<sup>17</sup>See Seierstad and Sydsaeter (1987), theorem 9, chapter 3 or Beavis and Dobbs (1990), theorem 7.6.



### Proof of Lemma 3

Given  $\{\delta, \omega(\cdot)\}$  and  $\{T(\cdot), R(\cdot)\}$  such that (3) and (4) hold, let us derive the conditions under which starting from  $e = 0$ ,  $\hat{x} = x$ , there does not exist any deviation which is profitable to the policyholder, given the reaction function of the auditor.

Obviously, for an admissible contract, only deviations such that  $\hat{x} \in M$  can be profitable to the policyholder. Assume first that the deviation  $\hat{x}$  is accepted by the auditor (i.e.  $\tilde{x} = \hat{x}$ ). Profitability of the deviation requires  $t(\hat{x}) > t(x)$ . Hence, let us consider a deviation  $\hat{x} = x' \neq x, x' \in M$ , with  $t(x') > t(x)$ ,  $e = e' \geq 0$ . Then, at stage 2, after auditing, the auditor may report  $\tilde{x} = x$  after gathering verifiable information which gives him a net gain  $R(x) - be'$ . He may also accept the policyholder's claim and then his net gain is  $R(x')$ . It is an optimal strategy for the auditor to accept the claim  $\hat{x} = x'$  if and only if  $be' \geq R(x) - R(x')$ . The upper bound of the policyholder's net gain induced by the deviation is obtained at

$$e' = \text{Sup}\{0, \frac{R(x) - R(x')}{b}\}$$

Hence, the deviation cannot be profitable if and only if

$$t(x) \geq t(x') - \text{Sup}\{0, \frac{R(x) - R(x')}{b}\}$$

or equivalently

$$R(x) + bt(x) \geq R(x') + bt(x')$$

This condition should hold for all  $x$  in  $[0, \bar{x}]$  and all  $x'$  in  $M$  such that  $t(x') > t(x)$ . Since  $t_0 = t(m_+) < t(x)$  if  $x \in M$ , we obtain :  $R(x) \geq R(m_+)$  if  $x \in M^c$ .

Let  $M = \bigcup_{i=1}^n (a_i, a_{i+1}]$ , with  $m = a_1 < a_2 \dots < a_{n+1} = \bar{x}$ , such that  $t(x)$  is increasing (respect. constant) over  $(a_i, a_{i+1}]$  if  $i$  is odd (respect. even). Let  $i$  be an odd number in  $\{1, \dots, n\}$ . Then, from the previous condition, we deduce that  $R(x) + bt(x)$  is nonincreasing over  $(a_i, a_{i+1}]$ . Let  $x^* \in (a_{i+1}, a_{i+2}]$ . Assume  $R(x^*) + bt(x^*) > R(a_{i+1}) + bt(a_{i+1})$ . Then  $R(x^*) + bt(x^*) > R(a_{i+1} - \varepsilon) + bt(a_{i+1} - \varepsilon)$  for  $\varepsilon$  positive, small enough and such that  $a_i < a_{i+1} - \varepsilon < a_{i+1}$ , which contradicts  $t(x) > t(a_{i+1} - \varepsilon)$ . We deduce that  $R(x) + bt(x)$  is nonincreasing over  $M = (m, \bar{x}]$ .

Assume now that the deviation  $\hat{x}$  is not accepted by the auditor (i.e.  $\tilde{x} = x$ ). It is profitable if and only if  $x \in M^c$  and  $T(x) > T(\phi) = t_0$ . Such profitable deviations do not exist if  $T(x) \leq t_0$  if  $x \in M^c$ .

In short, under A1-A2,  $\{T(\cdot), R(\cdot)\}$  implements  $\{\delta, \omega(\cdot)\}$ , if and only if (3) and (4) hold and

(i)  $R(x) + bt(x)$  is nonincreasing over  $M$

(ii)  $R(x) \geq R(m_+)$  and  $T(x) \leq t_0$  if  $x \in M^c$ .

Lemma 3 is obtained with  $r_0 = R(\phi)$  and  $a = R(m_+) + bt_0$ . ■

### Proof of Proposition 3

Letting  $\mu(x) = \dot{t}(x)$  and  $\nu(x) = \dot{\rho}(x)$ , the optimal solution maximizes  $EU$  given by (1) with respect to  $m \geq 0$ ,  $a, r_0, P, \{t(\cdot), \mu(\cdot), \rho(\cdot), x(\cdot)\} : (m, \bar{x}] \rightarrow R^4$  subject to (5), (6) and  $0 \leq \mu(x) \leq 1$ ,  $\nu(x) \leq 0$ ,  $t(m_+) = 0$ ,  $\rho(m_+) = 0$ .

