INSURANCE CLAIMS FRAUD: OPTIMAL AUDITING STRATEGIES IN INSURANCE COMPANIES

KATJA MÜLLER
HATO SCHMEISER
JOËL WAGNER

WORKING PAPERS ON RISK MANAGEMENT AND INSURANCE NO. 92

EDITED BY HATO SCHMEISER
CHAIR FOR RISK MANAGEMENT AND INSURANCE

SEPTEMBER 2011
Insurance Claims Fraud: Optimal Auditing Strategies in Insurance Companies

Katja Müller, Hato Schmeiser, Joël Wagner*

Abstract

Insurance claims fraud is one of the major concerns in the insurance industry. According to numerous estimates, excess payments due to fraudulent claims account for a large percentage of the total 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. In particular, optimization problems are formulated from both stakeholders’ positions considering that for each of them being willing to sign an insurance contract certain participation constraints need to be fulfilled. Besides deriving analytical solutions regarding optimal fraud and auditing strategies, we provide a numerical approach based on Monte Carlo methods. The simulation results illustrate the agreement range which consists of all valid fraud and auditing probability combinations both stakeholders are willing to accept. We discuss the impact of different valid probability combinations on the insurance company’s and the policyholder’s objective quantities respectively and analyze the sensitivity of the agreement range with respect to different input parameters. Furthermore, we take into account that insurance companies hold a wide variety of information on their policyholders which can be used as indicators for fraudulent behavior. Based on the policyholder’s estimated fraud probability and the relative fraud amount, we derive a threshold value which indicates whether an incoming claim filed by the corresponding policyholder should be audited or not.

1 Introduction

Insurance claims fraud arose to be one of the major concerns among the insurance industry. It occurs in all classes of insurance and accounts for a significant portion of the indemnity payments each year. Due to the nature of fraud, it is challenging to find an accurate estimate for its occurrence. Nevertheless, the Insurance Research Council (IRC) published the study "Fraud and Buildup in Auto Injury Insurance Claims: 2008 Edition" in 2008 according to which the excess payments due to fraudulent claims added up to an estimated total of $4.8 to $6.8 Mrd in the auto injury insurance sector in the U.S. during the year 2007. Conferred to the five main private passenger auto injury coverages this corresponds to 13 to 18 percent of the total payments.

The phenomenon of insurance claims fraud is based on 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. Consequently, the insurance companies can choose to audit incoming claims in order to determine their truthfulness. In case of detected fraudulent behavior, a penalty payment can be imposed on the policyholder. However, this verification process comes at some cost. Consequently, the costs for auditing have to be traded off against the savings resulting from detected fraud. Additionally,

*Katja Müller, Hato Schmeiser and Joël Wagner are with the Institute of Insurance Economics, University of St. Gallen, Kirchlistrasse 2, CH-9010 St. Gallen.

1See Insurance Research Council (2008).
the policyholder perspective needs to be considered as well when determining the optimal handling of insurance claims fraud. Excessive verification processes can lead to longer waiting periods for the indemnity payments. Especially in case of honest policyholders, potential delays can reduce the attractiveness of insurance.

There are various forms of insurance claims fraud. A distinction which is commonly used in the literature can be made between planned fraud and opportunistic fraud (see, e.g., Tennyson (2008) and Viaene and Dedene (2004)). Whenever a claim is filed even though no insured event occurred, we speak of planned fraud. The cases when attempts are made to exaggerate the loss amount in order to obtain higher indemnity payments are called opportunistic fraud. While it is assumed that only a small number of claims contain outright fraud, the more common and frequent type of claims fraud is the opportunistic one, also known as buildup. One of the reasons for this observation is that substantial evidence is needed in order to convict policyholders of committing planned fraud. According to Weisberg and Derrig (1991), insurance companies oftentimes tend to settle suspicious cases out of court. To avoid dealing with specific legal aspects, we will therefore focus on insurance claims fraud in the form of buildup in this paper.

Our model is based on a costly state verification environment, i.e., the insurance company has the possibility to verify the truthfulness of the incoming claims at some cost. This approach has been presented by Townsend (1979) as one of the first. In his paper, he discusses optimal insurance contracts between two agents holding different information on actual events which can be transferred at some cost. His analysis is based on deterministic auditing policies which are characterized by verifying incoming claims with a probability of either one or zero. Picard (2000) extends this approach and examines the impact of principal-agent relationships between the insurance company and agents who perform the auditing process on the design of optimal insurance contracts. Furthermore, the possibility of policyholders being able to manipulate auditing costs is included. Carlier and Dana (2003) study the existence of efficient contracts when introducing fixed costs per audit. Auditing can also be performed randomly, i.e., with a probability between zero and one depending on the amount claimed. Mookherjee and Png (1989) prove that under certain conditions random auditing is always the optimal strategy. Picard and Fagart (1999) pick up this topic and analyze questions regarding the form of the auditing probability function as well as the optimal coverage schedule. Furthermore, Picard (1996) adds the aspect of insurance companies not being able to commit to their audit strategy. He shows that this issue can be eliminated by introducing a third party to perform the auditing. This approach is also pursued by Dionne, Giuliano, and Picard (2009). According to them, the optimal auditing strategy consists of referring claims which show characteristic fraud signals to a "special investigative unit". Boyer (2000) however found that there are conditions under which centralizing the investigation and verification process with the help of insurance fraud bureaus would lead to an even higher amount of fraud.

Another way to approach the topic of insurance claims fraud is by applying costly state falsification models. In this case, the policyholder can manipulate at some cost the amount of the claimed loss in order to obtain higher indemnity payments. Crocker and Tennyson (2002) show that under these conditions claims auditing does not serve as a deterrent. This leads to the introduction of an optimal indemnification profile which consists of systematically underpaying claims. However, Hau (2008) argues that separating costly state verification and costly state falsification models does not map the reality. He proves that based on the assumption of imperfect auditing processes, i.e., fraudulent behavior is detected with a probability less than one in case of auditing, an optimal insurance contract consists of both falsification and verification. Similarly, Bond and Crocker (1997) have already found that the combination of overpaying claims which can be easily verified and underpaying claims which require a more costly auditing process can be used to effectively combat insurance claims fraud.

An alternative approach to eliminate insurance claims fraud is presented by Moreno, Vázquez, and Watt (2006). They introduce a bonus-malus system whose central idea is based on raising subsequent insurance premiums whenever a claim is filed.

One key element in the fight against insurance claims fraud is the detection process. Major and Riedinger
(2002) apply "Electronic Fraud Detection", a machine learning system based on statistical as well as expert information, to the field of health care provider fraud. Similarly, Brockett, Derrig, Golden, Levine, and Alpert (2002) use principal component analysis of RIDIT scores as a technique for a priori evaluating and ranking different fraud indicators. As a result, this detection process can be used as a basis for further investigation and is proven to be less costly than assigning the same task to human investigators. A comparison of several different classification techniques regarding expert automobile insurance claim fraud is performed by Viaene, Derrig, Baesens, and Dedene (2002). They base their evaluation on a data set which contains personal injury protection claims. Furthermore, Schiller (2006) proves that the application of detection systems in a costly state verification environment leads to more effective auditing processes and additionally reduces overcompensation.

Another line of research concerns the empirical analysis. Derrig, Weisberg, and Chen (1994) analyze a random sample of claims resulting from Massachusetts accidents aiming to prove the existence of personal injury insurance lotteries which inflate the number of claims filed per accident. Tennyson and Salsas-Forn (2002) bridge theoretical results on optimal auditing strategies and actual audit practices by examining a data set consisting of individual automobile insurance claims. Furthermore, Derrig, Johnston, and Sprinkel (2006) present their results from an analysis of closed auto injury insurance claims regarding appearance and different characteristics in the context of insurance claims fraud.

The model derived in this paper takes a different approach to the subject of claims auditing. Optimal strategies are presented and analyzed from both the insurance company and policyholder perspective aiming to find a common agreement range. For this purpose we introduce a model framework which considers the insurance company’s net present value of future cash flows as well as the policyholder’s expected utility of his terminal wealth position. In order to make sure that both stakeholders are willing to sign an insurance contract, we include participation constraints. The aims are to determine an auditing strategy such that the net present value is maximized and maximizing the final expected utility by applying the optimal fraud strategy respectively. Subsequently, some analytical solutions to the optimization problems are derived and interpreted. However, due to the complexity of the model it is not always possible to derive closed-form solutions. Therefore, we present a numerical approach using Monte Carlo methods. The simulation results and their implications for both stakeholders’ optimal strategies are discussed and illustrated graphically. Before concluding, we introduce one further application of the model. Up to this point of the paper, the insurance company adapted its optimal auditing strategy solely based on assumptions regarding the policyholder’s fraud behavior. However, in many cases the value of the claim will trigger the verification process. We therefore extend the numerical approach such that this additional information is considered as well when deriving the optimal auditing strategy from the insurance company perspective.

