



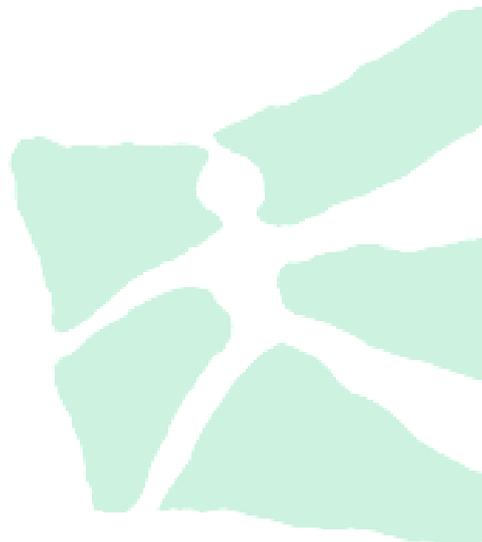
THE IMPACT OF AUDITING STRATEGIES ON INSURERS' PROFITABILITY

KATJA MÜLLER
HATO SCHMEISER
JOËL WAGNER

WORKING PAPERS ON RISK MANAGEMENT AND INSURANCE NO. 134

**EDITED BY HATO SCHMEISER
CHAIR FOR RISK MANAGEMENT AND INSURANCE**

JUNE 2013



The Impact of Auditing Strategies on Insurers' Profitability

Katja Müller, Hato Schmeiser, Joël Wagner*

Abstract

Insurance claims fraud is a major concern in the insurance industry. According to estimates, excess payments due to fraudulent claims account for somewhere between 15 and 25 percent of all claim payments, affecting all classes of insurance. In this paper, we develop a model framework based on a costly state verification setting in which - while policyholders observe the amount of loss privately - the insurance company can decide to audit incoming claims at some cost. The aim is to derive optimal auditing strategies from the insurance company perspective while maintaining contract attractiveness to policyholders who are willing to adhere to the insurance relationship. We present and analyze an auditing strategy which is triggered by the filed claims amount and also includes an option for each stakeholder to adapt its behavior based on a signal specific to its position. The impact of the optimal auditing strategy on the insurer's profitability is examined. Finally, practical implementations are discussed.

1 Introduction

Insurance claims fraud is one of the major industry concerns. It occurs in all classes of insurance and accounts for a substantial portion of the indemnity payments each year, yet due its nature, is difficult to estimate in total value. The Insurance Research Council (2008) estimated the excess payments due to fraudulent claims in 2007 as somewhere between \$4.8 and \$6.8 Mrd in the auto injury insurance sector in the U.S. alone, corresponding to 13 to 18 percent of total payments. These fraudulent activities, while undertaken by some individuals, have an impact on the policyholder population as a whole through higher insurance premiums (see Tennyson (2008)).

Previous research has explained the existence of fraud due to information being asymmetrically distributed between policyholder and the corresponding insurance company (see, e.g., Derrig (2002)). Since insureds may hold private information about the actual amount of the loss suffered, there exists the possibility of misrepresentation. Based on the approach of costly state verification (see, e.g., Townsend (1979), Mookherjee and Png (1989), Bond and Crocker (1997), Picard and Fagart (1999), Picard (2000)), the insurance company can consequently choose to audit incoming claims in order to determine their truthfulness. In cases when fraudulent activities haven been proven, a penalty payment can be imposed on the policyholder. However, these verification processes incur costs. From the insurance company perspective

*The authors are with the Institute of Insurance Economics, University of St. Gallen, Tannenstr. 19, 9000 St. Gallen, Switzerland.

this implies that the costs for auditing have to be traded off against the savings resulting from detected fraud. One of the costs may be on policyholder attitude, and therefore additionally, the policyholders' point of view needs to be considered as well. According to Viaene and Dedene (2004), individuals are more likely to develop an opportunistic attitude towards insurance fraud after having gained negative experiences with their insurance companies. Underpaid claims or the endurance of long waiting periods for indemnity payments may encourage fraudulent behavior in the future.

An opposing strategy to minimize the occurrence of insurance fraud comprises the implementation of bonus-malus systems. Dionne and Vanasse (1992) and Moreno, Vázquez, and Watt (2006) develop model frameworks where, instead of performing costly verification processes, the policyholder's insurance premium is increased whenever he or she files a claim in the previous period. This approach is applicable in the case of insurance contract renewals; however, in some countries policyholders may avoid this penalization mechanism by switching their insurance company (see, e.g., Dionne and Ghali (2005)).

An additional strand of research dealing with the occurrence and handling of insurance claims fraud is based on the costly state falsification approach introduced by Lacker and Weinberg (1989) and adapted to the insurance setting by Crocker and Morgan (1998) and Crocker and Tennyson (2002). While the premise with regard to the asymmetric distribution of information remains unchanged, this time the policyholder engages in costly manipulations such that a verification of the claims becomes impossible.

The term fraud has a rather negative connotation implying the engagement in illegal activities such as staged accidents. The range of actions, however, which is colloquially subsumed under this notion, is by far broader (see, e.g., Picard (2001), Derrig (2002), Tennyson (2008)). In general, a distinction is made between ex-ante and ex-post moral hazard, referring to the timing of fraudulent behavior. The former occurs by the time of insurance purchase, e.g., in case the potential policyholder fails to provide relevant information which might have resulted in unfavorable contract conditions or even in a rejection on part of the insurance company (see, e.g., Picard (2001)). The term ex-post moral hazard implies the engagement in fraudulent activities by the time of claims filing, i.e., after a potential loss has been suffered. In this context, there are several possible distinctions which can be made. A common one is to differ between criminal fraud, also called hard fraud, and soft fraud which is situated in an ethical gray area (see, e.g., Tennyson (2008)). Derrig (2002) defines criminal fraud as "the willful act of obtaining money or value from an insurer under false pretenses or material misrepresentations". While it is assumed that only a small number of claims contain outright fraud, the more frequent and more costly type of fraudulent behavior falls within the term of soft fraud (see, e.g., Weisberg and Derrig (1991), Viaene and Dedene (2004), Tennyson (2008)). Even though there does not exist an explicit definition of the term, it is associated with the misrepresentation of the loss magnitude after its occurrence, i.e., claimants exaggerate the amount to obtain higher indemnity payments. A notion often used in this context is build-up. In this paper, we will refer to fraud as soft fraud or build-up, i.e., we assume that some policyholders inflate the magnitude of loss after its occurrence if it appears profitable.

In addition to the policyholders, also third parties like service providers might be involved in fraudulent activities to exaggerate the loss amount (see, e.g., Tennyson (2008), Dulleck and Kerschbamer (2006) Derrig and Zicko (2002)). Prominent examples are known from automobile insurance. Derrig and Zicko

(2002) found that repair shops are likely to have developed insight about the insurance companies' prevalent auditing strategies due to repeated experiences in repairing cars for insured damages. As a consequence, there exists the possibility to adjust the defrauding strategy. Such actions are either undertaken without the knowledge of the policyholder, or performed by mutual agreement between the policyholder and the corresponding repair shop. The service providers' incentive to engage in fraudulent activities against the insurance company may stem from the hope to gain the policyholders' favor. As a result, the latter might request their services again at a later point in time. For the purpose of our study, we will assume that all fraudulent activities - even if committed by repair shops - to be advantageous for the policyholders only.

Considering its large prevalence, effective measures in dealing with the phenomenon of insurance claims fraud need to be found. As already mentioned, insurance companies have the possibility to perform audits. Many insurers have dedicated teams to identify and combat fraudulent claims. In interviews with fraud managers, Morley, Ball, and Ormerod (2006) find that verification processes are initiated whenever investigators detect anomalies or inconsistencies in the circumstances of the loss event or unusual behavior from the claimant. This approach seems to be suitable for filed claims that exhibit obvious signs indicating potential fraudulent behavior and/or the ones which are of significant magnitude. Apparently, high-magnitude claims are subject to verification right away. The majority of claims, however, appear to be legitimate at first sight, seeking low to medium indemnity payments, which do not trigger immediate investigation. Consequently, the key question is how to deal with the fraudulent cases that appear legitimate, and which account for the majority of inappropriate indemnity (see, e.g., Derrig (2002)).

We therefore set our focus on developing audit strategies which help to improve the efficiency of the claims settlement process with regard to inconspicuous claims. The key element in our model framework - which pursues the costly state verification approach - is the use of threshold values which indicate whether an incoming claim should be verified or not. This approach allows the insurer to determine the optimal auditing strategy based on the magnitude of the filed claim. For this purpose, we make use of information on policyholder claiming and, more importantly, defrauding behavior which insurance companies are expected to hold due to previous experiences and verification processes. In particular, the share of fraud-prone policyholders among the population has an impact on the actual threshold values for auditing. Interviews with experts as well as previous research (see, e.g., Belhadji, Dionne, and Tarkhani (2000), Bermúdez, Pérez, Ayuso, Gómez, and Vázquez (2008)) have shown that numerous criteria exist which help indicate the potential fraud behavior of an individual. Such factors may include gender, nationality or place of residence. Additionally, investigations in the past have shown that certain accidents are more prone to fraud than others. For example, wind screen and glass damages as well as thefts inside the passenger compartment are known to be cases where policyholders are likely to report exaggerated claim amounts. Combining this information on individual characteristics with fraud signals presented by Dionne, Giuliano, and Picard (2009)), one obtains accurate estimators for the share of defrauders among the claimants.

In what follows, a model framework is presented based on a costly state verification setting in which - while policyholders observe the amount of loss privately - the insurance company can decide to audit incoming claims at some cost. Our goal is to derive optimal auditing strategies from the insurance company perspective indicating which of the incoming claims are subject to verification. In any case, we take into account that contract conditions must still be attractive to policyholders such that they are willing to adhere to the insurance relationship. A key element of our study is the option of both stakeholders adapting their behavior respectively based on different signals is taken into account. Subsequently, the impact of the optimal auditing strategy on the insurers' profitability is analyzed.

Unlike other strands of research in this area (e.g., Townsend (1979), Picard and Fagart (1999)), which often set up the problem of optimizing insurance contracts in a way such that policyholders always report loss amounts truthfully in equilibrium, we do not develop an auditing scheme which completely prevents or deters insurance claims fraud. From a macroeconomic point of view, both the approach and the result are desirable. From a single company perspective, however, this does not necessarily hold true. Similarly as Dionne, Giuliano, and Picard (2009), our aim is to minimize the insurer's overall costs from fraud, in general including paying some portion of fraudulent claims, as well as incurring expenses to detect and address fraud, and in experiencing the reputational effects of investigations themselves. In fact, we can show that for low costs per audit process the insurance company's net present value increases in the presence of fraud compared to the case where no fraud exists. Our findings are in line with the assumption made by Watt (2003).

