Eberhard Feess*, Loan Cong To Nguyen and Ilan Noy

Is Investment in Prevention Correlated with Insurance Fraud? Theory and Experiment

Abstract: Policy holders who engage in loss inflation by reporting higher than actual losses are a significant challenge for the insurance market. Based on a behavioral game-theoretic model, we analyze in an online experiment whether prevention taken by policy holders can provide a signal on loss inflation. We argue that the willingness for loss inflation depends on lying costs, other-regarding preferences and moral licensing. We consider treatment groups where subjects themselves decide whether to invest in prevention and control groups where a random computer draw decides on investment. We thereby disentangle the impacts of the aforementioned factors. First, we find evidence that other regarding-preferences influence loss inflation. Second, the impact of moral licensing goes in the direction predicted by our model but is not statistically significant. Third, and aligned with our model, our data suggest that other-regarding preferences and moral licensing countervail each other. We find no impact of whether the experiment is framed neutrally or in an insurance context.

*Corresponding author: Eberhard Feess**, Victoria University of Wellington, Pipitea Campus, Lambton Quay, Wellington, New Zealand, E-mail: eberhard.feess@vuw.ac.nz
Loan Cong To Nguyen and Ilan Noy, Victoria University of Wellington, Pipitea Campus, Lambton Quay, Wellington, New Zealand, E-mail: loan.c.nguyen@vuw.ac.nz (L.C.T. Nguyen), ilan.noy@vuw.ac.nz (I. Noy)

1 Introduction

Insurance policy holders sometimes deliberately report losses which exceed the actual loss in order to receive higher compensation. The degree of this so-called loss inflation is difficult to quantify because insurance companies generally provide little information on their fraud detection systems and results, and it is likely that only a small part of the detected fraud ends up in courts. However, losses seem to be high (Georges 1997; Derrig 2002; FBI 2019; Villegas-Ortega, Bellido-Boza, and Mauricio 2021). Insurance companies therefore invest large amounts in sophisticated models to decide which insurance claims they investigate (Tumminello et al. 2022). These predictive models typically consist of a large variety of variables, including demographics, contract terms, the characteristics of accidents and claims, and the policy holder's history (Aslam, Hunjra, and Ftiti 2022). Insurance firms also frequently try to encourage precaution investments through monetary incentives and other promotional efforts. In this paper, we perform an online experiment to shed light on the question whether precaution measures could provide a noisy signal on whether the losses are inflated or not. We argue that there are two potentially countervailing channels why precaution and loss inflation could be correlated: On the one hand, people who invest in precaution might have higher moral standards, and might hence also be less susceptible to loss inflation. On the other hand, when an accident nevertheless occurs, people who (voluntarily) invested in precaution might feel entitled to get (part of) their money back by inflating actual losses. We view our paper as the first step for analyzing whether policy holders' investment in precaution might be correlated with fraud through loss inflation.

Due to a lack of reliable data on actual fraud or on preventative investments, we apply an experimental approach. Based on a simple behavioral game theoretical model, we perform an online experiment with 956 participants to analyze whether there is a correlation between the policy holders' investment in precaution and their loss inflation. In our experiment, the precaution of a participant, in the role of the policy holder (she), determines the probability of a fixed loss mostly covered by another participant in the passive role of an insurer (he).[1] Our design ensures that the policy holder's expected payoff is higher without precaution, while the joint payoff of the policy holder *and* the insurer is higher with precaution. Only if a loss occurs, policy holders learn that they can inflate the loss to get higher compensation. Policy holders know that there is no risk of punishment even if they decide

---

[1] We randomly allocated "he" and "she" to one type of players for simplicity.

to misreport their loss. Our assumptions imply that subjects with neoclassical standard preferences will choose no precaution, and will always inflate their loss in case of an event.

Following well-documented insights from behavioral economics, however, we assume that the policy holders differ in three respects from people with neoclassical standard preferences. First, they have other-regarding preferences, that is, they care also about the payoff of the participant they are matched with (Cooper and Kagel 2016). Our payoff structure is designed such that a risk neutral policy holder[2] takes precaution if she weighs the insurer's payoff with a percentage of at least $\frac{1}{3}$ of her own payoff.[3] Second, when policy holders inflate their losses, they face internal moral costs of lying. These costs, when high enough, may trigger honest reports even without the risk of punishment (Abeler, Nosenzo, and Raymond 2019). Third, we assume that investment reduces the policy holders' moral concerns about lying due to an entitlement effect (the perception that once you have invested in prevention, you are entitled to fuller compensation if a loss occurred). The literature refers to this as moral licensing or moral accounting (Gneezy, Imas, and Madarász 2014; Mullen and Monin 2016).

Our experimental design allows us to disentangle the impacts of other-regarding preferences and moral licensing. We consider two types of treatments. In the *random* treatments *R*, a random (computer) draw determines whether a policy holder is forced to invest in precaution or not. As these treatments mute self-selection in the investment stage, other-regarding preferences matter only when a loss has occurred. Due to moral licensing, our model then predicts that policy holders are more likely to inflate their losses if they have been forced to invest.

In the *own* investment treatments *O*, policy holders decide on investment. In our model, the correlation between investment and loss inflation is then ambiguous: On the one hand, people with high other regarding-preferences are more willing to invest in precaution and are, for any given lying cost, less likely to inflate their loss (the incentive for loss inflation decreases with the weight put on others' payoffs). On the other hand, investing in precaution may reduce moral concerns about misreporting due to moral licensing, similar to, but potentially more pronounced than for random investment.

