The Impact of Insurance Fraud Detection Systems

Jörg Schiller
Institute for Risk and Insurance
University of Hamburg, Germany
phone:+49 40 428386143, fax: +49 40 428385505
schiller@econ.uni-hamburg.de

First Draft: October 2002
This Draft: August 2004

Abstract

This paper analyzes the impact of detection systems in an insurance fraud context. In a non-commitment Costly State Verification setting insurers can only detect fraudulent claims by performing costly audits, and policyholders are over-compensated by the optimal insurance contract. We show that auditing becomes more effective and over-compensation can be reduced, when insurers are able to condition their audits on the information provided by detection systems.

Key words: Insurance fraud, Non-commitment, Auditing, Detection systems.

JEL-Classification: D 82, G 22

The Impact of Insurance Fraud Detection Systems

1 Introduction

Fraud is a well-known phenomenon in insurance markets, since policyholders usually have private information about the occurrence of an insured loss. The Costly State Verification approach (Townsend, 1979) concentrates on fraud situations where the principal is able to verify reports of the agent through costly audits. Consequently, the model focuses on two incentive devices: explicit contracts and costly audits. The range of economic applications is manifold. For instance, Gale and Hellwig (1985) consider credit relationships, whereas Mookherjee and Png (1989) refer to taxation problems.

Since the above-mentioned papers regard auditing problems with commitment, the revelation principle applies, and therefore, the agent truthfully reports his private information. As a consequence, insurance fraud can be prevented entirely. However, in many situations the principal is not able to credibly commit herself ex ante to an audit strategy. Although some papers, as for example, Melumad and Mookherjee (1989) or Picard (1996), propose solutions for the non-commitment problem, these approaches are not universally applicable. Consequently, the insurer's inability to credibly commit itself causes an inevitable market inefficiency, because policyholders will have an incentive to make fraudulent reports, as Picard (1996) and Boyer (2000) show. In such a situation the contract between the principal and the agent can partly serve as a strategic incentive device. More precisely, the over-compensation of agents is the only possibility to limit fraud, as was previously shown by Khahil (1997) and Boyer (1999).

This paper concentrates on the impact of fraud detection systems on the auditing procedure of insurers and on the over-compensation of policyholders. While some papers, like Belhadji et al. (2000) and Artís et al. (2002), analyzed fraud detection methods, to the best of our knowledge only Dionne et al. (2003) look at an auditing model with fraud detection. This is surprising, because detection systems are very popular and effective in existing insurance markets. For instance, in Germany a reinsurer and a software provider have jointly developed various detection systems, in cooperation with more than ten primary insurers, for different lines of insurance, like the property, liability, and auto insurance. Due to the system, the know-how of fraud experts can be duplicated and suspicious claims can be identified more easily. Furthermore, an insurer can concentrate on more suspicious claims and can therefore improve the effectiveness of its audits.[1] First of all,

[1] More detailed information about the German "Intelligent Claims Evaluation" systems can be obtained at http://www.genre.com. A similar system, "ISO ClaimSearch", in the

these lines in which detection systems are used are evidently susceptible to insurance fraud. But, more importantly, they usually contain a considerable number of policyholders. As we show, the market size, or more precisely, the number of policyholders are crucial for the use of a detection system that causes significant fixed cost.[2] In this paper, we consider a fraud detection system which provides soft information about the true state of the world. This information is modeled similar to Holmström (1979) as an additional signal that cannot be intentionally manipulated by the policyholder.[3]

Our main goal is to analyze the following problems: Which effects does a detection system have on the auditing game and the underlying insurance contract? How can a detection system with fixed cost be implemented in a competitive insurance market without any market intervention? As the first Proposition indicates, an informative system leads, ceteris paribus, to a lower fraud and audit probability, which is quite intuitive. As a further consequence, the over-compensation of agents can be reduced after the implementation of an informative detection system, because auditing becomes more effective. The welfare gains due to a system depend on the quality of the signal. Finally, we explore conditions under which a fraud detection system will be applied in a competitive insurance market. Extending the work by Picard (1996) and Boyer (2000), we give a new motivation for the use of an external party, such as a private supplier or an Insurance Fraud Bureau. The important role of the third party in our model is to transform the prevailing fixed implementation cost of a fraud detection system, which can lead to a market breakdown if firms compete in prices a la Bertrand, into variable cost.

The remainder of the paper is organized as follows. In section 2, some important results of Boyer (2000) are briefly summarized. His model represents the reference case without fraud detection. Section 3 concentrates on the effects of a fraud detection system on the equilibrium of the audit game and the underlying insurance contract. Furthermore, in section 4 we derive conditions under which a system will be implemented and show how the fraud detection should be organized. We conclude the analysis in section 5 with some final remarks and a brief outlook.

United States is operated by the Insurance Services Office, Inc. for auto, property and liability claims. For further information about this system, see http://www.iso.com.

[2]The implementation cost of the "ISO ClaimSearch" detection system is 10 to 12 million USD. The per annum operating cost is more than twice the implementation cost (see PANKO, Ron, 2001, "Making A Dent In Auto Insurance Fraud," *Best's Review*, October 2001).

[3]The presented model is in that way not a signaling game, because the signal is an exogenous statistical information generated from the fraud report. The policyholder is assumed to be unable to strategically influence the signal.