Allowing for the existence of some fraudulent activities accounts for the widespread attitude to consider insurance as an investment which is expected to yield a return. As evidence, a study from GDV (2011) reported that more than 20% of the Germans consider insurance fraud to be a "gentlemen's offence" which is committed by almost everyone at least once. This point of view can be found among all socio-demographic groups. Duffield and Grabosky (2001) compile different approaches individuals use to justify fraud. Among others, insurance fraud is assumed to cause no significant harm; insurers are accepted targets which can afford it; build-up is a way to recover past premium payments (see also Miyazaki (2008)). Given this line of reasoning, we assume therefore that there will always exist policyholders trying to defraud in terms of inflating their magnitude of loss. In this context, auditing processes can help to minimize the share of fraudulent activities among the policyholder population. But a complete eradication seems unlikely.

Including the option to change one's strategy based on signals provides another insightful result. We can show that there exist situations where - while the possibility to adapt one's behavior might be desirable for the insurance company - this option is disadvantageous from the policyholder perspective. This observation stands in contrast to the widespread opinion that the adaptation of the defrauding strategy, especially based on signals from service providers, would be favorable from the individuals point of view.

We assume that while the policyholders and the service providers, which might be involved, may obtain signals and information based on which they change the defrauding strategy, they do not know the exact auditing threshold values nor do they have enough information to derive them themselves. This

is a crucial assumption in our model which also seems to be realistic. Leaving out this assumption, any auditing strategy would be redundant since the policyholder would know how to adapt his fraud behavior in a way to avoid being caught. From the insurance company perspective, it would not make sense to verify any incoming claim in this case.

The remainder of this paper is organized as follows: We start by presenting the model framework and optimization problem in Section 2. Section 3 constitutes the introduction of the policyholder's and insurer's respective strategies as well as the behavioral adaptation process. The corresponding numerical results are presented in Section 4. In Section 5, we analyze practical implications, before we conclude in Section 6.

2 Model Framework and Stakeholders' Positions

We consider an insurance company with a fixed number of policyholders. The latter are assumed to differ from one another only in their willingness to defraud, i.e., the population consists of both honest individuals who never commit fraud and those who defraud if it appears to be profitable. We denote the share of fraud-prone policyholders among the population with p . Depending on the loss amount and the prevalent defrauding strategy, the share of claims which actually contain build-up might be lower.

Since we assume all individuals to belong to the existing policyholder population, they all pay an insurance premium P at the beginning of a period, i.e., in $t = t_0$. We assume this premium to equal the given market premium which is demanded by other existing insurance companies as well. At the same time, with probability $0 \leq \pi \leq 1$, they face some uncertain loss θ of stochastic amount which, by the time of occurrence, is observed privately. In case a policyholder suffers a loss, he chooses to file a claim of some size $\hat{\Theta}(\theta)$ during the period (t_0, t_1) . In case of honest behavior, the amount of the claim will equal the actual loss, i.e., $\hat{\theta} = \theta$. If the individual decides to defraud, he reports some finite $\hat{\theta} > \theta$. Equation (1) summarizes all values the policyholder can report to the insurance company in the claiming scheme $\hat{\Theta}$:

$$\hat{\theta} = \hat{\Theta}(\theta) \begin{cases} = 0, & \text{no loss occurred} \\ = \theta, & \text{loss but no fraud} \\ > \theta, & \text{loss and fraud} \end{cases} \quad (1)$$

In the course of this paper, we will use the notation

$$\mathcal{F} = \{\hat{\theta} | \hat{\theta} > \theta\} \quad (2)$$

to denote the set consisting of all fraudulent claims filed by the policyholders. The elements in \mathcal{F} are characterized by the individuals' defrauding strategy. We will analyze different scenarios for the fraud behavior in the course of this paper.

The insurance company however has no information about the true loss amount. It therefore has the opportunity to verify the truthfulness of incoming claims. However, this comes at some cost per audit k .

In this model framework, we take the audit process to be perfect, i.e., fraud is detected with probability of one anytime auditing is performed. As a consequence of detected fraud, the insurance company will reject any indemnification.¹ Figure 1 illustrates the interaction of the different processes introduced in this model framework.

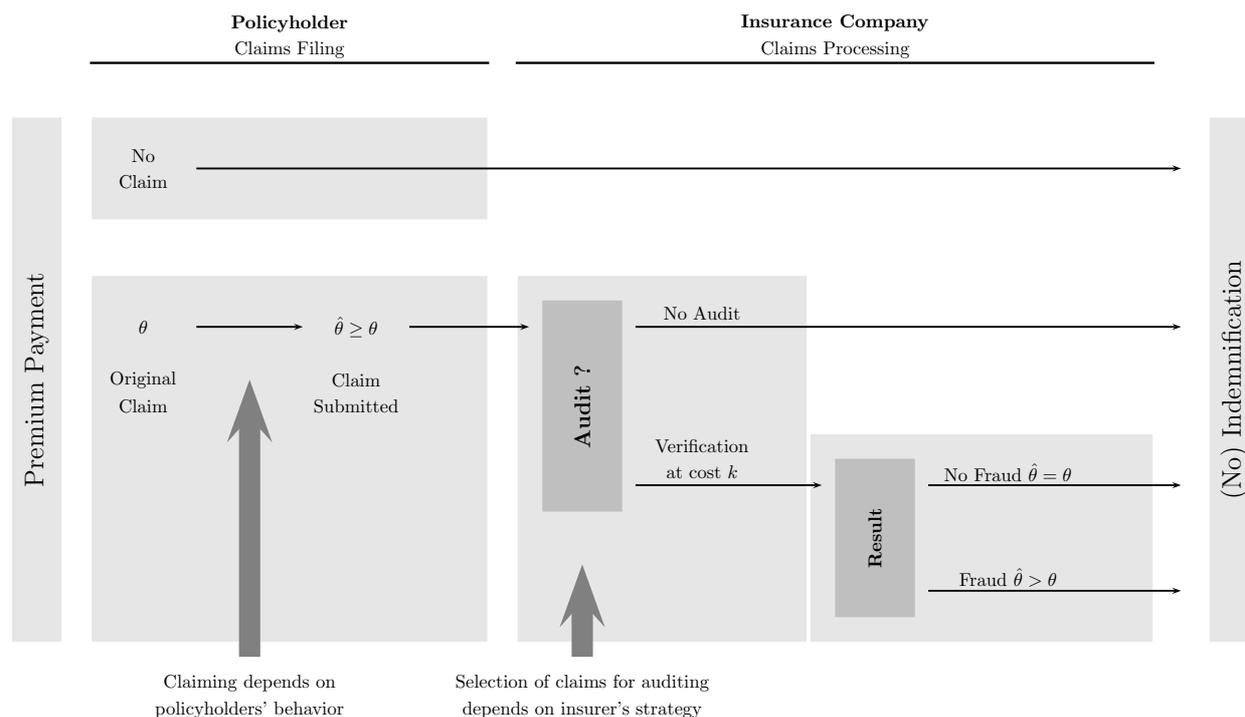


Figure 1: Overview of the processes associated with the filing and handling of insurance claims over the course of one period. An indemnification may or may not be paid out by the end of the period depending on the previous events.

Following Figure 1, the payment of an indemnification depends on several aspects. On the one hand, the policyholder needs to have suffered an insured loss² during the course of the observation period. Otherwise, he would not have a reason to file a claim. In the case when the claimant decides to defraud, his indemnification is dependent on whether auditing takes place or not: If the reported loss is not verified, he receives the payment of $\hat{\theta}$. If the filed claim undergoes an auditing process, however, the attempt to defraud will be revealed and any indemnification payment will be rejected, i.e., both excess and actual loss amount are denied. In case the policyholder belongs to the group of honest individuals, his loss θ will be indemnified no matter whether auditing took place or not.

¹In practice, gradations with regard to the indemnification are possible, i.e., the insurer might decide to pay the full loss amount or parts of it.

²Remember that we are considering only the situation where fraud will be committed through claim build-up and not through planned fraud.

Two key elements in the model framework which we have not elaborated on so far, are the behavioral strategies of both stakeholders: The amount of the reported claim in case of dishonesty is defined by the policyholders' defrauding behavior, while the decision whether or not to induce a verification process is dependent on the insurance company's auditing strategy. We present and analyze different examples for the respective strategies as well as behavioral adaptations in Sections 3 and 4.

2.1 Insurance Company: Contribution Margin and Participation Constraint

From the insurance company perspective, we observe the future cash flows at the beginning of a period in $t = t_0$ and at the time of loss realization and settling at the end of that period in $t = t_1$ and analyze the resulting contribution margin per contract CM in (t_0, t_1) . The insurance company receives a premium payment P from each policyholder in $t = t_0$. As already presented in the beginning of Section 2, the outflows of each period in $t = t_1$ depend on whether claims have been filed by the policyholders, and if so, whether they were audited or not and the result of the eventual verification process (see Figure 1).

A key element in this model framework is the insurance company's auditing strategy. It indicates which of the incoming claims are subject to verification and which of them are indemnified without verification. We use the following general notation to denote the auditing strategy

$$\mathcal{A} = \{\hat{\theta} | \hat{\theta} \text{ is audited}\}. \quad (3)$$

We distinguish four scenarios which lead to different values in the contribution margin for a single contract $CM(P, k, \hat{\theta}, \mathcal{A}, \mathcal{F})$. Using the notations introduced in Equations (2) and (3), we define

$$CM(P, k, \hat{\theta}, \mathcal{A}, \mathcal{F}) = \begin{cases} P & , \text{ no loss occurred, i.e., } \hat{\theta} = 0 \\ P - \hat{\theta} & , \text{ no auditing, i.e., } \hat{\theta} \in \mathcal{A}^c \\ P - \hat{\theta} - k & , \text{ auditing, but no fraud, i.e., } \hat{\theta} \in \mathcal{A} \cap \mathcal{F}^c \\ P - k & , \text{ auditing, fraud detected, i.e., } \hat{\theta} \in \mathcal{A} \cap \mathcal{F} \end{cases}, \quad (4)$$

which can also be stated as:

$$\begin{aligned} CM(P, k, \hat{\theta}, \mathcal{A}, \mathcal{F}) &= P - \hat{\theta} \cdot \mathbb{1}_{\mathcal{A}^c}(\hat{\theta}) - \hat{\theta} \cdot \mathbb{1}_{\mathcal{A} \cap \mathcal{F}^c}(\hat{\theta}) - k \cdot \mathbb{1}_{\mathcal{A}}(\hat{\theta}) \\ &= P - \hat{\theta} \left[1 - \mathbb{1}_{\mathcal{A} \cap \mathcal{F}}(\hat{\theta}) \right] - k \cdot \mathbb{1}_{\mathcal{A}}(\hat{\theta}), \end{aligned} \quad (5)$$

where $\mathbb{1}_X(y)$ denotes the indicator function, i.e., it takes the value 1 if y is in the set X or 0 if y is not in X .