The agreement range which we derive determines the optimal behavior of the insurance company and the policyholder to that effect that the respective other is willing to participate in the insurance contract. For which part of the agreement range both stakeholders settle depends on their market power. Assuming a highly competitive market, it is likely for those fraud and auditing probability combinations to be performed which maximize the policyholder’s objective while the insurance company is still willing to participate in the insurance contract. In case of an oligopoly, other probability combinations are of interest since the insurance company will try to maximize its own objective quantity while making sure that potential policyholders accept an insurance contract.

If the policyholder has accurate information on the prevalent auditing strategy in the insurance company, he has the chance to chose his corresponding optimal fraud probability from all valid ones which are given in the agreement range. However, it is also possible to assume that the policyholder can not adjust his fraud behavior. In this case, the insurance company has the chance to chose its optimal corresponding auditing strategy based on the information given in the agreement range.
The remainder of this paper is organized as follows: We start by presenting the model framework and first analytical results in Section 2. Thereafter, the numerical approach and the corresponding program are introduced in Section 3. In Section 4, we discuss the simulation results before providing further analysis of the model in Section 5. We conclude in Section 6.

2 Model framework

An individual with initial wealth \( W_0 \) is offered the possibility to sign an insurance contract with a fixed premium \( P \) due by the time of inception of insurance cover in \( t = t_0 \). At the same time, he faces some uncertain loss \( \theta \) of stochastic amount which, by the time of occurrence, is observed privately. In case he signed the insurance contract earlier, the policyholder can then choose to file a claim of some size \( \hat{\theta} \). In case of honest behavior the amount of that claim will equal the actual loss, i.e., \( \hat{\theta} = \theta \). If the policyholder decides to commit fraud, he reports some finite \( \hat{\theta} > \theta \). The probability of the policyholder choosing to report a fraudulent claim is denoted by \( p \).

The insurer on the other hand has no information about the actual occurred loss. He therefore audits incoming claims with some probability \( q \) and at the constant cost of \( k \) per audited claim. Depending on whether auditing took place or not and the outcome in case of an audit, a payment \( R \) is made from the insurance company to the policyholder. Figure 1 illustrates the different possible values of the payment \( R \) depending on whether fraud and auditing take place or not.

\[
R(\theta, \hat{\theta}) = (1 - p) \theta + p [(1 - q) \hat{\theta} + q (\theta - B)],
\]

with \( B \) being the penalty payment deducted from the claim amount \( \theta \).

Equation (1) can be interpreted as an indemnity payment if \( R \) is positive, whereas a negative \( R \) represents the payment made from the insured to the insurance company in case of detected fraud when \( B > \theta \). There are several possible cases: If the reported loss is not checked, the insured receives the payment of \( \hat{\theta} \). In case of auditing, the payment depends on whether the policyholder committed fraud.
or not. Proven honesty leads to a payment of $\theta = \hat{\theta}$. If a misrepresentation of loss is determined, the policyholder faces some penalty $B$. In practice, $B$ is mostly chosen such that $\theta - B = 0$.

In this framework setting, we take audits to be perfect, i.e., if a fraudulent claim is made it will surely be detected in case of auditing.

In Section 2.1 we introduce the setting as well as the objective quantity and the participation constraint from the insurance company perspective. The same is done from the policyholder point of view in Section 2.2. Based on these informations, we state the resulting optimization problems for both stakeholders in Section 2.3. Before presenting analytical results in Section 2.5, assumptions about the distribution of information among the policyholder and the insurance company are given in Section 2.4.

2.1 Insurance company: Cash flow, net present value and participation constraint

In the framework introduced above, we observe the future cash flows from the insurance company perspective at the time of insurance inception in $t = t_0$ and the time of loss realization and settling in $t = t_1$ and analyze their resulting present value.

In case of an insurance contract coming into existence, the insurance company receives the premium payment $P$ in $t = t_0$. An incoming claim in $t = t_1$ which is audited with probability $q$ and at some given cost $k (> 0)$ per analyzed claim, results in $-R(\theta, \hat{\theta}) - qk$.

The insurance company’s net present value of its future incoming and outgoing cash flows is denoted by

$$NPV = P - E(R(\theta, \hat{\theta})) - qk,$$

(2)

where $R(\theta, \hat{\theta})$ denotes the indemnity payment as defined in (1).

Condition 1 The insurance company is willing to participate in an insurance contract if its net present value is positive. Hence, one obtains the following participation constraint:

$$NPV \geq 0.$$  

(3)

Applying Equation (2), participation constraint (3) can be formulated as

$$P \geq E(R(\theta, \hat{\theta})) + qk.$$  

(4)

Apparently, the expression on the right-hand side represents a lower bound for the premium payments the insurance company is willing to accept. Its value depends on the expected value of the indemnity payments which will be made and a loading which reflects the auditing effort.

2.2 Policyholder: Wealth position, utility function and participation constraint

From the policyholder perspective, we analyze his wealth position and the corresponding expected utility at the time of inception of insurance cover $t = t_0$ and the time of loss realization and claiming $t = t_1$ for

---

2We consider the expected value of future cash flows discounted with the risk-free interest rate $r_f = 0$. 

---
the framework introduced above.

An individual initially owns some wealth $W_0$. Its consecutive development depends on whether he signs an insurance contract prior to the occurrence of loss or not. In a situation without insurance contract, the individual holds the unchanged wealth position

$$W^B_0 = W_0$$

at time $t = t_0$. At the time of loss occurrence in $t = t_1$, this amount decreases to

$$W^B_1 = W^B_0 - \theta = W_0 - \theta.$$  \hspace{1cm} (6)

The decision to sign an insurance contract is accompanied by the payment of an insurance premium $P$. Consequently, when signing the contract at time $t = t_0$ the individual owns the wealth position

$$W^A_0 = W_0 - P.$$ \hspace{1cm} (7)

Assuming a loss $\theta$ of some stochastic level occurs and therefore a claim is filed at time $t = t_1$, the policyholder’s wealth at that point in time is denoted by

$$W^A_1 = W^A_0 - \theta + R(\theta, \hat{\theta}) = W_0 - P - \theta + R(\theta, \hat{\theta})$$ \hspace{1cm} (8)

with $R(\theta, \hat{\theta})$ as defined in (1).

We assume the policyholder’s utility being described by a standard mean-variance utility function of his individual wealth. For a given wealth position $W$ and the risk aversion parameter $a (\geq 0)$ of the individual, this utility function is given by

$$U(W) = \mathbb{E}(W) - \frac{a}{2} \text{Var}(W),$$ \hspace{1cm} (9)

where $\mathbb{E}(W)$ denotes the expected value of the stochastic variable $W$.

In case no insurance contract was signed prior to the occurrence of loss, using Equation (6) and definition (9) the final utility is written as:

$$U(W^B_1) = \mathbb{E}(W_0 - \theta) - \frac{a}{2} \text{Var}(W_0 - \theta)$$

$$= W_0 - \mathbb{E}(\theta) - \frac{a}{2} \text{Var}(\theta).$$ \hspace{1cm} (10)

For the setting in which an insurance contract was signed by applying the definition in (9) to Equation (8), we obtain:

$$U(W^A_1) = \mathbb{E}(W_0 - P - \theta + R(\theta, \hat{\theta})) - \frac{a}{2} \text{Var}(W_0 - P - \theta + R(\theta, \hat{\theta}))$$

$$= W_0 - P - \mathbb{E}(\theta - R(\theta, \hat{\theta})) - \frac{a}{2} \text{Var}(\theta - R(\theta, \hat{\theta})).$$ \hspace{1cm} (11)

Comparing Equations (10) and (11), one sees the difference in influencing factors for the final expected utility for each situation: In a setting without the existence of an insurance contract the final expected utility $U(W^B_1)$ solely depends on the extend of the actual loss $\theta$ and the policyholder’s risk aversion parameter $a$. However, in a situation in which insurance coverage exists, the value of the corresponding expected utility $U(W^A_1)$ is not only influenced by $\theta$, $P$ and $a$. In fact, the policyholder’s fraud strategy $\hat{\theta}$


and the insurer’s auditing strategy \( q \) have an impact on that value due to the payment of \( R \) (see (1)). Moreover, if the insured decides to commit fraud, the size of \( \hat{\theta} \) which he choses to claim is relevant as well as the enforced penalty payment \( B \) (see (1)) in case the fraudulent claim gets detected.