Letting  $q_1(x)$  and  $q_2(x)$  be costate variables respectively associated with  $t(x)$  and  $\rho(x)$  and letting  $\lambda_1$  and  $\lambda_2$  denote the Lagrange multipliers that are respectively associated with (5) and (6), the Hamiltonian function is

$$H(t, \mu, \nu, q_1, q_2, x) = U(W - x - P + t)f(x) + q_1\mu + q_2\nu - \lambda_1[a + \rho + (1 - b)t + c]f(x) + \lambda_2V(a + \rho - bt)f(x)$$

with  $x \in (m, \bar{x}]$ .

Given  $m, P, a$  and  $r_0$ , the necessary optimality conditions are :

$$\dot{q}_1(x) = -\frac{\partial H}{\partial t} = -f(x)U'(W - x - P + t(x)) + \lambda_1(1 - b)f(x) + \lambda_2bV'(a + \rho(x) - bt(x))f(x) \quad (24)$$

$$\dot{q}_2(x) = -\frac{\partial H}{\partial \rho} = \lambda_1f(x) - \lambda_2V'(a + \rho(x) - bt(x))f(x) \quad (25)$$

$$q_1(\bar{x}) = 0 \quad (26)$$

$$q_2(\bar{x}) = 0 \quad (27)$$

$$\frac{\partial H}{\partial \mu} = q_1(x) = \begin{cases} \geq 0 & \text{if } \mu(x) = 1 \\ = 0 & \text{if } 0 < \mu(x) < 1 \\ \leq 0 & \text{if } \mu(x) = 0 \end{cases} \quad (28)$$

$$\frac{\partial H}{\partial \nu} = q_2(x) = \begin{cases} \geq 0 & \text{if } \nu(x) = 0 \\ = 0 & \text{if } \nu(x) < 0 \end{cases} \quad (29)$$

Furthermore, the optimality conditions on  $m$ ,  $P$ ,  $a$  and  $r_0$  can be respectively written as follows :

$$\left. \begin{aligned} & [U(W - m - P) + \lambda_2 V(r_0) - \lambda_1 r_0]f(m) \\ & - H(t(m), \mu(m), \rho(m), \nu(m), q_1(m), q_2(m), m) = 0 \quad \text{if } m > 0 \\ & [U(W - P) + \lambda_2 V(r_0) - \lambda_1 r_0]f(0_+) \\ & - H(t(0_+), \mu(0_+), \rho(0_+), \nu(0_+), q_1(0_+), q_2(0_+), 0_+) \leq 0 \quad \text{if } m = 0 \end{aligned} \right\} \quad (30)$$

$$\lambda_1 = \int_0^m U'(W - x - P)dF(x) + \int_{m_+}^{\bar{x}} U'(W - x - P + t(x))dF(x) \quad (31)$$

$$\lambda_2 V'(r_0)F(m) - \lambda_1 F(m) = 0 \quad (32)$$

$$\lambda_2 \int_{m_+}^{\bar{x}} V'(a + \rho(x) - bt(x))dF(x) - \lambda_1 [1 - F(m)] = 0 \quad (33)$$

and  $\lambda_1 \geq 0$ ,  $\lambda_2 \geq 0$ .

1. *Let us first show that  $\rho(x) \equiv 0$  at the optimum. Suppose that  $\nu(x) < 0$  in an interval  $[a, b] \subset (m, \bar{x}]$ ,  $b > a$ . Then from (29) we have  $q_2(x) = 0$  over  $[a, b]$ . Using (25), we deduce that  $a + \rho(x) - bt(x)$  is constant over  $[a, b]$  which gives  $\nu(x) = \dot{\rho}(x) = 0$ , hence a contradiction. We thus have  $\nu(x) \equiv 0$  which gives  $\rho(x) \equiv 0$ .*
2. *Let us establish that  $\mu(x) = 1$  if  $x \in (m, x_0]$  and  $0 < \mu(x) < 1$  if  $x \in (x_0, \bar{x}]$ , with  $m < x_0 \leq \bar{x}$ .*

Using (29) and  $t(m_+) = 0$  gives

$$q_1(m_+)\mu(m_+) \geq [\lambda_2(V(r_0) - V(a)) - \lambda_1(r_0 - a) + \lambda_1 c]f(m_+)$$

(32) and concavity of  $V(\cdot)$  then yield

$$\begin{aligned} q_1(m_+)\mu(m_+) & \geq [\lambda_2((a - r_0)V'(r_0) - V(a) + V(r_0)) + \lambda_1 c]f(m_+) \\ & \geq \lambda_1 c f(m_+) > 0 \end{aligned}$$

which gives  $q_1(m_+) > 0$ .

