Overcompensation as a Partial Solution to Commitment and Renegotiation Problems: The Case of Ex Post Moral Hazard[*]

M. Martin Boyer[†]

May 1999

Abstract

In a Costly State Verification world, an agent who has private information regarding the state of the world must report what state occurred to a principal, who can verify the state at a cost. An agent then has what is called ex post moral hazard: he has an incentive to misreport the true state to extract rents from the principal. Assuming the principal cannot commit to an auditing strategy, the optimal contract is such that: 1- the agent's expected marginal utility when there is an accident (high- and low-loss states) is equal to his marginal utility when there is no accident; 2- the lower loss is undercompensated, while the higher loss is overcompensated; 3- the expected benefit is greater than the expected loss if $U''' > 0$; and 4- the welfare of the agent is greater under commitment than under no-commitment. The second result is contrary to the results obtained if the principal can commit to an auditing strategy; under commitment higher losses are underpaid whereas lower losses are overpaid.

JEL classification: D82, G2, C72.

Keywords: Non-commitment, Ex post moral hazard, Asymmetric information, Contract theory.

[*] This paper draws extensively on a chapter of my doctoral dissertation at the Wharton School of the University of Pennsylvania. I would especially like to thank Sharon Tennyson, my advisor, for her encouragement. I would also like to thank Kodjovi Assoé, Marcel Boyer, Keith Crocker, Georges Dionne, Neil Doherty, Marc Duhamel, Nathalie Fombaron, Thomas Palfrey, Michel Poitevin, seminar participants at the Université de Montréal, University of Minnesota, American Risk and Insurance Association (1997), and Econometric Society North American Summer meeting (1998) for helpful discussions. Financial support from the S.S. Huebner Foundation and the SSHRC of Canada is gratefully acknowledged.

[†] Assistant Professor, Department of Finance, École des Hautes Études Commerciales, Université de Montréal. 3000 Côte-Sainte-Catherine, Montréal, QC H3T 2A7 CANADA. Tel.: 514-340-6704; Fax.: 514-340-5632; E-mail: martin.boyer@hec.ca

Overcompensation as a Partial Solution to Commitment and Renegotiation Problems: The Case of Ex Post Moral Hazard

Abstract. In a Costly State Verification world, an agent who has private information regarding the state of the world must report what state occurred to a principal, who can verify the state at a cost. An agent then has what is called ex post moral hazard: he has an incentive to misreport the true state to extract rents from the principal. Assuming the principal cannot commit to an auditing strategy, the optimal contract is such that: 1- the agent's expected marginal utility when there is an accident (high- and low-loss states) is equal to his marginal utility when there is no accident; 2- the lower loss is undercompensated, while the higher loss is overcompensated; 3- the expected benefit is greater than the expected loss if $U''' > 0$; and 4- the welfare of the agent is greater under commitment than under no-commitment. The second result is contrary to the results obtained if the principal can commit to an auditing strategy; under commitment higher losses are underpaid whereas lower losses are overpaid.

JEL classification: D82, G2, C72.

Keywords: Non-commitment, Ex post moral hazard, Asymmetric information, Contract theory.

1 Introduction

The traditionnal mechanism design literature has concentrated on designing contracts where it is optimal for all players to tell the truth. These mechanism do not seem to apply in reality as we observe agents who sometimes do not tell the truth. The two most blatant examples are income-tax fraud and insurance fraud.

The mechanism design literature appropriately points out that the revelation principle may not hold when there is no commitment. More specifically, the commitment of the principal to a strategy is in many situations not a reasonable assumption since the players may be better off if the principal deviates ex post from the strategy agreed upon ex ante. The case of ex post moral hazard is a clear example of the commitment problem. Suppose the agent's payoff depends on the state of the world he announces to the principal. Suppose also that the state of the world is private information to him. If it is costly for the principal to verify the state of the world, then the agent has an incentive to misrepresent it. Suppose the principal believes that the agent tells the truth always (since it is optimal for the agent to do so). It is then clear that the principal will not want to waste resources to verify the state of the world, since she knows that the agent has told the truth about it. Thus the principal wants to renegotiate ex post to save resources.

The goal of this paper is to address this commitment problem between the informed agent and the principal. An insurance-contract approach is used primarily because the ex post moral hazard problem in insurance, namely insurance fraud, is very important. Although many types of fraud exist, I shall restrict the analysis to the case of build-up because it appears to be the most costly to the economy.[1]

Although a policyholder/insurer approach to the problem is used, it could just as well have been an entrepreneur/financier approach à la Gale and Hellwig (1985), or a polluter/regulator approach à la Laffont and Tirole (1993) or Lewis (1996). In the first case the entrepreneur knows whether a project was a success or not, whereas the financier, who lent the money, does not. Thus there is an incentive for the entrepreneur to take the money and run.[2] In the second case, the polluter knows how much mess is left to be cleaned up, whereas the government does not unless it sends biochemists to the polluted site. In an insurance setting, the policyholder knows what loss he

[1] Insurance fraud build-up is characterized by a policyholder who is involved in an accident and who, upon learning of his real loss, engages in claims padding (exaggerates his loss) to collect more money from his insurer. Build-up appears to be the most costly type of fraud. The Rand Corporation estimates that in 1993, more than $18 billion dollars was paid for apparently fraudulent medical claims arising from automobile accidents in the United States (Research Brief, 1995). Another case of fraud known as outright fraud relates to the case where the agent invents an accident instead of just exaggerating one. Outright fraud can be seen as a special case of build-up with $\pi = 1$ and $\lambda_L = 0$ in the model presented in this paper. For more details on all types of insurance fraud, see Hoyt (1989).

[2] See, for example, Scheepens (1995), Persons (1997), Khalil and Parigi (1998) and Boyer (1998).

suffered in an accident, whereas the insurer knows only that an accident occurred.

The game is then for the agent to make a report to the principal concerning the severity of the accident (send a message) and for the principal to decide whether to verify the agent's report, at a cost. The results of the paper are four-fold.

First an agent's average marginal utility when there is an accident (high- and low-loss states) is equal to his marginal utility when he is not involved in an accident. The second result shows that an agent is overcompensated (undercompensated) in the event of an accident of high (low) severity. This means that in the event of a high (low) loss,[3] the benefit the agent receives is greater (smaller) than his loss. The third result shows that the expected benefit is greater than the expected loss. Finally, I show as my fourth result that there is a welfare loss from not being able to commit to an auditing strategy.

The contribution of these results is easier to understand if one inspects the related literature on ex post moral hazard and on non-commitment. The difference between ex post and ex ante moral hazard depends on whether the agent plays after (ex post) or before (ex ante) Nature. Spence and Zeckhauser (1971) were the first to characterize these two forms of moral hazard. They show that there is no real difference between the two problems owing to their assumption that auditing is either costless or impossible.

Townsend (1979) was the first to present a model where the principal can verify the realized state of the world at a cost. This approach is known as the costly state verification approach. In Townsend's model, the optimal insurance contract between the principal and the agent stipulates a no-auditing region where a fixed payment is made, and an auditing region (where audits are always performed) where the agent has full insurance less a deductible. Mookherjee and Png (1989) showed that truth telling can be achieved using stochastic audits (see also Bond and Crocker, 1997). They also showed that a bonus should be paid to those agents who are audited and who are foud to have told the truth. Moreover, they show that Townsend's debt contract is no longer optimal when agents are risk averse. Mookherjee and Png's optimal contract is Pareto suprior to Townsend's, since less money is wasted on audits.

Commitment is an important feature of contractual relationships between a principal and an agent. In a multiperiod setting we know that an agent and a principal would be better off if they could sign a long-term, full-commitment contract. The problem, however, is that when the time comes to implement the contract, both players find it in their best interests to renegotiate the contract to achieve ex post efficiency. The commitment assumption has typically been challenged

[3] The terms severity of accident and loss are synonymous in this paper.

in multiperiod settings.[4] In a one-period setting, the importance of commitment is also non-trivial.[5] This lack of commitment is the cornerstone of this paper.

The paper that relates the most to this one is Khalil (1997). Using the same basic framework as Baron and Myerson (1982), Khalil develops a model where the principal is unable to commit to an auditing strategy. This framework consists in a principal (regulator) who doesn't know the production cost function of the agent (firm) she seeks to regulate. I discuss the recent literature as it relates to my model more thoroughly in section 4.

The remainder of the paper is divided as follows. In the next section, the claiming game between the agent and the principal is presented. In this game the informed agent must report his type (severity of accident) to the principal, who must then decide whether to audit. Section 3 of the paper presents the optimal contract in such a setting as well as its implications. Finally, section 4 discusses the results and presents a conclusion.

2 The Claiming Game

Before the model is examined in detail, it seems appropriate to state some basic assumptions. The agent is risk-averse ($U'(W) > 0$, $U''(W) < 0$, $U'(0) = \infty$), where $U(W)$ is the agent's von Neumann-Morgenstern utility function over final wealth. The principal is risk-neutral and is making zero expected profits. Any message sent by the agent to the principal is costless, whether the message is truthful or not. The distribution of losses given by a quadruplet $\Xi = \{\pi, \rho, \lambda_L, \lambda_H\}$ is common knowledge. If an accident occurs (with probability π), then only two possible losses can occur, λ_H and λ_L, with $\lambda_H > \lambda_L \geq 0$. Conditional on an accident occurring, the agent suffers loss λ_H with probability $\rho < \frac{1}{2}$. Although the agent suffers loss $\lambda \in \{\lambda_H, \lambda_L\}$, the payment he receives is given by $\beta \in \{\beta_H, \beta_L\}$. Therefore, when the claiming game is played, the players must take into account not only the values of λ_H and λ_L, but also those of β_H and β_L. The only information that is private to the agent is the severity of the accident (the loss he suffered). The insurer does not know the severity of the accident unless she conducts a costly audit of the claim at a cost of c. Auditing is perfect if conducted, and the audit cost c is common knowledge.

[4] See Cooper and Hayes (1987), Hosios and Peters (1989), Dionne and Doherty (1994), and Fombaron (1997).

[5] See Graetz, Reinganum and Wilde (1986), Beaudry and Poitevin (1993, 1994), Picard (1996, 1997), Boyer (1997) and Khalil (1997). Melumad and Mookherjee (1989) attempt to address this commitment problem (see also Swierzbinski, 1994, Lewis and Sappington, 1995, and Picard, 1997). By allowing the principal to delegate the auditing to a third party, they contend that the commitment problem may be avoided. Unfortunately, that is not necessarily the case since the third party itself may be faced with commitment and/or agency problems. Suppose the principal can commit to an auditing budget ex-ante. This would yield the same truth-telling results as committing to an auditing strategy. The problem is then for the principal to determine whether the auditing has been conducted properly. In other words, Who audits the auditors? Melumad and Mookherjee assume this problem away by arguing that it can be done at no cost.

The sequence of the game is displayed in figure 1. The claiming game is played in stages 2 to 5. In stage 0 the contract is signed. Nature plays in stages 1 and 2. In stage 1 Nature reveals to all players whether there is an accident[6] or not, while in stage 2 it reveals the severity of the accident only to the agent. The agent then makes a report to the principal, who then decides whether to audit the report. Since the occurrence of an accident is common knowledge,[7] the only thing the principal does not know is whether the loss is equal to λ_H or to λ_L. I shall assume that if the principal verifies the state of the world, then she must compensate the agent according to the benefit corresponding to his true loss. Thus an agent caught sending a false message receives compensation equal to what he would have received had he told the truth. The agent must, however, incur a positive fixed monetary penalty k, which is common knowledge. k is a deadweight loss to the economy in the sense that it is only paid by the agent, and not pocketed by the principal.[8] Finally let Y be the exogenous wealth of the agent.

There are three choice variables for the principal in this model: the coverage in case of a high loss (β_H), the coverage in the case of a low loss (β_L), and the premium (α). The variables are chosen by the principal to maximize the agent's expected utility before the claiming game starts. They are thus fixed for the duration of the claiming game and can be manipulated as parameters. The game can end in two ways. First, there may be no accident, in which case there is no opportunity for the agent to exaggerate his claim. The payoff to the agent and the principal are then given by $U(Y - \alpha)$ and α. Second, there could be an accident, and the claiming game is played. Table 1 in the appendix summarizes all the variable used in this paper, while Table 2 lists all the possible payoffs to the players contingent on their actions. Figure 2 is an extensive-form representation of the claiming game.