Hence, in this context $\mathbb{1}_{\mathcal{A}}(\hat{\theta})$ represents the number of claims which were subject to verification whereas $\mathbb{1}_{\mathcal{A}^c}(\hat{\theta})$ states the number of claims which were indemnified without auditing. $\mathbb{1}_{\mathcal{A} \cap \mathcal{F}^c}(\hat{\theta})$ counts the number of cases when fraudulent claims underwent an auditing process, i.e., the number of cases when the attempt to defraud was unveiled while $\mathbb{1}_{\mathcal{A} \cap \mathcal{F}}(\hat{\theta})$ returns the number of cases when honest claims

where verified.

Overall, the insurance company is interested in its whole policyholder population and hence considers the average contribution margin per contract $\mathbb{E}(CM)$, i.e., the net present value of future incoming and outgoing cash flows NPV . We define :

$$NPV = NPV(P, k, \hat{\theta}, \mathcal{A}, \mathcal{F}) = \mathbb{E}(CM(P, k, \hat{\theta}, \mathcal{A}, \mathcal{F})), \quad (6)$$

where $\mathbb{E}(Y)$ denotes the expected value of the stochastic variable Y . Hence, we analyze the insurance company's net present value as the expected value of future cash flows discounted at the risk-free rate $r_f = 0$. This assumption $r_f = 0$ holds throughout the paper.

Based on Equation (6), we can formulate the insurance company's participation constraint.

Condition 1 *The insurance company is willing to maintain the insurance relationship with its policyholder population if the net present value per contract is non-negative:*

$$NPV \geq 0. \quad (7)$$

2.2 Policyholder: Expected Utility and Participation Constraint

From the policyholder perspective, we analyze their wealth positions and the corresponding average expected utility at the end of the period $t = t_1$ in the framework introduced above. We assume each individual initially to hold the same wealth position W^3 and to take out the same insurance contract. At the beginning of the observation period in $t = t_0$, the payment of the insurance premium P is due. Consequently, each policyholder is endowed with the wealth position $W_0^I = W - P$, where the superscript I indicates the existence of an insurance contract.

The consecutive development of this wealth position W_0^I depends on whether the policyholders suffer an insured loss and if so, whether they choose to report their loss truthfully or not and whether their fraud is being revealed in case of dishonesty or not. Taking the occurrence of loss into account, the wealth position at the end of the observation period in $t = t_1$ is given by

$$W_1^I = W - P - \theta + \hat{\theta} \left[1 - \mathbf{1}_{\mathcal{A} \cap \mathcal{F}}(\hat{\theta}) \right], \quad (8)$$

where W is invested riskless with $r_f = 0$. In the case when no insured loss has occurred throughout the observation period, Equation (8) simplifies to $W_1^I = W_0^I = W - P$.

In order to be able to formulate a participation constraint from the policyholder perspective, we consider the development of their wealth position throughout the observation period when not having signed

³Since the actual value of the initial wealth position will have no impact on the decision whether to maintain the insurance relationship or not, the policyholders might even be endowed with different wealth positions. We assume the latter to be invested safely with the risk-free rate $r_f = 0$.

an insurance contract prior to the occurrence of loss. In $t = t_0$, the individuals would not have to make a payment in the amount of the insurance premium. Hence, the corresponding wealth position is given by $W_0^N = W$, where the superscript N implies the absence of an insurance relationship. If some loss θ occurs up until $t = t_1$, this amount decreases to $W_1^N = W_0^N - \theta = W - \theta$, whereas, without loss during the observation period, we have $W_1^N = W_0^N = W$.

We assume the policyholders' expected utility to be described by a standard mean-variance utility function of the corresponding wealth position. The degree of risk aversion is expressed by the parameter $a (> 0)$. Generally, for a given stochastic wealth position Z , its expected utility for the individual is given by $U(Z) = \mathbb{E}(Z) - \frac{a}{2}\text{Var}(Z)$, where $\text{Var}(Z)$ denotes the variance of the stochastic variable Z .

Using the equations above and considering that the probability of loss occurrence is denoted by π , the final expected utility in case no insurance contract exists can be written as

$$U(W_1^N) = W - \pi\mathbb{E}(\theta) - \frac{a}{2}\pi^2\text{Var}(\theta). \quad (9)$$

Similarly, for the setting where an insurance relationship between policyholders and insurance company already exists, the final expected utility is given by

$$U(W_1^I) = W - P - \mathbb{E}\left(\pi\theta - \hat{\theta}\left[1 - \mathbb{1}_{\mathcal{A} \cap \mathcal{F}}(\hat{\theta})\right]\right) - \frac{a}{2}\text{Var}\left(\pi\theta - \hat{\theta}\left[1 - \mathbb{1}_{\mathcal{A} \cap \mathcal{F}}(\hat{\theta})\right]\right). \quad (10)$$

Comparing Equation (10) with Equation (9) results in the policyholders' participation constraint. For this purpose, we introduce the notion of the gain in expected utility from having signed an insurance contract

$$\Delta U = U(W_1^I) - U(W_1^N). \quad (11)$$

Condition 2 *The policyholders are willing to maintain the insurance relationship with the insurance company if their final expected utility is greater with having insurance than without it, i.e.,*

$$\Delta U \geq 0. \quad (12)$$

We want to point out that the policyholders' participation constraint functions the same in the case of a multi-period model, i.e., they do not quit the insurance relationship unless their gain in expected utility from having insurance coverage ΔU is negative. In particular, the insurance company is in the position to optimize its position while the policyholders would choose to cancel their insurance contracts only if $\Delta U < 0$.

2.3 Optimization Problem

Summing up the information we have presented so far with regard to the model framework as well as the insurance company's and the policyholders' participation constraints, we can formulate the resulting

optimization problem.

The insurance company is aiming to derive an auditing strategy \mathcal{A} such that its net present value of future cash flows NPV is maximized. At the same time, it needs to be made sure that the stakeholders are willing to maintain their insurance relationships, i.e., Equations (7) and (12) hold. Formally, this objective is given by

$$\begin{cases} \max_{\mathcal{A}} NPV(P, k, \hat{\theta}, \mathcal{F}, \mathcal{A}) \\ NPV \geq 0 \\ \Delta U \geq 0 \end{cases} . \quad (13)$$

We make the assumption that the former charges the given market premium P . In particular, this implies that an adaptation of the insurance premium as part of the fraud handling is not feasible. Furthermore, we assume no interaction with the number of acquired contracts.

3 Optimal Auditing Strategies

Section 2 constitutes the introduction of the model framework as well as the insurance company's and policyholders' behavioral strategies in general. In this section, we present a case for the individuals' defrauding behavior as well as a suitable approach to deriving the resulting optimal auditing strategy. Furthermore, the process of behavioral adaptation is included.

3.1 Policyholder Claiming Scheme

As discussed above, we consider insurance fraud in the form of build-up, i.e., in the case of occurrence of loss some of the policyholders among the population decide to file an exaggerated fraudulent claim if it appears to be profitable.

In the context of our study, we assume that the fraud amount by inflating the magnitude of a loss not be a decision variable for the policyholders. In case an individual decides to engage in fraudulent activities, he or she will consider the build-up as a percental "surcharge" on the actual loss amount instead of filing some value which deviates significantly from the actual loss amount. This approach increases the likelihood of a fraudulent claim to be perceived as legitimate by the insurance company, and not to be audited as a consequence. This assumption is in line with Viaene and Dedene (2004) who find that policyholders involved in soft fraud typically tend to file claims containing small fraud amounts. Further evidence can be found in the behavioral economic theory literature. Individuals weigh the consequences accompanied by losses stronger than the ones from a gain of the same size (see, e.g., Kahneman and Tversky (1979), Kerr (2012)). Applied to the context of our model framework this implies that the loss from being caught committing soft fraud (i.e., indemnification is waived completely) is perceived as a higher burden than the potential profit from a successful build-up attempt. Additionally,

regarding build-up as a surcharge on the loss suffered gives policyholders the opportunity to "back down and claim the appropriate amount" if investigated by the insurance company rather than in the case of outright fraud (see, e.g., Emerson (1992)).

From the perspective of the repair shops, which might also be involved in fraudulent activities, extensive inflations of the loss amount seem to be unlikely as well. According to Hubbard (2002), service providers in general have a reputational incentive to act in their clients favor since the latter tend to return more often if they are satisfied with previous services. These findings are in line with statements made by experts from an insurance company whose experiences have shown that repair shops orientate themselves on the actual loss amount when trying to charge too much for certain services. This way even exaggerated claims seem legitimate and are less likely to undergo an auditing process resulting in direct indemnification.

As part of the defrauding strategy, we assume that whenever those policyholders have suffered a loss during the observation period, they report back a multiple of the actual loss amount. We denote this constant multiplicative factor by $\alpha = \hat{\theta}/\theta$. However, it is well known that insurance companies do perform audits in order to verify the truthfulness of incoming claims. Consequently, fraud-prone policyholders can be expected to adjust their fraud strategy accordingly. We assume that all of these individuals have an inner threshold value for defrauding $\hat{\theta}_{\text{ph}}^*$. With regard to their fraud strategy this threshold value implies that they apply their fraud strategy α up to the preset threshold value $\hat{\theta}_{\text{ph}}^*$. To be more precise, if the amount of the actual loss θ is smaller than the threshold $\hat{\theta}_{\text{ph}}^*$, the minimum of $\alpha\theta$ and $\hat{\theta}_{\text{ph}}^*$ is reported to the insurance company. However, if the amount of the loss suffered already exceeds that threshold, the policyholder claims this amount truthfully. We can write this strategy as

$$\hat{\theta} = \hat{\Theta}(\theta, \alpha, \hat{\theta}_{\text{ph}}^*) = \begin{cases} 0 & , \text{ no loss occurred} \\ \theta & , \text{ no fraud or } \theta > \hat{\theta}_{\text{ph}}^* \\ \min\{\alpha\theta, \hat{\theta}_{\text{ph}}^*\} & , \text{ fraud and } \theta < \hat{\theta}_{\text{ph}}^* \end{cases} . \quad (14)$$

The set consisting of all fraudulent claims associated with this strategy $\hat{\Theta}(\theta, \alpha, \hat{\theta}_{\text{ph}}^*)$ is denoted by $\mathcal{F}_{\alpha, \hat{\theta}_{\text{ph}}^*}$.