2 Equilibrium without fraud detection

The following framework is assumed. There are N homogenous and risk-averse agents (policyholders) with the same attitude towards insurance fraud[4] and the same initial income $Y > 0$. Each policyholder possesses a continuous and twice differentiable utility function $u(W)$ of final wealth W, with $u'(W) > 0$, $u''(W) < 0$, and maximizes an expected utility function U. There are only two states of the world, "no accident" (ω_0) and "accident" (ω_1). The probability of an accident is π and the loss in the state of an accident is $L > 0$. It is assumed that the loss is not higher than the policyholder's initial income $L < Y$, and that insured policyholders have private information about the state of the world.

Additionally, we assume an insurance market with free entry and $I \geq 2$ risk-neutral principals (insurance companies) which offer insurance contracts C simultaneously. A contract consists of an insurance premium $\alpha \in \mathbb{R}_0^+$ and an indemnity $\beta \in \mathbb{R}_0^+$.[5] In the Bertrand equilibrium, premium offers correspond to the expected cost of a policyholder. Thus, in absence of any fixed cost, all insurance companies offer the same contract with a utility-maximizing indemnity at zero-expected profits and share the market equally.

The sequence of play is:

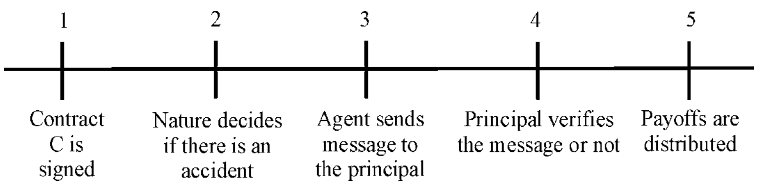

Figure 1: Sequence of play without fraud detection

After the policyholder has observed the state of nature, he decides whether to file a claim or not. Consequently, an insurance company has to decide whether to audit a received claim or not. The audit is assumed to be perfect. Thus, after an audit the insurer can observe the state of the world.

[4]Boyer (1999) analyzes a situation with two different types of policyholders. One type never commits insurance fraud and always reports the state of the world honestly. The second type is opportunistic and weighs up the costs and benefits of a fraudulent claim, if no accident occurred. In this paper we will concentrate on the second type of opportunistic policyholders.

[5]The considered insurance contracts are incomplete, because the indemnity is state-independent. Actually, insurers could condition the indemnity payment based on whether an audit took place or not and reward honest policyholders with a higher indemnity (see Fagart and Picard, 1999). But a contract with a state-dependent indemnity is incredible, because policyholders do not know whether their claim was audited or not, and they have no opportunity to determine this fact. In cases in which an insurer audits a claim and determines that it was honest, it keeps the audit secret and saves money. The credibility issue that applied to the auditing problem is also valid in this context.

The audit cost, which the insurer has to bear under any circumstances and which cannot be influenced by policyholders, is $c > 0$. Policyholders who are caught defrauding must pay a penalty k that is sunk and is not collected by the insurance company.[6]

The whole game can be solved by backward induction. First of all, we regard the auditing game of stages 2 - 4 for given contracts. The Perfect Bayesian Nash Equilibrium (PBNE) of this game consists of sequentially rational strategies given a belief system μ that is derived from the strategy profile through Bayes rule whenever possible. Furthermore, the equilibrium contract can be determined by maximizing the expected utility of the policyholder under the zero-profit constraint of insurers.

Policyholders can observe the state of the world. In equilibrium it is a dominant strategy for policyholders who suffered a loss to file a claim, if $\beta > 0$ holds. The other way around it will never be a best response for insurers to audit, if no claim was made.[7] Basically, the auditing game without commitment has only a mixed strategy equilibrium and the utility maximizing contract over-compensates the agent for his loss.

Lemma 1 *If $\beta > c$ and $\pi < \frac{1}{2}$ hold, the unique PBNE in mixed strategies and the equilibrium insurance contract have the following properties:*[8]

- *An insurer has an sequential consistent ex ante belief $\mu = \frac{c}{\beta}$, that a filed claim is fraudulent;*
- *A policyholder will always file a claim, if an insured loss occurred. Otherwise, if he has not suffered a loss, he will file a fraudulent claim with the probability*

$$\eta^* = \left(\frac{\pi}{1-\pi}\right)\left(\frac{c}{\beta - c}\right); \tag{1}$$

- *An insurer audits each claim with the probability*

$$\nu^* = \frac{u(Y-\alpha+\beta) - u(Y-\alpha)}{u(Y-\alpha+\beta) - u(Y-\alpha-k)}; \tag{2}$$

[6] There are two reasons why we assume the penalty k not to be part of the insurance contract. First of all, it is usually determined by law and/or courts. Therefore, insurers would have to renegotiate with fraudulent policyholders after an audit. Since negotiations cause significant transaction cost and the expected sanction for defrauders in reality is not very high (see Derrig and Zicko, 2002), it is questionable whether insurers can collect the penalty or not. In addition, when a penalty $k > c$ is paid to the insurer, auditing becomes a rent-extracting device and not a way to induce contract compliance of policyholders (see Picard, 1996, for a model where the insurer partially benefits from the penalty).

[7] This result is just technical, because insurers are not able to audit a claim that was not made.