For both treatment types, we distinguish between neutrally framed instructions and instructions framed in an insurance context. In the neutral framing, we mention only amounts and probabilities, and avoid any reference to insurance. Comparing the two types of framing (insurance or neutral) allows us to analyze if the reference to insurance reduces moral concerns. For the investment stage,

---

2 In the experiment, we control for the attitude towards risk.
3 We will discuss the assumptions in the concluding Section 6.

we observe that policy holders are about 4 percentage points less likely to invest with insurance framing, but the difference to the neutral framing is insignificant ($p = 0.233$).[4]

Our main results for loss inflation itself are threefold: First, and in line with our model, we find that subjects who did not invest misreport more often in the treatment groups (51.23 % compared to 43.52 %), which can be attributed to self-selection based on other-regarding preferences.[5] Thus, the difference we find is in the predicted direction, and is marginally significant with $p = 0.071$ in a Fisher's exact test. This result suggests that other-regarding preferences matter for the selection whether to invest in prevention, thereby indirectly also influencing the misreporting frequency.

Second, comparing the misreporting frequencies in treatment *R* with and without investment allows us to isolate the effect of moral licensing. As random assignment mutes self-selection, our model predicts a higher frequency of misreporting with investment. The effect goes in the predicted direction with misreporting frequencies of 47.37 % with compared to 41.43 % without investment, but the difference is not statistically significant ($p = 0.338$ in a Fisher's exact test). Third, we use a novel and particularly simple measure for the subjects' dishonesty (Grundmann, Spantig, and Schudy 2023), which predicts misreporting very well. Summing up, our results suggest that the behavioral preferences our model accounts for influence the frequency of loss inflation, but further research in this direction is needed (see Section 6).

The remainder of the paper is organized as follows: In Section 2, we relate to the literature. Section 3 presents a simple behavioral game-theoretic model to structure the countervailing effects just mentioned. Section 4 introduces the experimental design and procedure. Results are presented in Section 5. Section 6 discusses limitations and further research.

2 Related Literature

Our research aims at disentangling other-regarding preferences and moral licensing as two sources that may lead to a correlation between policy holders' investments in precaution and loss inflation. We are not aware of a study on this question, but previous research suggests that a related entitlement effect (i.e. moral licensing) is prevalent when policy holders have high deductibles (Tennyson 2008; Miyazaki

---

4 With 956 observations.
5 Considering all four treatments together, 47.13 % of our subjects misreport the loss, which is in line with other experiments on lying (Abeler, Nosenzo, and Raymond 2019).

2009; IRC 2013) or have paid insurance premiums for a long time without claiming a loss (Tseng and Kuo 2014). While deductibles and premiums are different to precaution as they are part of the contract, precaution might still yield similar effects because, in expectation, they redistribute money from policy holders to insurers.

There are only few experiments on insurance fraud. Focusing on contract design, Fiederling and von Bieberstein (2018) find that misreporting is more frequent with higher deductibles. This can be mitigated by bonus-malus contracts where insurance premiums depend on the claim history. Morrison and Ruffle (2020) use misreporting the outcome of a die as loss inflation and argue that their findings can best be interpreted as an entitlement effect, while explanations based on risk aversion, loss aversion, or self-selection are less supported by their experimental data. Our research is also related to experiments showing that subjects lie more frequently in case of bad luck, and when they have earned their endowment by performing well in real effort tasks (Galeotti, Kline, and Orsini 2017; Fries and Parra 2021).

Most papers on the individual determinants of insurance fraud use surveys (Dean 2004; Brinkmann and Lentz 2006; Dehghanpour and Rezvani 2015; Ribeiro et al. 2020). Mintchik and Knechel (2022) explore peoples' "fraud tolerance" with data from the World Value Survey. They find that fraud tolerance correlates positively with self-enhancing attitudes, lower work ethics, and more traditional gender stereotypes. von Bieberstein and Schiller (2018) perform an incentivized experiment resembling insurance fraud, and a survey about the subjects' attitudes towards insurance fraud. They find no significant correlation between the answers in the survey and the behavior in the incentivized experiment. In cooperation with insurance companies, Martuza et al. (2022) perform a natural experiment where individuals who take on travel insurance are assigned to different nudges for honesty. As the researchers cannot observe whether a claim is actually fraudulent or not, they use claimed amounts, the difference between claims and settlements, the frequency of claim rejections, and the length of the event description as outcome measures to potentially identify fraudulent claims. They find very little impact of nudges in changing these metrics.

From a more general perspective, our experiment is nested in the literature on moral licensing (also referred to as moral accounting), which means that subjects are more willing to behave immorally in stage 2 after a previous moral decision in stage 1 and vice versa. A positive correlation between the two decisions points to a dominance of (stable) preferences, while a negative correlation suggests that moral licensing dominates. The experimental results are mixed. Gneezy, Imas, and Madarász (2014) find experimental support for moral licensing. The meta study by Mullen and Monin (2016) finds that tangible outcomes tend to enhance moral licensing, and that consistent behavior is more often observed when initial actions

are perceived as ethical commitments. Blanken, Van de Ven, and Zeelenberg (2015) argue that moral licensing is more likely to occur in similar domains, which is supported by the meta-analysis of Dollan and Galizzi (2015). Mullen and Monin (2016), and Merritt, Effron, and Monin (2010), however, argue that moral licensing can be equally important in similar and in different domains.

Considering the general literature of moral licensing, the value added of our design is that, in the first stage, we distinguish between treatments with random assignment and self-selection. This enables us to isolate the effect of moral licensing (by comparing non-investors and investors in treatment R), to isolate the effect of other-regarding preferences (by comparing non-investors in treatments R and O), and to consider both effects together (by comparing non-investors and investors in treatment O, and investors in treatments R and O).