The question arises as to how policyholders determine an adequate threshold value for defrauding. Here we refer to the data presented in Section 1 regarding survey reports indicating that policyholders may use third parties, such as repair shops, to act as their partners in fraud. These repair shops are likely to have developed insight about the value of $\hat{\theta}^*$ due to repeated experiences in repairing cars for insured damages. Based on this information they can make proper assumptions concerning the insurance companies' prevalent auditing strategies. There are two potential ways repair shops can proceed: provide too much service or bill too much for a particular service (see Dulleck and Kerschbamer (2006)). A particular phenomenon is to include the cost to fix existing damages when repairing the ones which were caused by a current insured event. Such actions might often be undertaken even without the knowledge of the policyholder. On the one hand, the latter might not have the incentive to carefully

check the work performed by repair shops since the bill is passed on to the insurance company (see Tennyson (2008)). On the other hand, the policyholders might lack the necessary know-how to do so (see Dulleck and Kerschbamer (2006)). Policyholders, however, in particular the ones who are willing to defraud if it appears to be advantageous, may also seek information themselves to adapt their claiming behavior for purposes of defrauding the insurer. They may seek out others for ideas of successful fraud or perhaps be approached by third parties to participate in defrauding activities. Staying in the field of auto insurance, fraudulent activities may be performed by mutual agreement between the policyholder and the corresponding repair shop who share the excess insured payments with one another. In particular, there exists the possibility to adjust the policyholders' defrauding strategy as presented in Equation (14). Recall that we make the crucial assumption that neither the policyholders nor a third party other than the insurance company itself holds exact information concerning the prevalent auditing strategy. The contrary case would make the principal of auditing redundant.

3.2 Insurance Company Auditing Strategy

From the insurance company perspective, this fraud strategy implies the necessity to adjust its auditing strategy \mathcal{A} . After having performed audits for at least one period, the company should have gained enough information on policyholder fraud behavior to do so. Revealed fraud can serve as an especially helpful information base for improving the existing auditing strategy. Assuming that a sufficiently high number of verification processes has been performed (which has revealed sufficiently many defrauding attempts), the insurer will note that fraudulent claims do not exceed a certain threshold value. Hence, it is unnecessary to audit claims above that value, allowing insurers instead to verify medium sized claims. Consequently, in this case the aim is to derive the optimal auditing range \mathcal{A}_R . The upper bound of this range would have to be the policyholders' persistent inner threshold value $\hat{\theta}_{ph}^*$. Unfortunately, the insurance company does not hold this information. However, an adequate estimate for this value is the maximum fraud amount which was detected during the last period. We denote this value by $\hat{\theta}_{max}$ and formally define it as

$$\hat{\theta}_{max} = \max\{\hat{\theta} \cdot \mathbb{1}_{\mathcal{A} \cap \mathcal{F}}\}. \quad (15)$$

Setting the upper bound of the auditing range \mathcal{A}_R to $\hat{\theta}_{max}$ practically implies that no incoming claim above that value will be verified. However, since one cannot be absolutely sure whether the actual maximum of all fraudulent claims has been determined, i.e., whether or not fraud is being committed beyond $\hat{\theta}_{max}$, it is reasonable to include a safety margin $s > 0$. Consequently, the updated upper bound of the auditing range is set $(1 + s) \cdot \hat{\theta}_{max}$.

The second parameter that characterizes the audit range \mathcal{A}_R is its lower bound which is expressed by $\hat{\theta}_R^*$. Summing up, we can formulate this auditing strategy as

$$\mathcal{A}_R = \mathcal{A}_R(\hat{\theta}_R^*, s, \hat{\theta}_{max}) = \{\hat{\theta} | \hat{\theta}_R^* \leq \hat{\theta} \leq (1 + s) \cdot \hat{\theta}_{max}\}. \quad (16)$$

We can specify the optimization problem given in Equation (13) with regard to this new setting. The aim here is to derive the optimal lower bound $\hat{\theta}_R^*$ of the auditing range such that the insurance company's net present value NPV is maximized:

$$\begin{cases} \max_{\hat{\theta}_R^*} NPV(P, k, \hat{\theta}, \mathcal{F}_{\alpha, \hat{\theta}_{ph}^*}, \mathcal{A}_R) \\ NPV \geq 0 \\ \Delta U \geq 0. \end{cases} \quad (17)$$

As already mentioned above, the auditing strategy \mathcal{A}_R requires the availability of information with regard to the prevalent defrauding behavior, i.e., $\hat{\theta}_{max}$ needs to be determined. Consequently, \mathcal{A}_R is applicable after the first observation period the earliest when enough audits have been performed to specify the value of $\hat{\theta}_{max}$. However, this implies the need for a different verification scheme for the very first observation period. In our case, we assume that the insurance company will revert to an initial strategy for auditing during the first period. This one is characterized by a threshold value $\hat{\theta}_{init}^*$, i.e., all incoming claims which exceed this threshold are subject to verification whereas claims whose amount is below this indicator will be indemnified right away. We denote this specific auditing strategy by \mathcal{A}_{init} and define it formally as

$$\mathcal{A}_{init} = \mathcal{A}_{init}(\hat{\theta}_{init}^*) = \{\hat{\theta} | \hat{\theta}_{init}^* \leq \hat{\theta}\}. \quad (18)$$

Then, when having gathered enough information throughout the first observation period, the insurer switches to the auditing range \mathcal{A}_R in the consecutive periods.

3.3 Behavioral Adaptation

So far, the policyholders in the population who are likely and willing to defraud were assumed to adhere to a constant fraud strategy. They chose a constant multiplicative factor α and/or the same threshold value for defrauding $\hat{\theta}_{ph}^*$ over the course of several periods. This constancy results from the assumption that the policyholder population does not obtain any information on the insurance company's prevalent verification process. Hence, there was no basis to give occasion for an adjustment of their behavior. That also had an impact on the insurance company's corresponding optimal auditing strategy. As soon as the optimal verification process with respect to the prevalent defrauding behavior was found, i.e., right after the first observation period, there was no need for the insurance company to perform any adjustments to it. Summarizing, no one of the participants in the insurance relationship had to change their behavior.

However, it is realistic to assume that there might exist signals indicating whether the insurance company changes its auditing strategy in the subsequent period or not. Especially in the case of automobile insurance, signaling is issued by (authorized) repair shops. Since dealing with a large number of insured events, they are able to estimate changes in the auditing behavior of different insurance companies. In these cases, the policyholders who are prone to defrauding would be given a chance to change their fraud

behavior and react to the new verification scheme. In case the insurance company announces strengthened controls in order to combat insurance claims fraud, the fraudulent part of the policyholder population would choose to act more carefully in terms of their defrauding strategy, i.e., lower the multiplicative factor α and/or their threshold value for defrauding $\hat{\theta}_{\text{ph}}^*$. In the opposite case, knowing that the insurance company will relax their auditing scheme, it can be assumed that attempts are made to exaggerate the actual loss amount θ even more and obtain higher indemnification payments, i.e., by increasing α and/or $\hat{\theta}_{\text{ph}}^*$.

Figure 2 gives an overview of the interaction between insurance company and (defrauding) policyholders and the resulting adjustment processes of the respective behavioral strategies.

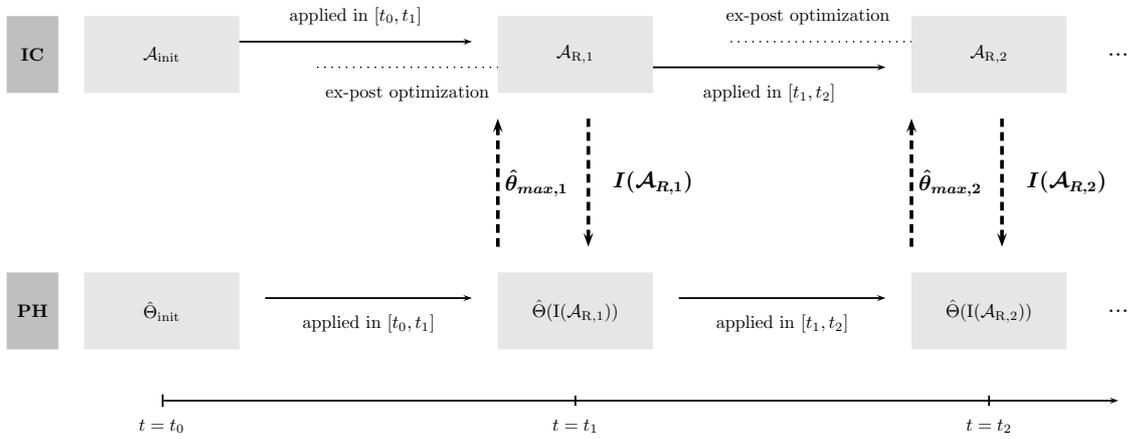


Figure 2: Interaction between insurance company (IC) and policyholders (PH) over the course of the first two periods in an insurance relationship.

Both stakeholders define their respective initial strategy at the beginning of the very first observation period in $t = t_0$. For the insurance company, this is its initial auditing strategy denoted by $\mathcal{A}_{\text{init}}$. Since by that point in time, no information regarding the policyholders' defrauding strategy is available, it will be characterized by an initial threshold value for auditing $\hat{\theta}_{\text{init}}^*$. The policyholders themselves choose a claiming scheme $\hat{\Theta}_{\text{init}}$.⁴ In particular, the initial fraud strategy is defined by the multiplicative factor α and an initial threshold value for defrauding $\hat{\theta}_{\text{ph},0}^*$ (see Equation (14)). These two strategies are applied throughout the first observation period $[t_0, t_1]$. At its end, in $t = t_1$, all information with regard to the actual distribution of the claimed losses throughout that period as well as indications on the policyholders' defrauding scheme, i.e., the maximum value of detected fraud $\hat{\theta}_{\text{max},1}$ in $[t_0, t_1]$, are available to the insurance company.⁵ Based on this information, the insurer determines its optimal auditing strategy $\mathcal{A}_{R,1}$ for the first observation period ex-post. The corresponding optimization problem is defined in Equation (17). This adjusted auditing scheme will then be applied in the second observation

⁴The function $\hat{\Theta}$ includes the defrauding strategy for the dishonest policyholders (see Equation (14)).