Let  $\dot{q}_1(x) = h(x)f(x)$  for all  $x \in (m, \bar{x}]$  with

$$h(x) = \lambda_1(1 - b) + \lambda_2 b V'(a - bt(x)) - U'(W - x - P + t(x))$$

Let  $\Delta_0 = \{x \in (m, \bar{x}] \text{ s.t. } q_1(x) > 0\}$ , with  $m + \varepsilon \in \Delta_0$  for  $\varepsilon$  positive and small. We have from (28) :

$$\mu(x) = 1 \text{ and } \dot{h}(x) = -\lambda_2 b^2 V''(a - bt(x)) > 0 \text{ for all } x \in \Delta_0$$

$q_1(x)$  has no local maximum in  $\Delta_0$  (since  $\frac{d\dot{q}_1}{dx}(x) = \dot{h}(x)f(x) > 0$  if  $\dot{q}_1(x) = 0$  and  $x \in \Delta_0$ ). Using  $q_1(\bar{x}) = 0$  then implies  $\dot{q}_1(x) < 0$  over  $\Delta_0$ . We thus have  $\Delta_0 = (m, x_0]$  with  $x_0 \in (m, \bar{x}]$ .

Let  $\Delta_1 = \{x \in (m, \bar{x}] \text{ s.t. } q_1(x) < 0\}$ . Assume  $\Delta_1 \neq \emptyset$ . Then there exists  $x' \in \Delta_1$  such that  $\dot{q}_1(x') < 0$ . We have from (28) :

$\mu(x) = 0$  and  $\dot{h}(x) = U''(W - x - P + t(x)) < 0$  for all  $x \in \Delta_1$ .  $q_1(x)$  has no local minimum in  $\Delta_1$  (since  $\frac{d\dot{q}_1}{dx}(x) = \dot{h}(x)f(x) < 0$  if  $\dot{q}_1(x) = 0$  and  $x \in \Delta_1$ ). Hence  $[x', \bar{x}] \subset \Delta_1$  which contradicts  $q_1(\bar{x}) = 0$ .

Hence  $\Delta_1 = \emptyset$  and  $q_1(x) = 0$  for all  $x$  in  $[x_0, \bar{x}]$ , which gives

$$U'(W - x - P + t(x)) - \lambda_2 b V'(a - bt(x)) = \lambda_1(1 - b)$$

Hence

$$\mu(x) = \dot{t}(x) = \frac{U''(W - x - P + t(x))}{U''(W - x - P + t(x)) + \lambda_2 b V''(a - bt(x))} \in (0, 1)$$

if  $x \in (x_0, \bar{x}]$ .

3 Finally, we will find a sufficient condition for  $m > 0$ . Integrating (24) and using (26), (32) and (33), give

$$\begin{aligned} q_1(m_+) = & -[1 - F(m) + bF(m)] \int_0^m U'(W - P - x) dF(x) \\ & + (1 - b)F(m)[(F(x_0) - F(m))U'(W - P - m) \\ & + \int_{x_0}^{\bar{x}} U'(W - P - x + t(x)) dF(x)] + \lambda_2 b V'(r_0)F(m) \end{aligned} \quad (34)$$

Furthermore, we have  $\dot{q}_1(x) = 0$  for all  $x \in [x_0, \bar{x}]$ , which gives

$$\begin{aligned} \int_{x_0}^{\bar{x}} U'(W - P - x + t(x))dF(x) &= (1 - b)[1 - F(x_0)]\lambda_1 \\ &+ \lambda_2 b \int_{x_0}^{\bar{x}} V'(a - bt(x))dF(x) \end{aligned} \quad (35)$$

Using (32), (33), (34) and (35) then yields

$$\begin{aligned} q_1(m_+) &= \frac{[F(x_0) - F(m)]F(m)}{F(x_0)} [U'(W - P - m) - \frac{1}{F(m)} \int_0^m U'(W - P - x)dF(x) \\ &+ \lambda_2 b V'(r_0) - \frac{\lambda_2 b}{F(x_0) - F(m)} \int_{m_+}^{x_0} V'(a - bt(x))dF(x)] \end{aligned} \quad (36)$$

Assume  $m = 0$ . We would then have

$$q_1(0_+) = \frac{\lambda_2 b f(0)[F(x_0) - f(0)]}{F(x_0)} [V'(r_0) - \frac{\int_0^{x_0} V'(a - bx)dF(x)}{F(x_0) - f(0)}]$$

which implies -using (32)-

$$q_1(0_+) < \frac{\lambda_2 b f(0)[F(x_0) - f(0)]V'(r_0)}{F(x_0)} = \frac{\lambda_1 b f(0)[F(x_0) - f(0)]}{F(x_0)}$$

This contradicts  $q_1(0_+) \geq \lambda_1 c f(0_+)$  if  $c > b f(0)[1 - f(0)]/f(0_+)$  ■

*A sufficient condition for  $x_0 < \bar{x}$  is obtained as follows.* Let  $\bar{m} \in (0, \bar{x})$ . Let us first show that at the optimum  $m \in [0, \bar{m}]$  if  $b$  and  $c$  are not too large.