**Condition 2** The individual’s decision to get insurance coverage in the first place depends on whether his utility by the time of loss occurrence is greater with having insurance than without it, i.e.,

\[
U(W^A_1) \geq U(W^B_1).
\]  
(12)

Using (11) and (10) this participation constraint (12) can be written as

\[
-P + \mathbb{E}(R(\theta, \hat{\theta})) - \frac{a}{2} \text{Var}(\theta - R(\theta, \hat{\theta})) \geq -\frac{a}{2} \text{Var}(\theta)
\]

\[\iff P - \mathbb{E}(R(\theta, \hat{\theta})) \leq -\frac{a}{2} \text{Var}(R(\theta, \hat{\theta})) + a \text{Cov}(\theta, R(\theta, \hat{\theta})).\]  
(13)

Based on the representation \( P \leq \mathbb{E}(R(\theta, \hat{\theta})) - \frac{a}{2} \text{Var}(R(\theta, \hat{\theta})) + a \text{Cov}(\theta, R(\theta, \hat{\theta})) \), the inequality in (13) can be interpreted as an upper bound for the insurance premium the potential policyholder is willing to pay for his insurance coverage. It depends on the utility of the payment \( R \) and the covariance between actual loss \( \theta \) and \( R \). Furthermore, the individual’s risk aversion parameter \( a \) has an influence on his willingness to pay.

### 2.3 Optimization of positions

So far, the model framework as well as the participation constraints for both the policyholder and the insurance company have been presented. Based on this information, we state the corresponding optimization problems.

The insurance company is aiming to maximize the net present value of the incoming and outgoing future payments with respect to its audit strategy such that both stakeholders are still willing to participate, i.e., Equations (12) and (3) hold. Again, it is assumed that the other parameters are given. This objective function can be written as

**Insurance company’s optimization problem**

Find the optimal audit strategy \( q \) s.t. \( NPV \) is maximized and Equations (12), (3) hold.  
(14)

At the same time, the policyholder’s aim is to maximize his final expected utility with respect to his fraud strategy such that both participation constraints (12) and (3) hold, i.e., an insurance contract exists. It is assumed that all the other parameters are given. We will denote this optimization problem by the following

**Policyholder’s optimization problem**

Find the optimal fraud behavior \( p \) s.t. \( U(W^A_1) \) is maximized and Equations (12), (3) hold.  
(15)

Both stakeholders try to optimize their own position respectively. Our aim is to analyze these conflicting objectives and participation constraints from both insured’s and insurer’s perspective and find a common agreement range for the resulting fraud and auditing strategies.
2.4 Assumptions about the distribution of information

Before presenting the results of our analytical analyses, we summarize the assumptions regarding the distribution of information among the stakeholders.

Insurance company perspective

The insurance company is assumed to have full information about the distribution of the reported losses $\hat{\theta}$ due to observing the incoming claims. Furthermore, we expect that it has an adequate estimate for the distribution of the actual losses $\theta$ based on the outcomes of previous auditing processes. Consequently, the insurer knows the relative fraud amount $\alpha := \hat{\theta}/\theta$ in case fraudulent behavior occurs. Since the optimal auditing strategy also depends on the prevalent fraud probability $p$, the insurance company has to estimate this value. It then chooses the corresponding optimal audit probability which maximizes its NPV in response.

Policyholder perspective

The policyholder has no reliable information about the behavior of the insurance company regarding the verification process. He can only make assumptions about its auditing probability $q$. Based on this estimate, the policyholder decides upon his optimal fraud strategy $p$ which maximizes $U(W^A)$.

It needs to be possible that once a choice regarding the fraud behavior $p$ was made no further adjustments are possible.

2.5 Analytical results

In the course of this subsection, we derive analytical results for the presented optimization problems assuming different conditions. The proofs can be found in the Appendix.

In the following first proposition, we derive optimal fraud and auditing strategies $p$ and $q$ for a special setting of the presented model framework. The crucial assumption in this case is concerning the policyholder’s risk aversion parameter $a$ which is set $a = 0$, i.e., we assume the policyholder to be risk-neutral. This implies optimizing the insured’s objective function from a present value perspective.

**Proposition 1.** For $a = 0$ and $\theta$, $\hat{\theta}$ such that $0 \leq \theta < \hat{\theta}$, the optimal fraud and auditing strategies from both stakeholders perspectives are $p = 1$ and $q = 0$. This results in $P = E(\hat{\theta})$.

This result implies that under the given assumptions, the insurance company should waive auditing incoming claims and allow fraudulent behavior instead. In return, the expected amount of fraud will be added to the insurance premium. Furthermore, the proposition confirms a characteristic behavior risk-neutral policyholders show. They are assumed to have no interest in insuring a potential loss at a premium which exceeds its expected value.\(^3\) Since in this specific setting, all policyholders claim the fraudulent amount $\hat{\theta}$ at all times, the premium can not be set higher than the expected value of $\hat{\theta}$. On the other hand, for the insurance company to be willing to participate in the insurance agreement, this premium can not exceed this value. Hence, the insurance premium equals exactly $E(\hat{\theta})$.

In the remainder of this section, optimal fraud and auditing strategies will be derived for the policyholder and the insurance company respectively in a more general setting. First of all, the policyholder’s risk aversion parameter $a$ is supposed to be non-zero, i.e., $a > 0$. This means, his objective function is actually given as an expected utility function, i.e., the variance of the difference between indemnity

\(^3\)See, e.g., Kirstein (2000).
payment $R(\theta, \hat{\theta})$ and actual loss $\theta$, denoted by $\text{Var}(\theta - R(\theta, \hat{\theta}))$, has an impact on the final result. Furthermore, whenever the policyholder decides to make a fraudulent claim, he reports $\hat{\theta} = \alpha \theta$ for some given finite $\alpha \geq 1$ to the insurance company. This setting implies that the amount of fraud is constant.

We will derive optimal fraud and auditing strategies, namely $p^{\text{opt}}$ and $q^{\text{opt}}$, for the setting introduced above. Other than in Proposition 1, the potential policyholder is assumed to be risk-averse.

**Proposition 2** Assume $p$, $q$ to be in the agreement range, i.e., an insurance contract exists. For $a > 0$, $B = \theta$, $\theta = \alpha \theta$ with some given $\alpha \geq 1$, the respective optimal strategies $p^{\text{opt}}$, $q^{\text{opt}}$ are given by:

(i) Insurance company perspective

Let some $p$ be given. In order for the net present value $\text{NPV}$ to be maximized, choose

$$q^{\text{opt}} = \begin{cases} \text{as large as possible if } & p > p^* \\ \text{as small as possible if } & p \leq p^* \end{cases},$$

where $p^* := \frac{k}{\alpha \text{E}(\theta)}.$

(ii) Policyholder perspective

Let some $q$ be given. In order for the final expected utility $U(W^A_1)$ to be maximized, choose

$$p^{\text{opt}} = \begin{cases} \text{as large as possible such that } & \frac{-\text{E}(\theta)}{ap(1 - \alpha(1 - q)) \text{Var}(\theta)} \geq 1 \\ \text{as small as possible such that } & \frac{-\text{E}(\theta)}{ap(1 - \alpha(1 - q)) \text{Var}(\theta)} \leq 0 \end{cases}.$$  

For the case $0 < \frac{-\text{E}(\theta)}{ap(1 - \alpha(1 - q)) \text{Var}(\theta)} < 1$ no general statement can be made.

Proposition 2(i) looks at the optimization problem from the insurance company perspective. It states the optimal auditing strategy with respect to a given fraud probability. The insurance company has two general strategies to chose from. It can either decide to audit the incoming claims with the maximal probability possible, i.e., such that the participation constraints of both policyholder and insurance company hold true, or the auditing probability can be chosen as minimal as possible. This decision depends on an estimate of the policyholder’s behavior $p$. Based on whether it exceeds or deceeds the threshold $\frac{k}{\alpha \text{E}(\theta)}$, the insurance company opts for a high or low auditing probability respectively. According to Proposition 2(ii), the exceed of the threshold is influenced by the costs per audit $k$. The lower these are, given some fixed $\alpha$ and $\theta$, the more likely it is for the fraud probability to exceed the resulting threshold. In this case, it becomes optimal for the insurance company to verify the incoming claims with a high probability. The opposite relationship holds true for the expected loss amount $\theta$ and the degree of fraud which is represented by $\alpha$. The higher their values are, the lower the threshold becomes and the more likely it is for the estimated fraud probability to exceed the latter. For the insurance company this implies auditing the incoming claims with the highest probability possible as well. For an illustration of the results obtained in Proposition 2(ii) see Figure 2a and the discussions in Section 4.1.