⁵We assume that a sufficiently high number of audits has been performed during the course of the first period which has resulted in the detection of fraudulent claims, i.e., $\hat{\theta}_{\text{max},1}$ is observed.

period $[t_1, t_2]$. At the same time, some signal $I(\mathcal{A}_{R,1})$ concerning the adjustment of the verification process is communicated to the policyholder population. They themselves now have the opportunity to adapt their behavior accordingly. While the honest policyholders adhere to reporting the actual loss amount when having suffered an insured loss, the ones who are willing to defraud adjust their threshold value for defrauding to $\hat{\theta}_{ph,1}^*$. This new claiming scheme will be denoted by $\hat{\Theta}(I(\mathcal{A}_{R,1}))$ and is applied in the second observation period $[t_1, t_2]$. However, the change in the defrauding behavior will be registered by the insurance company in the form of a different maximum value of detected fraud $\hat{\theta}_{max,2}$ in the course of that second period. This new piece of information on the policyholders' behavior induces once again an ex-post optimization of the prevalent auditing strategy to $\mathcal{A}_{R,2}$ at the end of this very period in $t = t_2$. Again, the adjusted verification scheme $\mathcal{A}_{R,2}$ is applied in the following observation period after having provided the policyholder population with a signal $I(\mathcal{A}_{R,2})$ concerning the change in auditing. The interaction and adaptation processes can be repeated in the same fashion over the course of several periods.

It needs to be emphasized that while we focus on the derivation of the optimal auditing strategies $\mathcal{A}_{R,n}$, we also assure that all participants are willing to maintain the insurance relationship, i.e., all participation constraints as defined in Equations (7) and (12) need to hold when applying the optimal verification process.

3.4 Numerical Implementation of Iterative Optimization

In this subsection, we present the iterative approach regarding the optimization of the auditing range $\mathcal{A}_{R,n}$ in the n^{th} period from the insurance company perspective when interaction and hence adaptation is observable.⁶

Step 1: Adjustment of claiming scheme We first consider the policyholders' claiming scheme $\hat{\Theta}_n(\theta_n, \alpha, \hat{\theta}_{ph,n}^*)$ is applied at the beginning of each iteration. It is characterized by the defrauders' strategy indicated by a constant multiplicative factor α and a threshold value for defrauding $\hat{\theta}_{ph,n}^*$ which is adjusted each period based on the signal $I(\mathcal{A}_{R,n-1})$. In accordance to Equation (14), we obtain for the policyholders' claiming scheme of the n^{th} period

$$\hat{\theta}_n = \hat{\Theta}_n(\theta_n, \alpha, \hat{\theta}_{ph,n}^*) = \begin{cases} 0 & , \text{ no loss occurred} \\ \theta_n & , \text{ no fraud or } \theta_n > \hat{\theta}_{ph,n}^* , \\ \min\{\alpha\theta_n, \hat{\theta}_{ph,n}^*\} & , \text{ fraud and } \theta_n < \hat{\theta}_{ph,n}^* \end{cases} \quad (19)$$

where θ_n represents a realization of the loss variable in the n^{th} period. The set consisting of all fraudulent claims associated with this strategy is denoted by $\mathcal{F}_{\alpha, \hat{\theta}_{ph,n}^*}$.

Step 2: Determining maximum value of detected fraud The next step to determining the optimal auditing strategy $\mathcal{A}_{R,n}$ is the identification of the maximum fraud value which was actually

⁶The subscripts n in the course of this subsection indicate the n^{th} iteration process, i.e., the respective quantities for the n^{th} observation period. Hereby, we consider all $n \geq 2$. The special case of $n = 1$ corresponds to the initial period.

detected in the n^{th} iteration, i.e., period. Apparently, it depends on the auditing strategy $\mathcal{A}_{R,n-1}$ which was derived to be optimal in the $(n-1)^{\text{th}}$ period and is then applied in the n^{th} period. Using the notation presented in Equation (5), the maximum value of detected fraud in the n^{th} period can be defined as

$$\hat{\theta}_{\max,n} = \max\{\hat{\theta}_n \cdot \mathbb{1}_{\mathcal{A}_{R,n-1} \cap \mathcal{F}_{\alpha, \hat{\theta}_{\text{ph},n}^*}}(\hat{\theta}_n)\}. \quad (20)$$

Step 3: Ex-post optimization of auditing strategy The value of $\hat{\theta}_{\max,n}$ forms the basis for the actual optimization process. Considering Equation (16), the insurance company's auditing strategy for the n^{th} period is given by

$$\mathcal{A}_{R,n} = \left\{ \hat{\theta}_n \mid \hat{\theta}_{R,n}^* \leq \hat{\theta}_n \leq (1+s) \cdot \hat{\theta}_{\max,n} \right\}, \quad (21)$$

with s being the safety margin.

The aim is now to find the lower bound $\hat{\theta}_{R,n}^*$ of this audit range such that the insurance company's net present value of future cash flows NPV is maximized and the stakeholders' participation constraints hold:

$$\begin{cases} \max_{\hat{\theta}_{R,n}^*} NPV(P, k, \hat{\theta}, \mathcal{F}_{\alpha, \hat{\theta}_{\text{ph},n}^*}, \mathcal{A}_{R,n}) \\ NPV_n \geq 0 \\ \Delta U_n \geq 0, \end{cases} \quad (22)$$

where ΔU_n denotes the policyholders' gain in expected utility from having insurance coverage in the n^{th} period.

Step 4: Communication of signal After having found the optimum value for the lower bound $\hat{\theta}_{R,n}^*$ of the auditing range, i.e., the optimal auditing strategy $\mathcal{A}_{R,n}$, at the end of observation period n , a signal $I(\mathcal{A}_{R,n})$ is communicated to the policyholder population informing about the adjustment of the prevalent verification scheme. For this purpose, we define this signal as follows:

$$I(\mathcal{A}_{R,n}) = \left(\frac{\hat{\theta}_{R,n}^*}{\hat{\theta}_{R,n-1}^*} + \frac{\hat{\theta}_{\max,n}}{\hat{\theta}_{\max,n-1}} \right) / 2. \quad (23)$$

This signal is used to adapt the claiming scheme $\hat{\Theta}_{n+1}$ at the beginning of period $n+1$. In this context, the signal I is an average of two ratios: The first one represents the change of the upper bound from period $n-1$ to n whereas the second one constitutes the change of the maximum value of detected fraud from period $n-1$ to n .

4 Simulation Results

In Sections 2 and 3, we have presented a model framework and potential optimal auditing strategies when interaction between policyholder and insurance company and consequently behavioral adaptation over the course of several observation periods is possible. In this section, we present and analyze the corresponding numerical solutions for the optimal auditing strategy from the insurance company perspective.

4.1 Parametrization of the Reference Setting

In our simulations, we consider a policyholder population consisting of $M = 2'500'000$ individuals. An assumed probability of loss occurrence of $\pi = 0.2$ leads to $N = 500'000$ loss realizations of θ per observation period. Hereby, the latter follow a log-normal distribution. This assumption is commonly used as mentioned in Marlin (1984) since it guarantees positive values for the realizations of the random variable. In particular, the expected value $\mathbb{E}(\theta)$ is set $\mu = 1$ and the variance $Var(\theta) = \sigma^2 = 0.4$. All individuals among the population are assumed to be risk averse. For instance, their risk aversion parameter a is considered to be 6. Furthermore, the policyholders' initial wealth position is set $W = 0$. At the same time, the individuals have to pay an insurance premium P at the beginning of each observation period. The latter can be split up into the fair premium and an appropriate loading factor. The fair premium corresponds to the expected loss. Hence, having set the expected value of the loss variable θ to $\mu = 1$ and considering the probability of suffering a loss to be $\pi = 0.2$, this implies a fair premium of 0.2. However, since the insurance company faces additional costs, it will add a corresponding loading factor to the fair premium. As mentioned in Cummins and Mahul (2004), the loading factor can not be too large since potential policyholders would not sign the insurance contract under such conditions. For our analyses, we will assume the total insurance premium to be $P = 0.3$. Furthermore, the cost per audit is set $k = 0.05$ which corresponds to 16.67% of the insurance premium P . For the purpose of our analyses, we will disregard costs other than the ones due to auditing. The share of fraud-prone policyholders among the population is assigned the value $p = 0.2$, i.e., 20% of all policyholders who suffer an insured loss may exaggerate that amount if it is in accordance with their defrauding strategy. Their defrauding strategy is accompanied by the choice of the relative fraud amount α and/or an appropriate threshold value for defrauding $\hat{\theta}_{ph}^*$. To start with, we take α to be 2, i.e., the policyholders who decide to defraud report back an amount twice as high as the actual loss amount. However, in case the individuals have a threshold value for defrauding as described in Section 3.1, they never claim more than that value $\hat{\theta}_{ph}^*$. For the purpose of our analyses, we assume $\hat{\theta}_{ph}^* = 1.1$ which is 10% higher than the expected value of the loss variable. Finally, the insurance company needs to decide on the parameters concerning its auditing strategy \mathcal{A} . During the very first period, it opts for a verification process \mathcal{A}_{init} which is characterized by a threshold value for auditing $\hat{\theta}_{init}^*$, we set its initial value to 1 which corresponds to the expected value of the loss variable θ . As already explained at the end of Section 3.2, information with regard to the policyholders' defrauding behavior needs to be gathered first before being able to apply this verification scheme \mathcal{A}_R . Hence, we initially set $\hat{\theta}_{init}^* = 1$, and determine the parameters $\hat{\theta}_{max}^*$ and $\hat{\theta}_R^*$ based on the information obtained based on the first program run. With regard to the upper bound of the auditing

range, the value for the safety margin s is assumed to be 0.1, i.e., the upper bound of the auditing range is 10% higher than maximum of detected fraud $\hat{\theta}_{\max}$. The resulting auditing range \mathcal{A}_R is then applied in a consecutive program run.

Table 1 sums up the choices for the input parameters for the reference setting as introduced above. In the course of this and the following sections, we base our simulations and studies on these values.

Input parameter		Reference level
Total number of policyholders	M	2'500'000
Number of loss realizations	N	500'000
Loss distribution	θ	$\ln\mathcal{N}(1, 0.4)$
Insurance premium	P	0.3
Share of fraud-prone policyholders	p	0.2
Relative fraud amount	α	2
Initial threshold value for auditing	$\hat{\theta}_{\text{init}}^*$	1
Safety margin	s	0.1
Policyholder's initial threshold	$\hat{\theta}_{\text{ph},0}^*$	1.1
Auditing cost	k	0.05
Risk aversion parameter	a	6

Table 1: Input parameters for the reference setting.