Let  $\Phi(m, b, c)$  be the optimal expected utility level, written as a function of  $m$  (considered here as a parameter) and of  $b$  and  $c$ . The optimal solution entails full insurance when  $b = c = 0$ . We also have

$$\frac{\partial \Phi}{\partial m}(m, 0, 0) < 0$$

which gives

$$\Phi(m, 0, 0) < \Phi(\bar{m}, 0, 0) = \Phi(0, 0, 0) - \varepsilon$$

for all  $m > \bar{m}$ , with  $\varepsilon > 0$ .

From the Maximum Theorem,  $\Phi(m, b, c)$  is a continuous function. Hence, there exist  $m' > 0$  and a neighbourhood  $\nu \subset R^2$  of  $(0, 0)$  such that

$$\Phi(m, b, c) > \Phi(0, 0, 0) - \varepsilon$$

if  $0 \leq m \leq m'$  and  $(b, c) \in \nu$ .

Furthermore, we have

$$\Phi(m, b, c) \leq \Phi(m, 0, 0) \quad \text{for all } m \geq 0 \text{ and for all } (b, c) \in R_+^2$$

Assume  $m > \bar{m}$  and  $(b, c) \in \nu^+ \cap R_+^2$ . Then, for all  $m'' \in [0, m']$ , we have:

$$\Phi(m'', b, c) > \Phi(m, b, c)$$

Hence, the optimal threshold  $m$  belongs to the interval  $[0, \bar{m}]$  if  $(b, c) \in \nu^+$ .

Assume  $m > \bar{m}$  and  $(b, c) \in \nu^+ = \nu \cap R_+^2$ . Then for  $0 \leq m'' \leq m'$ , we have  $\Phi(m'', b, c) > \Phi(m, b, c)$ . We deduce that the optimal threshold  $m$  belongs to the interval  $[0, \bar{m}]$  if  $(b, c) \in \nu^+$ .

Assume  $x_0 = \bar{x}$ . We would then have

$$\dot{q}(\bar{x}) = f(\bar{x})[\lambda_2 b V'(a - b(\bar{x} - m)) + \lambda_1(1 - b) - U'(W - P - m)] \leq 0$$

Using (32) and (33), this may be written as

$$b \left[ \frac{V'(a - b(\bar{x} - m))[1 - F(m)]}{\int_{m+}^{\bar{x}} V'(a - b(x - m))dF(x)} - 1 \right] \leq \frac{U'(W - P - m) - \lambda_1}{\lambda_1} \quad (37)$$

Assume that  $V(\cdot)$  and  $U(\cdot)$  are CARA utility functions. Let  $\Gamma_1(m, b)$  and  $\Gamma_2(m)$  be respectively the left-hand side and the right-hand side of (38), with  $\partial\Gamma_1/\partial m < 0$  and  $\partial\Gamma_2/\partial m > 0$ . Note that we use (31),  $t(x) = x - m$  for all  $x$  in  $[m, \bar{x}]$  and the fact that  $U(\cdot)$  is CARA to write  $\Gamma_2$  as a function of  $m$  only. Likewise  $\Gamma_1$  does not depend on  $a$  when  $V(\cdot)$  is CARA.

Assume  $(b, c) \in \nu^+$ . Then  $\Gamma_1(m, b) > \Gamma_1(\bar{m}, b)$  and  $\Gamma_2(m) < \Gamma_2(\bar{m})$ . Consequently, a sufficient condition for (37) not to hold is  $\Gamma_1(\bar{m}, b) > \Gamma_2(\bar{m})$ , which gives  $x_0 < \bar{x}$  in that case. Given  $\bar{m}$  and  $b$ , this will be satisfied when the auditor's (respect. policyholder's) absolute risk aversion is large (respect. low) enough.