Proposition 2(ii) considers the policyholder point of view in this optimization problem. In this case, the decision whether to chose the fraud probability as large or small as possible given a certain auditing strategy, is not as clear as in the previous situation described in Proposition 2(i) especially since there are situations for which no forecast can made. Furthermore, difficulties arise when trying to interpret the
impact of single model parameters on the value of the threshold which determines the optimal auditing behavior in the known cases. However, see Figure 2b and the discussions in Section 4.1 for an illustration of the optimal fraud probability from the policyholder perspective.

The challenges which occur with finding a closed-form analytical solution to the introduced maximization problem emphasize the need for a numerical approach. In Section 3, we therefore present a method for deriving the agreement range for both policyholder and insurance company. Furthermore, the impact of valid \( p - q \) combinations on the objective quantities \( U(W^A) \) and \( NPV \) is analyzed and illustrated.

Before presenting the numerical solution to the optimization problems, we want to point out that there is one further possible approach to looking at the given problem. From the insurance company perspective, we can assume that it commits itself to a constant auditing strategy \( q \). As a consequence, the company then has to select those claims for auditing which have been reported by those policyholders whose fraud probability \( p \) is such that it maximizes the net present value \( NPV \).

The analytical results regarding the choice of the optimal prevalent fraud strategy \( \tilde{p} \) are summarized in the following corollary which is based on Proposition 2.

**Corollary 1** Assume \( p, q \) to be in the agreement range, i.e., an insurance contract exists, and \( a > 0, B = \theta, \hat{\theta} = \alpha \theta \) with some finite \( \alpha \geq 1 \). For any given auditing probability \( q \), in order for the net present value \( NPV \) to be maximized, choose

\[
\tilde{p} = \begin{cases} 
\text{as large as possible if } & q > q^* \\
\text{as small as possible if } & q \leq q^*
\end{cases}
\]

(18)

where \( q^* := \frac{\alpha - 1}{\alpha} \).

The insurance company’s choice regarding the optimal prevalent fraud strategy \( \tilde{p} \) when having fixed its own auditing strategy \( q \), is solely influenced by the fraud amount \( \alpha \). For small and medium values of \( \alpha \), the resulting threshold \( \frac{\alpha - 1}{\alpha} = 1 - \frac{1}{\alpha} \) is relatively small as well. Consequently, it is likely for the insurance company’s auditing strategy \( q \) to exceed this level. It then has to choose those claims for auditing which are reported by policyholders whose fraud probability is high in order to maximize the \( NPV \). The opposite holds true for extremely large values of \( \alpha \). In this case, the value of the threshold increases to a level which is unlikely to be exceeded by the fixed auditing strategy \( q \). Therefore, the insurance company would have to choose the incoming claims from those policyholders who are assumed to commit fraud with a low probability \( p \) in order to achieve an optimal \( NPV \). Again, see Figure 2a for an illustration of the results.

### 3 Computational aspects

As seen in the previous section, simple analytical solutions to the optimization problem cannot be derived for all general settings. Moreover, the results may be hard to interpret both graphically and economically. In this section, we will approach these challenges by using numerical methods and Monte Carlo simulation. The aim is to compute the agreement range with respect to the fraud and auditing strategies for various parameterizations of the model. After having introduced the procedure, the results of the simulations will be analyzed and presented graphically.

**Monte Carlo simulation and numerical methods**

We use the Monte Carlo technique to find the optimal agreement range regarding the fraud and auditing strategies of the policyholder and the insurance company respectively. The main idea behind this approach
is to generate a sufficiently large number of realizations \( N \) of the random variable \( \theta \). Furthermore, we consider all fraud and auditing probabilities \( p \) and \( q \) which are represented by \( l \cdot \frac{1}{M} \) for \( l = 0, 1, ..., M \) where \( M \) denotes the number of discretization points on the interval \([0,1]\). Based on these assumptions, the resulting indemnity payments \( R \), the policyholder’s wealth positions with and without having signed the insurance contract \( W^k_A \) and \( W^k_B \) as well as the insurance company’s value \( V \) are calculated for each outcome of the simulation and each fraud and auditing probability combination. Using Equations (1), (6) and (8) for \( R \), \( W^B \) and \( W^A \) respectively, this can written as follows

\[
R[n, i, j] = (1 - p[i])\theta[n] + p[i](1 - q[j])\alpha\theta + q[j](\theta[n] - B[n])
\]  
(19)

\[
W^B[n, i, j] = W_0 - \theta[n]
\]  
(20)

\[
W^A[n, i, j] = W_0 - P\theta[n] + R[n, i, j],
\]  
(21)

where \( \theta[n] \) denotes the \( n^{th} \) realization of the random variable \( \theta \) and \( p[i] \) and \( q[j] \) are the considered fraud and auditing probability represented by \( i \cdot \frac{1}{M} \) and \( j \cdot \frac{1}{M} \) for \( i, j = 0, 1, ..., M \) respectively. Consequently, the term \([n, i, j]\) indicates for which combination of loss realization and fraud and auditing probabilities the quantities \( R, W^B \) and \( W^A \) are evaluated.

The next step to determining the agreement range is to derive the objective quantities, i.e., the policyholder’s final utility depending on whether he signed the insurance contract prior to the loss or not and the insurance company’s present value based on the corresponding wealth and value positions calculated before. For this purpose, we use arithmetic averaging with respect to the realizations of the random variable \( \theta \) for each possible combination of \( p \) and \( q \). Regarding the individual’s final utility when having decided against insurance coverage, we use the following formula, derived from Equation (10)

\[
U(W^B[i, j]) = \hat{\mu}_n(W^B[n, i, j]) - \frac{a_n^2}{2} = \hat{\sigma}_n(W^B[n, i, j])
\]  
(22)

where \( \hat{\mu}_n \) denotes the estimator for the expected value with respect to all realizations \( n = 1, ..., N \) and \( \hat{\sigma}_n \) the estimator for the variance with respect to all realizations \( n = 1, ..., N \). The same procedure applies for the case when an insurance contract was signed, this time using Equation (11):

\[
U(W^A[i, j]) = \hat{\mu}_n(W^A[n, i, j]) - \frac{a_n^2}{2} = \hat{\sigma}_n(W^A[n, i, j])
\]  
(23)

From the insurance company point of view, the net present value of its future incoming and outgoing cash flows depending on the fraud and auditing probability can be derived as follows, based on Equation (2)

\[
NPV[i, j] = P - \hat{\mu}_n(R[n, i, j]) - q[j]k
\]  
(24)

We are now in the position to check for the participation constraints of both policyholder and insurance company. Only if these hold true, an insurance contract comes into existence and merely in this case, the optimization problems are well defined. The idea here is to systematically analyze the participation constraints given in Equations (12) and (3) for each combination of fraud and auditing probabilities. In case these are verified, we consider the corresponding \( p - q \) combination as valid. At the end of this procedure, we obtain the agreement range.

The actual aim is to find the optimal strategies \( p \) and \( q \) such that the objective quantities, i.e., the policyholder’s final wealth position \( U(W^A) \) and the present value of the insurance company’s future incoming and outgoing cash flows \( NPV \), are maximized from each of the participants perspectives. In order for these to be determined, we calculate the results for \( U(W^A) \) and \( NPV \) evaluated with respect to the valid \( p - q \) combinations respectively. Once the maximal values have been found, we can retrace the corresponding fraud and auditing probabilities under which the maximum was attained. This procedure is performed separately for the two participants.
Choice of parameters
We analyze the implementation of the model for different parameterizations. The aim here is to study the influence of certain model parameters on the agreement range regarding the valid fraud and auditing probabilities.

We make assumptions concerning the distribution of the loss variable \( \theta \), the policyholder’s initial wealth position \( W_0 \) as well as the penalty payment \( B \) which remain fixed throughout the whole analysis. For instance, the policyholder’s wealth position is set to \( W_0 = 0 \). Since his participation constraint which is given by Equation (12) is independent of this parameter, our choice will not have any influence on the fact whether he signs the insurance contract or not. Furthermore, we assume the random variable \( \theta \) to 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 is set \( \mu = 1 \) and the variance \( \sigma^2 = 0.4 \). Regarding the penalty payment \( B \), we take it to be of the same value as the corresponding realization of the loss \( \theta \) such that in the case of detected fraudulent behavior the indemnity payment is 0. Additionally, we will not consider exogenously given penalties. Viaene and Dedene (2004) claim that in practice insurance companies tend to negotiate with allegedly suspicious policyholders since substantial legal evidence is needed in order to prosecute insurance claim fraud successfully.