[8] The conditions $\beta > c$ and $\pi < \frac{\beta - c}{\beta}$ ensure that the auditing game has a mixed strategy equilibrium, with $\eta^* \in (0, 1)$. Since $\beta > 2c$ is a necessary condition that the utility maximization problem has an interior solution, the condition $\pi < \frac{1}{2}$ is a sufficient condition to get a mixed strategy equilibrium (see Boyer, 1999, for further details).

- *The unique utility maximizing contract* $C^* = (\pi\beta^* \frac{\beta^*}{\beta^*-c}, \beta^* > L)$ *entails over-insurance.*

Proof. See Boyer [2000] ■

The auditing game only has a PBNE in mixed strategies. Therefore, both parties choose a randomization, η^* and ν^* respectively, in order to make their opponent indifferent between the available actions. Additionally, the insurer has to design a contract as a combination of coverage and premium which maximizes the policyholder's expected utility

$$\begin{aligned} U &= \pi u (Y - \alpha - L + \beta) + (1 - \pi)(1 - \eta) u (Y - \alpha) \\ &\quad + (1 - \pi)\eta [(1 - \nu) u (Y - \alpha + \beta) + \nu u (Y - \alpha - k)] . \end{aligned} \tag{3}$$

The zero-profit insurance premium $\alpha(\beta)$ is given by

$$\alpha(\beta) = \pi\beta + (1 - \pi)(1 - \nu)\eta\beta + \nu [\pi + (1 - \pi)\eta] c. \tag{4}$$

The designed contract, in particular, the indemnity β, alters the payoffs of the policyholder and influences the equilibrium randomization of both parties. Using (1) and (4) leads to

$$\alpha^* = \pi\beta \frac{\beta}{\beta - c}. \tag{5}$$

Over-compensation is optimal for policyholders, because the slope of the zero-profit premium function

$$\frac{\partial \alpha^*}{\partial \beta} = \pi \frac{\beta [\beta - 2c]}{[\beta - c]^2} \tag{6}$$

is positive and always smaller than the slope of the actuarially fair premium π, if $\beta > 2c$ holds.

Since the second order derivative

$$\frac{\partial^2 \alpha^*}{\partial \beta^2} = \pi \frac{2c^2}{[\beta - c]^3}. \tag{7}$$

is positive for $\beta \in (c, \infty)$, the premium function α^* is convex in the relevant interval $\beta \in (c, \infty)$, with a local minimum at $\beta = 2c$, as displayed in Figure 2.

For $\beta < L$, an increase in coverage raises the expected payoff and reduces the income risk of a policyholder. Due to the risk-neutrality of the policyholder at $\beta = L$, an increase of the expected payoff leads to a positive marginal utility. Therefore, the optimal indemnity must be greater than the loss. For $\beta > L$, there is a trade-off between increasing the expected payoff and risk taking. Since the premium α^* entails an implicit absolute fraud loading q, the interior solution with $\beta^* > L$ is a global maximum, if the policyholder is sufficiently risk-averse.

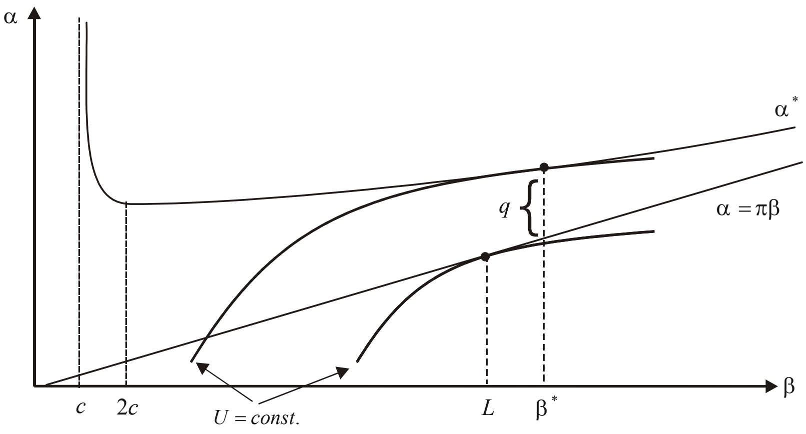

Figure 2: Equilibrium contract without fraud detection

3 Equilibrium with fraud detection

Let us now assume that insurance companies are able to implement a fraud detection system with a given technology that assigns an exogenous signal s to each reported loss. Since the signal of the system is private information of an insurer, it is non-contractible. For the moment, we desist from the cost of the fraud detection system and specify its properties. But later on, in section 4, we examine the consequences that the cost of the system might have.

The earlier stated sequence of play must be modified with respect to the fraud detection system. After the implementation of a system, insurers are able to audit contingent on the signal. Consequently, after the observation of the signal they have to decide whether to audit a claim or to pay the indemnity immediately without any audit.

The sequence of play with fraud detection is as follows:

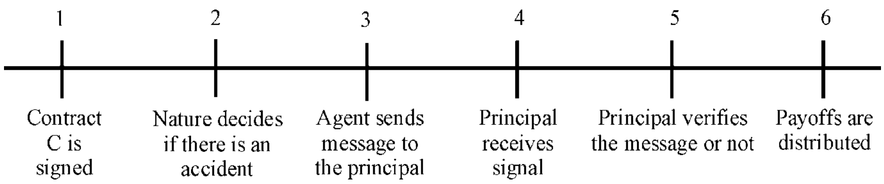

Figure 3: Sequence of play with fraud detection

We can still solve the whole game by backward induction. Thus, we start looking at the auditing game of stages 2 - 6 for given contracts.