The solution to the claiming game gives us a Perfect Bayesian Nash Equilibrium (PBNE). By

[6] The term accident is used as a generic term to represent an action of Nature that is common knowledge. It can be viewed as an automobile accident where there are witnesses, a flood, a hurricane, a fire, etc... Everyone knows if a flood occurs or if a hurricane hits. It is also easy (read costless) to verify that a fire damaged a house: was the fire department called?

[7] This allows to push aside the observability issue raised by Crocker and Snow (1989). They show that being able to observe if the accident occurred or not may affect the shape of the optimal contract.

[8] Another way to view this penalty is as an opportunity cost of being in autarchy. Suppose an agent who is caught cheating is shut out of the insurance market, and thus receives the expected utility of remaining in autarchy. If $\overline{V}$ represent the agent's expected utility of participating in the market, and if $\underline{V}$ represents the agent's expected utility of remaining in autarchy, then $k = \left(\frac{\delta}{1-\delta}\right)\left(\overline{V} - \underline{V}\right)$ is the implicit penalty for being shut out of the market (where δ is the discount factor). This assumption that the penalty cannot be collected by the principal eliminates the problem of auditing as a rent-extracting device for the principal as in Khalil (1997). Since the principal does not profit from audits (he can only reduce his payout at best), audits become only a means of reducing the number of false messages sent in the economy. Picard (1996) studies the outright fraud case where the penalty is paid partially to the principal.

The penalty is set exogenously in my model. Endogenizing the penalty has been done by Mookhejee and Png (1989). But since in equilibrium my optimal contract is independent of the choice of penalty, endogenizing the penalty would not change any of the results in the model (except the equilibrium probability of audit).

definition the PBNE of this game is a sextuplet.

Definition 1 A PBNE is defined in this game as

$$
PBNE = \begin{pmatrix}
\text{Agent's strategy if Nature chose } \lambda = \lambda_L, \\
\text{Agent's strategy if Nature chose } \lambda = \lambda_H; \\
\text{Principal's strategy if the Agent reported } \lambda' = \lambda_L, \\
\text{Principal's strategy if the Agent reported } \lambda' = \lambda_H; \\
\text{Principal's beliefs in the information set resulting from } \lambda' = \lambda_L, \\
\text{Principal's beliefs in the information set resulting from } \lambda' = \lambda_H
\end{pmatrix}
$$

$$
= (\theta_L; \theta_H; \delta_L; \delta_H; \gamma_L; \gamma_H)
$$

I shall use η to represent the probability that an agent announces a loss λ_H when his loss is in fact λ_L. ν then represents the probability that an agent who reported a loss of λ_H is audited. This brings me to the first result of the paper.

Lemma 1 If $\beta_H > \beta_L$ and if $\rho < \frac{1}{2}$, then the unique PBNE in mixed strategy of this game[9] is

$$
\begin{array}{lll}
\theta_H = \lambda'_H & \delta_L = N & \gamma_L = 0 \\
\theta_L = \eta\lambda'_L + (1-\eta)\lambda'_H & \delta_H = \nu A + (1-\nu)N & \gamma_H = \frac{\rho}{\rho+(1-\rho)\eta}
\end{array}
$$

with

$$
\eta = \left(\frac{c}{\beta_H - \beta_L - c}\right)\left(\frac{\rho}{1-\rho}\right) \tag{1}
$$

$$
\nu = \frac{U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L)}{U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L-k)} \tag{2}
$$

Proof: All the proofs are in the appendix. ■

What lemma 1 does is model the strategic behavior of each player depending on the relationship between the contingent payoff in the case of a high-severity loss and in the case of a low-severity loss. Although there are other equilibriums, I shall concentrate on this one because it is the most interesting and the most plausible.[10] The lemma tells us that the agent 1- always reports λ'_H if he suffers a high loss ($\lambda = \lambda_H$); and 2- plays a mixed strategy between reporting losses λ'_L and λ'_H

[9] A necessary condition for the mixed strategy to exist is that the agent's probability of committing fraud be between zero and one (i.e. $\eta \in [0;1]$). This occurs only if $\rho < \frac{\beta_H-\beta_L-c}{\beta_H-\beta_L}$. A sufficient condition for $\rho < \frac{\beta_H-\beta_L-c}{\beta_H-\beta_L}$ to hold is that $\rho < \frac{1}{2}$ since, as I will show later in the paper, $\beta_H - \beta_L$ is always greater than 2c. This means that $\frac{\beta_H-\beta_L-c}{\beta_H-\beta_L}$ is always greater than $\frac{1}{2}$. As for uniqueness, Gibbons (1992) and Myerson (1991) state that in a 2×2 game (two players each with two possible actions) there will be at most one mixed-strategy equilibrium.

[10] The other two possibilities occur when $\beta_L = \beta_H$ or when $\beta_H < \beta_L$. The flat payment is not interesting since there will be no fraud in the economy. As for the case where $\beta_H < \beta_L$, Proposition 5 shows that it is always dominated by the case where $\beta_H = \beta_L$. There is therefore no loss in generality to concentrate on the case where $\beta_H \geq \beta_L$. These two other cases are presented in section 3.5.

if the real loss is low ($\lambda = \lambda_L$). Also, the principal 1- never audits (N) if the message is that the severity is low ($\lambda' = \lambda'_L$); and 2- plays a mixed strategy between auditing (A) and not auditing (N) if $\lambda' = \lambda'_H$.

Given the players' optimal strategies, it is possible to find the price that gives zero expected profits to the principal. This price is implicitly given by

$$\alpha = \pi[\rho\beta_H + (1-\rho)\beta_L] + \pi(1-\rho)(\beta_H - \beta_L)\eta(1-\gamma) + \pi c\gamma[\rho + (1-\rho)\eta] \tag{3}$$

On the right, $\pi[\rho\beta_H + (1-\rho)\beta_L]$ represents the expected benefits paid when the agent tells the truth. The second term, $\pi(1-\rho)(\beta_H - \beta_L)\eta(1-\gamma)$, is the rent an agent can expect to extract from the principal. Finally, $\pi c\gamma[\rho + (1-\rho)\eta]$ is the expected cost of the auditing strategy.[11] By substituting in (3) for the equilibrium value of η given in (1), the price of the contract becomes

$$\alpha = \pi\left[\beta_L + \rho\frac{(\beta_H - \beta_L)^2}{\beta_H - \beta_L - c}\right] \tag{4}$$

This zero-profit constraint has the characteristic that the principal's auditing strategy does not directly affect the price of the contract. Therefore the price function is not related to the agent's utility function, the agent's initial wealth (Y) or the penalty (k). The reason is that Y and k only appear in the price of the contract by way of the audit probability as we can see in (2) and (3). But since the audit probability washes out once we substitute for η, all dependence of α on Y and k are discarded.

It is then possible to study the impact the optimal contract has on the Nash Equilibrium behavior of the players. More to the point, the optimal contract affects the probability that a false message is sent given by $F = \pi(1-\rho)\eta$. I am then able to state my first proposition.

Proposition 1 The greater is $(\beta_H - \beta_L)$, the smaller the probability that a false message is sent.

By promising an overpayment of the high loss and an underpayment of the low loss, the principal reduces fraud in the economy. This proposition is a direct consequence of the mixed strategy played by the agent in the event of a low severity accident. We know from how mixed equilibriums are constructed that one player's strategy depends only on the other player's payoffs, not his own. This means that the greater the difference between β_H and β_L, the more the principal has to lose by not auditing (and the more he has to gain by auditing). Therefore, in order for the principal

[11] It is interesting to note that if $\eta = 0$ and $\gamma = 0$, then the price is given only by the first term of (3). Thus if the agent never lies and the principal never audits, then the price of the contract collapses to what is known as the actuarially fair price of the contract, or the pure premium. In such a setting, it is well known - and easily proven - that the agent would choose coverage equal to his possible loss in each state, i.e. $\beta_L = \lambda_L$ and $\beta_H = \lambda_H$.

to remain indifferent between auditing and not auditing, the agent must reduce his probability of filing a fraudulent claim.

The optimal contract also affects the probability that a successful false message, given by $S = F(1-\gamma)$, is sent. A sufficient condition for S to decrease when the difference between β_H and β_L increases would be for $\frac{\partial(1-\gamma)}{\partial(\beta_H-\beta_L)} < 0$. Unfortunately, it is not possible to find a functional form for $\frac{\partial(1-\gamma)}{\partial(\beta_H-\beta_L)}$. If, however, $\frac{\partial(1-\gamma)}{\partial\beta_H} < 0$ and $\frac{\partial(1-\gamma)}{\partial\beta_L} > 0$, then we may infer that $\frac{\partial(1-\gamma)}{\partial(\beta_H-\beta_L)} < 0$. This gives us proposition 2.

Proposition 2 Conditional on a false message being sent, the probability that it is successful decreases as β_H increases ($\frac{\partial(1-\gamma)}{\partial\beta_H} < 0$) if and only if

$$\frac{U'(Y-\alpha-\lambda_L+\beta_H)}{\alpha'_H} > \left[U'(Y-\alpha-\lambda_L+\beta_H) - U'(Y-\alpha-\lambda_L+\beta_L-k)\right](1-\xi) \qquad (5)$$

while it decreases as β_L decreases ($\frac{\partial(1-\gamma)}{\partial\beta_L} > 0$) if and only if

$$\frac{U'(Y-\alpha-\lambda_L+\beta_H)}{(1-\pi+\alpha'_H)} > \left[U'(Y-\alpha-\lambda_L+\beta_H) - U'(Y-\alpha-\lambda_L+\beta_L-k)\right](1-\xi) \qquad (6)$$

where $\alpha'_H \in (0, \pi\rho)$ is given in 12 and

$$\xi = \frac{\dfrac{U(Y-\alpha-\lambda_L+\beta_H)-U(Y-\alpha-\lambda_L+\beta_L-k)}{U'(Y-\alpha-\lambda_L+\beta_H)-U'(Y-\alpha-\lambda_L+\beta_L-k)}}{\dfrac{U(Y-\alpha-\lambda_L+\beta_L)-U(Y-\alpha-\lambda_L+\beta_L-k)}{U'(Y-\alpha-\lambda_L+\beta_L)-U'(Y-\alpha-\lambda_L+\beta_L-k)}} \qquad (7)$$

Although the conditions for $\frac{\partial(1-\gamma)}{\partial\beta_H} < 0$ and $\frac{\partial(1-\gamma)}{\partial\beta_L} > 0$ seem a bit messy (especially equation 7), it is easy to show that they are always satisfied if the agent's utility function takes the form $U(W_i) = -e^{-aW_i}$, because then $\xi = 1$. Since the marginal utility is positive, then the conditions stated in the proposition are satisfied.[12]

We now know that as $\beta_H - \beta_L$ increase the probability of committing fraud decreases, and the probability of a fraud being successful decreases as well. What is left to derive is the optimal insurance contract in such a setting.

3 The Optimal Contract

The problem for the principal is to choose a coverage pair and a premium that maximize the agent's expected utility over final wealth. The principal anticipates rationally the optimal strategies of the two players in the event of an accident. The principal must include those strategies in her contract design.

[12] It is obvious that if $\xi < 1$, then the proposition always holds. The question is whether there are utility functions such that ξ is so large that the inequalities in (5) and (6) are reversed; I have not found any.