In the course of this section, we analyze the influence of the policyholder’s risk aversion \( a \), the amount of fraud which is represented by \( \alpha \), the insurance premium \( P \) as well as the cost per audit \( k \) on the agreement range respectively. For this purpose, we use the ceteris paribus assumption in the analysis, i.e., we study the change in the agreement range caused by one isolated factor while keeping all the others constant. Unless noted otherwise, the policyholder is taken to be risk averse. Hence, to start with, his risk aversion parameter \( a \) is set 6. Furthermore, we firstly assume that in the case of fraudulent behavior the policyholder always decides to claim an amount which is 20% higher than the actual loss. According to Derrig, Johnston, and Sprinkel (2006), this value seems reasonable. In an auto injury insurance claim study from 2002, they revealed that the average payments which were made related to bodily injury claims added up to approximately $7872 if no buildup or fraud was detected whereas in case when fraudulent behavior appeared the amount rose up to $9559 on average. The last assumption which we have to make concerns the insurance premium. It 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 \( \theta \) to \( \mu = 1 \) implies a fair premium of 1 as well. However, the loss ratio in the automobile insurance in many industrialized countries over the last years averaged out to approximately 70%. Using this observation and the assumption of \( \mu = 1 \), we set the fair premium to 1.4. Furthermore, since the insurance company faces additional costs due to the auditing process with positive probability, it will add a corresponding loading factor to the fair premium. However, as mentioned in Cummins and Mahul (2004), the loading factor can not be chosen too big since potential policyholders would not sign the insurance contract under such conditions. For the purpose of starting our analysis, we will assume the total insurance premium to be \( P = 1.45 \). The last parameter whose influence on the agreement range will be analyzed is the cost per audit \( k \). It is set \( k = 0.1 \) which corresponds to 10% of the expected value of the loss \( \theta \). For the purpose of our analysis and in order to keep focused, we will disregard costs other than the ones due to auditing.

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

Unless otherwise noted, the simulation results are based on \( N = 100,000 \) realizations of the loss variable \( \theta \) and \( M = 50 \) discretization points in the interval \([0, 1]\).

---

4See, e.g., U.S., German or Swiss market supervisory data.
Table 1: Input parameters for the reference setting.

<table>
<thead>
<tr>
<th>Input parameter</th>
<th>Reference level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial wealth position</td>
<td>$W_0$</td>
</tr>
<tr>
<td>Insurance premium</td>
<td>$P$</td>
</tr>
<tr>
<td>Occurred loss</td>
<td>$\theta$</td>
</tr>
<tr>
<td>lnN(1, 0.4)</td>
<td>$\ln N(1, 0.4)$</td>
</tr>
<tr>
<td>Fraud amount</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>Risk aversion parameter</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>Auditing cost</td>
<td>$k$</td>
</tr>
<tr>
<td>Penalty payment</td>
<td>$B$</td>
</tr>
</tbody>
</table>

4 Simulation results

This section contains the results based on the numerical simulation. First, we discuss the reference setting and the impacts on the objective quantities and the corresponding optimal strategies. Furthermore, a sensitivity analysis of the relevant parameters is performed.

4.1 The reference setting

Before discussing the effects of different parameterizations regarding the policyholder’s risk aversion, the amount of fraud, the insurance premium as well as the cost per audit on the agreement range, we first want to illustrate the results given the input parameters as summarized in Table 1.

Figure 2 shows the agreement range from both the policyholder as well as the insurance company perspective based on the values for the input parameters which were presented above. Each point in the graphic represents a valid fraud and auditing probability combination.

In order to illustrate the dimension of the objective quantities $U(W_1^A)$ and $NPV$ which result from the current parameter choice and a certain $p - q$ combination, the points are displayed in different colors according to the value. For this purpose, given that the input parameters are fixed, the $p - q$ combinations which lead to the lowest third of outcomes are presented in the lightest color whereas those combinations which result in the highest third of outcomes are shown in the darkest color. The remaining points are displayed in a medium color. This implies that the darker the color of a point, the higher is the relative value of the corresponding $U(W_1^A)$ or $NPV$.

Insurance company perspective

The insurance company’s net present value can be maximized based on two different perspectives. On the one hand, we can consider some given fraud probability $p$ which is applied constantly by the policyholder. The insurer then has to choose the optimal corresponding auditing strategy. On the other hand, we can assume that the insurance company commits itself to a constant auditing probability $q$. Consequently, it has to select the claims reported from those policyholders whose fraud strategy is such that the net present value $NPV$ is maximized. In the following, we discuss both points of view:

Let some constant fraud behavior which is characterized by $p$ be given. The choice regarding the optimal corresponding auditing strategy $q$ from the insurance company perspective depends on the value of the fraud probability $p$. As already proven in Proposition 2, there exists a threshold $p^*$ which determines whether it is optimal to audit the incoming claims with the highest probability possible or the lowest valid probability, i.e., the highest and lowest $q$ respectively contained in the agreement range. Considering the choice of the input parameters for the reference setting, the value of this threshold is given by
Figure 2: Agreement range from both stakeholders’ perspectives respectively. All parameters are chosen as presented in Table 1. $p - q$ combinations which result in the highest third of $NPV$ and $U(W^A_1)$ respectively are displayed in the darkest color, the ones which result in the lowest third of $NPV$ and $U(W^A_1)$ respectively in the lightest color and the remaining ones in a medium color.

$p^* = k/\alpha E(\theta) = 0.083$. This implies that in case $p > 0.083$, it is best for the insurance company to audit the incoming claims with the highest valid probability whereas if $p \leq 0.083$, the optimal strategy is to chose $q$ as minimal as possible. These relationships can be observed in Figure 2.

Assuming a fixed auditing strategy $q$, the insurance company has to select the claims depending on the policyholder’s fraud probability $p$. As we have proven in Corollary 1, this choice is influenced by a threshold $q^*$. In the reference setting, its value is $q^* = \alpha - 1/\alpha = 0.167$. Using Corollary 1, this result can be interpreted the following way: If $q > 0.167$, it is optimal from the insurance company perspective to audit the claim from policyholders with the highest fraud probability $p$ possible. In case $q \leq 0.167$, the insurer’s net present value is maximized by choosing the reported claims from policyholders who commit fraud with the lowest probability possible. This result is confirmed by the observations made in Figure 2.

Policyholder perspective

The maximization problem from the policyholder perspective can also be analyzed from two different starting points. The first possibility is to assume some auditing strategy $q$ performed by the insurance company. The policyholder then has to decide upon the optimal corresponding fraud probability $p^{opt}$. Another approach is to consider the policyholder himself committing fraud with a constant probability $p$. In this case, the aim is to find the insurance company whose auditing strategy $q$ is such that it maximizes the insured’s expected utility of his final wealth position. Both points of view are discussed in the following:

Let the insurance company be committed to some constant auditing strategy $q$. From the policyholder perspective it is always optimal to correspond with reporting fraudulent claims at the highest valid probability $p$. Figure 2 supports this result.

We assume the policyholder to report fraudulent claims with a constant probability $p$. In this case, it is always optimal for the claims to be audited with the lowest probability $q$ possible.

As already mentioned in Section 2.4, the policyholder can not adjust his fraud strategy $p$ once he made his decision. Consequently, he has no possibility to react on the insurance company’s chosen auditing probability $q$. This assumption guarantees that we can derive a solution to the optimization problems as
defined in Equations (14) and (15) without applying game theoretical approaches given that the agreement range is non-empty.

In the course of this section, we will waive the separate representation of the agreement range from both stakeholders’ perspectives. Since the effect of the different valid \( p - q \) combinations on the policyholder’s final utility position \( U(W^A) \) as well as on the insurance company’s present value \( NPV \) has been presented and analyzed above, we restrict ourselves to showing the agreement range itself without the impacts on the objective quantities.

4.2 Sensitivity analysis of relevant parameters

In the remainder of this section, we present and discuss the resulting agreement ranges, i.e., all valid \( p - q \) combinations based on different choices regarding the input parameters risk aversion \( a \), fraud amount \( \alpha \), insurance premium \( P \) and cost per audit \( k \).

Influence of policyholder’s risk aversion

In this subsection, we aim to analyze the impact of different risk aversion parameters on the agreement range of the fraud and auditing probabilities. For this purpose, we chose different values for \( a \) while keeping all the other input parameters as given in Table 1. In particular, Figure 3 shows the agreement range for the risk aversion parameters \( a = 5 \) and \( a = 10 \).