#### Proof of Proposition 4

Under infinite risk-aversion or under a limited liability constraint, the constraints

$$a + \rho(x) - bt(x) \geq \bar{v} \quad \text{for all } x \in (m, \bar{x}]$$

(with  $\bar{v}$  an exogenously given constant) should be substituted to (6). Since  $t(\cdot)$  is nondecreasing and  $\rho(x)$  is nonincreasing, this is equivalent to

$$a + \rho(\bar{x}) - bt(\bar{x}) \geq \bar{v} \quad \text{and} \quad r_0 \geq \bar{v} \quad (38)$$

(5) and (38) are equivalent to

$$P \geq F(m)\bar{v} + \int_{m_+}^{\bar{x}} [\bar{v} - \rho(\bar{x}) + bt(\bar{x}) + \rho(x) + (1-b)t(x) + c]dF(x)$$

Noting that  $\rho(x) = \rho(\bar{x}) = 0$  is optimal, we obtain

$$P \geq \bar{v} + [1 - F(m)]bt(\bar{x}) + \int_{m_+}^{\bar{x}} [(1-b)t(x) + c]dF(x) \quad (39)$$

At the optimum,  $t(\cdot) : (m, \bar{x}] \rightarrow \mathbf{R}_+$ ,  $P$  and  $m \geq 0$  maximize  $EU$  given by (1) subject to (39),  $0 \leq dt/dx \leq 1$ , and  $t(m_+) = 0$ .

Letting  $q(x)$  be the costate variable associated with  $t(x)$  and letting  $\lambda$  denote the Lagrange multiplier associated with (39), the Hamiltonian function is

$$\begin{aligned} H(t, \mu, q, x) = & U(W - x - P + t)f(x) + q\mu \\ & - \lambda[(1-b)t + c]f(x) \end{aligned}$$

We have

$$\dot{q}(x) = -\frac{\partial H}{\partial t} = -f(x)U'(W - x - P + t(x)) + \lambda(1-b)f(x) \quad (40)$$

$$q(\bar{x}) + \lambda b[1 - F(m)] = 0 \quad (41)$$

and (20), (21) and (22), which implies  $q(m_+) \geq \lambda[c + bt(\bar{x})]f(m_+) > 0$ .

1. Let us establish that  $\mu(x) = 1$  if  $x \in (m, x_0]$  and  $\mu(x) = 0$  if  $x \in (x_0, \bar{x}]$ , with  $m < x_0 < \bar{x}$ .

Let  $\dot{q}(x) = g(x)f(x)$  for all  $x \in (m, \bar{x}]$  with

$$g(x) = \lambda(1 - b) - U'(W - x - P + t(x))$$

Let  $\Delta_0 = \{x \in (m, \bar{x}] \text{ s.t. } q(x) > 0\}$  with  $m + \varepsilon \in \Delta_0$  for  $\varepsilon$  positive and small. For all  $x \in \Delta_0$ , we have  $\mu(x) = 1$ . Consequently  $g(x)$  remains constant over  $\Delta_0$  and we may note  $g(x) = \bar{g}$  for all  $x$  in  $\Delta_0$ .

The function  $q(x)$  is continuously differentiable over  $(m, \bar{x}]$  with  $q(\bar{x}) = -\lambda b[1 - F(m)] < 0$ . Hence  $\bar{g} < 0$ . We deduce that  $\Delta_0 = (m, x_0)$ , with  $m < x_0 < \bar{x}$ . Let  $\Delta_1 = \{x \in (m, \bar{x}] \text{ s.t. } q(x) < 0\}$ . For all  $x \in \Delta_1$  we have  $\mu(x) = 0$  and  $\dot{q}(x) = U''(W - x - P + t(x)) < 0$ . Hence,  $q(x)$  has no local minimum in  $\Delta_1$ . Furthermore  $x_0 + \varepsilon \in \Delta_1$  for  $\varepsilon$  small enough (since  $\dot{q}(x_0) = g < 0$  and  $q(x_0) = 0$ ). Using  $q(\bar{x}) < 0$ , we deduce that  $\Delta_1 = (x_0, \bar{x}]$ .

2. Let us find a sufficient condition for  $m > 0$ . Integrating (40) and using (41) gives

$$q(m_+) = \left. \begin{aligned} & -[1 - F(m)] \int_0^m U'(W - P - x) dF(x) \\ & + F(m) \int_{m^+}^{\bar{x}} U'(W - P - x + t(x)) dF(x) \end{aligned} \right\} \quad (42)$$

Furthermore,  $q(x_0) = 0$  gives

$$-\lambda b[1 - F(m)] = (\bar{x}) - q(x_0) = \int_{x_0}^{\bar{x}} [\lambda(1 - b) - U'(W - x - P + x_0 - m)] dF(x) \quad (43)$$