4.2 Simulation Results and Sensitivity Analyses

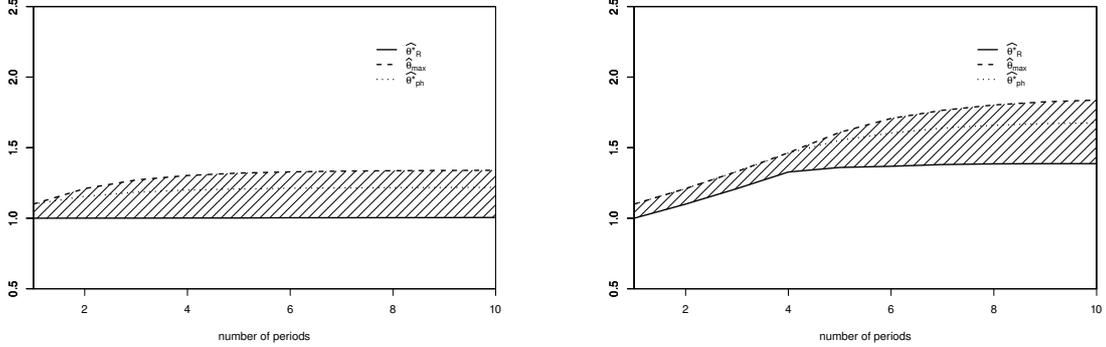
The remainder of this section constitutes the presentation and discussion of the simulation results. For the simulations, we adhere to the parameter choice presented in Table 1 unless noted otherwise.

4.2.1 Development of Optimization Results Over Several Iterations

In this subsection, we present and discuss the development of the optimal auditing range \mathcal{A}_R over the course of several iterations. Furthermore, we analyze its impact on quantities like the number of performed audits, the amount of fraudulent claims, the net present value NPV and the gain in utility ΔU . To get a better insight of the effects, we consider both the parametrization of the reference setting with costs per audit $k = 0.05$ as well as the case when the costs per audit are raised to $k = 0.3$ and compare the results.

Development of Optimal Auditing Range \mathcal{A}_R

Comparing Figure 3(a) with Figure 3(b), we can see that higher costs per audit k result in a slightly broader optimal auditing range \mathcal{A}_R , i.e., the share of values which may be verified becomes greater. At



(a) Development of optimal auditing strategy, $k = 0.05$ (b) Development of optimal auditing strategy, $k = 0.3$

Figure 3: Development of the optimal auditing strategy throughout the course of several iterations for two different choices of the cost per audit k . The remaining parameters are chosen as presented in Table 1.

the same time, the optimal auditing range shifts in an upward direction, i.e., the value of the claims which may be subject to auditing becomes higher. For the insurance company this implies that, in case of high expenditures per case, they should focus their investigations on those claim which exhibit high saving potential whenever an engagement in fraudulent activities is detected.

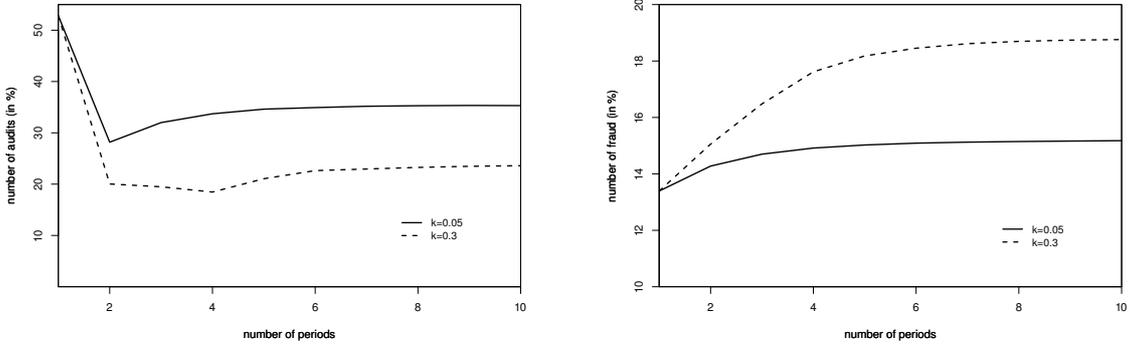
From the policyholder perspective, we find that higher costs per audit k lead to a higher threshold of defrauding $\hat{\theta}_{ph}^*$, i.e., the value up to which policyholders take build-ups into consideration increases. This observation shows that the cost per audit does not only have an impact on the insurance companies auditing strategy but indirectly also on the policyholders' behavior, in particular on their defrauding strategy. Since the insurer signals an upward shift in his verification behavior, policyholders (and the corresponding service providers) are left with the impression that inflating a loss amount up to some value is more likely to remain undetected than in the previous period. In return, they then raise their threshold value for defrauding.

Development of Number of Performed Audits and Fraudulent Claims

We measure both the number of performed audits and the number of fraudulent claims in relation to the total number of filed claims.

Figure 4(a) confirms the intuition that higher costs per audit k result in a lower share of incoming claims which are subject to verification, i.e., fewer auditing processes are performed. As mentioned above, this restriction in the number of verification processes leads the insurance companies to focus on those claims which, in case of detected fraud, have a higher saving potential, i.e., higher valued claims.

From the policyholder point of view, the number of fraudulent claims consequently increases (see Figure 4(b)). This finding is in line with our results presented in Figure 3. Due to receiving a signal indicating an upward shift, the policyholders themselves raise their threshold value for defrauding. As a



(a) Development of the number of audits performed by the insurance company when applying the optimal auditing strategy for the respective period.

(b) Development of the number of fraudulent claims when the insurance company applies the optimal auditing strategy for the respective period.

Figure 4: Development of number of audits and the number of fraudulent claims over the course of several iterations measured in relation to total number of losses, i.e., filed claims. The development is illustrated for two different choices of the cost per audit k . The remaining parameters are chosen as presented in Table 1.

consequence of this elevation, the number of losses below the threshold increases resulting in more cases where actions are taken to inflate the claim amount.

It needs to be noted that even though 20% of the policyholder population are willing to exaggerate their loss amount, the real numbers lie below that value (see Figure 4(b)). This phenomenon can be explained with the existence of the threshold value for defrauding $\hat{\theta}_{ph}^*$ up to which fraud is actually taken into consideration. The actual share of build-up among all claims may depend on different factors. As already seen in Figure 3, higher costs per audit k result in an increased defrauding threshold $\hat{\theta}_{ph}^*$. This, however, implies a higher likelihood of the actual loss amount being below the threshold value. As a consequence, the amount of fraud may increase up to the maximum of 20%.

Development of Net Present Value NPV

Looking at Figure 5, it strikes attention that both NPV and ΔU are positive for both chosen values of the cost per audit k . Especially with regard to gain in utility ΔU , this implies that the policyholders among the population are willing to adhere to the insurance relationship (see Equation (12)). This observation proves that the derived optimal auditing schemes are feasible from both stakeholders' perspectives.

From the insurance company point of view, higher costs per audit k lead to a lower net present value NPV . This phenomenon may be explained by the fact that higher auditing costs result in fewer auditing processes (see Figure 4(a)) but in a higher share of fraudulent claims (see Figure 4(b)). As a consequence, more exaggerated claim amounts remain undetected resulting in higher, unjustified expenses from the insurance company perspective.

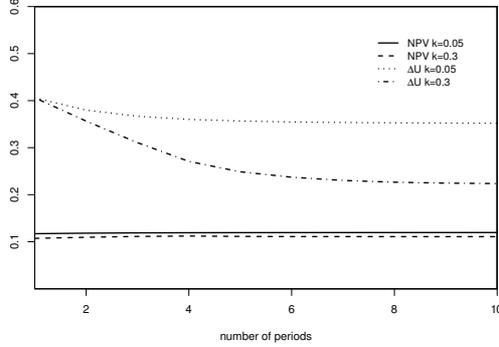


Figure 5: Development of the insurance company’s net present value and the policyholders’ gain in utility respectively over the course of several iterations when applying the optimal auditing strategy for the respective period. The development is illustrated for two different choices of the cost per audit k . The remaining parameters are chosen as presented in Table 1.

Another intriguing result can be taken from Figure 5 recalling the case where behavioral adaption is not possible. Since in that simplified setting signals are not exchanged, no changes in both the auditing and the claiming strategy would be possible, resulting in a stable solution after period two. Comparing the values from the second observation period with the consecutive ones in Figure 5, we see that, given the current setting, from the policyholder point of view, the gain in utility in the second observation period is considerably higher than the ones in all the consecutive observation periods, implying that the option to change ones strategy is disadvantageous to them. This result is very enlightening since it contradicts the widespread opinion that adapting the defrauding strategy based on signals especially from third parties like service providers is favorable from the individuals perspective.

4.2.2 Sensitivity Analyses

The remainder of this section constitutes the presentation and discussion the impact of relevant input parameters have on the insurance company’s optimal auditing strategy \mathcal{A}_R and the resulting effects on the net present value NPV . In particular, the influence of the cost per audit k , the relative fraud amount α and the policyholder’s initial threshold value $\hat{\theta}_{ph}^*$ will be analyzed respectively.

For this purpose, we consider the final values of the optimal auditing range \mathcal{A}_R , net present value NPV and gain in utility ΔU after 12 iterations each when stable results are achieved.

Cost Per Audit

We take $k \in [0.05, 0.5]$ and illustrate the results for the auditing range as well as the corresponding values for the insurance company’s net present value NPV in Figure 6 keeping the remaining parameters as presented in Table 1.

Figure 6(a) shows that a change in the cost per audit k merely has an impact on the width of the auditing range. However, the values themselves which trigger the verification process shift upwards for

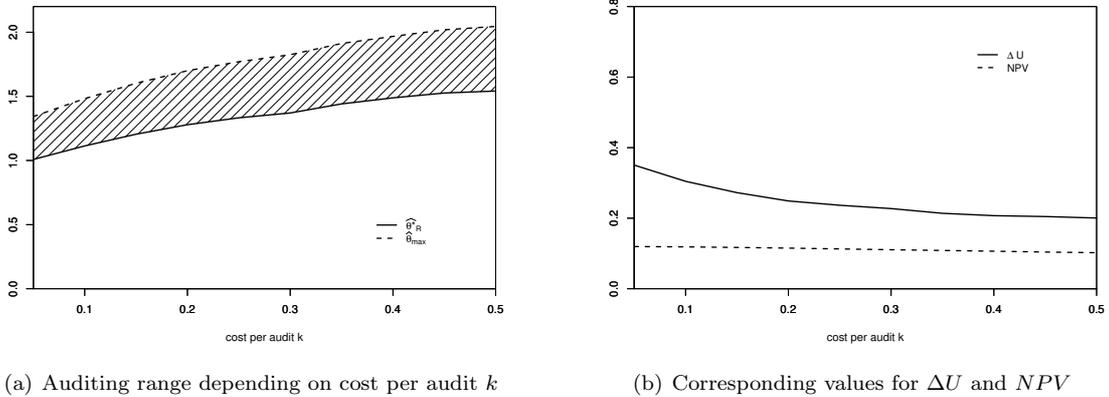


Figure 6: Auditing range and the corresponding objective quantities from insurance company and policyholder depending on the cost per audit k . The remaining parameters are chosen as presented in Table 1.