![Figure 3: Agreement range for different risk aversion parameters \( a \). The remaining parameters are chosen as presented in Table 1.](image)

Comparing the two graphics for the agreement range, we see that the upper bound shifts in an upward direction when increasing the policyholder’s risk aversion parameter. This implies that the higher the risk aversion of the policyholder is, the broader the agreement range becomes assuming all the other input parameters to be constant.

In other words, the more risk averse the policyholder is, the higher the auditing probability \( q \) can be chosen for each fraud strategy \( p \) while the policyholder is still willing to participate in the insurance contract.
**Influence of fraud amount**

In Figure 4 the agreement range is displayed for the fraud amounts $\alpha = 1.1$ and $\alpha = 1.8$ respectively, i.e., in case of fraudulent behavior the claimed loss is given by $\hat{\theta} = 1.1 \cdot \theta$ or $\hat{\theta} = 1.8 \cdot \theta$. Again, the remaining input parameters are chosen as displayed in Table 1.

![Figure 4: Agreement range for different fraud amounts $\alpha$. The remaining parameters are chosen as presented in Table 1.](image)

Comparing the graphics for the different choices of $\alpha$, we see that this time the upper bound of the agreement range as well as part of the lower bound shift in an upward direction when increasing the fraud amount. To be more precise: While the number of valid $p-q$ combinations with high auditing probabilities $q$ increases for all fraud strategies $p$, the change in the lower bound occurs only in the area of high fraud probabilities $p$. Summing up these effects, we can say that the higher the amount of fraud the broader the agreement range becomes. However, a change from $\alpha = 1.2$ in the reference setting to $\alpha = 1.1$ results in marginal modifications within the agreement range.

This outcome can be interpreted the following way: The higher the amount of fraud $\alpha$ per claim, the more likely it is for the policyholder to accept higher auditing probabilities $q$ given that the own fraud probability $p$ is fixed. In these cases, even though the auditing activity increased, the gain in final utility $U(W^A_1)$ due to excessive claiming is still positive despite the higher chance of being convicted and imposed with a penalty payment. On the other hand, it becomes unattractive from the insurance company perspective to audit the incoming claims with a low probability $q$ when the amount of fraud is increased assuming a high fixed fraud behavior $p$. Such a strategy would imply that the majority of fraudulent claims remained undetected which consequently leads to an increase in outgoing cash flows due to excessive fraud amounts which is not covered by incoming positions like the insurance premiums or penalty payments. Therefore, if the fraud amount $\alpha$ goes up, $p-q$ combinations with higher values for $q$ become acceptable to both stakeholders whereas no insurance contract will come into existence with individuals who are expected to commit excessive fraud frequently.

**Influence of insurance premium**

The insurance premium is another common way to influence to the willingness of both potential policyholder and insurance company in participating in an insurance contract. In Figure 5 the agreement range is presented for two different values for the insurance premium, i.e., $P = 1.35$ and $P = 1.55$ while...
the other input parameters are chosen as in the reference setting.

![Graph](image)

**Figure 5:** Agreement range for different insurance premiums $P$. The remaining parameters are chosen as presented in Table 1.

A comparison of the agreement ranges when choosing $P = 1.35$ and $P = 1.55$ respectively shows that the upper bound shifts in a downward direction when increasing the value of the insurance premium. This means that the higher the insurance premium is, the smaller the agreement range gets while keeping the remaining input parameters unchanged.

In other words, the lower the insurance premium $P$ is, the more willing the policyholder is to accept higher audit probabilities $q$ when keeping his own fraud probability $p$ constant. However, if the insurance premium is set too high, i.e., it exceeds the expected loss amount by far, potential policyholders will have no benefit from signing such an insurance contract.

The effect of shrinking agreement ranges due to high insurance premiums can be weakened by offering such contracts to potential policyholders whose risk aversion is assumed to be high as well. As we have seen in Figure 3, the increase in risk aversion has the opposite effect on the agreement range as the choice of the insurance premium.

It needs to be pointed out that the insurance premium seems to have a significant impact on the agreement range. Even though the values of $P$ have been varied only marginally throughout the analysis, i.e., $\approx \pm 7\%$ of the reference level, the resulting number and positions of the valid $p - q$ combinations differ markedly.

**Influence of cost per audit**

The last input parameter which can be adjusted easily is the cost per audit $k$. Its value can give an indication of what type of auditing is being performed by the insurance company. Auditing procedures in which standard techniques are applied require minor costs whereas investigative processes which are initialized to verify major claims result in high costs.

Figure 6 displays the agreement ranges when the cost per audit is chosen to be $k = 0.01$ and $k = 1.0$ respectively while keeping the remaining input parameters as in the reference setting.

When comparing the graphic where the cost per audit is set $k = 0.01$ to the one with $k = 1.0$, we clearly see that the upper boundary of the agreement range shifts in a downward direction in case of low
fraud probabilities $p$ while there appears to be no change in the remaining valid $p - q$ combinations. This implies, that the higher the cost per audit, the smaller the agreement range becomes when keeping the other input parameters constant. However, only marginal changes within the agreement range can be observed when choosing $k = 0.01$ instead of $k = 1.0$ as given in the reference setting.

This observation can be explained the following way: The higher the cost per auditing process, the less willing the insurance company becomes to verify those incoming claims for which a low fraud probability is assumed. Such a strategy would lead to high expenses for the insurer which are not likely to be covered. The policyholder rarely commits fraud and even in the cases he does, the additional amount claimed is not excessive. Therefore, relatively high auditing costs and comparably low expenses resulting from undetected fraudulent claims are opposing each other. As a consequence, no insurance contracts will come into existence with policyholders whose fraud probability $p$ and amount $\alpha$ are expected to be low while the cost per audit $k$ is set on a high level.

A way to avoid this effect is to adjust the effort put into the auditing process to each specific case. Depending on the type of loss and the corresponding amount claimed, the insurance company can decide whether to apply a basic procedure at low cost or a extensive process which leads to high expenses.

As indicated by the very extreme choice of the parameters, i.e., in the first case $k = 0.01$ corresponds to 1% of the expected loss and in the second one $k = 1.0$ equals the expected loss, the cost per audit $k$ does not have a significant influence on the agreement range. However, the results imply that extensive auditing in form of high values for $q$ is not sustainable for the insurance company if the cost per audit $k$ is high.

5 Practical implementation and decision guidance

In the previous sections, we have derived and analyzed the given optimization problem from both stakeholders’ perspectives at the same time. The result is an agreement range which represents all valid fraud and auditing probability combinations $p$ and $q$ such that the policyholder as well as the insurance company are willing to participate in the insurance contract. From the insurance company point of view, this implied in particular that it chooses its auditing strategy depending on the assumed prevalent fraud probability.
However, in practice the insurance company may also choose its adequate auditing probability based on the claimed losses $\hat{\theta}$ and the corresponding loss category. For each incoming claim, it has to be decided whether to proof the correctness of the alleged loss or whether an indemnity payment is made without any verification process. Hence, the model presented above needs to be augmented to these new requirements. The aim is to determine the optimal auditing strategy $q$ from the insurance company point of view based on the amount of the incoming claim $\theta$.

The insurance company holds a variety of information on its policyholders which is obtained from their respective claiming history as well as general statistical claiming data and customer relationship management data. Based on this information, the insurance company has the possibility to profile its insureds. Such a classification is especially interesting with regard to parameters which indicate the policyholders’ potential fraudulent strategy and therefore have an impact on $\theta$. One of the main factors is the prevalent fraud probability $p$. As already shown in the previous sections of this paper, its value has an impact on the insurance company’s optimal auditing behavior. Another considerable factor is the relative fraud amount $\alpha$. Since the insurer has full information on the distribution of the claimed losses $\hat{\theta}$, he can calculate the resulting relative fraud amount $\alpha$. Clearly, this value influences the decision regarding whether to verify the truthfulness of the reported claim or not. Due to the amount of information held by the insurance company, adequate estimators for the fraud probability $p$ and the relative fraud amount $\alpha$ can be derived for each policyholder.

As seen in the analysis of the reference setting, the insurance company either has to audit the incoming claims with the highest probability $q$ possible or the lowest valid $q$ in order to maximize the net present value of its future incoming and outgoing cash flows. Apparently, there exists a threshold which gives an indication of the optimal strategy. Determining this threshold corresponds to having derived the optimal auditing probability: For the net present value $NPV$ to be maximized, any incoming claim whose value is below the threshold has to be verified which the smallest valid probability whereas those claims who exceed the threshold should be audited with the highest probability possible. We will denote the threshold value for auditing a reported claim by $\hat{\theta}^*$.