The whole set of claims $\frac{N}{I}\left[\pi + (1-\pi)\eta\right]$ of each insurance company consists of two different types of reports:

- truthful claims with the fraction $\frac{\pi}{\pi+(1-\pi)\eta}$ and
- fraudulent claims with the fraction $\frac{(1-\pi)\eta}{\pi+(1-\pi)\eta}$.

An insurer privately receives one of two possible signals, s_0 and s_1, when a policyholder files a claim. The properties of the fraud detection system are common knowledge. The conditional probabilities of the signal s given the state of world ω are:

	$p(s \mid \omega_1)$	$p(s \mid \omega_0)$
s_1	δ	ϕ
s_0	$(1-\delta)$	$(1-\phi)$

Table 1: Contingent probabilities

The displayed probabilities in Table 1 imply that the fraction δ (resp. $(1-\delta)$) of all honest claims get the signal s_1 (s_0) and that the fraction ϕ (resp. $(1-\phi)$) of all fraudulent claims get the signal s_1 (s_0). Since insurers update their fraud beliefs μ according to Bayes rule, the posterior fraud beliefs after observing the signal s are

$$\mu_1 = \frac{\phi(1-\pi)\eta}{\delta\pi + \phi(1-\pi)\eta} \tag{8}$$

and

$$\mu_0 = \frac{(1-\phi)(1-\pi)\eta}{(1-\delta)\pi + (1-\phi)(1-\pi)\eta}. \tag{9}$$

The case $\delta = \phi$ corresponds to a situation with an uninformative signal which is treated in section 2. In the following, we assume without loss of generality that the signal is valuable or informative in the sense of Holmström (1979) and $\phi > \delta$ holds. Hence, the fraction of fraudulent claims in the set of claims with the signal s_1 is higher than in set of claims with the signal s_0. This assumption directly implies that fraud beliefs of insurers are higher after the observation of the signal s_1 than after the observation of s_0 and $\mu_1 > \mu_0$ holds.

In a PBNE in mixed strategies, insurers choose their audit strategy in order to make policyholders without a loss indifferent between their actions. Therefore, the audit probability ν^* without fraud detection corresponds to the critical detection probability κ^c, which makes policyholders without a loss indifferent. Since no payoffs change compared to the case without fraud detection, the indifference condition is still given by

$$u(Y-\alpha) = \kappa^c u(Y-\alpha-k) + (1-\kappa^c)u(Y-\alpha+\beta) \tag{10}$$

with

$$\kappa^c = \frac{u(Y-\alpha+\beta) - u(Y-\alpha)}{u(Y-\alpha+\beta) - u(Y-\alpha-k)}. \tag{11}$$

Now, with a detection system the expected detection probability $\bar{\kappa}$ of a fraudulent claim is a convex combination of the conditional audit probabilities ν_1 and ν_0.

$$\bar{\kappa} = \phi\nu_1 + (1-\phi)\,\nu_0 \tag{12}$$

In equilibrium the expected detection probability $\bar{\kappa}$ must correspond to the critical detection probability of policyholders without a loss κ^c. Using (11) and (12) yields

$$\phi\hat{\nu}_1 + (1-\phi)\,\hat{\nu}_0 = \bar{\kappa} = \kappa^c = \frac{u(Y-\alpha+\beta) - u(Y-\alpha)}{u(Y-\alpha+\beta) - u(Y-\alpha-k)}. \tag{13}$$

From $\phi > \delta$, (8) and (9) follow that the insurer's fraud beliefs are higher for claims with the signal s_1 than for those with the signal s_0. Consequently, insurers begin to audit claims with the signal s_1. In the considered context, an optimal audit strategy a has the following properties:

$$a = \begin{cases} (\hat{\nu}_1 \geq 0, \hat{\nu}_0 = 0) & if \quad \phi \geq \kappa^c \\ (\hat{\nu}_1 = 1, \hat{\nu}_0 > 0) & if \quad \phi < \kappa^c \end{cases}. \tag{14}$$

In equilibrium the conditional audit probability $\hat{\nu}_0$ can only be zero, if insurers reach the critical detection probability by exclusively auditing claims with the signal s_1. Without loss of generality, we will concentrate on this case with $\phi \geq \kappa^c$, because it is straightforward to extend the results to the case $\phi < \kappa^c$. In the latter case an insurer has to audit all claims with the signal s_1 and some claims with the signal s_0.