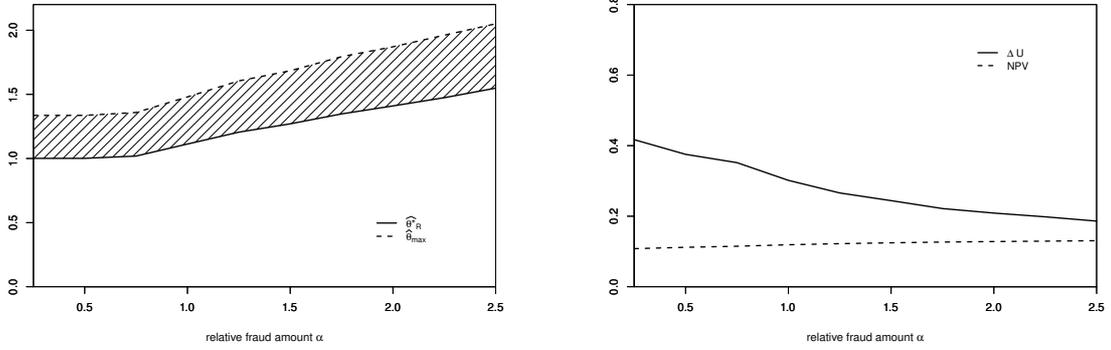
higher values of this specific input parameter. The higher the cost per audit k is, the higher are the claim amounts which will be subject to auditing. These findings are in line with the ones presented in Figure 3. The lower the costs per audit are, the more verification processes the insurance company can perform assuming a given budget. This allows the insurer to review a larger number of incoming claims and consequently enhances the probability of revealing the fraudulent ones, in particular those which are close to the policyholders' threshold value. Such an approach enables the insurance company to adjust its auditing strategy optimally to the prevalent defrauding behavior at an early stage. As a consequence, potential escalations with regard to the fraud strategies, i.e., expansion of the policyholder's individual threshold value, can be prevented. In this context, note that the upper bound of the audit range in Figure 6 which is defined by the maximum values of detected fraud, is decreasing when lowering the costs per audit k .

From the policyholder perspective, we observe that the gain in utility ΔU is always positive, implying that the auditing strategies discussed above are feasible.

Relative Fraud Amount

Considering $\alpha \in [0.25, 2.5]$, we illustrate the effects on the optimal auditing strategy \mathcal{A}_R in Figure 7 and discuss them afterwards. Again, the remaining input parameters are chosen as given in Table 1.

As can be seen in Figure 7(a), the relative fraud amount α has a considerable impact on the insurance company's optimal auditing strategy. The width of the auditing range increases (slightly) for greater values of this input parameter. At the same time, the claim values which indicate the necessity of verification shift in an upward direction. As a result, the relative fraud amount α also has an impact on both stakeholders' objective quantities. Figure 7 illustrates that the insurance company's net present value NPV is increasing when raising the values of this input parameter. The reason for this observation is that in case of a successful verification process, i.e., the detection of fraudulent behavior, the relative



(a) Auditing range depending relative fraud amount α

(b) Corresponding values for ΔU and NPV

Figure 7: Auditing range and the corresponding objective quantities from insurance company and policyholder depending on the relative fraud amount α . The remaining parameters are chosen as presented in Table 1.

fraud amount α will become known. The higher the relative fraud amount α is - while assuming the loss distribution itself has not changed - the more profitable it is from the insurance company perspective to audit claims which demand high indemnity payments. Hence, the relative fraud amount α has a direct influence on the upper bound of the auditing range. This observation gets even clearer when keeping in mind that the latter is determined by the maximum of all detected fraudulent claims during one period.

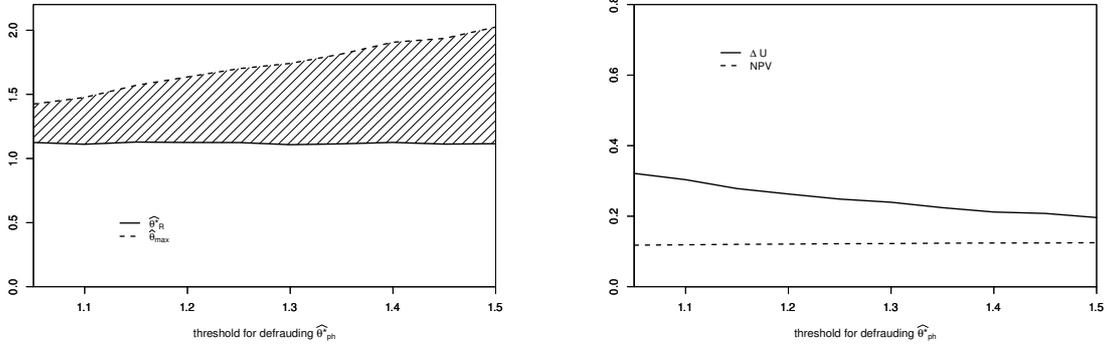
Furthermore, Figure 7(b) shows that the insurance company profits from raising and widening the auditing range whenever the relative fraud amount α is increased. In this current scenario, the relative fraud amount α is raised while keeping the probability for fraudulent behavior p constant, i.e., fraud is not committed more often but more severely. Since the loss distribution remains unchanged, this implies that fraudulent claims are more likely to be the higher valued ones. From the insurance company perspective this means that auditing becomes more profitable when shifting its auditing range into this area. Since in the case of detected fraudulent behavior no indemnity payments to the policyholder are made, this auditing strategy has a positive effect on the insurance company's net present value NPV .

Again, we see that the policyholder's gain in utility ΔU is positive throughout all observation periods guaranteeing that the corresponding auditing strategies are feasible.

Policyholder's Initial Threshold

The final input parameter whose influence on the insurance company's optimal auditing strategy we aim to analyze is the policyholders' initial threshold value $\hat{\theta}_{ph}^*$. This value serves as an upper bound for the potential fraud amount. As already introduced in Section 3.1, policyholders who decide to commit fraud exaggerate their loss amount by some constant factor α up to that threshold value $\hat{\theta}_{ph}^*$. For the purpose of our sensitivity analysis, we consider $\hat{\theta}_{ph}^* \in [1.05, 1.5]$ and present the results in Figure 8.

Figure 8 illustrates that the choice of the policyholders' threshold value for defrauding $\hat{\theta}_{ph}^*$ has a significant impact on the optimal auditing range \mathcal{A}_R . Higher values of $\hat{\theta}_{ph}^*$ imply an increasing discrepancy



(a) Auditing range depending on policyholder's initial threshold $\hat{\theta}_{ph}^*$

(b) Corresponding values for ΔU and NPV

Figure 8: Auditing range and the corresponding objective quantities from insurance company and policyholder depending on the policyholder's initial threshold value $\hat{\theta}_{ph}^*$. The remaining parameters are chosen as presented in Table 1.

with regard to the insurance company's initial threshold for auditing $\hat{\theta}_{init}^*$ which is kept at a constant value of 1.0. As a result, the optimal auditing range \mathcal{A}_R becomes broader for increasing values of $\hat{\theta}_{ph}^*$. In particular, the upper bound of the auditing range is continuously increasing while the lower bound remains almost constant. The explanation for this phenomenon is that a raise in the policyholders' defrauding threshold $\hat{\theta}_{ph}^*$ results in an increase in the share of fraudulent claims among all filed claims. In particular, the amount of exaggerated claims among the higher-valued ones will increase. Since the upper bound of the optimal auditing range is determined as the maximum of detected fraud $\hat{\theta}_{max}$ and some safety margin s , its value will become higher as well.

This explanation can also be used for understanding the resulting marginal increase in the insurance company's net present value NPV . Since higher values of $\hat{\theta}_{ph}^*$ lead to an increased likelihood of a loss amount being below this threshold, the number of cases where fraud-prone policyholders engage in build-up rises, especially in the high-value segment. Whether this development has an impact on the insurance company's net present value depends on the prevalent auditing strategy \mathcal{A}_R and its detection success. Interpreting Figure 8(b), the number of detected fraudulent activities increases compared to the number of unjustifiably paid out claims resulting in fewer indemnification which in turn leads to a slightly higher net present value.

Like in the previous analyses, the gain in utility ΔU from having signed an insurance contract prior to the occurrence of loss is always positive implying that all participation constraints are met.

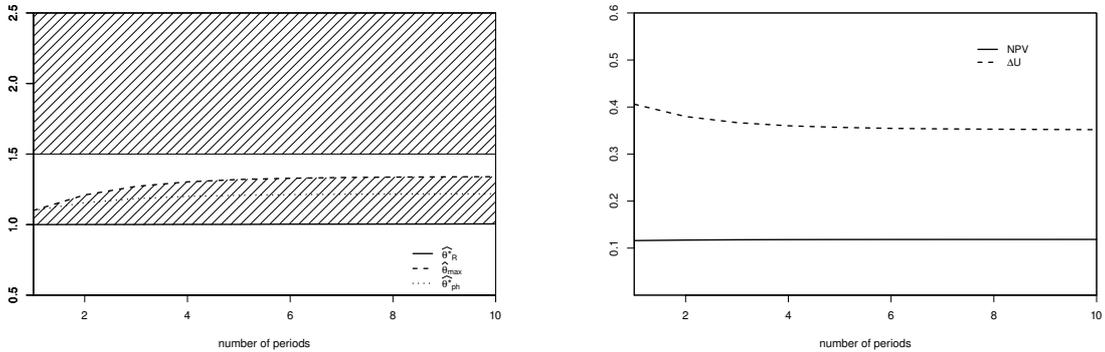
5 Critical Discussion

The insurance company's optimal auditing strategies derived in this paper are characterized by two threshold values indicating the range of claimed values which should be subject to verification. All the other

incoming claims are indemnified without further particular proof of their truthfulness. In particular, the presented approach results in examining claims especially from the medium segment leaving out small and high-valued ones. This strategy is based on the assumption that policyholders avoid any engagement in fraudulent activities whenever the actual loss amount is above some personal threshold since they fear the probability of being caught to be particularly high in this segment. As a consequence, theoretically there is no need to verify incoming claims of higher magnitude.