3.1 Results

The problem faced by the principal is

$$
\begin{aligned}
\max_{\alpha,\beta_L,\beta_H} V = \; & (1-\pi)U(Y-\alpha) + \pi\rho U(Y-\alpha-\lambda_H+\beta_H) \\
& + \pi(1-\rho)(1-\delta)U(Y-\alpha-\lambda_L+\beta_L) \\
& + \pi(1-\rho)\delta\gamma U(Y-\alpha-\lambda_L+\beta_L-k) \\
& + \pi(1-\rho)\delta(1-\gamma)U(Y-\alpha-\lambda_L+\beta_H)
\end{aligned} \tag{MP}
$$

subject to the constraints

$$\alpha = \pi\left[\beta_L + \rho\frac{(\beta_H-\beta_L)^2}{\beta_H-\beta_L-c}\right] \tag{8}$$

$$\delta = \left(\frac{c}{\beta_H-\beta_L-c}\right)\left(\frac{\rho}{1-\rho}\right) \tag{9}$$

$$\gamma = \frac{U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L)}{U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L-k)} \tag{10}$$

$$\text{Participation Constraint} \tag{11}$$

Equation (8) represents the zero-profit constraint of the principal in which the two Nash equilibrium strategies (9 and 10) were substituted. Let's discard the participation constraint[13] for now as it does not bind in equilibrium. By substituting (9), (10) and (8) into MP, the problem simplifies to

$$
\begin{aligned}
\max_{\beta_L,\beta_H} V = \; & (1-\pi)U\left(Y-\pi\left[\beta_L+\rho\frac{(\beta_H-\beta_L)^2}{\beta_H-\beta_L-c}\right]\right) \\
& + \pi\rho U\left(Y-\pi\left[\beta_L+\rho\frac{(\beta_H-\beta_L)^2}{\beta_H-\beta_L-c}\right]-\lambda_H+\beta_H\right) \\
& + \pi(1-\rho)U\left(Y-\pi\left[\beta_L+\rho\frac{(\beta_H-\beta_L)^2}{\beta_H-\beta_L-c}\right]-\lambda_L+\beta_L\right)
\end{aligned} \tag{SP}
$$

It is interesting to note that the penalty for committing fraud is never a factor in the design of the optimal contract: the parameter k is nowhere to be found in SP. Therefore it cannot be a factor when the optimal contract is chosen. A related point is that it will not matter whether the penalty imposed on the agent caught committing fraud is a decision variable in the model. As long as the decision to audit comes last in the sequence of play, the penalty parameter will disappear

[13] The participation constraint is such that the agent must be better off buying the contract than in autarchy. Algebraically, it means that

$$EV^* \geq (1-\pi)U(Y) + \pi\rho U(Y-\lambda_H) + \pi(1-\rho)U(Y-\lambda_L)$$

where EV^* is the agent's expected utility if he purchases the optimal contract.

from the simplified problem. There is no loss in generality here to use a fixed penalty rather than a penalty that would be a function of the reported loss, or the difference between the reported loss and the real loss since there is only one possible report that may be fraudulent. As long as the penalty is not pocketed by the principal, letting the penalty equal $k(\lambda_H)$, $k(\lambda_H;\lambda_L)$, or k has no impact on the contracting problem.

Another interesting point to mention is that the audit probability γ drops off completely from the optimization problem. The reason is that the principal is indifferent between auditing and not auditing, which means that changes in the audit probability do not affect the zero-profit constraint.

I am then able to show the following lemma.

Lemma 2 Letting

$$\alpha'_H = \frac{\partial\alpha}{\partial\beta_H} = \pi\rho\frac{(\beta_H-\beta_L)(\beta_H-\beta_L-2c)}{(\beta_H-\beta_L-c)^2} \tag{12}$$

$$\alpha'_L = \frac{\partial\alpha}{\partial\beta_L} = \pi\left[1-\rho\frac{(\beta_H-\beta_L)(\beta_H-\beta_L-2c)}{(\beta_H-\beta_L-c)^2}\right] \tag{13}$$

and

$$EV' = (1-\pi)U'(Y-\alpha) + \pi\rho U'(Y-\alpha-\lambda_H+\beta_H) + \pi(1-\rho)U'(Y-\alpha-\lambda_L+\beta_L) \tag{14}$$

where EV' is defined as the **expected marginal utility** with respect to wealth, the optimal $(\beta_L^*;\beta_H^*)$ solves

$$\pi\rho\frac{U'(Y-\alpha-\lambda_H+\beta_H)}{EV'} = \alpha'_H \tag{NC1}$$

$$\pi(1-\rho)\frac{U'(Y-\alpha-\lambda_L+\beta_L)}{EV'} = \alpha'_L = \pi-\alpha'_H \tag{NC2}$$

It is clear that the left-hand sides of NC1 and NC2 are positive. Therefore, α'_H and α'_L must also be positive.[14] It then becomes clear that the participation constraint does not bind. Suppose the opposite: the agent prefers to be in autarchy. This is the same as choosing a contract where

[14] α'_H is positive only if $\beta_H<\beta_L$ or $\beta_H-\beta_L>2c$. The first case can be discarded because we assumed that $\beta_H>\beta_L$. We are left with $\beta_H-\beta_L>2c$. Recall from footnote 10 that $\rho<\frac{\beta_H-\beta_L-c}{\beta_H-\beta_L}$ was a necessary condition for a mixed-strategy equilibrium to exist in the claiming game. Assume that $\rho<\frac{1}{2}$, we determine that $\frac{\beta_H-\beta_L-c}{\beta_H-\beta_L}>\frac{1}{2}$. it follows that $\beta_H-\beta_L>2c$. Thus $\rho<\frac{1}{2}$ is a sufficient condition for $\eta\in[0;1]$. α'_L is always positive since $\alpha'_H<\pi$. To see why, note that

$$\frac{(\beta_H-\beta_L)(\beta_H-\beta_L-2c)}{(\beta_H-\beta_L-c)^2}<1$$

if and only if $c^2>0$. Therefore, it follows that

$$1-\rho\frac{(\beta_H-\beta_L)(\beta_H-\beta_L-2c)}{(\beta_H-\beta_L-c)^2}<1$$

Thus $\alpha'_L<\pi$.

$\beta_H = \beta_L = 0$. Suppose the agent does choose such a contract. Then $\alpha'_H = 0$. This means, from FOC_H, that $U'(Y - \lambda_H) = 0$, which is not possible. Thus $\beta_H = \beta_L = 0$ is not an optimal choice.[15] What do these two necessary conditions imply for the economy? With some manipulation, I can prove my first theorem.

Theorem 1 At the optimum, the agent's expected marginal utility with respect to wealth is the same in the accident state as in the no-accident state.

The statement of theorem 1 is not surprising; this is a classic result in the literature (see Winter, 1992, and Bond and Crocker, 1997). The theorem only tells us that the marginal utilities accross accident states are the same; it says nothing about the marginal utility in a given loss state conditional on an accident occurring. More specifically, it does not tell us whether there is perfect income smoothing (marginal utilities equal in every state), nor whether the agent's marginal utility in the high-loss state is greater (or smaller) than in the low-loss state. The reason this result is obtained is that there is no information asymmetry between the principal and the agent regarding whether there was an accident or not. We know from classical results that under perfect information an agent will choose to be fully insured. In our case, there is perfect information concerning whether an accident occurred or not.

Solely on the basis of theorem 1, it seems possible to envision a contract where perfect income smoothing is obtained ($\beta_H = \lambda_H$ and $\beta_L = \lambda_L$). Such a contract is not optimal, however. In fact the optimal coverage for the agent will be such that coverage is better (i.e., the agent's out-of-pocket amount is smaller) in the event of a high-severity accident than in the event of a low-severity accident. This is proven as corollary 1.

Corollary 1 The optimal contract is such that the agent does not smooth his income perfectly. In fact, $\beta_H - \lambda_H > \beta_L - \lambda_L$.

This corollary not only shows that the optimal contract does not give the agent the same final wealth in every state of the world, it also tells us that the agent gets greater utility in the high-loss state than in the low-loss state. This means that the agent is better covered if he suffered a high-loss accident, which has the strange implication that the agent would rather be involved in a high-severity accident than in a low-severity one. How can this be explained?

Looking back to the zero-profit constraint (equation 8), we see that by increasing the difference between β_H and β_L one reduces the implicit loading included in the premium. This load is given

[15] In the proof of proposition 4, I show that $\beta^* > 0$ even by assuming $\beta_H = \beta_L = \beta$.

by $m = \pi\rho(\beta_H - \beta_L)\left(\frac{c}{\beta_H - \beta_L - c}\right)$. Clearly $\frac{\partial m}{\partial(\beta_H - \beta_L)} < 0$. This means that if the difference between β_H and β_L is increased, the amount of money wasted because of moral hazard decreases. Therefore, letting $\beta_H - \lambda_H > \beta_L - \lambda_L$ is a way to reduce the load induced by the ex post moral hazard.

Another reason that larger losses may be better covered is that they give the insurer greater incentive to audit, and thus reduce the agent's probability of lying. We know from the construction of mixed equilibria that one player's strategy must take into account only the parameters the other player cares about. Thus the agent must consider what effect a greater difference between β_H and β_L has on the principal's payoff. It is straightforward to see that the greater it is, the more the principal has to lose by not auditing. Knowing this, the agent must reduce his probability of sending a false signal.

The question then becomes whether the compensation in the high-loss state may be so large that the agent prefers to be in it rather than in the no-accident state. In other words, is the agent underinsured? Or does he choose coverage that is greater than his loss? This question is answered in the following theorem, which is the most striking result of the paper.

Theorem 2 The agent's higher loss is overinsured while his lower loss is underinsured. In other words, $\beta_H > \lambda_H$ and $\beta_L < \lambda_L$.

The implications of this theorem are very interesting. Because of ex post moral hazard and non-commitment, the optimal contract is such that compensation is greater than the actual loss if the loss is high. Also, compensation is smaller than the loss in the event of a low-severity accident. The players' willingness to reduce the load inherent to the presence of ex post moral hazard therefore induces the principal to make the difference between the higher and the lower compensation so great that the agent becomes overcompensated in the case of the high loss, and undercompensated in the case of the low loss.

This contract resembles what is called a replacement-cost-new insurance contract. These contracts stipulate that if a good is totally destroyed, the insurer will replace it with another similar, but brand-new good. For example, if a brand-new car is totally destroyed within two years of its original purchased date, the owner of the car receives another brand new car as compensation for his loss.[16] There is no provision for depreciation. Thus the owner of a $25,000 Ford Taurus bought in January 1997 and destroyed (or stolen) in December 1998 receives a brand new $25,000 Ford Taurus as compensation for his loss. Presumably, the destroyed Ford Taurus is worth a lot less than $25,000 after two years of wear and tear. If the car is only partially damaged (or if only the radio is

[16] In other words, for the first two years the benefit paid to the policyholder is in terms of the replacement cost of a new vehicle.

stolen), then the insurance company will pay for the repairs, minus the deductible. This contract therefore stipulates some overcompensation in the case of the higher loss, and undercompensation in the case of the lower loss (because of the deductible).

We can see what the contract looks like by examining figure 3. Any point on the 45-degree line means that the loss is perfectly compensated. We see that the lower loss is undercompensated (point L), while the higher loss is overcompensated (point H). The dotted line between points H and L gives us a way to compare expected benefits with expected losses. If the expected loss is smaller (greater) than λ^*, then the expected benefit will be smaller (greater) than the expected loss. λ^* is set so that the expected benefit is equal to the expected loss.

When the agent receives the same expected marginal utilities in the accident state as in the no-accident state, this does not mean that he receives the same marginal utility in every loss state. From theorem 2 we know that the optimal contract entails overpayment of the higher loss and underpayment of the lower loss, which means that the marginal utility of wealth is greatest in the low-loss state. The question is then to compare the expected wealth in the accident state with the wealth in the no-accident state. This yields proposition 3.

Proposition 3 $U''' > 0$ ($U''' < 0$) is a sufficient condition for the agent's expected wealth to be larger (smaller) in the accident state than in the no-accident state.

The proof is perhaps clearer when the agent's marginal utility is plotted as a function of his wealth (see figure 4). We see that the average of marginal utilities of wealth is always greater (smaller) than the marginal utility of the average wealth if the marginal utility function is convex (concave). A convex marginal utility seems to make more sense as it is implied by a decreasing absolute risk-aversion utility function, which was shown to be a reasonable assumption by Arrow (1970) and Pratt (1964).

Proposition 1 means that when the agent chooses his optimal contract, he agrees to receive lower expected wealth when there is no accident than when there is an accident. This is the same as saying that the expected benefit is greater than the expected loss, as I prove in the following corollary.

Corollary 2 If $U''' > 0$, then the expected benefit is greater than the expected loss.

Point E in figure 3 compares the expected benefit with the expected loss. If E is above the 45-degree line, then the expected benefit is greater than the expected loss. The result makes intuitive sense. The agent's expected wealth is greatest in the state where he faces some uncertainty. In other words, the agent needs to be compensated for incurring some risk in the accident state.

3.2 Commitment

In this section, the commitment concern is addressed. What I want to do is compare the welfare of the agent when the principal can commit to an auditing strategy to his welfare when the principal cannot commit.