Proposition 1 *If $\beta > c$, $\pi < \frac{1}{2}$ and $\phi \geq \kappa^c$ hold, the unique PBNE in mixed strategies for a given fraud detection system with $\phi > \delta > 0$ has the following properties:*

- *An insurance company has an ex ante sequential consistent fraud belief $\mu = \frac{\delta c}{\phi\beta - [\phi-\delta]c}$;*
- *The equilibrium fraud probability is*

$$\hat{\eta} = \left(\frac{\pi}{1-\pi}\right)\left(\frac{\delta c}{\phi\,[\beta - c]}\right); \tag{15}$$

- *An insurer audits only claims with the signal s_1 with the probability*

$$\hat{\nu}_1 = \frac{1}{\phi}\frac{u(Y-\alpha+\beta) - u(Y-\alpha)}{u(Y-\alpha+\beta) - u(Y-\alpha-k)} = \frac{\kappa^c}{\phi}, \tag{16}$$

which leads to an overall audit probability

$$\hat{\nu} = \frac{\delta\beta}{\phi\beta - [\phi-\delta]\,c}\kappa^c. \tag{17}$$

Proof. See Appendix. ■

An PBNE in mixed strategies only exists if a policyholder is able to make the insurer indifferent between its action. With a perfect fraud detection system ($\delta = 0$), this is not the case, because the insurer's fraud belief (8) after observing the signal s_1 is $\mu_1 = 1$, irrespective of the strategy of the policyholder. Auditing all claims with the signal s_1 is therefore a dominant strategy for insurers. Clearly, it is a best response for policyholders to report the state of the world honestly, which leads to $\hat{\eta} = 0$. Thus, in equilibrium the insurance company does not receive any claim with the signal s_1 and no audit is necessary. In any other case, with $\delta > 0$, the system is not perfect, and an equilibrium in mixed strategies exists, because policyholders are able to fulfill the indifference condition of insurers.

An increase (decrease) of ϕ leads to a decreasing (increasing) fraud, conditional audit and overall audit probability. An increase (decrease) of δ raises (diminishes) the fraud and overall audit probability, since the fraction of honest claims in the set of claims with the signal s_1 increases (decreases). At first glance, it is surprising that the conditional probability $\hat{\nu}_1$ is unaffected by a variation of δ. Policyholders without a loss choose their fraud probability in order to make insurers after the observation of the signal s_1 indifferent between their strategies. Hence, they compensate changing incentives of insurers to audit claims with the signal s_1 by a variation of the fraud probability. The overall audit probability rises (declines), ceteris paribus, with an increasing (decreasing) fraction δ, because insurers get the signal s_1 more (less) often and audit all claims with that signal s_1 with the same conditional probability as before.

The signal is non-contractible, because it is private information of the insurer. For that reason, we can derive the optimal insurance contract $\hat{C}$ in the same manner as in section 2.

Corollary 1 *If there is an informative fraud detection system with $\phi > \delta$, the equilibrium contract $\hat{C} = \left(\pi\hat{\beta}\frac{\hat{\beta}-c+\frac{\delta}{\phi}c}{\hat{\beta}-c}, \hat{\beta}\right)$ has the following properties:*[9]

- *As long as $\delta > 0$, the system is imperfect, and the optimal contract entails over-insurance with $\hat{\beta} > L$;*
- *If and only if $\delta = 0$, the system is perfect, and the optimal contract entails a fair premium and full insurance with $\hat{\beta} = L$;*
- *The equilibrium contract with fraud detection entails less over-insurance than the contract without fraud detection.*

[9] Since the premium function $\hat{\alpha}$ and, therefore, the implicit fraud loading q strictly decline in the system's quality $\frac{\delta}{\phi}$, the interior solution with fraud detection $\hat{\beta}$ is a global maximum, if the interior solution $\beta^* > L$ is a global maximum in the case without fraud detection.

Proof. See Appendix. ■

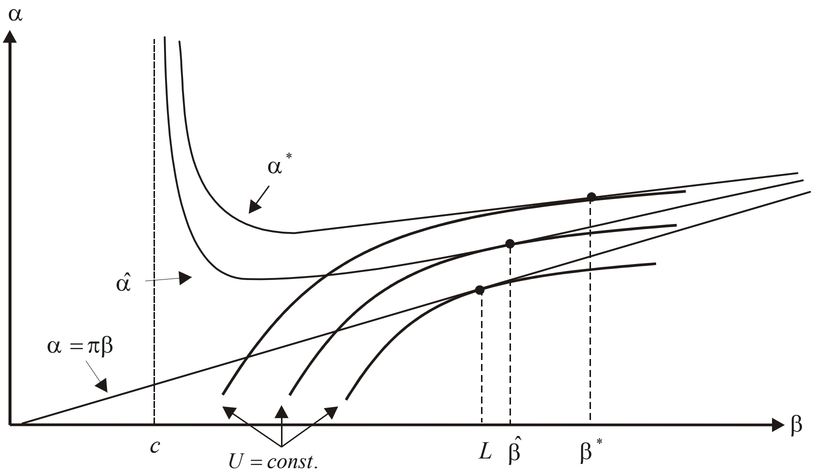

Figure 4: Equilibrium contract with fraud detection

The motivation of over-compensation as given in section 2 still holds. After the implementation of an informative detection system the extent of over-compensation can be reduced. For that reason, over-compensation and fraud detection are two strategic complements that help to reduce insurance fraud. When an informative signal is available, auditing becomes more attractive and the monetary incentives can be reduced.

4 The organizational design of fraud detection

The effectiveness of fraud combat is very sensitive to its organizational design. There are many approaches addressing whether or how a third party can help to mitigate the market inefficiency caused by the non-commitment problem of the principal. In Picard's view, the non-commitment problem can (partly) be solved by a common agency created for all insurance companies, that does not audit, but (partly) takes charge of the audit expenditures and is financed by the whole market through partition fees (Picard, 1996, p. 45).