3 The Model

3.1 Set-Up

We consider a risk-neutral insurance policy holder (she) who faces a fixed loss of L with probability p_i, $i \in \{I, N\}$. I (N) expresses that the policy holder invests (does not invest) into precaution, so $p_I < p_N$. Costs of precaution are $C(I) = I$. Precaution is socially efficient, $I < L(p_N - p_I)$. The policy holder has a fixed absolute deductible D, so the insurer (he) pays $L - D$.

There are two treatments. In the *random investment* treatment R, a random draw determines whether the policy holder invests or not. By contrast, in the *own investment* treatment O, the policy holder herself decides whether to invest.

In both treatments, there are at most three stages: In stage 1, a random draw or the policy holder decides on investment. In stage 2, a random draw determines whether the loss occurs or not. In case of no loss, the game ends. In case of a loss, the policy holder learns in stage 3 that she can inflate her report by claiming a loss $H > L$, and decides on her report $r \in \{L, H\}$. The policy holder knows that the insurer cannot observe the actual loss size, and that there is no risk of detection.

Policy holders with neoclassical standard preferences would inflate the loss in stage 3, as this maximizes their payoff, which is then $H - (L - D)$. However, we assume that the policy holder's preferences differ from these preferences in three respects:

- First, she puts weight one on her own payoff, and weight $\alpha \in (0,1)$ on the insurer's payoff. The higher α is, the higher is the degree of the policy holder's other-regarding preferences.

Table 1: Policy holder's costs in stage 3.

$$
C_3 = \begin{cases}
D + \alpha(L - D) & \text{with report } L \\
D - (H - L) + \alpha(H - D) + \theta & \text{with report } H \text{ and no investment} \\
D - (H - L) + \alpha(H - D) + (1 - \phi_J)\theta & \text{with report } H \text{ and investment}
\end{cases}
$$

- Second, when inflating her report, she faces lying costs of $\theta \in (0, \theta^{\max})$.[6]
- Third, lying costs decline from θ to $(1 - \phi_J)\theta$, $\phi_J \in (0, 1)$, for policy holders who have invested I in treatment $J \in \{R, O\}$. ϕ_J is the "discount" on lying costs after investment, and thus captures the possibility of moral licensing as outlined in the introduction.

We denote the distribution functions of the three parameters by $F(\alpha)$, $G(\theta)$, and $H(\phi_J)$, and assume that all distributions have positive weight on positive values. For simplicity, we assume that the three distribution functions are independent from each other.[7]

This given, the policy holder's expected costs in stage 3, C_3, are as summarized in Table 1.

With honest report L, the policy holder bears the deductible D, and puts weight α on the part of the loss borne by the insurer ($L - D$). This is independent of whether she has invested or not. If she inflates the report, she again bears D, but now gets the difference between H and L. She puts weight α on the insurer's cost $H - D$, and has lying costs of θ (line 2). The only difference when she has invested is that lying costs decrease to $(1 - \phi_J)\theta$ (line 3).

Comparing the policy holder's respective costs shows that the policy holder misreports after investment if $\theta \leq \widetilde{\theta}^I_J = \frac{(1-\alpha)(H-L)}{1-\phi_J}$, while she misreports with no investment if $\theta \leq \widetilde{\theta}^N_J = (1 - \alpha)(H - L)$. Thereby, $\widetilde{\theta}^i_J$, $i\{I, N\}$ are the threshold lying costs such that the policy holder misreports if $\theta \leq \widetilde{\theta}^i_J$. For both thresholds, the lying costs for which the policy holder is indifferent between an honest and an inflated report decrease with the degree of other-regarding preferences, $\frac{\partial \widetilde{\theta}^N_J}{\partial \alpha} = -(H - L) <$

---

6 The experimental literature on lying reveals a large heterogeneity of lying behavior. Some people don't lie at all, some lie to the maximum possible degree, and some lie, but only partially (see the meta study by Abeler, Nosenzo, and Raymond 2019; Grundmann, Spantig, and Schudy 2023 for an elicitation of lying costs and social image concerns at an individual level).

7 All of our results would hold if other-regarding preferences and lying costs were correlated. Our data, however, suggest that they are uncorrelated.

$0, \frac{\partial \widetilde{\theta}_J^I}{\partial \alpha} = \frac{-(H-L)}{1-\phi_J} < 0$. The reason is that the incentive to re-distribute money to the own account decreases with the weight put on the other participant's payoff.

3.2 Random Investment

As there is no self-selection in treatment R, we can restrict attention to stage 3. Comparing the two thresholds for misreporting with and without investment, we get:

Proposition 1. *In the random investment treatment R, the misreporting probability is higher with investment.*

Proof. Without self-selection to investment, it suffices to compare the thresholds for lying costs in stage 3, and we get $\widetilde{\theta}_R^I - \widetilde{\theta}_R^N = \frac{(1-\alpha)(H-L)\phi_R}{\phi_R} > 0$ for all distributions $F(\alpha)$ and $H(\phi_R)$. ☐

Note that Proposition 1 is only driven by moral licensing, while other-regarding preferences do not differ with and without investment.

3.3 Own Investment

For analyzing the misreporting behavior in stage 3 of the own investment treatment O, we need to take into account that policy holder's self-select in the investment stage 1. As the policy holder is unaware of the possibility to inflate her report when she decides on investment, there is no need for backward induction. This ensures that differences in lying costs cannot influence the investment decision in stage 1. For ease of exposition, we consider the investment stage first.