In the remainder of this section, we derive a threshold matrix and analyze its implications for the insurance company’s optimal audit strategy.

**Computational aspects**

Some adjustments need to be made to the original simulation program presented in Section 3. For a given fraud probability $p$ and fraud amount $\alpha$, we calculate the indemnity payments $R$ and the resulting wealth positions $W^A_1$ and $W^B_1$ using Equations (19), (21) and (20). It needs to be noted that since the fraud probability $p$ is fixed for each simulation process, it becomes a constant in all equations we use. Consequently, the only variables in this current simulation program are the auditing probability $q$ as well as the loss variable $\theta$. After that, again we calculate $U(W^A_1)$, $U(W^B_1)$ and $NPV$ in order to be able to determine the valid auditing probabilities $q$ corresponding to the current choice of the input parameters. This way we make sure to only consider audit strategies $q$ which are within the agreement range.

The next steps differ from the original simulation program presented in Section 3. At first, we now calculate the amount claimed $\hat{\theta}$ for each realization of the loss variable $\theta$ as follows

$$\hat{\theta}[n] = (1 - p)\theta[n] + p\alpha\theta[n].$$  \hspace{1cm} (25)

In case of honest behavior which occurs with a probability of $1 - p$, the policyholder reports the actual loss amount $\theta$ whereas in the case of fraudulent behavior he reports $\alpha\theta$.

We then determine the optimal valid auditing probability $q$, i.e., that valid $q$ which results in the highest net present value $NPV$ for the given realizations of $\hat{\theta}$ respectively. Hence, we obtain a list of the
optimal auditing strategies $q$ for the corresponding realizations of $\hat{\theta}$. Since the optimal $q$ only attains the values 0 and 1, i.e., either an incoming claim is audited or not, we need to determine that value of $\hat{\theta}$ which triggers the change in the insurance company’s optimal behavior. This value is searched for threshold value for auditing $\hat{\theta}^*$. 

Simulation results and decision guidance

Table 2 shows the threshold values for auditing $\hat{\theta}^*$ for different combinations of the policyholder’s fraud probability $p$ and relative fraud amount $\alpha$. The results are based on $N = 1000$ realizations of the random variable $\theta$ and $M = 50$ discretization points in the interval $[0, 1]$.

<table>
<thead>
<tr>
<th>$p$</th>
<th>$\alpha = 1.1$</th>
<th>$\alpha = 1.2$</th>
<th>$\alpha = 1.4$</th>
<th>$\alpha = 1.6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>1.84</td>
<td>1.69</td>
<td>1.46</td>
<td>1.29</td>
</tr>
<tr>
<td>0.10</td>
<td>0.92</td>
<td>0.85</td>
<td>0.74</td>
<td>0.66</td>
</tr>
<tr>
<td>0.15</td>
<td>0.62</td>
<td>0.57</td>
<td>0.51</td>
<td>0.46</td>
</tr>
<tr>
<td>0.20</td>
<td>0.46</td>
<td>0.43</td>
<td>0.38</td>
<td>0.34</td>
</tr>
<tr>
<td>0.25</td>
<td>0.36</td>
<td>0.36</td>
<td>0.36</td>
<td>0.31</td>
</tr>
<tr>
<td>0.30</td>
<td>0.31</td>
<td>0.30</td>
<td>0.30</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 2: Thresholds $\hat{\theta}^*$ depending different fraud probabilities $p$ and the fraud amount $\alpha$ for the normalized loss category with $E(\theta) = 1$. The remaining parameters are chosen as presented in Table 1.

The results presented in Table 2 can be interpreted as a decision guidance to whether an incoming claim should be audited or not. The insurance company profiles its policyholders according to their estimated fraud probability $p$ and the relative fraud amount $\alpha$ in case of fraudulent behavior. For each of these groups, we obtain different threshold values for auditing $\hat{\theta}^*$. The procedure is the following: Whenever a claim $\hat{\theta}$ is filed, the insurance company can look up the policyholders estimated fraud probability $p$ as well as the relative fraud amount $\alpha$. Based on Table 2, the corresponding threshold value for auditing $\hat{\theta}^*$ can be determined for the specific case. If the value of an incoming claim $\hat{\theta}$ is below the $\hat{\theta}^*$ of the group with which the policyholder is associated, it is best from the insurance company perspective to not perform any auditing. On the other hand, if the policyholder reports an amount $\hat{\theta}$ which is higher than the corresponding threshold $\hat{\theta}^*$ of his group, it is optimal for the insurer to audit that claim.

Apparently, for very low fraud probabilities $p$ in combination with small relative fraud amounts $\alpha$, the threshold for auditing $\hat{\theta}^*$ becomes so high that it seems very unlikely for the incoming claims to be audited. On the other hand, in case the policyholder commits fraud with a high probability $p$ and/or choses to report claims with a high relative fraud amount $\alpha$, it is optimal to perform auditing from the insurance company point of view.

Again, the assumption which was introduced in Section 2.4 regarding the policyholder’s lack of ability to adapt his fraud strategy $p$ after the initial choice is essential for the interpretation of the results. Furthermore, it is realistic to assume that the insured does not have reliable information about the threshold value for auditing $\hat{\theta}^*$. Waiving these assumptions would mean that the policyholder had the possibility to consequently adjust his fraud behavior to his advantage. The insured could decide to constantly report small claims $\hat{\theta}$ for which $\hat{\theta} < \hat{\theta}^*$ holds true. This way, auditing would never take place which would guarantee unlimited fraud possibilities given that the reported amounts are always below the threshold.
6 Conclusive remarks

In this paper we build and analyze a model framework which depicts the handling of insurance claims fraud based on a costly state verification approach. We present analytical solutions as well as numerical methods for solving the resulting optimization problems which take both the insurance company and the policyholder perspective into account. Our focus is set on deriving an agreement range consisting of all valid fraud and auditing probability combinations and analyzing their optimality regarding both stakeholders’ objective quantities respectively. In addition, we discuss the impact of different relevant input parameters on the size of the agreement range. Furthermore, we are able to calculate a threshold value for incoming claims based on which the insurance company can decide whether to perform auditing or not.

One of our main findings is the derivation of optimal auditing and fraud strategies from stakeholders’ perspectives respectively. Especially from the insurance company point of view, they seem intuitive: Summarizing, it can be said that the optimal answer to low fraud probabilities is to perform auditing with a small probability as well whereas medium and high fraud probabilities require the largest valid audit probability in order to maximize the net present value. However when audits are performed with a small constant probability, it is optimal to chose those policyholders whose claims contain fraud with a low probability. If the insurance company commits itself to a high audit probability, it is optimal to chose policyholders who report fraudulent claims with a high probability.

Based on our numerical approach, we are able to present and analyze the agreement range for different parametrizations as well as the optimality of different auditing and fraud probability combinations regarding the stakeholders’ respective objective quantities. While a relatively high risk aversion as well as a high relative amount of fraud and low insurance premiums result in broadening the agreement range, the latter becomes smaller whenever the value of these input parameters is chosen the opposite way. We also find that the cost per audit merely influences the number of valid fraud and auditing probability combinations. Furthermore, the simulation results support and illustrate our analytical findings regarding optimal fraud and auditing strategies.

We make use of the fact that insurance companies hold a broad range of information regarding their policyholders and factors indicating their potential fraudulent behavior in particular. Based on a classification of policyholders with respect to their assumed fraud strategy and the relative fraud amount, we are able to derive thresholds for auditing incoming claims for each of these groups respectively. These thresholds can be interpreted as a decision guidance to whether a filed claim is ought to be audited or not.

The model which we present in this paper can be extended for future research. On the one hand, one further type of auditing could be introduced which while less costly than the perfect one detects fraud only with some probability less than one. On the other hand, insurance premiums could depend on the auditing probability since more strict auditing policies require a longer period to process incoming claims and policyholders might not be willing to pay the original insurance premium due to possible delays in indemnity payments. Another topic for further research is to back test the results derived in this paper with insurance company data and profiling experience.
Appendix

In the Appendix, we state the propositions and the corollary presented in the main part of the paper once again and provide their proofs respectively.

**Proposition 1** For \( a = 0 \) and \( \theta, \hat{\theta} \) such that \( 0 \leq \theta < \hat{\theta} \), the optimal fraud and auditing strategies from both stakeholders perspectives are \( p = 1 \) and \( q = 0 \). This results in \( P = \mathbb{E}(\hat{\theta}) \).