From a practical point of view, however, this approach appears to be incomplete. It seems unimaginable that insurance companies indemnify loss amounts which are far above the corresponding expected value without further examination of their legitimization. For this purpose, we once more extend our model framework to accommodate this aspect. We therefore introduce an additional threshold value for auditing, $\hat{\theta}_{\text{high}}^*$. In addition to verifying all incoming claims whose value fall within the auditing range \mathcal{A}_R , the insurance company also audits those which are above the new threshold value $\hat{\theta}_{\text{high}}^*$.

In order to depict the impact of adding an additional threshold value to the existent auditing range \mathcal{A}_R , we consider $\hat{\theta}_{\text{high}}^* = 1.5$ which equals one and a half the expected value of the loss amount θ . The remaining parameters are chosen as in the reference setting. Figure 9 illustrates the results.



(a) Auditing range depending with additional auditing threshold $\hat{\theta}_{\text{high}}^* = 1.5$

(b) Corresponding values for ΔU and NPV

Figure 9: Auditing range including an additional auditing threshold $\hat{\theta}_{\text{high}}^* = 1.5$ and the corresponding objective quantities from insurance company and policyholder. The remaining parameters are chosen as presented in Table 1.

As can be seen from Figure 9, in the particular setting of our model framework, an additional auditing threshold identifying high-valued claims for verification does not generate benefit for the insurance company. The optimal auditing range \mathcal{A}_R including the additional threshold $\hat{\theta}_{\text{high}}^*$ results in a net present value of $NPV = 0.118$ after the tenth observation period, whereas the optimal auditing range \mathcal{A}_R alone in $NPV = 0.120$, i.e., we obtain a change of 2%. From the policyholder perspective, the introduction of the new threshold value $\hat{\theta}_{\text{high}}^*$ has no impact. The gain in utility ΔU remains positive.

This observation can be explained by the underlying defrauding behavior of the policyholders. Since fraud-prone individuals do not inflate loss amounts in this segment, the insurer performs costly auditing processes without ever detecting any fraudulent activities, i.e., additional costs arise without ever leading to savings due to refusal of indemnification. As a consequence, the insurance company's net present value *NPV* decreases when introducing the additional threshold value for auditing high-valued claims. From the policyholder perspective, no changes in the gain in utility arise since the new auditing scheme has no impact on the indemnification payments.

In practice, however, this observation does not necessarily have to hold true. On the one hand, policyholders and/or service providers might indeed engage in fraudulent activities in case of high-valued losses or even if no insured loss occurred at all. As an example, Emerson (1992) recapitulates the case "State vs. Book" in which the policyholder exaggerated the value of his stolen luxury class automobile by 20%. On the other hand, insurance companies might profit from performing verification processes even in those cases when no fraudulent activities are detected. Their investigations might have a deterrent effect discouraging policyholders and service providers to dare fraud attempts in the future. The corresponding monetary benefits, however, are almost impossible to measure (see, e.g., Viaene and Dedene (2004)).

6 Conclusion

In this paper, we develop a model framework which depicts an optimal auditing scheme with regard to inflated insurance claims. The key element in this context is the auditing range which - triggered by the filed amount - selects those claims which should be subject to verification. Its actual configuration is chosen in way that maximizes the insurance company's position while at the same time maintaining contract attractiveness such that policyholders are willing to adhere to the insurance relationship. In addition, we incorporate the possibility for each stakeholder to adapt its behavioral strategy over the course of several periods. By this means, we take into consideration that changes in the policyholder defrauding behavior have a crucial impact on the optimal corresponding auditing strategy and vice versa. Insurance companies may use their experiences from previous verification processes as a source of information whereas policyholders often rely on third parties like service providers.

One of our main findings is the derivation of the optimal auditing scheme characterized by a range whose exact boundaries we are able to calculate. We come to the conclusion that given some constant cost per audit it is optimal to verify the truthfulness of claims from the mid-value segment. In particular, it is not reasonable from the insurance company point of view to examine small claims since the accompanying costs outweigh the savings potential in case of detected fraud. Not verifying high-valued claims results from the assumption that fraud-prone policyholders do not inflate the magnitude of their losses above some personal threshold value since they fear that the likelihood of getting caught in this segment to be above-average. Omitting this assumption, however, may require the introduction of an additional threshold value for auditing.

Furthermore, we are able to show that while the option to adapt one's strategy might be favorable from the insurance company perspective, it actually has a negative impact on the policyholders' position compared to the situation where no signals are exchanged based on which one could change its behavior. This result is astonishing since it disproves the common believe that adapting the defrauding strategy with the help of signal from service providers would be advantageous from the policyholder point of view.

Using a numerical approach based on Monte Carlo simulations, we are able to illustrate and analyze the impact of different parametrizations on the optimal auditing range. High costs per audit as well as a high relative fraud amount result in an upward shift of the auditing range, whereas an increase in the policyholders' defrauding threshold leads to broadening the range. With regard to the insurance company's objective quantity, the relative fraud amount has a particularly strong impact on its result.

References

- Belhadji, B., G. Dionne, and F. Tarkhani, 2000, A Model for the Detection of Insurance Fraud, *The Geneva Papers on Risk and Insurance*, 25(4):517–538.
- Bermúdez, L., J. Pérez, M. Ayuso, E. Gómez, and F. Vázquez, 2008, A Bayesian dichotomous model with asymmetric link for fraud in insurance, *Insurance: Mathematics and Economics*, 42(2):779–786.
- Bond, E. and K. Crocker, 1997, Hardball and the Soft Touch: The Economics of Optimal Insurance Contracts with Costly State Verification and Endogenous Monitoring Costs, *Journal of Public Economics*, 63(2):239–264.
- Crocker, K. and J. Morgan, 1998, Is Honesty the Best Policy? Curtailing Insurance Fraud through Optimal Incentive Contracts, *Journal of Political Economy*, 106(2):355–375.
- Crocker, K. and S. Tennyson, 2002, Insurance Fraud and Optimal Claims Settlement Strategies, *Journal of Law and Economics*, 45(2):469–507.
- Cummins, D. and O. Mahul, 2004, The Demand for Insurance with an Upper Limit on Coverage, *Journal of Risk and Insurance*, 71(2):253–264.
- Derrig, R., 2002, Insurance Fraud, *Journal of Risk and Insurance*, 69(3):271–287.
- Derrig, R. and V. Zicko, 2002, Prosecuting Insurance Fraud - A Case Study of the Massachusetts Experience in the 1990s, *Risk Management and Insurance Review*, 5(2):77–104.
- Dionne, G. and O. Ghali, 2005, The (1992) Bonus-Malus System In Tunisia: An Empirical Evaluation, *Journal of Risk and Insurance*, 72(4):609–633.
- Dionne, G., F. Giuliano, and P. Picard, 2009, Optimal Auditing with Scoring: Theory and Application to Insurance Fraud, *Management Science*, 55(1):58–70.
- Dionne, G. and C. Vanasse, 1992, Automobile Insurance Ratemaking In The Presence Of Asymmetrical Information, *Journal of Applied Econometrics*, 7(2):149–165.

- Duffield, G. and P. Grabosky, 2001, The Psychology of Fraud, In *Trends and Issues in Crime and Criminal Justice*, 199. Australian Institute of Criminology.
- Dulleck, U. and R. Kerschbamer, 2006, On Doctors, Mechanics, and Computer Specialists: The Economic of Credence Goods, *Journal of Economic Literature*, 44(1):5–42.
- Emerson, R., 1992, Insurance Claims Fraud Problems and Remedies, *University Of Miami Law Review*, 46:907–973.
- GDV, 2011, Versicherungsbetrug: aktuelle Entwicklungen, Muster und ihre Abwehr, Technical Report.
- Hubbard, T., 2002, How Do Consumers Motivate Experts? Reputational Incentives in an Auto Repair Market, *Journal of Law and Economics*, 45(2):437–468.
- Insurance Research Council, 2008, Fraud and Buildup in Auto Injury Insurance Claims, Technical Report.
- Kahneman, D. and A. Tversky, 1979, Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, 47(2):263–292.
- Kerr, D., 2012, Exploring the Role of Pseudodeductibles in Auto Insurance Claims Reporting, *Journal of Insurance Issues*, 35(1):44–72.
- Lacker, J. and J. Weinberg, 1989, Optimal Contracts under Costly State Falsification, *Journal of Political Economy*, 97(6):1345–1363.
- Marlin, P., 1984, Fitting the Log-Normal Distribution to Loss Data Subject to Multiple Deductibles, *Journal of Risk and Insurance*, 51(4):627–701.
- Miyazaki, A. D., 2008, Perceived Ethicality of Insurance Claim Fraud: Do Higher Deductibles Lead to Lower Ethical Standards?, *Journal of Business Ethics*, 87(4):589–598.
- Mookherjee, D. and I. Png, 1989, Optimal Auditing, Insurance, and Redistribution, *Quarterly Journal of Economics*, 104(2):399–415.
- Moreno, I., F. Vázquez, and R. Watt, 2006, Can Bonus-Malus Alleviate Insurance Fraud?, *Journal of Risk and Insurance*, 73(1):123–151.
- Morley, N., L. Ball, and T. Ormerod, 2006, How the detection of insurance fraud succeeds and fails, *Psychology, Crime & Law*, 12(2):163–180.
- Picard, P., 2000, On the Design of Optimal Insurance Policies Under Manipulation of Audit Cost, *International Economic Review*, 41(4):1049–1071.
- Picard, P., 2001, Economic Analysis of Insurance Fraud, In *Handbook of Insurance*, 1997. Springer.
- Picard, P. and M.-C. Fagart, 1999, Optimal Insurance Under Random Auditing, *Geneva Papers on Risk and Insurance Theory*, 24(1):29–54.

- Tennyson, S., 2008, Moral, Social, and Economic Dimensions of Insurance Claims Fraud, *Social Research*, 74(4):1181–1204.
- Townsend, R., 1979, Optimal Contracts and Competitive Markets with Costly State Verification, *Journal of Economic Theory*, 21(2):265–293.
- Viaene, S. and G. Dedene, 2004, Insurance Fraud: Issues and Challenges, *Geneva Papers on Risk and Insurance Theory*, 29(2):313–333.
- Watt, R., 2003, Curtailing Ex-Post Fraud in Risk Sharing Arrangements, *European Journal of Law and Economics*, 16(2):247–263.
- Weisberg, H. and R. Derrig, 1991, Fraud and Automobile Insurance: A Report on Bodily Injury Liability Claims in Massachusetts, *Journal of Insurance Regulation*, 9(4):497–541.