Stage 1. Investment. Recalling that the policy holder puts weight α on the insurer's cost and is unaware of the possibility to inflate the loss, her overall expected costs in stage 1 with no investment and investment are $C_N^1 = p_N[D + \alpha(L - D)]$ and $C_I^1 = p_I[D + \alpha(L - D)] + I$, respectively. The policy holder hence invests if her other-regarding preferences are above the threshold $\widetilde{\alpha}$:

$$C_I^1 \leq C_N^1 \Rightarrow \alpha \geq \widetilde{\alpha} = \frac{I - D(p_N - p_I)}{(L - D)(p_N - p_I)}.$$

The higher the benefit of investment as expressed by $p_N - p_I$ is, the lower is the minimum degree of social preferences $\widetilde{\alpha}$ that only just triggers investment. We henceforth assume that an interior solution for $\widetilde{\alpha}$ exists.

Stage 3. Report. If a loss occurs in stage 2, the game proceeds to stage 3. We know that the thresholds for misreporting are $\widetilde{\theta}_O^N = (1 - \alpha)(H - L)$ without

investment, and $\widetilde{\theta}_O^I = \frac{(1-\alpha)(H-L)}{1-\phi_O}$ with investment. Comparing the thresholds $\widetilde{\theta}_O^N$ and $\widetilde{\theta}_O^I$ shows the trade-off between self-selection and moral licensing outlined in the introduction.

Proposition 2. *In the own investment treatment O, the misreporting probability is higher with investment if and only if the discount factor of moral licensing, ϕ_O, is sufficiently large.*

Proof. Recall that a policy holder invests if $\alpha \geq \widetilde{\alpha} = \frac{I - D(p_N - p_I)}{(L-D)(p_N - p_I)}$. Denote by α^I some arbitrarily chosen $\alpha > \widetilde{\alpha}$ for an investor, and by α^N some arbitrarily chosen $\alpha < \widetilde{\alpha}$ for a non-investor. Hence, $\alpha^I > \alpha^N$. The difference in the critical lying costs is $\widetilde{\theta}_O^I - \widetilde{\theta}_O^N = \frac{(1-\alpha^I)(H-L)}{1-\phi} - (1-\alpha^N)(H-L)$. Suppose first that there is no moral licensing, $\phi_O = 0$. Then, $\widetilde{\theta}_O^I(\phi_O = 0) - \widetilde{\theta}_N = (H-L)(\alpha^N - \alpha^I) < 0$, that is, the lying frequency is higher without investment due to self-selection in the investment stage. Suppose next that $\phi_O \mapsto 1$. Then, $\widetilde{\theta}_O^I(\phi_O \mapsto 1) \geq \theta^{\max}$, and all investors misreport. Thus, $\widetilde{\theta}_O^I(\phi_O \mapsto 1) - \widetilde{\theta}_O^N > 0$. Finally, $\frac{\partial(\widetilde{\theta}_O^I - \widetilde{\theta}_O^N)}{\partial \phi_O} = \frac{(1-\alpha^I)(H-L)}{(1-\phi_O)^2} > 0$. The existence of $\widetilde{\phi}$ such that $\widetilde{\theta}_O^I - \widetilde{\theta}_O^N > 0$ if $\phi_O > \widetilde{\phi}$ for policy holders α^I and α^N then follows from the intermediate value theorem. Of course, the overall impact of investment depends on the type distributions $F(\alpha)$ and $G(\theta)$. □

To see the intuition, assume first that there is no moral licensing. As voluntary investors have higher other-regarding preferences, they also care more about the insurer's payoff in stage 3, and, therefore, have lower incentives to re-distribute money to their own account (for any θ given). The critical $\widetilde{\theta}$ is then lower with investment, as the impact of social preferences goes in the same direction in the investment- and the misreporting stage. However, if lying costs decrease sufficiently strongly after investment (that is, if ϕ_O is sufficiently large), then the result is reversed. Policy holders who have invested still have higher other-regarding preferences, but this effect is outweighed by moral licensing.

3.4 Comparison of Random and Own Investment

To derive predictions for the comparison of the behavior in the two treatments, note first that there is no moral licensing for policy holders who did not invest. We get:

Proposition 3. *The misreporting probability for policy holders who did not invest is higher in the own investment treatment O than in the random investment treatment R.*

Proof. Recall that a policy holder who has not invested misreports if $\theta \leq \widetilde{\theta}_J^N = (1 - \alpha)(H - L)$. The claim then follows from the fact that other regarding preferences for those who did not invest are chosen from $\alpha \in (0, 1)$ for random investment, and from $\alpha \in (0, \widetilde{\alpha}), \widetilde{\alpha} < 1$ for own investment. □

As moral licensing is muted for policy holders who did not invest, and because the distribution of lying costs is identical in stage 3 for all situations, the only difference between the two treatments for those who did not invest is self-selection in the investment stage, and hence the impact of other-regarding preferences. With own investment, only policy holders with low other-regarding preferences do not invest, and these policy holders are then also more likely to misreport. In the control group R, the degree of other-regarding preferences is statistically just determined by the original subject pool, while this degree is systematically lower in the treatment group O due to self-selection. While Proposition 1 isolates the impact of moral licensing, Proposition 3 isolates the impact of other-regarding preferences.