**Proof of Proposition 1**

By setting \( a = 0 \), the policyholder’s participation constraint given in Equation (13) can be written as

\[
P \leq \mathbb{E}(R(\theta, \hat{\theta})).
\]

Since both participation constraints have to be met for an insurance contract to come into existence, (4) and (26) result in

\[
\mathbb{E}(R(\theta, \hat{\theta})) + qk \leq P \leq \mathbb{E}(R(\theta, \hat{\theta})) \iff q = 0 \ \forall k > 0,
\]

i.e., for any \( k > 0 \), \( q = 0 \) is the only solution.

In this case, the policyholder’s objective function (15) can be written as

\[
U(W^{\alpha}_1) = W_0 - P + p[\mathbb{E}(\hat{\theta}) - \mathbb{E}(\theta)].
\]

Due to the assumption of \( \theta < \hat{\theta} \), it attains its maximum at \( p = 1 \).

Furthermore, setting \( q = 0 \) and \( p = 1 \) in Equation (1), we get \( \mathbb{E}(R(\theta, \hat{\theta})) = \mathbb{E}(\hat{\theta}) \). At the same time, one can conclude from (27) that \( P = \mathbb{E}(R(\theta, \hat{\theta})) \). This leads to \( P = \mathbb{E}(\hat{\theta}) \).

\[\square\]

**Proposition 2** Assume \( p, q \) to be in the agreement range, i.e., an insurance contract exists. For \( a > 0, B = \theta, \hat{\theta} = \alpha \theta \) with some \( \alpha \geq 1 \), the respective optimal strategies \( p^{opt}, q^{opt} \) are given by:

(i) **Insurance company perspective**

Let some \( p \) be given. In order for the net present value \( NPV \) to be maximized, choose

\[
q^{opt} = \begin{cases} 
\text{as large as possible if } & p > p^* \\
\text{as small as possible if } & p \leq p^* 
\end{cases},
\]

where \( p^* := \frac{k}{\alpha \mathbb{E}(\theta)} \).

(ii) **Policyholder perspective**

Let some \( q \) be given. In order for the final expected utility \( U(W^{\alpha}_1) \) to be maximized, choose

\[
p^{opt} = \begin{cases} 
\text{as large as possible such that } & \frac{-\mathbb{E}(\theta)}{\alpha p(1 - \alpha(1 - q)) \text{Var}(\theta)} \geq 1 \\
\text{as small as possible such that } & \frac{-\mathbb{E}(\theta)}{\alpha p(1 - \alpha(1 - q)) \text{Var}(\theta)} \leq 0
\end{cases}.
\]

For \( 0 < \frac{-\mathbb{E}(\theta)}{\alpha p(1 - \alpha(1 - q)) \text{Var}(\theta)} < 1 \) no general statement can be made.

**Proof of Proposition 2**
(i) Using Equations (1), (2) and the assumptions $B = \theta$, $\hat{\theta} = \alpha \theta$ with $\alpha \geq 1$, we get

$$NPV = P - (1 - p)E(\theta) - p[(1 - q)E(\hat{\theta}) + qE(\theta - B)] - qk$$
$$= P - (1 - p)E(\theta) - p(1 - q)\alpha E(\theta) - qk$$
$$= P - E(\theta) + p(1 - \alpha)E(\theta) + q[\alpha pE(\theta) - k]. \quad (30)$$

Deriving (30) with respect to $q$ leads to

$$\frac{\partial}{\partial q} NPV = \alpha pE(\theta) - k, \quad (31)$$

which can be distinguished into two cases with respect to its sign.

(a) If for the given fraud strategy $p > \frac{k}{\alpha E(\theta)}$ holds, the $NPV$ as defined in (2) has a positive slope with respect to the parameter $q$. Consequently, the optimal auditing strategy $q^{opt}$ has to be chosen as large as possible in order to maximize the value of $NPV$.

(b) If the given fraud strategy $p$ is given such that $p \leq \frac{k}{\alpha E(\theta)}$ holds, the $NPV$ has a negative slope with respect to the parameter $q$. Hence, the optimal auditing strategy $q^{opt}$ has to be chosen as small as possible for the $NPV$ to be maximized.

(ii) Applying the assumptions $a \neq 0$, $B = \theta$, $\hat{\theta} = \alpha \theta$ with $\alpha \geq 1$ to Equations (1) and (11), we obtain

$$U(W^A_1) = W_0 - P - E(\theta) + (1 - p)E(\theta) + p[(1 - q)E(\hat{\theta}) + qE(\theta - B)] - \frac{\alpha}{2} \text{Var}[-\theta + (1 - p)\theta + p(1 - q)\hat{\theta} + pq(\theta - B)]$$
$$= W_0 - P - E(\theta) + (1 - p)E(\theta) + p(1 - q)E(\theta) - \frac{\alpha}{2} \text{Var}[-p\theta + p\alpha \theta - pq\alpha \theta]$$
$$= W_0 - P - pE(\theta) + p\alpha E(\theta) - p\alpha E(\theta) - \frac{\alpha}{2} \text{Var}[-p\theta + p\alpha \theta - pq\alpha \theta]$$
$$= W_0 - P - p(1 - \alpha + q\alpha)E(\theta) - \frac{\alpha}{2} p^2(1 - \alpha + q\alpha)^2 \text{Var}(\theta). \quad (32)$$

Deriving (32) with respect to $p$ results in

$$\frac{\partial}{\partial p} U(W^A_1) = -(1 - \alpha + q\alpha)E(\theta) - \alpha p(1 - \alpha + q\alpha)^2 \text{Var}(\theta). \quad (33)$$

Based on (33), three cases can be identified:

(a) For $\frac{-E(\theta)}{ap(1 - \alpha(1 - q)) \text{Var}(\theta)} \geq 1$, the policyholder can choose any fraud strategy $p \in [0, 1]$, especially any $p$ in the agreement range, such that $p \leq \frac{-E(\theta)}{ap(1 - \alpha(1 - q)) \text{Var}(\theta)}$. Applying this inequality to Equation (33), we obtain $\frac{\partial}{\partial p} U(W^A_1) \geq 0$. From this can be concluded that $U(W^A_1)$ has a positive slope. Consequently, the optimal fraud strategy $p^{opt}$ has to be chosen as large as possible in order to maximize the value of $U(W^A_1)$. 

23
(b) Similarly, for $\frac{-E(\theta)}{\alpha \rho p (1 - \alpha (1 - q)) \text{Var}(\theta)} \leq 0$, the policyholder can choose any fraud strategy $p \in [0, 1]$, especially any $p$ in the agreement range, such that $p \geq \frac{-E(\theta)}{\alpha (1 - \alpha (1 - q)) \text{Var}(\theta)}$. For Equation (33) this implies that $\frac{\partial}{\partial p} U(W_1^A) \leq 0$. This means that in this case $U(W_1^A)$ has a negative slope and hence, the optimal fraud strategy $p^{opt}$ needs to be chosen as small as possible for $U(W_1^A)$ to be maximized.

(c) For $0 < \frac{-E(\theta)}{\alpha (1 - \alpha (1 - q)) \text{Var}(\theta)} < 1$, no general statement about the corresponding optimal fraud strategy $p^{opt}$ can be made. □

Corollary 1 Assume $p, q$ to be in the agreement range, i.e., an insurance contract exists, and $a > 0$, $B = \theta, \tilde{\theta} = \alpha \theta$ with some finite $\alpha \geq 1$. For any given auditing probability $q$, in order for the net present value $NPV$ to be maximized, choose

$$p = \begin{cases} \text{as large as possible} & \text{if } q > q^* \\ \text{as small as possible} & \text{if } q \leq q^* \end{cases}. \quad (34)$$

Proof of Corollary 1

Based on Equation (30) from the proof of Proposition 2, we can calculate the derivative of the net present value with respect to the fraud probability $p$. This results in

$$\frac{\partial}{\partial p} NPV = (1 - \alpha) \mathbb{E}(\theta) + \alpha \mathbb{E}(\theta) q. \quad (35)$$

Like in Proposition 2(i), we can distinguish two cases with respect to the sign.

(a) If for the given auditing probability $q > \frac{a - 1}{\alpha}$ holds, the $NPV$ as defined in (2) has a positive slope with respect to $p$. Hence, the optimal corresponding fraud strategy $\tilde{p}$ has to be chosen as large as possible in order to maximize the value of $NPV$.

(b) If for the given auditing probability $q \leq \frac{a - 1}{\alpha}$ holds, the $NPV$ has a negative slope with respect to $p$. Therefore, the optimal corresponding fraud strategy $\tilde{p}$ has to be chosen as small as possible in order for $NPV$ to be maximized. □

References