To analyze the problem under commitment, a new decision variable is needed: the principal's auditing strategy, γ. There is no Nash equilibrium constraint, but there is an incentive constraint that stipulates that the agent is (weakly) better off telling the truth (second constraint in CP below). The zero-profit constraint (first constraint in PC) is messier, however. In the no-commitment case the Nash strategy of the agent was such that the Nash strategy of the principal disappeared from the zero-profit constraint. By setting $\delta = 0$, it is no longer possible to cancel the audit strategy of the principal. Thus the premium can no longer be written as an explicit function. The Lagrange function to maximize is then

$$\begin{aligned}\max_{\beta_L,\beta_H,\gamma,\alpha} \mathcal{L} = \; & (1-\pi)U(Y-\alpha) + \pi\rho U(Y-\alpha-\lambda_H+\beta_H) + \pi(1-\rho)U(Y-\alpha-\lambda_L+\beta_L) \\ & - \mu_1\left[\alpha - \pi[\rho\beta_H + (1-\rho)\beta_L] - \pi\rho\gamma c\right] \\ & - \mu_2\left[\gamma - \frac{U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L)}{U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L-k)}\right] \end{aligned} \qquad \text{(CP)}$$

It is interesting to see that the no-commitment case is much easier to analyze than the commitment case. Without commitment the simplifed maximization problem (SP) was unconstrained with two variables. In the commitment case, there are four variables and two constraints. This adds two Lagrange multipliers, which means that there are six first-order conditions instead of only two. Theorem 3 follows.

Theorem 3 If the principal can commit to an auditing strategy, then the optimal contract is such that $\beta_H < \lambda_H$ and $\beta_L > \lambda_L$.

If the principal can commit to an auditing strategy, then the insurance contract is such that the agent is overcompensated for his loss in the low-loss state, and undercompensated for his loss in the high-loss state. This result is the complete opposite of the one under no-commitment. The reason is that the principal does not need to signal ex ante an ex post auditing strategy if she can fully commit credibly ex ante. Without commitment, the principal's signal was achieved by overcompensation of the high loss and undercompensation of the low loss. If on the other hand the principal can commit to an auditing strategy ex ante, then she will design a contract that minimizes her audit costs. The principal does this by reducing her probability of auditing (second constraint in CP), which is achieved by increasing the payment in the case of a low-severity accident and reducing the payment in the case of a high-severity accident.

It is then clear that the premium paid by the agent is greater under no-commitment, and that there is a welfare loss if the principal is unable to commit. This is shown in the following corollary

Corollary 3 The premium is greater under no-commitment than under commitment.Furthermore, there is a welfare loss from being unable to commit to an auditing strategy.

It is logical to obtain such a result since the revelation principal states that amongst all feasible optimal contracts, at least one induces the agent to tell the truth. Therefore the contract under no-commitment cannot be better than the contract under commitment. Corollary 4 shows that the agent's utility is in fact strictly greater if the principal can commit than it will be if she cannot commit.

3.3 Over-Compensation

It seems odd that the optimal contract in the presence of ex post moral hazard entails overpayment of the higher loss. If fraud is so common in the economy, why is it that we do not systematically observe contracts similar to the one we found? Two reasons come to mind: insurer expenses and ex ante moral hazard.

Suppose that the insurer's expenses, such as underwriting, marketing and taxes, can be modeled as a proportional loading factor τ on the premium. The simplified maximization problem then becomes

$$\begin{aligned} \max_{\beta_L,\beta_H} V &= (1-\pi)U\left(Y-(1+\tau)\pi\left[\beta_L+\rho/\frac{(\beta_H-\beta_L)^2}{\beta_H-\beta_L-c}\right]\right) \\ &+\pi\rho U\left(Y-(1+\tau)\pi\left[\beta_L+\rho/\frac{(\beta_H-\beta_L)^2}{\beta_H-\beta_L-c}\right]-\lambda_H+\beta_H\right) \\ &+\pi(1-\rho)U\left(Y-(1+\tau)\pi\left[\beta_L+\rho/\frac{(\beta_H-\beta_L)^2}{\beta_H-\beta_L-c}\right]-\lambda_L+\beta_L\right) \end{aligned} \tag{15}$$

Solving for this problem yields the following two necessary conditions

$$\pi\rho\frac{U'(Y-(1+\tau)\alpha-\lambda_H+\beta_H)}{EV'}=(1+\tau)\alpha'_H \tag{16}$$

$$\pi(1-\rho)\frac{U'(Y-(1+\tau)\alpha-\lambda_L+\beta_L)}{EV'}=(1+\tau)\alpha'_L \tag{17}$$

where α'_H and α'_L are as before, and

$$EV' = (1-\pi)U'(Y-(1+\tau)\alpha) + \pi\rho U'(Y-(1+\tau)\alpha-\lambda_H+\beta_H) + \pi(1-\rho)U'(Y-(1+\tau)\alpha-\lambda_L+\beta_L) \quad (18)$$

In accordance with standard results in the literature, proportional premium loading reduces insurance coverage in the sense that the agent receives greater expected marginal utility in the accident state than in the no-accident state. With total derivatives it is possible to show that β_H decreases as τ increases, as in Boyer (1997). It is left to the reader to show that there exists a $\hat{\tau} > 0$ so that $\beta_H = \lambda_H$, and $\beta_L < \lambda_L$. Therefore, for any proportional loadings τ greater than $\hat{\tau}$ overinsurance should not be observed.

The other reason that overinsurance is not observed more often in reality is the presence of ex ante moral hazard. If the agent knows that in the event of a high loss he will be overcompensated, then he may want to increase the probability that he will find himself in that state of the world. In other words, he may cause a high loss on purpose. Consequently, it may no longer be optimal to overcompensate an agent's higher loss if the agent is indeed able to increase the probability that it will occur. It is possible to imagine that the agent will want to help nature in choosing this higher loss. For example, an agent whose house has just caught fire may want to wait a few minutes before calling the fire fighters so that way he can increase the probability of suffering the higher loss. We can eliminate this problem by including a constraint in the maximization problem which states that the agent cannot receive more utility in the high-loss state than in the low-loss state.

There are two ways to do so. First, I could assume that in an accident the agent suffers some disutility that is proportional to the loss (loss of family heirlooms and pictures) so that he will not want to be in the high-loss state. This is a kind of state-dependent utility function approach. Second, I could restrict the contract so that $U(Y-\alpha-\lambda_H+\beta_H) \leq U(Y-\alpha-\lambda_L+\beta_L)$. In this case the optimal contract is such that $U(Y-\alpha-\lambda_H+\beta_H) = U(Y-\alpha-\lambda_L+\beta_L)$, which means that $\lambda_H-\beta_H = \lambda_L-\beta_L = D > 0$. This contract is then nothing more than a pure deductible contract.

3.4 The other two cases

I have concentrated in this paper on the case that was the most interesting and the most plausible. The other two cases, $\beta_L = \beta_H$ and $\beta_L > \beta_H$, were put aside to make the paper more readable. This section of the paper addresses those two cases briefly and presents the optimal contracts.

Lemma 3 *There are two other possible equilibriums in the claiming game*

1- if $\beta_H = \beta_L$, then the unique type of PBNE of this game is

$$\begin{array}{lll} \theta_L = \eta_L \lambda'_L + (1-\eta_L)\lambda'_H & \delta_L = N & \gamma_L \in [0;1] \\ \theta_H = \eta_H \lambda'_L + (1-\eta_H)\lambda'_H & \delta_H = N & \gamma_H \in [0;1] \end{array}$$

where

$$\eta_L \in [0;1] \text{ and } \eta_H \in [0;1];$$

and 2- if $\beta_H < \beta_L$, then the unique PBNE in mixed strategy of this game is

$$\begin{array}{lll} \theta_L = \lambda'_L & \delta_L = \nu A + (1-\nu)N & \gamma_L = \frac{\rho}{\rho+(1-\rho)\eta} \\ \theta_H = \eta\lambda'_L + (1-\eta)\lambda'_H & \delta_H = N & \gamma_H = 0 \end{array}$$

with

$$\eta = \left(\frac{c}{\beta_L - \beta_H - c}\right)\left(\frac{1-\rho}{\rho}\right) \tag{19}$$

$$\nu = \frac{U(Y-\alpha-\lambda_H+\beta_L) - U(Y-\alpha-\lambda_H+\beta_H)}{U(Y-\alpha-\lambda_H+\beta_L) - U(Y-\alpha-\lambda_H+\beta_H-k)} \tag{20}$$

Suppose first that $\beta_H = \beta_L = \beta$. The advantage of this type of contract is that audit costs are saved. If the payment is fixed whatever the agent's report, there is no need to audit since the principal has nothing to gain from it. Moreover, the agent never pays the penalty because he is never caught cheating. The problem faced by the principal in this case is

$$\max_{\beta} V = (1-\pi)U(Y-\pi\beta) + \pi\rho U(Y+(1-\pi)\beta-\lambda_H) + \pi(1-\rho)U(Y+(1-\pi)\beta-\lambda_L) \tag{21}$$

Since there is no reason to commit fraud or to audit, and since choosing $\beta = 0$ is equivalent to autarchy, the Nash equilibrium constraint and the participation constraint disappear. Solving for β yields

$$\begin{aligned} \frac{\partial V}{\partial \beta} = 0 = & -\pi(1-\pi)U'(Y-\pi\beta) + (1-\pi)\pi\rho U'(Y+(1-\pi)\beta-\lambda_H) \\ & +(1-\pi)\pi(1-\rho)U'(Y+(1-\pi)\beta-\lambda_L) \end{aligned} \tag{22}$$

The necessary condition for an optimum is then

$$U'(Y-\pi\beta) = \rho U'(Y+(1-\pi)\beta-\lambda_H) + (1-\rho)U'(Y+(1-\pi)\beta-\lambda_L) \tag{23}$$

which says only that the optimal coverage is chosen to equalize the marginal utility in the no-accident state to the expected marginal utility in the accident state. This necessarily means that $\lambda_H > \beta > \lambda_L$. Therefore, the higher loss is undercompensated and the smaller loss is over-compensated. If one compares the expected utility at their respective optimal coverage schedule, propositions 4 and 5 follow

Proposition 4 A contract where $\beta_L = \beta_H$ gives greater utility to the agent than a contract where $\beta_H > \beta_L$ if the cost of auditing is larger than some $\mathbf{e}$ that solves

$$U'\left(Y - \pi\left[\beta^*_L + \rho\frac{(\beta^*_H - \beta^*_L)^2}{\beta^*_H - \beta^*_L - \mathbf{e}}\right]\right) > U'(Y - \pi\beta^*) \tag{24}$$

Proposition 5 A contract where $\beta_L = \beta_H$ gives greater utility to the agent than a contract where $\beta_H < \beta_L$.

Propositions 5 tells us is that we need only consider the cases where $\beta_H = \beta_L = \beta$ and $\beta_H > \beta_L$ in deciding what coverage schedule is optimal since $\beta_H < \beta_L$ is dominated by $\beta_H = \beta_L$.

Proposition 4 makes intuitive sense. We can see two reasons that a flat payment may be preferable to a variable payment where the higher loss is overcompensated. First, when audit costs are high, the amount of money wasted through auditing increases since the greater the cost of auditing, the greater the probability of committing fraud. This invariably increases the number of audits in the economy (if the probability of auditing remains constant), which increases the total amount of money wasted. It is therefore logical to expect that there is a limit to the cost of auditing so that a variable payment is preferable to the fixed payment. The second reason has to do with the insurance concept of the contract. If the cost of auditing increases, the distance between β_H and β_L increases. The difference may become so large that the payment schedule no longer gives any insurance to the agent; it becomes a lottery.

4 Discussion and Conclusion

This paper examines from the insurance standpoint which contract a principal, who cannot commit credibly to an auditing strategy, would offer an informed agent who is faced with ex post moral hazard. In a two-point distribution of losses conditional on the occurrence of an accident, the optimal contract entails overpayment of the larger loss and underpayment of the lower loss. Furthermore, the optimal contract is such that the agent receives the same expected marginal utility when an accident occurs as when no accidents occur, even though the agent's expected final wealth is greater in the accident state. Finally, there is a loss of welfare resulting from the principal's being unable to commit.