Finally, we compare the behavior in the two treatments for policy holders who invested:

Proposition 4. (i) *The misreporting probability for policy holders who invested may be higher with own or with random investment. (ii) If the moral licensing effect is weakly lower with own investment, $\phi_O \leq \phi_N$, the misreporting frequency for those invested is lower with own investment.*

Proof. We know that, for any distribution of lying costs given, the misreporting frequency increases with moral licensing, ϕ_J, and decreases with other-regarding preferences, α. As $\alpha \in (0, 1)$ for random investment, and $\alpha \in (\widetilde{\alpha}, 1), \widetilde{\alpha} > 0$ for own investment, self-selection to investment ceteris paribus leads to less misreporting in treatment O. For $\phi_O \leq \phi_N$, the moral licensing effect alone also leads to less lying in treatment O than R. Both effects then go in the same direction, which proves *part (ii)*. However, if $\phi_O > \phi_N$, the moral licensing effect goes in the opposite direction and may dominate (*part (i)*). □

If only other regarding-preferences mattered, then the misreporting frequency would be lower with own investment, as the pool of policy holders in stage 3 consists only of those with high other-regarding preferences. However, the misreporting frequency also depends on moral licensing, and our model contains no assumption

Table 2: Structure of misreporting frequencies for our Propositions.

	Invest	Don't invest
Own investment	F_I^{Own}	F_N^{Own}
Random investment	F_I^{Random}	F_N^{Random}

on whether the moral licensing parameter ϕ_J is larger for the own or the random investment treatment, as there may be countervailing effects: On the one hand, one might think that $\phi_O > \phi_R$, as the policy holder feels that she deliberately did something that supports the insurer, and that this entitles her to behave more selfishly in the reporting stage 3. But on the other hand, after being forced into an investment by a random draw, one might feel even more entitled to get the money back, as one did not even have the right to decide in stage 1.[8]

Table 2 illustrates the underlying logic for our four Propositions:

In Table 2, F_I^{Own} and F_N^{Own} denote the frequency of subjects who misreport losses after they have invested (have not invested) in the treatments with own investment. The interpretation of F_I^{Own} and F_N^{Own} refers analogously to the treatments with random investment.

– The only difference between subjects who were forced and were not forced into investment in the random investment treatments (second row) is moral licensing, as there is no self-selection. We therefore predict that $F_I^{Random} > F_N^{Random}$ (Proposition 1).

– For the comparison of those who invested and did not invest in the treatments with own investment (first row), there are countervailing effects: Moral licensing alone predicts $F_I^{Own} > F_N^{Own}$, but self-selection predicts the opposite. The overall effect is hence ambiguous (Proposition 2).

– The only difference between those who did NOT invest in the treatments with own and random investment (second column) is that there is no self-selection with random investment, as there is no moral licensing. We therefore predict that $F_N^{Own} > F_N^{Random}$ (Proposition 3).

– The main difference between those who invested in the treatments with own investment versus random investment (first column) is self-selection with own investment. Thus, if moral licensing is the same, our model predicts $F_I^{Own} < F_I^{Random}$. This, however, may be different if moral licensing is larger with own investment (Proposition 4).

---

8 There is mixed experimental evidence on whether moral licensing also plays a role under these circumstances, so we did not want to exclude it by assumption; see e.g. the overview by Mullen and Monin (2016).

3.5 Discussion

There are at least two aspects of our model set-up that deserve a brief discussion. First, we assume that other-regarding preferences (α) and lying costs (θ) are uncorrelated, and one might suppose that less selfish people are also more honest. If so, then investment in the own investment treatment O would further reduce the lying frequency, as the lying frequency in stage 3 would not only (indirectly) decrease through the higher weight put on the payoff of other players, but also (directly) due to higher lying costs. In our experiment, we use a proxy for lying costs that has recently been developed by Grundmann, Spantig, and Schudy (2023), and find that this proxy is highly correlated with the misreporting probability in stage 3, but not with the investment decision in stage 1. This supports our assumption.

Second, we acknowledge that there are several other ways to model the trade-off between other-regarding preferences and moral licensing we are interested in. For instance, it wouldn't make a difference if we assumed that, instead of a reduction in lying costs, investment leads to a reduction in the weight α that policy holders put on the insurer's payoff. As investment is costly, this might be triggered by equity concerns, because the insurer's wealth relative to the policy holder's wealth is higher with investment.[9] However, all that counts for our trade-off is (i) that the investment incentive in stage 1 increases with some kind of social preferences, (ii) that these social preferences ceteris paribus reduce the incentive to misreport in stage 2, and (iii) that this is countervailed by some kind of entitlement effect from investing, which ceteris paribus increases the incentive to misreport.

4 The Experiment

4.1 Design and Treatments

As shown in Table 3, we apply a two-by-two design with "Own investment" (O) versus "Random investment" (R) and "'Insurance framing" (I) versus "Neutral framing" (N).

In all treatments, subjects were informed that the experiment consists of two potentially payoff-relevant parts and a short survey that has no impact on their payoff. They knew that exactly one of the first two parts would be paid out; each with 50 % probability. For the first part, subjects knew that they had to make

---

9 See the large literature following the seminal paper by Fehr and Schmidt (1999) on inequity aversion.

Table 3: Treatments.

	Insurance	Neutral
Own investment	OI	ON
Random investment	RI	RN

decisions that would not only influence their own payoff, but also the payoff of another (passive) participant who would be randomly drawn from the subject pool after all decisions have been made.

4.1.1 Own Investment and Insurance Framing (Treatment OI)

We describe the treatment with own investment and insurance framing (treatment OI) in detail, and then explain how the other treatments differ from OI.

In insurance framing, we refer to active players explicitly as an insurance policy holder, and to the role of the passive participant as insurer. We informed policy holders that they would be asked for at most two decisions, and that these decisions influence the payoff of themselves and the insurer they would eventually be matched with.

In addition to the fixed payment of $1.50 for participation, policy holders got an endowment of $5, and were informed that there might be a loss of $4. The original loss probability was 90 %, but policy holders could reduce it to 50 % by investing $0.80 of their endowment into precaution. Policy holders were informed that, in case of a loss, we would deduct $1 from their endowment of $5, and that the remaining part of the loss (hence $3) would be paid by the insurer they are matched with.