A different approach employed by Melumad and Mookherjee (1989) suggests, in the case of a tax audit, that the authority can delegate the auditing job to an independent investigation agency. This solution is not really convincing for two reasons that the authors concede. Firstly, it is not renegotiation-proof, because the insurer has an incentive to privately renegotiate the contract with the audit agency after the audit contract is publicly signed. Secondly, Melumad and Mookherjee abstract from the moral hazard problem between the insurer and the agency.

In this section, we present how a fraud detection system can and will be implemented in insurance markets when it mainly causes fixed cost. In contrast to Boyer (2000), we do not consider different audit cost of insurers, because we have assumed that they are completely homogenous. In addition to the information about the probability of a fraudulent claim, a fraud detection system will usually generate relevant information for the audit itself. Therefore, we could assume that the audit cost after the implementation of the systems would be lower than before. Since a system will cause some cost of data input, we suppose that the two effects lead to the same audit cost c.

The main advantage of the implementation of a detection system is that a single insurer with a detection system can always undercut insurers without a system, since $\hat{\alpha} < \alpha^*$ holds for any indemnity level β. If the invention and implementation of a detection system cause fixed costs, the existence of a competitive market equilibrium is questionable, because fixed costs are sunk for insurance companies at the competition stage. For the remainder of the section we will concentrate on this arising problem.

Suppose that a single insurer can invent and implement a given detection system at fixed cost D. In the considered situation this insurer can make positive expected profits by offering a contract C_M which yields the same expected utility as the contract C^*. In order to simplify the analysis, we assume that policyholders always choose the contract with fraud detection, when they are indifferent between a contract with and one without detection. Consequently, the insurer with a detection system maximizes its profits by offering insurance coverage at a premium $\alpha_M = \hat{\alpha} + z$. The premium α_M includes an absolute loading z which corresponds to the per policy profit of an insurer and makes policyholders indifferent between the contracts C^* and C_M. Since the considerations concerning the optimal coverage are unaffected by z, the expected utility is maximized by $\hat{\beta}$. Such a market configuration is only feasible in the sense of Baumol et al. (1982) if the market has a considerable size and no capacity constraint is present. In equilibrium the monopoly insurer makes non-negative expected profits:

$$\Pi = N \cdot z - D \geq 0, \tag{18}$$

for some N.

For the remainder of the section, we will concentrate on the socially efficient case, where a single fraud detection system with a given technology (δ, ϕ) is implemented by an external supplier. The supplier offers every insurer the same participation contract $P = (F, f)$ that entails a fixed fee $F \geq 0$ for using the system and/or a variable fee $f \geq 0$ for every signal s to which the system refers. We can incorporate a two-stage subgame in the sequence of play in section 3 as follows.

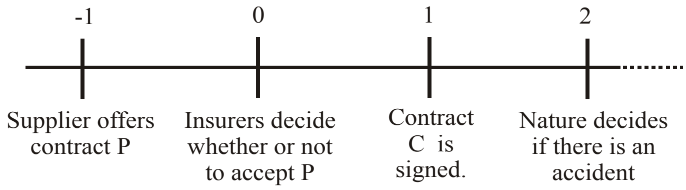

Figure 5: Sequence of play with participation contract

Subsequently, we only have to consider the first three stages of the game displayed in Figure 5, since they determine whether the system can be implemented or not. Due to the Bertrand competition between insurers at stage 1, the participation contract P cannot entail any fixed fee $F > 0$. Insurance companies that accept such a contract will make expected losses in the amount of the fixed cost. Thus, the supplier has to transform the fixed cost D into variable cost. Hence, the external supplier offers participation contracts $P = (0, f > 0)$, where insurers solely pay a variable fee f for each received signal.

Insurance companies that accept a participation contract P can only offer zero-profit and utility maximizing contracts C_P on stage 1 which entail a premium $\alpha_P = \frac{\beta - c + \frac{\delta}{\phi} c}{\beta - c} [\beta + f]$ and an indemnity $\beta_P = \hat{\beta}$. A market equilibrium with fraud detection can only be reached if:

i. The market configuration with fraud detection is feasible and consequently the expected contract with fraud detection is weakly preferred by policyholders

$$U(C_P) \geq U(C^*) \tag{C1}$$

and

ii. The external supplier makes non-negative profits.

$$f \geq \frac{D}{N\pi \frac{\beta - c + \frac{\delta}{\phi}}{\beta - c}} \tag{C2}$$

If condition (C1) holds, there will be no market configuration where an insurance company can make non-negative expected profits by not participating in the fraud detection system and offering a contract C that policyholders weakly prefer to C^*. On the other hand, as can be seen from (C2), if the market size (number of policyholders) does not exceed a critical value, no fraud detection will be introduced.

5 Conclusions

The over-compensation of agents is a possible strategic commitment device to reduce fraud in the non-commitment context, since it commits the insurer to audit more often. In this paper, we have characterized the impact of a fraud detection system on the equilibrium of the audit game and the insurance contract. We show that, in addition to insurance contracts with over-compensation, informative fraud detection systems are another possible strategic commitment device to reduce the fraud-induced welfare losses in the economy. The main advantage of these systems is that insurers can condition their audits on the received signal. Due to the properties of Nash equilibria in mixed strategies, the detection system affects the randomization of the policyholders and insurers as well. The fraud-induced welfare losses strictly decrease in the signal quality. Yet, the non-commitment problem of insurers remains as long as the signal is imperfect.