The contribution of this paper is that it is the first where the optimal contract stipulates that the higher loss must be overcompensated, whereas the lower loss must be undercompensated (as opposed to Bond and Crocker, 1997). Also, overcompensation does not depend on an agent's being audited (as in Mookherjee and Png, 1989). The reason that the players agree to this counterintuitive scheme is that it reduces the cost of fraud in the economy. The reduction in the cost of fraud does not come primarily from the reduction of false messages that are not detected, which are merely transfers from the principal to the agent. It comes from the smaller number of messages that are audited. Since the agent is less likely to send a false message when there is overcompensation of large losses, the total number of high-loss claims is reduced. This means that the total audit cost may go down. And since the cost of auditing is a deadweight cost to the economy, a reduction in this deadweight can only make the agent better o¤. Another gain from overcompensation is that agents are less likely to pay the deadweight penalty. Since the agent sends fewer false messages, he is less likely to be caught having sent a false message. This means that less money is wasted in deadweight penalties.

To what extent can this contract be applied to the real world? There are many instances where smaller losses are undercompensated; the preponderance of deductibles is a blatant example. The overcompensation of large losses is less frequent. There are cases where we observe some type of overcompensation, however. For example, in France some hospital patients appear to be overcompensated for their loss: not only are the patients treated for whatever ails them, they also receive some monetary compensation when they are discharged from the hospital.[17]

Another example, which is probably more intuitive, comes from what are called replacement-cost-new insurance contracts. These contracts stipulate that if a good is totally destroyed, the insurer will replace it with another good that is similar, but brand-new. If the good is partially damaged, the insurance company will pay only to have it repaired (minus a deductible presumably). A case where this happens is homeowner insurance. Suppose the house is faced with some …re hazard. Whether there is a …re or not is perfectly observable: was the …re department called? If the …re is minor, say it a¤ects only the kitchen, then the insurer may pay only to repair the …re damage in the kitchen. The insurer will replace the damaged goods with the same common appliances, and the ‡oor tiling with the same easily found pattern. Given that the policyholder must pay a deductible, he is monetarily worse o¤ than before the …re. Suppose, however, that the house is a total loss; it burned to the ground. The insurance company may be willing to rebuild the house according to current standards (install central heat and air, a gas stove, new insulation, etc.) instead of rebuilding the house as it was. This may increase the value for the policyholder since the house is now worth more.

Jost (1996) and Khalil (1997) treat problems similar to the one considered here, but with some major di¤erences. Khalil studies the case of an informed risk-neutral agent who must make a report to the risk-averse principal. The basic framework he uses is that of Baron and Myerson

[17] I am indebted to Nathalie Fombaron for this anecdote.

(1982), i.e. the information the agent has concerns his production cost. Khalil …nds that the agent overproduces in comparison with the case where there is full information. This is similar to my result of overinsurance in the high-loss state. Surprisingly Khalil obtains these results using a completely di¤erent set of assumptions. The major di¤erences between the Khalil paper and the present one are as follows: 1- Khalil has the penalty paid by the agent to the principal (as in Picard, 1996), so that auditing is used by the principal to extract rents from the agent, 2- the agent is risk-neutral and the principal is risk-averse, and 3- there is no state of the world where the game is not played.

It is interesting to note in Khalil (1997), Boyer (1997) and Khalil and Parigi (1998), and in this paper as well that the traditional predictions of the contract literature where the principal can perfectly commit are reversed. Khalil (1997) …nds that the agent overproduces when the principal cannot commit, whereas Baron and Myerson (1982) …nd the opposite. The same applies to Khalil and Parigi (1998) who …nd that the entrepreneur should overinvest (or overborrow) when the …nancier cannot commit to an auditing strategy, compared with underinvestment in Gale and Hellwig (1985). Here I obtain that the agent is overinsured in the high-loss state when the principal cannot commit. This contradicts the general literature on commitment, which shows that higher losses should be undercompensated (see Winter, 1992).

Can these results be generalized? In other words, will non-commitment of the principal to an action always lead to a prediction that is the opposite of the prediction under full commitment: overinsurance vs. underinsurance, overproduction vs. underproduction, overinvestment vs. underinvestment, etc. Although there appears to be a pattern to that e¤ect, no one has shown any direct relationship or the conditions for a direct relationship. Another avenue for generalization would be to look at T possible losses given the occurrence of an accident. In some preliminary work, it seems that there is some overcompensation of losses, and that the expected marginal utilities are still equalized across accident states.

5 References

1. Arrow, K (1970). Essays in the Theory of Risk-Bearing, North Holland, Amsterdam and London.
2. Baron, D.P. and R.B. Myerson (1982). Regulating a Monopolist with Unknown Costs. Econometrica, 50:911-930.
3. Beaudry, P. and M. Poitevin (1993). Signalling and Renegotiation in Contractual Relationships. Econometrica, 61:745-782.
4. Beaudry, P. and M. Poitevin (1994). The Commitment Value of Contracts under Dynamic Renegotiation. Rand Journal of Economics, 25:501-517.
5. Beaudry, P. and M. Poitevin (1995). Competitive Screening in Financial Markets when Borrowers can Recontract. Review of Economic Studies, 62:401-423.
6. 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. Journal of Public Economics, 63:239-254.
7. Boyer, M. M. (1997). Insurance Taxation and Insurance Fraud. Working Paper 97-10, Risk Management Chair, HEC-Université de Montréal.
8. Boyer, M. M. (1998). A Rationale for Borrowing More than Needed. Working Paper 98-05, Risk Management Chair, HEC-Université de Montréal.
9. Crocker, K. J. and A. Snow (1985). The Efficiency Effect of Categorical Discrimination in the Insurance Industry. Journal of Political Economy, 94:321-344.
10. Cooper, R. and B. Hayes (1987). Multi-Period Insurance Policies. International Journal of Industrial Organization, 5:211-231
11. Dionne, G. and N. A. Doherty (1994). Adverse Selection, Commitment, and Renegotiation: Extension to and Evidence from the Insurance Markets. Journal of Political Economy, 102:209-235.
12. Dionne, G. and P. Viala (1992). Optimal Design of Financial Contracts and Moral Hazard. Working paper 9219, Université de Montréal.
13. Fombaron, N. (1997). No-Commitment and Dynamic Contracts in Competitive Insurance Markets with Adverse Selection. Mimeo, THEMA-Université de Paris-X.
14. Gale, D. and M. Hellwig (1985). Incentive-Compatible Debt Contracts: The One-Period Problem. Review of Economic Studies, 52:647-663.
15. Gibbons, R. (1992). Game Theory for Applied Economists. Princeton University Press, 1992.
16. Graetz, M. J., J. F. Reinganum and L. L. Wilde (1986). The Tax-Compliance Game: Toward an Interactive Theory of Law Enforcement. Journal of Law, Economics and Organization, 2:1-32.
17. Grossman, S. J. and O. Hart (1986). The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. Journal of Political Economy, 94:691-719.
18. Holmstrom, B. (1979). Moral Hazard and Observability. Bell Journal of Economics, 10:74-91.

19. Hosios, A. J. and M. Peters (1989). Repeated Insurance Contracts with Adverse Selection and Limited Commitment. Quarterly Journal of Economics, 104:229-253.

20. Hoyt, R. E. (1989). The E¤ect of Insurance Fraud on the Economic System. Journal of Insurance Regulation, 8:304-315.

21. Khalil, F. (1997). Auditing without Commitment. Rand Journal of Economics, 28:629-640.

22. Khalil, F. and B. M. Parigi (1998). Loan Size as a Commitment Device. International Economic Review, forthcoming.

23. La¤ont, J.-J. and J. Tirole (1993). A Theory of Incentives in Procurement and Regulation. MIT Press, Cambridge, MA.

24. Lewis, T. R. (1996). Protecting the Environment When Costs and Bene…ts are Privately Known. Rand Journal of Economics, 27:819-847.

25. Lewis, T. R. and D. E. M. Sappington (1995). Using Markets to Allocate Pollution Permits and Other Scarce Resource Rights under Limited Information. Journal of Public Economics, 57:431-455.

26. Melumad, N. D. and D. Mookherjee (1989). Delegation as Commitment: The Case of Income Tax Audits. Rand Journal of Economics, 20:139-163.

27. Mookherjee, D. and I. Png (1989). Optimal Auditing, Insurance and Redistribution. Quarterly Journal of Economics, 104:205-228.

28. Myerson, R. B. (1979). Incentive Compatibility and the Bargaining Problem. Econometrica, 47:61-73.

29. Myerson, R. B. (1991). Game Theory. Harvard University Press, Cambridge, MA.

30. Picard, P. (1996). Auditing Claims in the Insurance Market with Fraud: The Credibility Issue. Journal of Public Economics, 63:27-56.

31. Picard, P. (1997). On the Design of Optimal Insurance Policies under Manipulation of Audit Cost. Mimeo, Université de Paris-X.

32. Pratt, J. (1964). Risk Aversion in the Small and the Large. Econometrica, 32:122-136.

33. Rand Corporation Institute for Civil Justice. Research Brief, May 1995.

34. Reinganum, J. F. and L. L. Wilde (1985). Income Tax Compliance in a Principal-Agent Framework. Journal of Public Economics, 26:1-18.

35. Scheepens (1995). Bankruptcy Litigation and Optimal Debt Contract. European Journal of Political Economy, 11:535-556.

36. Shavell, S. (1979). On Moral Hazard and Insurance. Quarterly Journal of Economics, 11:541-562.

37. Spence, A. M. and R. Zeckhauser (1971). Insurance, Information, and Individual Action. American Economic Review, 61:380-387.

38. Swierzbinski, J. E. (1994). Guilty until Proven Innocent - Regulation with Costly and Limited Enforcement. Journal of Environmental Economics and Management, 27:127-146.

39. Townsend, R. M. (1979). Optimal Contracts and Competitive Markets with Costly State Verification. Journal of Economic Theory, 21:265-293.

40. Winter, R. A. (1992). Moral Hazard and Insurance Contracts. In Contributions to Insurance Economics, G. Dionne ed., Kluwer Academic Publishers, Boston.

6 Appendix

6.1 Proofs

Proof of lemma 1. For ease of notation, define the perfect bayesian nash equilibrium as $PBNE = (\theta_L; \theta_H; \delta_L; \delta_H; \gamma_L; \gamma_H)$. From the left-hand side of figure 2, it is clear that $\gamma_L = 0$. Suppose Nature chooses the severity of the accident to be high (λ_H). Sending message λ'_H then always dominates sending message λ'_L, whatever the principal does. Also, if the principal hears message λ'_L, then she knows for sure that she is not playing at the upper node of information set 1.L. Consequently the only meaningful strategy for the principal is never to audit since she collects $-c - \beta_L$ if she does, and $-\beta_L$ if she does not. We have now found θ_H, δ_L and γ_L. Let's now move to the right side of figure 2.

Let η be the probability (in the mixed-strategy sense) that the agent sends message λ'_H when Nature chooses $\lambda = \lambda_L$. By Bayes' rule we can find γ_H, the principal's posterior belief that the true loss is high given that the agent sent message λ'_H.

$$\gamma_H = \frac{\rho}{\rho + (1 - \rho)\eta} \tag{25}$$

Only one strategy on the part of the agent makes the principal indifferent as to whether to audit or not. That strategy must be such that γ_H solves

$$(-c - \beta_H)\gamma_H + (-c - \beta_L)(1 - \gamma_H) = -\beta_H \tag{26}$$

and

$$\gamma_H = \frac{(\beta_H - \beta_L) - c}{\beta_H - \beta_L} \tag{27}$$

Substituting γ_H in (25) yields[18]

$$\eta = \left(\frac{c}{\beta_H - \beta_L - c}\right)\left(\frac{\rho}{1 - \rho}\right) \tag{28}$$

All that is left to calculate is the principal's strategy at information set 1.H. Her strategy must be such that the agent is indifferent between telling the truth and lying, given that $\lambda = \lambda_L$. Let ν be the probability (in a mixed-strategy sense) of auditing a λ'_H message. ν must then solve

$$U(Y - \alpha - \lambda_L + \beta_L) = \nu U(Y - \alpha - \lambda_L + \beta_L - k) + (1 - \nu)U(Y - \alpha - \lambda_L + \beta_H) \tag{29}$$

which means that

$$\nu = \frac{U(Y - \alpha - \lambda_L + \beta_H) - U(Y - \alpha - \lambda_L + \beta_L)}{U(Y - \alpha - \lambda_L + \beta_H) - U(Y - \alpha - \lambda_L + \beta_L - k)} \tag{30}$$

[18] Note that we need to assume that $\rho < \frac{\beta_H - \beta_L - c}{\beta_H - \beta_L}$ for $\eta \in [0; 1]$. If not, then the agent always commits fraud if he suffers loss λ_L. A sufficient condition for $\rho < \frac{\beta_H - \beta_L - c}{\beta_H - \beta_L}$ is to assume $\rho < \frac{1}{2}$ since, as we can see in the first-order condition, $\beta_H - \beta_L > 2c$.