Recall that, in our model, we assume that policy holders have other-regarding social preferences and care about the insurer's payoff with a percentage α of their own payoff. With our numbers, the policy holder's expected utilities without and with investment are then

$$U = \begin{cases} 0.5 \cdot 5 + 0.5(4 - 3\alpha) - 0.80 & \text{with investment} \\ 0.1 \cdot 5 + 0.9(4 - 3\alpha) & \text{without investment} \end{cases}$$

To avoid that decisions are influenced by mistakes when thinking about payoffs, we showed policy holders the probabilities for the possible payoff distributions. We did not inform policy holders about the insurer's endowment to avoid that equity concerns influence decisions. We did inform them, however, that the

insurer they will be matched with knows the numbers and can observe whether they invest or not. We then asked three comprehension questions about the payoff distributions with and without investments. Active players (zero recorded) were excluded from the study if they made more than one mistake even in the second attempt. Note that, for subjects without other-regarding preferences, our numbers ensure that non-investment does not only have the higher expected payoff (4.1 compared to 3.7) but also the lower variance (0.09 compared to 0.89). Thus, investment can hardly be rationalized with neoclassical standard preferences.[10]

After policy holders had made their investment decisions in stage 1, a random draw determined whether the loss occurred. If no loss occurred, the game ended and policy holders were directed to the survey (see below). In case of a loss, policy holders proceeded to stage 2. We only then informed them that they would now need to decide whether they report the actual loss of $4 or an inflated loss of $5. As policy holders have a deductible of $1, they received $3 or $4 from the insurer's account when they reported the actual or the inflated loss, respectively. Policy holders knew that there was no risk of punishment, but also that the insurer knows the actual loss size, and hence whether the report is inflated.

4.1.2 Other Treatments

The neutrally framed treatment *ON* differs from the treatment *OI* only by the framing, that is, there is no reference to insurance. This allows us to analyze if subjects behave differently when they are nudged into an insurance context.

The two treatments with random investment differed from those with own investment solely by substituting the investment decision in stage 1 with a random computer choice. In our pilot, the frequency of players who invested was 60.8 %. In order to have an (almost) balanced sample, we let the computer choose "invest" with the same probability of 0.6.[11]

With treatments *R*, we eliminated self-selection in stage 1 to ensure that the distribution of other-regarding preferences in stage 3 does not systematically differ between subjects with and without investment in stage 1. This allows us to isolate the effect of moral licensing in case it exists even with random assignment to investment.

---

10 We acknowledge, though, that we cannot separate other-regarding preferences from other behavioral motivations such as following a social norm or efficiency concerns.

11 However, conversely to the pilot, the frequency of own investments turned out to be about 80 %, so we got less observations with investment in the random investment treatment.

4.1.3 Survey

In the survey, we asked our participants for their age, gender, country of residence, the number of studies performed on Prolific, and their attitude towards risk on a scale from 0 to 10. Most importantly, we elicited the subjects' lying costs by their answers to a simple hypothetical question developed by Grundmann, Spantig, and Schudy (2023), which have proven to be a good predictor for the subjects' behavior in standard lying games where they can misreport the outcome of a lottery. The question reads as follows: "Imagine the host of a live radio show calls you to participate in a radio raffle. Your task is the following: you must flip a coin 4 times and report the number of "tails" you flipped. You will receive US$10 for each reported "tail". You know that there is no way the host can verify whether you reported the true number of "tails" you flipped…. Imagine the unlikely case occurs that you flip 0 tails. How many tails will you report to the host?"

4.2 Procedures

Our study received ethical permission by Victoria University of Wellington Human Ethics Committee,[12] and was preregistered at the American Economic Association's registry for randomized controlled trials.[13] All sessions were carried out on Prolific between October 19 and 20, 2023. Prolific is a large online platform where participants can sign up to participate in research and business studies. Our study announced a scientific study and a survey on individual decision making. We informed the subjects that the study takes about 10 min and involves filling out a short survey and a maximum of two decisions. To ensure a high data quality, we required subjects to be fluent in English, reside in either UK or USA, be 18 or older, and to have an approval rate of at least 95 %. Each subject participated only in one treatment, that is, we applied a between-subject design. We implemented measures to prevent restarting of the survey and self-selection into treatments. 956 subjects participated in total.[14] On average, the experiment took subjects 5.8 min, and they earned $15.7 per hour.

Recall that, in the beginning of the experiment, subjects were informed that there are two parts of the experiment (except for the survey), and that just one of these parts would be paid out. In part 1, we assigned subjects to one of the four treatments (according to their time of arrival). When the quota of observations required

---

12 Approval number 0000031122, dated 31 July 2023.

13 https://www.socialscienceregistry.org/trials/11628.

14 Power calculation is based on the comparison between inflated reports with and without investment, aggregated over all O treatments in our pilot study, and using Chi2-test with a power of 80 % and alpha of 5 %.

for each treatment was reached, the randomization referred only to the remaining treatments. In part 2, each subject played the role of the passive player, and hence no decision was to be made. To determine the bonus, subjects were randomly paired up, and a random computer draw (with equal probability of 50 %) decided whether a participant's active or passive role determined their payoff. This procedure enabled us to collect the decisions from all participants. Alternatively, we could have applied a strategy method where subjects would first be paired, then informed about the roles of active and passive players, and would then have been asked for their decisions in case they would be randomly assigned to the role of active players. In contrast to the strategy method, however, our procedure avoids the not unlikely possibility that the subject's behavior is influenced by knowing that they also take the role as passive players.[15]

5 Results

5.1 Summary Statistics

Table 4 provides summary statistics, separated by the four treatments (columns 1–4), and aggregated over all treatments (column 5). Randomization worked generally well, but there are some differences for risk tolerance and our dishonesty. Risk tolerance is highest for treatment NO (5.18) and lowest for treatment IO (4.66), but the difference is not significant ($p = 0.193$) in a double-sided t-test. The mean for our dishonesty measure is also highest for NO (1.14), and lowest for treatment IR (0.94), and the difference is marginally significant at $p = 0.098$.