The amount and structure of the cost of a detection system are crucial in our context. Fraud detection systems are based on computer software and so have a high implementation cost. Operating a system will only be remunerative if the benefits from the declining fraud and audit probability exceed the relative cost of the system related to an individual consumer. For that reason, the number of policyholders in the market is critical for the effectiveness of fraud detection. Thus, we can confirm the fact that fraud detection is only performed in insurance markets with a considerable size, such as the property, liability, and auto insurance. Moreover, we deduce that due to the fixed cost of the system and the Bertrand competition in the considered insurance market, the contract between insurers and the external supplier of the system can only consist of a variable transfer from the insurer to the supplier.

Our analysis concerning the auditing game is partly limited by the fact that we neglect to consider any renegotiation aspects. As Macho-Stadler and Pérez-Castrillo (2004) for the commitment audit case show, a fraud detection system gives the insurer incentives for renegotiations with the policyholder. After observing the signal, it is profitable for insurers to settle with policyholders in order to save audit expenditures. As a consequence, future research should concentrate on the renegotiation incentives in a non-commitment environment.

Appendix

Proof of Proposition 1. The equilibrium fraud probability $\hat{\eta}$ must solve the indifference condition of insurers, which is given by

$$-\beta = \mu_1 [-c] + (1 - \mu_1) [-\beta - c] . \tag{19}$$

Substituting μ_1 in (19) by (8) and rearranging the term yields

$$\hat{\eta} = \left(\frac{\pi}{1 - \pi} \right) \left(\frac{\delta c}{\phi [\beta - c]} \right) . \tag{20}$$

Given (20) the resulting ex ante fraud beliefs μ of insurers are

$$\mu = \frac{(1 - \pi)\hat{\eta}}{\pi + (1 - \pi)\hat{\eta}} = \frac{\delta c}{\phi\beta - [\phi - \delta] c}. \tag{21}$$

After the observation of the signal s_1 the posterior fraud beliefs μ_1 are

$$\mu_1 = \frac{\phi(1 - \pi)\hat{\eta}}{\delta\pi + \phi(1 - \pi)\hat{\eta}} = \frac{c}{\beta}. \tag{22}$$

The equilibrium audit probability $\hat{\nu}_1$ satisfies the following indifference condition of policyholders

$$u(Y - \alpha) = \phi\hat{\nu}_1 u(Y - \alpha - k) + (1 - \phi\hat{\nu}_1) u(Y - \alpha + \beta). \tag{23}$$

After some manipulations one obtains

$$\hat{\nu}_1 = \frac{1}{\phi} \frac{u(Y - \alpha + \beta) - u(Y - \alpha)}{u(Y - \alpha + \beta) - u(Y - \alpha - k)} = \frac{\kappa^c}{\phi}. \tag{24}$$

The overall audit probability is given by the proportion of audited claims to all received claims, which is

$$\hat{\nu} = \frac{\delta\pi + \phi(1 - \pi)\hat{\eta}}{\pi + (1 - \pi)\hat{\eta}} \hat{\nu}_1. \tag{25}$$

By using (20) and (24) the overall probability simplifies to

$$\hat{\nu} = \frac{\delta\beta}{\phi\beta - c [\phi - \delta]} \frac{u(Y - \alpha + \beta) - u(Y - \alpha)}{u(Y - \alpha + \beta) - u(Y - \alpha - k)}. \tag{26}$$

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

Proof of Corollary 1. The zero-profit premium is given by

$$\alpha(\beta) = \pi\beta + (1-\nu)(1-\pi)\eta\beta + \nu\left[\pi + (1-\pi)\eta\right]c. \tag{27}$$

Using (20) and rearranging (27) leads to

$$\hat{\alpha} = \pi\beta\frac{\beta - c + \frac{\delta}{\phi}c}{\beta - c}. \tag{28}$$

If $\phi > \delta$, the zero-profit premium with fraud detection (28) is strictly lower than that without fraud detection (5), as (29) shows.

$$\pi\beta\frac{\beta}{\beta - c} > \pi\beta\frac{\beta - c + \frac{\delta}{\phi}c}{\beta - c} \quad \forall\beta \tag{29}$$

Given that $\hat{\alpha}$ is twice differentiable with respect to β, the partial derivatives of (28) with respect to β are

$$\frac{\partial\hat{\alpha}}{\partial\beta} = \pi\frac{[\beta - c]^2 - \frac{\delta}{\phi}c^2}{[\beta - c]^2} \tag{30}$$

and

$$\frac{\partial^2\hat{\alpha}}{\partial\beta^2} = \pi\frac{2\frac{\delta}{\phi}c^2}{[\beta - c]^3}. \tag{31}$$

For $\phi > \delta > 0$, the premium function (28) has only one local minimum at $\beta_0 = \left(1 + \sqrt{\frac{\delta}{\phi}}\right)c$ in the open interval (c, ∞). Furthermore, it is convex throughout this interval, because (31) is positive for $\beta \in (c, \infty)$ and $\delta > 0$. In the case $\delta = 0$, (28) corresponds to the fair premium, which is linear increasing in β.