Using (22) and (43) gives

$$\begin{aligned} & \int_{x_0}^{\bar{x}} U'(W - x - P + x_0 - m) dF(x) \\ &= \lambda[b(1 - F(m)) + (1 - b)(1 - F(x_0))] \\ &= [b(1 - F(m)) + (1 - b)(1 - F(x_0))] \left[ \int_0^m U'(W - x - P) dF(x) \right. \\ & \quad \left. + (F(x_0) - F(m))U'(W - P - m) \right. \\ & \quad \left. + \int_{x_0}^{\bar{x}} U'(W - x - P + x_0 - m) dF(x) \right] \end{aligned}$$



which yields

$$\int_{x_0}^{\bar{x}} U'(W - x - P + x_0 - m) dF(x) = \frac{b(1 - F(m)) + (1 - b)(1 - F(x_0))}{F(x_0) - b[F(x_0) - F(m)]} \times \\ [\int_0^m U'(W - P - x) dF(x) + (F(x_0) - F(m))U'(W - P - m)] \quad (44)$$

Then using (42) and (44) gives

$$q(m_+) = \frac{[F(x_0) - F(m)][F(m)U'(W - P - m) - (1 - b) \int_0^m U'(W - P - x) dF(x)]}{F(x_0) - b[F(x_0) - F(m)]}$$

Assume  $m = 0$  ; we would then have

$$q(0_+) = \frac{[F(x_0) - f(0)]bf(0)U'(W - P)}{F(x_0) - b[F(x_0) - f(0)]}$$

Furthermore, using (22) and (43) yields

$$U'(W - P) = \frac{\lambda[F(x_0) - b(F(x_0) - f(0))]}{F(x_0)}$$

We deduce

$$q(0_+) = \frac{[F(x_0) - f(0)]bf(0)\lambda}{F(x_0)}$$

which contradicts  $q(0_+) \geq \lambda[c + bt(\bar{x})]f(0_+) = \lambda(c + bx_0)f(0_+)$  if  $c > bf(0)[1 - f(0)]/f(0_+)$  or if  $f(\cdot)$  is nonincreasing over  $(0, \bar{x}]$ . ■

## APPENDIX 2

This appendix provides sufficient conditions for  $m = 0$  to be optimal when the audit cost may be manipulated by the policyholder and the auditor is infinitely risk-averse. Given proposition 4, the optimal insurance contract is characterized by  $P$ ,  $m \geq 0$  and  $x_0 < \bar{x}$ , with

$$EU = \int_0^m U(W - P - x) dF(x) + [F(x_0) - F(m)]U(W - P - m) \\ + \int_{x_0}^{\bar{x}} U(W - P - m - x + x_0) dF(x)$$

Using (39), the non-negative profit constraint may be written as

$$P \geq \bar{v} + b[1 - F(m)](x_0 - m) + \int_{m_+}^{x_0} [(1 - b)(x - m)] dF(x) \\ + (1 - b)[1 - F(x_0)](x_0 - m) + c[1 - F(m)]$$

Let  $\lambda$  denote the Lagrange multiplier associated with the constraint. Let us consider  $m$  as a given parameter. The first-order optimality conditions on  $x_0$  and  $P$  are respectively

$$\int_{x_0}^{\bar{x}} U'(W - P - m - x + x_0) dF(x) - \lambda[(1 - F(m))b + (1 - F(x_0))(1 - b)] = 0 \quad (45)$$

$$\int_0^m U'(W - P - x) dF(x) + [F(x_0) - F(m)]U'(W - P - m) \\ + \int_{x_0}^{\bar{x}} U'(W - P - m - x + x_0) dF(x) + \lambda = 0 \quad (46)$$

Let  $\Phi(m)$  be the optimal expected utility level, written as a function of  $m$ . The envelope theorem gives

$$\frac{\partial \Phi}{\partial m} = -[F(x_0) - F(m)]U'(W - P - m) - \int_{x_0}^{\bar{x}} U'(W - P - m - x + x_0) dF(x) \\ + \lambda[f(m_+)(c + b(x_0 - m)) + 1 - F(m)]$$

Using (45) then yields

$$\frac{\partial \Phi}{\partial m} = [F(x_0) - F(m)][\lambda(1 - b) - U'(W - P - m)] + \lambda f(m_+)[c + b(x_0 - m)]$$