5.2 Investments

Recall that the investment probability in the two computer treatments was 60 %, which led to investment frequencies of 59.75 % with neutral framing (144 of 241 players) and 59.92 % with insurance framing (142 of 237 players). In the two treatments where policy holders decided about investment, 82.86 % (203 of 245 players) and 75.74 % (176 of 233 players) invested in the neutral and the insurance framing, respectively. Given the amounts in our experiment, investment is optimal for policy holders who put at least 1/3 weight on the passive player's payoff. The investment

---

15 Note that our procedure bears no risk of deception: First, participants knew that there is a second part of the experiment, and that the first part of the experiment would be paid out only with 50 % probability. Second, subjects were randomly matched with different players for the active and the passive role.

Table 4: Summary statistics.

Variable	IO	NO	IR	NR	All treatments
Age	40.26 (14.12)	37.13 (12.59)	37.68 (11.33)	38.79 (13.07)	38.44 (12.84)
Female	0.56 (0.51)	0.57 (0.54)	0.62 (0.50)	0.56 (0.51)	0.58 (0.51)
Country (UK)	0.48 (0.50)	0.51 (0.50)	0.49 (0.50)	0.50 (0.50)	0.50 (0.50)
Risk tolerance	4.66 (2.27)	5.18 (2.46)	4.88 (2.37)	4.96 (2.49)	4.92 (2.41)
Dishonesty	1.03 (1.38)	1.14 (1.41)	0.94 (1.35)	1.03 (1.34)	1.04 (1.37)
Observations	233	245	237	241	956

Notes: All numbers show means and standard deviations (in brackets). The Female dummy takes the value 1 if the participant is female. Age is measured in years. Risk tolerance is self-reported on a scale from 0 to 10 (0 means: not at all willing to take risks, 10 means: very willing to take risks). Dishonesty captures the answers to the hypothetical question spelled out in the end of Section 4.1.

frequency is higher than we expected from our pilots, which suggests pronounced other-regarding preferences.

Table 5 shows results for OLS regressions with the investment decision as dependent variable. Intuitively, we thought that insurance framing reduces the subjects' other-regarding preferences, and thereby the investment frequency. We observe a small effect in this direction with reductions between three and four percentage points, but the differences are not statistically significant ($p = 0.166$ in Column III where personal controls are included). Our dishonesty measure is also not significant ($p = 0.141$ and 0.139 in Column III and IV respectively), so we can argue that those who invested and did not invest do not systematically differ with regards to their lying costs.

5.3 Inflated Reports

5.3.1 Descriptive and Non-Parametric Tests

Figure 1 shows the frequencies of inflated reports for our four treatments, separated by those who invested and those who did not invest. Absolute numbers are in brackets.

Recall that our model yields four Propositions: One for the comparison of misreporting with and without investment with random assignment, one for the same comparison when policy holders decide on investment, and two for comparing the treatments with own and random investment. However, only Propositions 1 and 3 yield predictions in one direction, whereas Propositions 2 and 4 just express countervailing effects.

Table 5: Investment decisions.

Variable	I	II	III	IV
Insurance	−0.0369 (0.0309)	−0.0376 (0.309)	−0.0429 (0.0309)	−0.0337 (0.0461)
Female		0.0352 (0.0301)	0.0260 (0.0315)	0.0336 (0.0425)
Age			0.0020 (0.0012)	0.0020 (0.0012)
Risk tolerance			−0.0051 (0.0067)	−0.0051 (0.0067)
Country (UK)			−0.0063 (0.0316)	−0.0065 (0.0316)
Dishonesty			−0.0170 (0.0116)	−0.0171 (0.0116)
Insurance × female				−0.0159 (0.0591)
Observations	956	956	956	956
R^2	0.0015	0.0029	0.0099	0.01

Notes: All columns report results for OLS regressions. The investment decision as dependent variable takes the value 1 (0) for investment (no investment). The main independent variable of interest takes the value 1 (0) for insurance (neutral) framing. Column II adds all controls except our dishonesty measure based on the answers to the hypothetical question spelled out in the end of Section 4.1. *, **, and *** Indicate statistical significance at the 10 %, 5 %, and 1 % level, respectively. Robust standard errors are in parentheses.
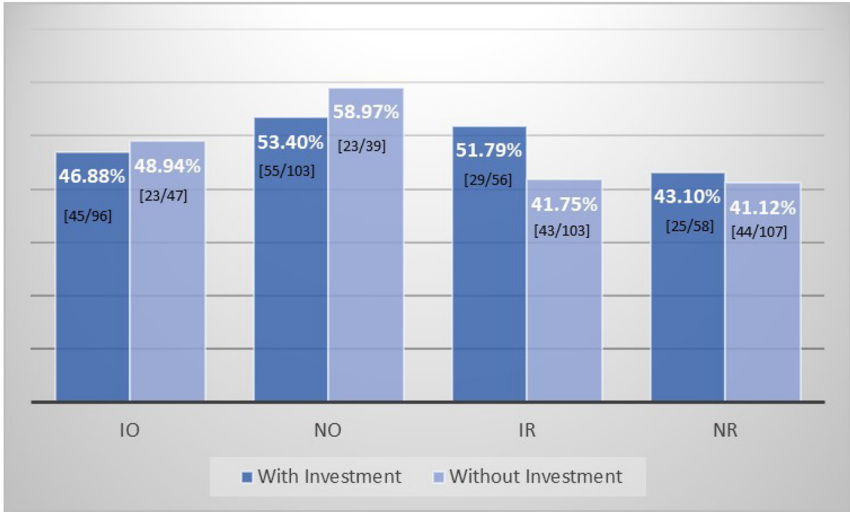

Figure 1: Misreporting by treatments and investment decisions. Source: Calculations by authors.