The simplified maximization problem with a fraud detection system is

$$\max_{\beta}\ U = \pi u\left(Y - \pi\beta\frac{\beta - c + \frac{\delta}{\phi}c}{\beta - c} - L + \beta\right) + (1-\pi)u\left(Y - \pi\beta\frac{\beta - c + \frac{\delta}{\phi}c}{\beta - c}\right). \tag{32}$$

Over-insurance will be optimal if and only if the slope of the indifference curve at $\beta = L$ is positive and

$$\pi u'\left(Y - \pi L\frac{L - c + \frac{\delta}{\phi}c}{L - c}\right) - u'\left(Y - \pi L\frac{L - c + \frac{\delta}{\phi}c}{L - c}\right)\left[\pi\frac{[L-c]^2 - \frac{\delta}{\phi}c^2}{[L-c]^2}\right] > 0 \tag{33}$$

holds.

Rearranging (33) leads to

$$\delta > 0. \tag{34}$$

If and only if $\delta = 0$ holds, the detection system is perfect and the zero-profit premium with fraud detection is actuarially fair, which leads to an optimal full-insurance contract with $\hat{\beta} = L$.

Finally, we show the premium and the slope of the premium function are reduced when the signal is more informative. The impact of $\frac{\delta}{\phi}$ on the premium (28) is given by

$$\frac{\partial \hat{\alpha}}{\partial \frac{\delta}{\phi}} = \pi\beta \frac{c}{\beta - c} > 0. \tag{35}$$

Obviously, the insurance premium and, therefore, the implicit fraud loading q are reduced when the signal is more informative. Moreover, the impact of the signal's informativeness on the premium is linear as $\frac{\partial^2 \hat{\alpha}}{\partial \left(\frac{\delta}{\phi}\right)^2} = 0$.

The impact of the signal quality on the slope of the premium function is given by

$$\frac{\partial^2 \hat{\alpha}}{\partial \beta \, \partial \frac{\delta}{\phi}} = -\pi \frac{c^2}{[\beta - c]^2} < 0. \tag{36}$$

As (36) indicates, the slope of the premium function strictly decreases in $\frac{\delta}{\phi}$. Therefore, the over-compensation decreases in the signal quality. ■

References

ARTÍS, M., AYUSO, M. and GUILLÉN, M. (2002): "Detection of Automobile Insurance Fraud with discrete Choice Models and misclassified Claims," *Journal of Risk and Insurance*, 69, 325-340.

BAUMOL, W.J., PANZAR, J.C. and WILLIG, R.D. (1982), *Contestable Markets and the Theory of Industry Structure*, Hartcourt Brace Jovanovic, New York.

BELHADJI, E.B., DIONNE, G. and TARKHANI, F. (2000): "A Model for the Detection of Insurance Fraud," *Geneva Papers on Risk and Insurance – Issues and Practice*, 25, 517-538.

BOYER, M.M. (1999): "When is the Proportion of Criminal Elements irrelevant? A Study of Insurance Fraud when Insurers cannot commit," DIONNE, G. and LAMBERGE-NADEAU, C. (Eds.): *Automobile-Insurance: Road Safety, New Drivers, Risks, Insurance Fraud and Regulation*, Kluwer Academic Publishers, Boston, 151-173.

BOYER, M.M. (2000): "Centralizing Insurance Fraud Investigation," *Geneva Papers on Risk and Insurance Theory*, 25, 159-178.

DERRIG, R.A. and ZICKO, V. (2002): "Prosecuting Insurance Fraud – A Case Study of the Massachusetts Experience in the 1990s", *Risk Management and Insurance Review*, 5, 77-104.

DIONNE, G., GIULIANO, F. and PICARD, P. (2003): "Optimal Auditing for Insurance Fraud," mimeo, CIRPÉE, Working Paper 03-29.

FAGART, M.-C. and PICARD, P. (1999): "Optimal Insurance Under Random Auditing," *Geneva Papers on Risk and Insurance Theory*, 24, 29-54.

GALE, D. and HELLWIG, M. (1985): "Incentive Compatible Debt Contracts: The One-Period Problem," *Review of Economic Studies*, 52, 647-663.

HOLMSTRÖM, B. (1979): "Moral hazard and observability," *Bell Journal of Economics*, 10, 74-91.

KHALIL, F. (1997): "Auditing without commitment," *Rand Journal of Economics*, 28, 629-640.

MACHO-STADLER, I. and PÉREZ-CASTRILLO, D. (2004): "Settlement in Tax Evasion Prosecution," *Economica*, 71, 349-368.

MELUMAD, N.D. and MOOKHERJEE, D. (1989): "Delegation as commitment: the case of income tax audits," *Rand Journal of Economics*, 20, 139-163.

MOOKHERJEE, D. and PNG, I. (1989): "Optimal Auditing Insurance, and Redistribution," *Quarterly Journal of Economics*, 104, 399-415.

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

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

Acknowledgments

I would like to thank Klaus Bender, Richard Derrig, Uwe Focht, Jennifer Kirkland, Henrike Lesch, Martin Nell, Martin Peitz, Pierre Picard, Andreas Richter, Andrea Seul, two anonymous referees and the participants of the 2003 annual ARIA meeting in Denver for their very helpful comments and suggestions on earlier versions of the paper.