Since all six elements of the PBNE have been found, the proof is done. ∎

<u>Proof of proposition 1</u>. Straightforward: $\frac{\partial F}{\partial(\beta_H-\beta_L)} = \pi(1-\rho)\frac{\partial \eta}{\partial(\beta_H-\beta_L)} < 0$. ∎

<u>Proof of proposition 2</u>. Straightforward if one takes $\frac{\partial(1-\gamma)}{\partial\beta_H}$ and $\frac{\partial(1-\gamma)}{\partial\beta_L}$ and rearranges the terms. ∎

<u>Proof of lemma 2</u>. The first-order conditions of SP are

$$\begin{aligned}\frac{\partial V}{\partial\beta_H} = 0 = & -\alpha'_H(1-\pi)U'\left(Y-\pi\left[\beta_L+\rho\frac{(\beta_H-\beta_L)^2}{\beta_H-\beta_L-c}\right]\right) \\ & +\left(1-\alpha'_H\right)\pi\rho U'\left(Y-\pi\left[\beta_L+\rho\frac{(\beta_H-\beta_L)^2}{\beta_H-\beta_L-c}\right]-\lambda_H+\beta_H\right) \\ & -\alpha'_H\pi(1-\rho)U'\left(Y-\pi\left[\beta_L+\rho\frac{(\beta_H-\beta_L)^2}{\beta_H-\beta_L-c}\right]-\lambda_L+\beta_L\right)\end{aligned} \tag{FOC$_H$}$$

and

$$\begin{aligned}\frac{\partial V}{\partial\beta_L} = 0 = & -\alpha'_L(1-\pi)U'\left(Y-\pi\left[\beta_L+\rho\frac{(\beta_H-\beta_L)^2}{\beta_H-\beta_L-c}\right]\right) \\ & -\alpha'_L\pi\rho U'\left(Y-\pi\left[\beta_L+\rho\frac{(\beta_H-\beta_L)^2}{\beta_H-\beta_L-c}\right]-\lambda_H+\beta_H\right) \\ & +\left(1-\alpha'_L\right)\pi(1-\rho)U'\left(Y-\pi\left[\beta_L+\rho\frac{(\beta_H-\beta_L)^2}{\beta_H-\beta_L-c}\right]-\lambda_L+\beta_L\right)\end{aligned} \tag{FOC$_L$}$$

From FOC_H, we obtain

$$\pi\rho U'(Y-\alpha-\lambda_H+\beta_H) = \alpha'_H\left[\begin{array}{c}(1-\pi)U'(Y-\alpha)+\pi\rho U'(Y-\alpha-\lambda_H+\beta_H)\\ +\pi(1-\rho)U'(Y-\alpha-\lambda_L+\beta_L)\end{array}\right] \tag{31}$$

Substituting for EV' and dividing everywhere by it yields the desired result. The same technique is used with FOC_L, except that we find that

$$\pi(1-\rho)U'(Y-\alpha-\lambda_L+\beta_L) = \alpha'_L\left[\begin{array}{c}(1-\pi)U'(Y-\alpha)+\pi\rho U'(Y-\alpha-\lambda_H+\beta_H)\\ +\pi(1-\rho)U'(Y-\alpha-\lambda_L+\beta_L)\end{array}\right] \tag{32}$$

It is left to the reader to verify that $\alpha'_L = \pi - \alpha'_H$. ∎

<u>Proof of theorem 1</u>. By substituting NC1 into NC2,

$$\pi(1-\rho)\frac{U'(Y-\alpha-\lambda_L+\beta_L)}{EV'} = \pi - \pi\rho\frac{U'(Y-\alpha-\lambda_H+\beta_H)}{EV'} \tag{33}$$

Simplifying π and multiplying everywhere by EV' yields

$$(1-\rho)U'(Y-\alpha-\lambda_L+\beta_L) = EV' - \rho U'(Y-\alpha-\lambda_H+\beta_H) \tag{34}$$

Expanding EV', regrouping terms and simplifying yield

$$(1-\rho)U'(Y-\alpha-\lambda_L+\beta_L)+\rho U'(Y-\alpha-\lambda_H+\beta_H)=U'(Y-\alpha) \tag{35}$$

The left-hand side is the expected marginal utility in the accident state, while the right-hand side is the marginal utility in the no-accident state. $\square$

<u>Proof of corollary 1.</u> From NC1, it is clear that $\frac{\alpha'_H}{\pi\rho}<1$ since

$$\frac{(\beta_H-\beta_L)(\beta_H-\beta_L-2c)}{(\beta_H-\beta_L-c)^2}<1 \tag{36}$$

when $c>0$. Thus

$$\frac{U'(Y-\alpha-\lambda_H+\beta_H)}{EV'}<1 \tag{37}$$

It is also clear from NC2 that $\frac{\pi-\alpha'_H}{\pi(1-\rho)}>1$. Thus

$$\frac{U'(Y-\alpha-\lambda_L+\beta_L)}{EV'}>1 \tag{38}$$

Combining (37) and (38), we find that $\beta_H-\lambda_H>\beta_L-\lambda_L$. It follows that $\beta_H>\beta_L$. $\square$

<u>Proof of theorem 2</u>: From theorem 1 we know that expected marginal utilities are equalized across accident states. Thus

$$\rho=\frac{U'(Y-\alpha)-U'(Y-\alpha-\lambda_L+\beta_L)}{U'(Y-\alpha-\lambda_H+\beta_H)-U'(Y-\alpha-\lambda_L+\beta_L)} \tag{39}$$

It is clear that the denominator on the right is negative from corollary 1. Thus the numerator must also be negative. This occurs if and only if $\beta_L<\lambda_L$. Similarly, from theorem 1 we have

$$1-\rho=\frac{U'(Y-\alpha)-U'(Y-\alpha-\lambda_H+\beta_H)}{U'(Y-\alpha-\lambda_L+\beta_L)-U'(Y-\alpha-\lambda_H+\beta_H)} \tag{40}$$

Here, it is clear that the denominator on the right is positive. Thus the numerator must also be positive. This occurs if and only if $\beta_H>\lambda_H$. $\square$

<u>Proof of proposition 3</u>. The graphic proof is straightforward. We know that the agent is risk-averse ($U'>0$, $U''<0$). If U''' is positive (negative), then the marginal utility function is decreasing and convex (concave). Since U' is convex (concave), then the average of the marginal utilities of wealth is larger (smaller) than the marginal utility of the average wealth. Thus the wealth in the accident state is larger (smaller) than the expected wealth in the no-accident state. $\square$

<u>Proof of corollary 2</u>. Expected wealth in the accident state is given by

$$\rho(Y-\alpha-\lambda_L+\beta_L)+(1-\rho)(Y-\alpha-\lambda_H+\beta_H) \tag{41}$$

Wealth in the no-accident state is given by $Y - \alpha$. Expected wealth in the accident state is greater if and only if

$$\rho(-\lambda_L + \beta_L) + (1-\rho)(-\lambda_H + \beta_H) > 0 \tag{42}$$

This inequality holds if and only if

$$\rho\beta_L + (1-\rho)\beta_H > \rho\lambda_L + (1-\rho)\lambda_H \tag{43}$$

Thus the expected benefit is greater than the expected loss. □

Proof of theorem 3. We get the following six first-order conditions when the principal can commit

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \beta_L} &= 0 = \pi(1-\rho)U'(Y-\alpha-\lambda_L+\beta_L) + \mu_1\pi(1-\rho) \\ &\quad - \mu_2 \frac{U'(Y-\alpha-\lambda_L+\beta_L)[U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L-k)]}{[U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L-k)]^2} \\ &\quad + \mu_2 \frac{U'(Y-\alpha-\lambda_L+\beta_L-k)[U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L)]}{[U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L-k)]^2}\end{aligned} \tag{44}$$

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \beta_H} &= 0 = \pi\rho U'(Y-\alpha-\lambda_H+\beta_H) + \mu_1\pi\rho \\ &\quad + \mu_2 \frac{U'(Y-\alpha-\lambda_L+\beta_H)[U(Y-\alpha-\lambda_L+\beta_L) - U(Y-\alpha-\lambda_L+\beta_L-k)]}{[U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L-k)]^2}\end{aligned} \tag{45}$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = 0 = \mu_1\pi\rho c - \mu_2 \tag{46}$$

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \alpha} &= 0 = -(1-\pi)U'(Y-\alpha) - \pi\rho U'(Y-\alpha-\lambda_H+\beta_H) \\ &\quad - \pi(1-\rho)U'(Y-\alpha-\lambda_L+\beta_L) - \mu_1 \\ &\quad - \mu_2 \frac{\left[\begin{array}{c} U'(Y-\alpha-\lambda_L+\beta_H) \\ -U'(Y-\alpha-\lambda_L+\beta_L)\end{array}\right]\left[\begin{array}{c} U(Y-\alpha-\lambda_L+\beta_H) \\ -U(Y-\alpha-\lambda_L+\beta_L-k)\end{array}\right]}{[U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L-k)]^2} \\ &\quad + \mu_2 \frac{\left[\begin{array}{c} U(Y-\alpha-\lambda_L+\beta_H) \\ -U(Y-\alpha-\lambda_L+\beta_L)\end{array}\right]\left[\begin{array}{c} U'(Y-\alpha-\lambda_L+\beta_H) \\ -U'(Y-\alpha-\lambda_L+\beta_L-k)\end{array}\right]}{[U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L-k)]^2}\end{aligned} \tag{47}$$

$$\frac{\partial \mathcal{L}}{\partial \mu_1} = 0 = -[\alpha - \pi[\rho\beta_H + (1-\rho)\beta_L] - \pi\rho\gamma c] \tag{48}$$

$$\frac{\partial \mathcal{L}}{\partial \mu_2} = 0 = -\left(\gamma - \frac{U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L)}{U(Y-\alpha-\lambda_L+\beta_H) - U(Y-\alpha-\lambda_L+\beta_L-k)}\right) \tag{49}$$

Making substitutions, we obtain

$$(1-\rho)U'(A-\alpha-\lambda_L+\beta_L)+\rho U'(A-\alpha-\lambda_H+\beta_H)=U'(A-\alpha) \tag{50}$$

at the optimum. This means that expected marginal utilities are equalized across accident states, just as in the no-commitment case. Therefore, whether the principal can commit or not does not affect the fact that the agent will be fully insured (although he may not smooth his income perfectly). We then have that

$$\frac{U'(A-\alpha-\lambda_H+\beta_H)}{U'(A-\alpha)}=(1+c\Sigma) \tag{51}$$

and

$$\frac{U'(A-\alpha-\lambda_L+\beta_L)}{U'(A-\alpha)}=\frac{1}{(1-\rho)}-\frac{\rho}{(1-\rho)}(1+c\Sigma) \tag{52}$$

where

$$\Sigma=\frac{U'(A-\alpha-\lambda_L+\beta_H)[U(A-\alpha-\lambda_L+\beta_L)-U(A-\alpha-\lambda_L+\beta_L-k)]}{[U(A-\alpha-\lambda_L+\beta_H)-U(A-\alpha-\lambda_L+\beta_L-k)]^2} \tag{53}$$

We then find that

$$U'(A-\alpha-\lambda_L+\beta_L)<U'(A-\alpha-\lambda_H+\beta_H) \tag{54}$$

if and only if

$$\frac{1}{(1-\rho)}-\frac{\rho}{(1-\rho)}(1+c\Sigma)<(1+c\Sigma) \tag{55}$$

which occurs whenever $c\Sigma>0$ which it is since Σ and c are positive. We can therefore conclude that the agent gets greater marginal utility in the high-loss state. Thus $\beta_H-\lambda_H<\beta_L-\lambda_L$. Finally using theorem 2 we find that $\beta_H<\lambda_H$ and $\beta_L>\lambda_L$. $\square$