At  $m = 0$ , (46) gives

$$\lambda = F(x_0)U'(W - P) + \int_{x_0}^{\bar{x}} U'(W - P - x + x_0)dx$$

and (45) implies

$$U'(W - P)F(x_0) = \lambda[b f(0) + (1 - b)F(x_0)] \quad \text{if } m = 0$$

we deduce that :

$$\frac{\partial \Phi}{\partial m} \Big|_{m=0} = -\lambda \frac{b f(0)[F(x_0) - f(0)]}{F(x_0)} + \lambda f(0_+)(c + b x_0)$$

and  $\frac{\partial \Phi}{\partial m} \Big|_{m=0} < 0$  if

$$\frac{c}{b} < \frac{f(0)[F(x_0) - f(0)]}{F(x_0)f(0_+)} - x_0 \quad (47)$$

where  $x_0 \in (0, \bar{x})$  corresponds to the optimal solution when  $m = 0$ . When (47) holds, there exists  $\bar{m} \in (0, \bar{x})$  such that  $\Phi(m) < \Phi(0)$  for all  $m$  in  $[0, \bar{m}]$ . Furthermore, an argument developed in appendix 1 (after the proof of proposition 3) shows that the optimal threshold  $m$  belongs to  $[0, \bar{m}]$ , when  $(b, c) \in \nu^+ = \nu \cap R_+^2$ , where  $\nu$  is neighbourhood of  $(0, 0)$ . Hence (47) and  $(b, c) \in \nu^+$  are sufficient conditions for  $m = 0$  to be optimal.

## APPENDIX 3

In section 3, we have restricted attention to implementation schemes in which the policyholder does not manipulate the audit cost at equilibrium. A more general definition of implementation is as follows.

Let  $\delta = \{t(\cdot), M, P\}$  be an admissible contract and let  $\omega(\cdot) : [0, \bar{x}] \rightarrow \mathbb{R}$ . Consider the three-stage game presented in section 3, characterized by  $T(\cdot) : [0, \bar{x}] \cup \{\phi\} \rightarrow \mathbb{R}$  and  $R(\cdot) : [0, \bar{x}] \cup \{\phi\} \rightarrow \mathbb{R}$  but where claims are audited if  $\hat{x} \in M' \subset [0, \bar{x}]$ . The set  $M'$  may differ from  $M$ . Let  $e(x) \in \mathbb{R}_+$ ,  $\hat{x}(x) \in [0, \bar{x}]$ ,  $\tilde{x}(x) \in [0, \bar{x}] \cup \{\phi\}$  be strategies played at a perfect equilibrium of this game.

**Definition 1**  $\{T(\cdot), R(\cdot)\}$  implements  $\{\delta, \omega(\cdot)\}$  with cost manipulation if :

- (i) For all  $x \in [0, \bar{x}]$ ,  $t(x) = T[\tilde{x}(x)]$ ,  $\omega(x) = R(\tilde{x}(x)) - be(x)$  if  $\tilde{x}(x) \neq \hat{x}(x)$  and  $\omega(x) = R(\tilde{x}(x))$  if  $\tilde{x}(x) = \hat{x}(x)$
- (ii)  $\tilde{x}(x) \in [0, \bar{x}]$  -i.e.  $\hat{x}(x) \in M'$  - if  $x \in M$  and  $\tilde{x}(x) = \phi$  -i.e.  $\hat{x}(x) \in (M')^c$  - if  $x \in M^c$ .

Thus, if  $\{T(\cdot), R(\cdot)\}$  implements  $\{\delta, \omega(\cdot)\}$  with cost manipulation, then the payment from the insurer to the policyholder is  $t(x)$  and the equilibrium net gain of the policyholder is  $t(x) - e(x)$ . The net payment from the insurer to the auditor (after deduction of the standard audit cost  $c_a$ ) is  $\omega(x) + be(x)$ . Furthermore, the claim is audited if and only if  $x \in M$ .

Without loss of generality, we assume that, at equilibrium, the claim is always accepted by the auditor, i.e.  $\hat{x}(x) = \tilde{x}(x)$  for all  $x \in M$ <sup>18</sup>. Assume that  $\{T(\cdot), R(\cdot)\}$  implements  $\{\delta, \omega(\cdot)\}$  with cost manipulation. For simplicity's sake, assume that  $t(\cdot)$  is increasing over  $M$ . Let  $\tilde{\omega}(x) \equiv R(\tilde{x}(x))$ . We have  $\tilde{\omega}(x) \geq \omega(x)$  for all  $x$  and  $\tilde{\omega}(x) = R(\phi)$  if  $x \in M^c$ . Let us show that  $\{\delta, \tilde{\omega}(\cdot)\}$  can be implemented without cost manipulation at equilibrium (and thus with a higher net gain both for the policyholder and for the auditor).

---

<sup>18</sup>Manipulating the audit cost (by choosing  $e > 0$ ) in such a way that nevertheless the claim is rejected, is obviously a suboptimal strategy of the policyholder.

Given the reaction function of the auditor, a policyholder who suffers a loss  $x$  has to incur a manipulation cost at least equal to  $\text{Sup}\{0, \frac{R(x) - R(\hat{x})}{b}\}$  for the claim  $\hat{x}$  to be accepted by the auditor.

Let  $x \in M$ . Then  $\hat{x}(x) = \tilde{x}(x)$  gives

$$\tilde{x}(x) \in \underset{\hat{x} \in M'}{\text{ArgMax}} \{T(\hat{x}) - \text{Sup}\{0, \frac{R(x) - R(\hat{x})}{b}\}\}$$

Since  $\tilde{x}(x') \in M'$  for all  $x' \in M$ , -from (ii) and  $\hat{x}(x') = \tilde{x}(x')$ - we deduce

$$x \in \underset{x' \in M}{\text{ArgMax}} \{T(\tilde{x}(x')) - \text{Sup}\{0, \frac{R(x) - R(\tilde{x}(x'))}{b}\}\}$$

and (i) gives

$$x \in \underset{x' \in M}{\text{ArgMax}} \{t(x') - \text{Sup}\{0, \frac{R(x) - \tilde{\omega}(x')}{b}\}\} \text{ if } x \in M$$

We deduce that  $bt(x) + \tilde{\omega}(x)$  is locally constant if  $\tilde{\omega}(x) < R(x)$  and that  $bt(x) + \tilde{\omega}(x)$  is locally nonincreasing if  $\tilde{\omega}(x) = R(x)$ . These conditions should hold for all  $x \in M$ . Hence  $bt(x) + \tilde{\omega}(x)$  is locally nonincreasing over  $M$ .

The proof of lemma 3 shows that in such a case,  $\{t(\cdot), \tilde{\omega}(\cdot)\}$  can be implemented without cost manipulation by  $\{\tilde{T}(\cdot), \tilde{R}(\cdot)\}$  such that  $\tilde{T}(x) \equiv t(x)$ ,  $\tilde{R}(x) = \tilde{\omega}(x)$  if  $x \in M$ ,  $\tilde{T}(x) \leq \tilde{T}(\phi) = t_0$ ,  $\tilde{R}(x) \geq \tilde{R}(m_+)$  if  $x \in M^c$  and  $\tilde{R}(\phi) = R(\phi)$ . Hence, in this model, allowing for cost manipulation at equilibrium is a weakly dominated strategy for the insurer.

## References

- [1] Arrow, K. (1971), *Essays in the Theory of Risk Bearing*, North-Holland, Amsterdam.
- [2] Baron D. and D. Besanko (1984), "Regulation, asymmetric information and auditing", *Rand Journal of Economics*, 15, 447-470.
- [3] Beavis, B. and I. Dobbs (1990), *Optimization and Stability Theory for Economic Analysis*, Cambridge University Press, Cambridge.
- [4] Bond, E.W. and K.J. Crocker (1995), "Hardball and the soft touch : the economics of optimal insurance contracts with costly state verification and endogenous monitoring costs", Department of Economics - The Pennsylvania State University, mimeo.
- [5] Crocker, K.J. and J. Morgan (1995), "The optimality of costly state falsification : sharecropping and insurance fraud", Department of Economics - the Pennsylvania State University, mimeo.
- [6] Derrig, R.A., H.I. Weisberg and X. Chen (1994), "Behavioral factors and lotteries under no-fault with a monetary threshold : a study of Massachusetts automobile claims", *Journal of Risk and Insurance*, Vol. 61 (2), 245-275.
- [7] Dionne, G. and P. St-Michel (1991), "Workers' compensation and moral hazard", *Review of Economic and Statistics*, vol. 83 (2), 236-44.
- [8] Gollier, C. (1987), "Pareto-optimal risk sharing with fixed costs per claim", *Scandinavian Actuarial Journal*, 62-73.
- [9] Holmström, B. (1979), "Moral hazard and observability", *Bell Journal of Economics*, Vol. 10 (1), 74-91.
- [10] Huberman, G., D. Mayers and C. W. Smith Jr. (1983), "Optimum insurance policy indemnity schedules", *Bell Journal of Economics*, 14, Autumn, 415-426.
- [11] Maggi, G. and A. Rodriguez-Clare (1995), "Costly distortion of information in agency problems", *Rand Journal of Economics*, vol. 26 (4), 675-689.

- [12] Mookherjee D. and I. Png (1989), "Optimal auditing insurance and re-distribution", *Quarterly Journal of Economics*, CIV, 205-228.
- [13] Picard, P. (1994), "Auditing claims in insurance market with fraud : the credibility issue", *Journal of Public Economics*, forthcoming.
- [14] Raviv, A. (1979), "The design of an optimal insurance policy", *American Economic Review*, 69 (1), 854-96.
- [15] Seierstad, A. and K. Sydsaeter (1987), *Optimal Control Theory with Economic Applications*, North Holland.
- [16] Townsend, R. (1979), "Optimal contracts and competitive markets with costly state verification", *Journal of Economic Theory*, XXI, 265-293.