Proof of corollary 3. Denote by **c** the case under commitment and by **f** the case under no-commitment. From the zero-profit constraints, we know that

$$\Pi^{c}=\pi\left[\rho\bar{\beta}^{c}_H+(1-\rho)\bar{\beta}^{c}_L\right]+\pi\rho\bar{\beta}c \tag{56}$$

$$\Pi^{f}=\pi\left[\bar{\beta}^{f}_L+\rho\frac{\left(\bar{\beta}^{f}_H-\bar{\beta}^{f}_L\right)^2}{\bar{\beta}^{f}_H-\bar{\beta}^{f}_L-c}\right] \tag{57}$$

The goal is to show that $\Pi^{c}<\Pi^{f}$. This occurs if and only if

$$\bar{\beta}<\frac{1}{\rho c}\left[\bar{\beta}^{f}_L+\rho\frac{\left(\bar{\beta}^{f}_H-\bar{\beta}^{f}_L\right)^2}{\bar{\beta}^{f}_H-\bar{\beta}^{f}_L-c}-\rho\bar{\beta}^{c}_H-(1-\rho)\bar{\beta}^{c}_L\right] \tag{58}$$

Evaluating ⓐ at $\left(\hat{e}_L; \hat{e}_H\right) = \left(\hat{b}_L; \hat{b}_H\right) = (\beta_L; \beta_H)$, ⓑ < ⓐ if and only if

$$\text{ⓑ} < \frac{\beta_H - \beta_L}{\beta_H - \beta_L - c} \tag{59}$$

which is always true since $\frac{\beta_H - \beta_L}{\beta_H - \beta_L - c} > 1$. Thus ⓑ < ⓐ.

The second part of the uses the fact that ⓑ < ⓐ. Since marginal utilities are equalized across accident states in both the commitment and the no-commitment cases, the agent's expected marginal utility when the principal can commit is given by $U^0(Y - \text{ⓑ})$, whereas his expected marginal utility when the principal cannot commit is given by $U^0(Y - \text{ⓐ})$. Since ⓑ < ⓐ, it follows that $U^0(Y - \text{ⓑ}) < U^0(Y - \text{ⓐ})$, which means that $U(Y - \text{ⓑ}) > U(Y - \text{ⓐ})$. □

Proof of lemma 3. PART 1: $\beta_H = \beta_L$. Since the compensation is the same in every state of the world, there is no need for the agent to cheat since he gains nothing by doing so. The insurer on the other hand has nothing to gain by auditing quite the contrary. By auditing the insurer incurs a cost and she gains nothing. Therefore her only rational strategy is never to audit. Finally, since she knows that the agent is indifferent about revealing the truth, she assigns any belief to the information set.

PART 2: $\beta_H < \beta_L$. This part of the proof is done exactly as the proof of lemma 1 and in the same fashion as part 1. However, instead of having the agent exaggerate his loss, we have him understate his loss.[19] The principal on the other hand will never audit a report of a high loss, because that is the lowest payment she can ever make in the case of an accident. Going through the same steps as for part 1 yields the result presented in the statement of the theorem. □

Proof of proposition 4. The proof of the proposition has three parts. First, I show for some value of the auditing cost that the agent is better off with the variable payment. I then show for some other value that the agent is better off with the fixed payment. I complete the proof by showing that the utility function is continuous in the cost of auditing, which means, from the mean value theorem, that there is a critical value where the agent's preferences change from one system to the next.

The goal in the first part is to show there exist a value of c such that the agent prefers a variable

[19] As for part 1 of the proof, for the reporting strategy to make any sense we need to assume that $\rho > \frac{c}{\beta_L - \beta_H}$; otherwise the agent will always cheat.

payment to a fixed payment. In other words, we want to find a c such that

$$\begin{array}{c}(1-\pi)U(Y-\pi\beta^*) \\ +\pi\rho U\left(Y+(1-\pi)\beta^*-\lambda_H\right) \\ +\pi(1-\rho)U\left(Y+(1-\pi)\beta^*-\lambda_L\right)\end{array} < \begin{array}{c}(1-\pi)U\left(Y-\pi\left(\beta^*_L+\rho\frac{(\beta^*_H-\beta^*_L)^2}{\beta^*_H-\beta^*_L-c}\right)\right) \\ +\pi\rho U\left(Y-\pi\left(\beta^*_L+\rho\frac{(\beta^*_H-\beta^*_L)^2}{\beta^*_H-\beta^*_L-c}\right)-\lambda_H+\beta^*_H\right) \\ +\pi(1-\rho)U\left(Y-\pi\left(\beta^*_L+\rho\frac{(\beta^*_H-\beta^*_L)^2}{\beta^*_H-\beta^*_L-c}\right)-\lambda_L+\beta^*_L\right)\end{array} \tag{60}$$

It is clear that the inequality holds if the cost of auditing is very small. Specifically, when $c \to 0$, $\pi\left(\beta^*_L+\rho\frac{(\beta^*_H-\beta^*_L)^2}{\beta^*_H-\beta^*_L-c}\right) = \pi[(1-\rho)\beta^*_L+\rho\beta^*_H]$, which is the premium paid under full information. It is then easy to show that the optimal contract is such that $\beta^*_H = \lambda_H$ and $\beta^*_L = \lambda_L$. Clearly, the agent is then better off with a variable compensation rather than a fixed compensation.

For the second part of the proof, I want to show that there exist a value of c such that the agent prefers a fixed payment to a variable payment. Instead of going directly through the expected utility, I will use marginal utilities. If the expected utility when $\beta_L = \beta_H$ is greater than the expected utility when $\beta_L < \beta_H$, then the expected marginal utility with $\beta_L = \beta_H$ is smaller than the expected marginal utility when $\beta_L < \beta_H$. Using the fact that in each case the expected marginal utility is the same in the accident state as in the no-accident state, I want to find the conditions on c such that the following inequality holds

$$U'(Y-\pi\beta^*) < U'\left(Y-\pi\left[\beta^*_L+\rho\frac{(\beta^*_H-\beta^*_L)^2}{\beta^*_H-\beta^*_L-c}\right]\right) \tag{61}$$

It is straightforward to show that as $c \to \infty$, $\beta^*_H - \beta^*_L \to \frac{c}{1-\rho}$. Taking the limit on the right-hand side yields that the agent will prefer the fixed payment to the variable payment when the cost of auditing is extremely high if and only if

$$U'(Y-\pi\beta^*) < U'\left(Y-\pi\left(\beta^*_L+\frac{c}{1-\rho}\right)\right) \tag{62}$$

But taking the limit on the right-hand side as $c \to \infty$, we find that

$$U'(Y-\pi\beta^*) < U'(0) \tag{63}$$

Since by assumption $U'(0) = \infty$, it is clear that this inequality holds. We can thus conclude that as $c \to \infty$, the expected utility of the fixed payment is greater than the expected utility of the variable payment.

To conclude the proof, we need to show that the expected utility is continuous in c and monotone: i.e., $\text{sign}\left(\frac{\partial EU(\cdot)}{\partial c}\right) = \text{cst}$. Letting

$$\Omega = (1-\pi)U(Y-\alpha) + \pi\rho U(Y-\alpha-\lambda_H+\beta^*_H) + \pi(1-\rho)U(Y-\alpha-\lambda_L+\beta^*_L) \tag{64}$$

and taking the total derivative with respect to the cost of auditing yields

$$\frac{d\Omega}{dc} = -(1-\pi)\frac{d\alpha}{dc}U'(Y-\alpha) + \pi\rho\left(\frac{\partial\beta_H}{\partial c} - \frac{d\alpha}{dc}\right)U'(Y-\alpha-\lambda_H+\beta_H) + \pi(1-\rho)\left(\frac{\partial\beta_L}{\partial c} - \frac{d\alpha}{dc}\right)U'(Y-\alpha-\lambda_L+\beta_L) \quad (65)$$

Expanding $\frac{d\alpha}{dc}$ and letting

$$\Phi = \frac{\partial\alpha}{\partial c} + \frac{\partial\alpha}{\partial\beta_H}\frac{\partial\beta_H}{\partial c} + \frac{\partial\alpha}{\partial\beta_L}\frac{\partial\beta_L}{\partial c} \quad (66)$$

we obtain

$$\frac{d\Omega}{dc} = -(1-\pi)\Phi U'(Y-\alpha) + \pi\rho\left(\frac{\partial\beta_H}{\partial c} - \Phi\right)U'(Y-\alpha-\lambda_H+\beta_H) + \pi(1-\rho)\left(\frac{\partial\beta_L}{\partial c} - \Phi\right)U'(Y-\alpha-\lambda_L+\beta_L) \quad (67)$$

Combining terms yields

$$\frac{d\Omega}{dc} = -\Phi(1-\pi)U'(Y-\alpha) - \Phi\pi\left[\rho U'(Y-\alpha-\lambda_H+\beta_H) + (1-\rho)U'(Y-\alpha-\lambda_L+\beta_L)\right] + \pi\rho\frac{\partial\beta_H}{\partial c}U'(Y-\alpha-\lambda_H+\beta_H) + \pi(1-\rho)\frac{\partial\beta_L}{\partial c}U'(Y-\alpha-\lambda_L+\beta_L) \quad (68)$$

Recall from theorem 1 that

$$U'(Y-\alpha) = \rho U'(Y-\alpha-\lambda_H+\beta_H) + (1-\rho)U'(Y-\alpha-\lambda_L+\beta_L) \quad (69)$$

and from NC1 and NC2 that

$$U'(Y-\alpha-\lambda_H+\beta_H) = U'(Y-\alpha)\frac{\partial\alpha}{\partial\beta_H}\frac{1}{\pi\rho} \quad (70)$$

$$U'(Y-\alpha-\lambda_L+\beta_L) = U'(Y-\alpha)\frac{\partial\alpha}{\partial\beta_L}\frac{1}{\pi(1-\rho)} \quad (71)$$

we can simplify the first two lines of $\frac{d\Omega}{dc}$ to $-\Phi U'(Y-\alpha)$ and find

$$\frac{d\Omega}{dc} = -\Phi U'(Y-\alpha) + \frac{\partial\beta_H}{\partial c}U'(Y-\alpha)\frac{\partial\alpha}{\partial\beta_H} + \frac{\partial\beta_L}{\partial c}U'(Y-\alpha)\frac{\partial\alpha}{\partial\beta_L} \quad (72)$$

Recall that $\frac{\partial\alpha}{\partial\beta_L} = \pi - \frac{\partial\alpha}{\partial\beta_H}$. Factoring out $U'(Y-\alpha)$ we obtain

$$\frac{d\Omega}{dc} = U'(Y-\alpha)\left[-\Phi + \left(\frac{\partial\beta_H}{\partial c} - \frac{\partial\beta_L}{\partial c}\right)\frac{\partial\alpha}{\partial\beta_H} + \pi\frac{\partial\beta_L}{\partial c}\right] \quad (73)$$

Expanding Ψ and substituting $\frac{\partial\alpha}{\partial\beta_L} = \pi - \frac{\partial\alpha}{\partial\beta_H}$ in it yields

$$\frac{d\Psi}{dc} = U'(Y - \alpha)\left[\begin{array}{c} -\frac{\partial\alpha}{\partial c} + \frac{\partial\alpha}{\partial\beta_H}\left(\frac{\partial\beta_L}{\partial c} - \frac{\partial\beta_H}{\partial c}\right) - \pi\frac{\partial\beta_L}{\partial c} \\ + \left(\frac{\partial\beta_H}{\partial c} - \frac{\partial\beta_L}{\partial c}\right)\frac{\partial\alpha}{\partial\beta_H} + \pi\frac{\partial\beta_L}{\partial c} \end{array}\right] \quad (74)$$

Simplifying finally yields

$$\frac{d\Psi}{dc} = -U'(Y - \alpha)\frac{\partial\alpha}{\partial c} \quad (75)$$

Since $\frac{\partial\alpha}{\partial c}$ is positive for all c, we conclude that the agent's expected utility is monotone and decreasing for all c. The limit $\tilde{c}$ is such that

$$U'\left(Y - \pi\left[\beta^*_L + \rho\frac{(\beta^*_H - \beta^*_L)^2}{\beta^*_H - \beta^*_L - \tilde{c}}\right]\right) > U'(Y - \pi\beta^*) \quad (76)$$

This completes the proof. $\square$

Proof of proposition 5. This proof is done in a similar manner to that of proposition 4. When $\beta_H < \beta_L$, the optimal contract is the solution to[20]

$$\max_{\alpha,\beta_L,\beta_H} V = (1-\pi)U(Y-\alpha) + \pi\rho U(Y - \alpha - \lambda_H + \beta_H) + \pi(1-\rho)U(Y - \alpha - \lambda_L + \beta_L) \quad (77)$$

subject to

$$\alpha = \pi\left[\beta_H + (1-\rho)\frac{(\beta_L - \beta_H)^2}{\beta_L - \beta_H - c}\right] \quad (78)$$

What we want to show is that

$$\begin{array}{c} (1-\pi)U(Y - \pi\beta^*) \\ +\pi\rho U(Y + (1-\pi)\beta^* - \lambda_H) \\ +\pi(1-\rho)U(Y + (1-\pi)\beta^* - \lambda_L) \end{array} > \begin{array}{c} (1-\pi)U\left(Y - \pi\left[\beta^*_H + (1-\rho)\frac{(\beta^*_L - \beta^*_H)^2}{\beta^*_L - \beta^*_H - c}\right]\right) \\ +\pi\rho U\left(Y - \pi\left[\beta^*_H + (1-\rho)\frac{(\beta^*_L - \beta^*_H)^2}{\beta^*_L - \beta^*_H - c}\right] - \lambda_H + \beta^*_H\right) \\ +\pi(1-\rho)U\left(Y - \pi\left[\beta^*_H + (1-\rho)\frac{(\beta^*_L - \beta^*_H)^2}{\beta^*_L - \beta^*_H - c}\right] - \lambda_L + \beta^*_L\right) \end{array} \quad (79)$$

for any value of c. Suppose $c \to \infty$. Using marginal utilities as in the proof of proposition 4, and using the fact that the expected marginal utility is the same in the accident state as in the no-accident state, the (79) holds if and only if

$$U'(Y - \pi\beta^*) < U'\left(Y - \pi\left[\beta^*_H + (1-\rho)\frac{(\beta^*_L - \beta^*_H)^2}{\beta^*_L - \beta^*_H - c}\right]\right) \quad (80)$$

It is straightforward to show that as $c \to \infty$, $\beta^*_L - \beta^*_H \to \frac{c}{\rho} \to \infty$. Taking the limit on the right-hand side, the agent prefers the fixed payment to the variable payment as $c \to \infty$ if and only if

$$U'(Y - \pi\beta^*) < U'(0) \quad (81)$$

[20] I have already simplified the Nash equilibrium constraints and the participation constraint.

But since by assumption $U'(0) = \infty$, we can conclude that the expected utility of the fixed payment is greater than the expected utility of the variable payment. Now, suppose that $c \to 0$ the premium becomes

$$\alpha = \pi\left[\beta^*_H + (1-\rho)\left(\beta^*_L - \beta^*_H\right)\right] \tag{82}$$

In this case we know that the agent prefers $\lambda_H = \beta^*_H > \beta^*_L = \lambda_L$. However, this option is not possible since $\beta^*_H < \beta^*_L$ was assumed. The best the agent can do is $\beta^*_H = \beta^*_L$, which is the definition of a flat payment.

The last thing to show is that

$$\begin{aligned} EU(\cdot) \;=\; & (1-\pi)U\left(Y - \pi\left[\beta^*_H + (1-\rho)\frac{(\beta^*_L - \beta^*_H)^2}{\beta^*_L - \beta^*_H - c}\right]\right) \\ & +\pi\rho U\left(Y - \pi\left[\beta^*_H + (1-\rho)\frac{(\beta^*_L - \beta^*_H)^2}{\beta^*_L - \beta^*_H - c}\right] - \lambda_H + \beta^*_H\right) \\ & +\pi(1-\rho)U\left(Y - \pi\left[\beta^*_H + (1-\rho)\frac{(\beta^*_L - \beta^*_H)^2}{\beta^*_L - \beta^*_H - c}\right] - \lambda_L + \beta^*_L\right) \end{aligned} \tag{83}$$

is continuous and monotone in c. Going through equations (64) to (75) again shows that $\frac{\partial EU(\cdot)}{\partial c} < 0$ for all c. $\square$

6.2 Tables and Figures

Table 1

Summary of all the variables used in this paper

Variable name	Signification
Y	Agent's initial wealth
β_L	Benefit in case of a low loss
β_H	Benefit in case of a high loss
α	Premium
λ_L	Size of the low loss
λ_H	Size of the high loss
π	Probability of an accident occurring
ρ	Probability that the accident results in a high loss
c	Cost of auditing
k	Penalty for committing fraud
δ	Agent's probability of committing fraud given the loss is low
γ	Principal's probability of auditing given that the reported loss is high

Table 2

Monetary payoffs to the agent and the principal contingent on their actions and the state of the world.

State of the world	Action of Agent	Action of Principal	Payoff to Agent	Payoff to Principal
No accident	-	-	$Y - \alpha$	α
Low loss	Tell truth	Audits	$Y - \alpha - \lambda_L + \beta_L$	$\alpha - \beta_L - c$
Low loss	Tell truth	Does not audit	$Y - \alpha - \lambda_L + \beta_L$	$\alpha - \beta_L$
Low loss	Lie	Audits	$Y - \alpha - \lambda_L + \beta_L - k$	$\alpha - \beta_L - c$
Low loss	Lie	Does not audit	$Y - \alpha - \lambda_L + \beta_H$	$\alpha - \beta_H$
High loss	Tell truth	Audits	$Y - \alpha - \lambda_H + \beta_H$	$\alpha - \beta_H - c$
High loss	Tell truth	Does not audit	$Y - \alpha - \lambda_H + \beta_H$	$\alpha - \beta_H$
High loss	Lie	Audits	$Y - \alpha - \lambda_H + \beta_H$	$\alpha - \beta_H - c$
High loss	Lie	Does not audit	$Y - \alpha - \lambda_H + \beta_L$	$\alpha - \beta_L$

The contingent states in italics never occur in equilibrium: they represent actions that are off the equilibrium path.

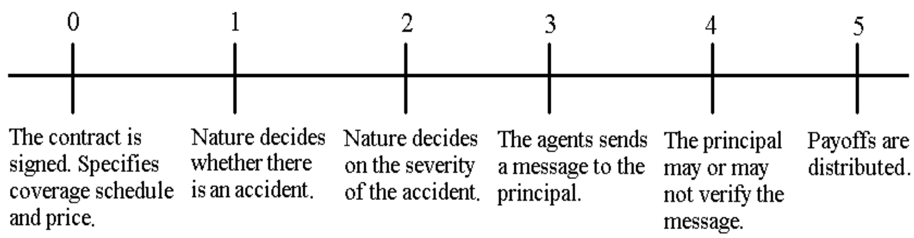

Figure 1: Sequence of play.

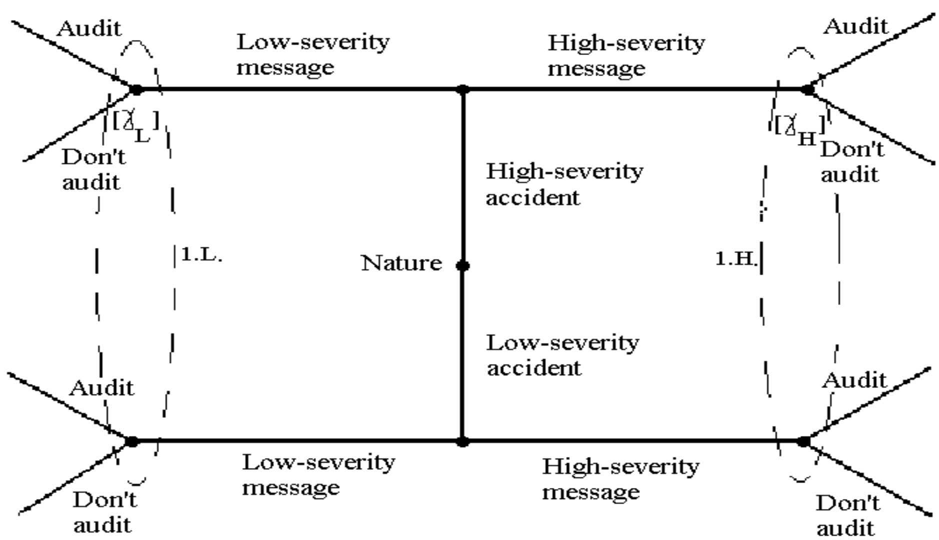

Figure 2: Extensive form of the claiming game

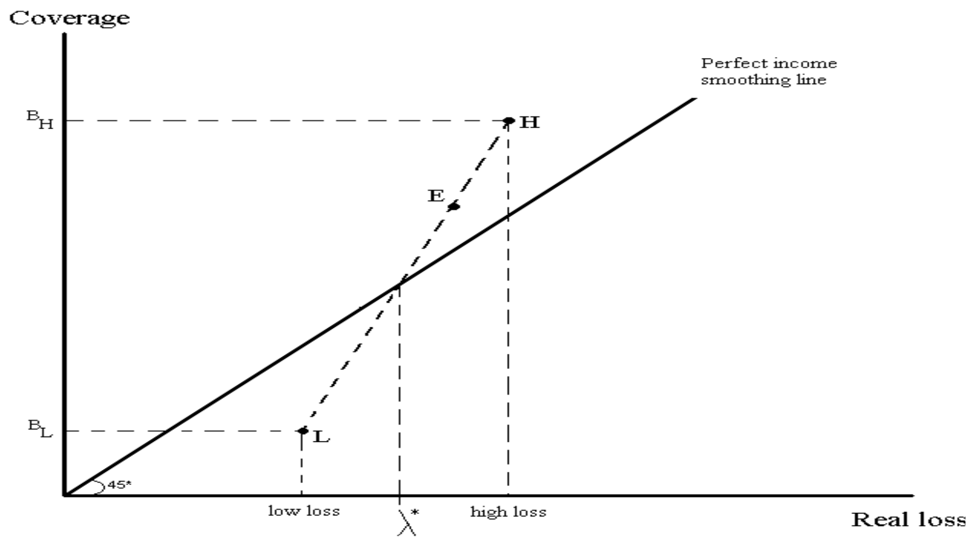

Figure 3: Optimal contract form

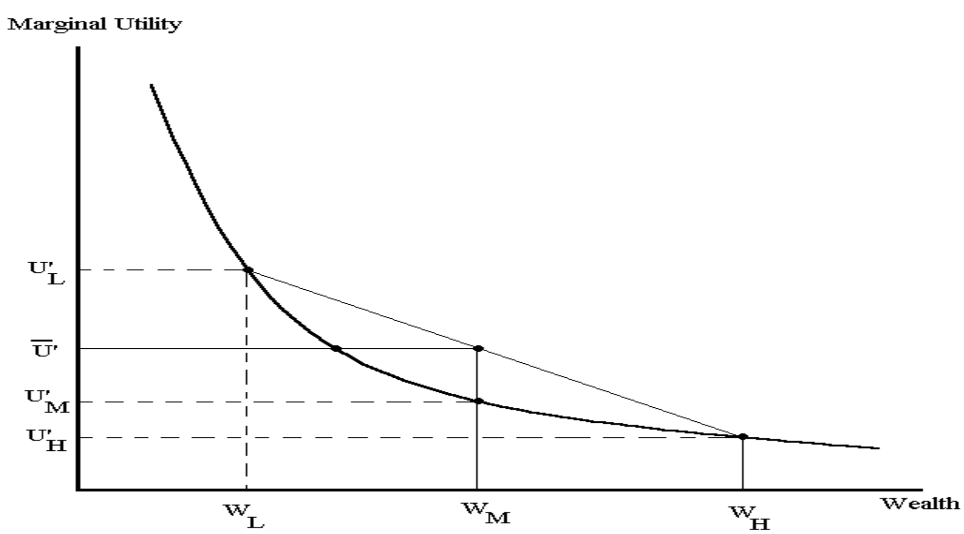

Figure 4: Marginal utility of wealth