For all subsequently reported Fisher's exact tests, we pool the data over neutral and insurance framing, as the insurance dummy is insignificant in all regressions.

Proposition 1 predicts that the misreporting frequency in random treatments is higher with than without investment. The effect goes in the predicted direction with

misreporting frequencies of 41.4 % compared to 47.3 % with investment, but the difference is insignificant ($p = 0.348$). Recall that random assignment to investment mutes differences in other-regarding preferences and lying costs, so differences are likely to be driven by moral licensing.[16]

Next, Proposition 3 predicts that, for those who did not invest, the misreporting frequency is lower in treatment R (random investment) than in treatment O (own investment). In line with this prediction, we observe misreporting frequencies of 41.4 % and 53.4 % in treatment O. This difference is marginally significant at $p = 0.071$. As there is no moral licensing for subjects who did not invest, other-regarding preferences are now the only difference between the two situations. Considering the results for Propositions 1 and 3 together hence suggests that other-regarding preferences are somewhat more important than moral licensing.

For the other two comparisons, our model does not yield clear predictions due to the countervailing effects (see Propositions 2 and 4) of other-regarding preferences and moral licensing. Our data does not lend support for the dominance of either factor. Comparing the misreporting frequencies for those who did and did not invest in treatment O (see Proposition 2) yields $p = 0.699$ with 285 observations, and comparing the frequencies for those who invested in treatment O and in treatment R (see Proposition 4) yields $p = 0.640$ with 313 observations. Given that we do find a marginally significant difference for those who did not invest in treatment R (that is, without the self-selection effect of other-regarding preferences), these results suggest that the countervailing effects of other-regarding preferences and moral licensing offset each other.

5.3.2   Regression Analysis

We now analyze the robustness of the results with respect to our control variables. In all tables, we first include only the dummy of interest and the framing dummy (with neutral as reference category). In the second columns, we add all controls except the dishonesty measure, which is added in third columns. All potentially interesting interactions are insignificant and not reported.

In Table 6, the dependent variable is a dummy that takes the value 1 for misreporting and 0 for no misreporting. The first (last) three columns refer to the treatments with random (own) investment, and the investment dummy is the independent variable of interest. In line with the Fisher's exact test, the dummy has a

---

[16] However, since investment reduces the policy holders' wealth, and thus makes an honest report more expensive, we cannot separate the moral licensing effect from income effect in case of risk aversion. We share this problem with all papers about moral licensing where pro-social behavior in the first stage reduces the wealth.

Table 6: Misreporting in random investment and own investment.

Variables	Random investment			Own investment		
	I	II	III	I	II	III
Investment	0.0594 (0.0578)	0.0788 (0.0578)	0.0633 (0.0525)	−0.0372 (0.0647)	−0.0366 (0.0646)	0.0493 (0.0553)
Insurance	0.0346 (0.0552)	0.0293 (0.0546)	0.0403 (0.0495)	−0.0758 (0.0594)	−0.0545 (0.0595)	−0.0391 (0.0504)
Female		0.0647 (0.0574)	0.1176** (0.0525)		−0.0442 (0.0571)	0.0026 (0.0486)
Age		−0.0046** (0.0023)	−0.0015 (0.0021)		−0.001 (0.0022)	0.0009 (0.0019)
Risk-tolerance		0.0307** (0.0119)	0.0257** (0.0108)		0.0350*** (0.0125)	0.0216** (0.0106)
Country (UK)		−0.0560 (0.0572)	−0.1066** (0.0523)		0.1005* (0.0595)	0.0577 (0.0506)
Dishonesty			0.1561*** (0.0188)			0.1909*** (0.0181)
Observations	324	324	324	285	285	285
R^2	0.0045	0.0424	0.2135	0.0066	0.052	0.3229

Notes: All columns report results for OLS regressions. The dependent variable takes the value 1 (0) for misreporting (no misreporting). The first (last) three columns include only observations with random (own) investment. The insurance dummy takes the value 1 (0) for insurance framing (neutral framing). Column II adds all controls except our dishonesty measure based on the answers to the hypothetical question spelled out in the end of Section 4.1. *, **, and ***Indicate statistical significance at the 10 %, 5 %, and 1 % level, respectively. Robust standard errors are in parentheses.

Table 7: Misreporting of subjects who invested and who did NOT invest.

Variables	Without investment			With investment		
	I	II	III	I	II	III
Own	0.1220* (0.0637)	0.1285** (0.0637)	0.0471 (0.0566)	0.0287 (0.0590)	0.0287 (0.0588)	0.0311 (0.0523)
Insurance	−0.0245 (0.0578)	−0.0203 (0.0574)	−0.0029 (0.0504)	−0.0098 (0.0568)	0.0139 (0.0565)	0.0188 (0.0502)
Female		−0.0137 (0.0584)	0.0547 (0.0517)		0.0214 (0.0021)	0.0579 (0.0507)
Age		−0.004 (0.0024)	−0.001 (0.0022)		−0.0024 (0.0021)	−0.0001 (0.0019)
Risk-tolerance		0.0253** (0.0125)	0.0192* (90.0109)		0.0371*** (0.0120)	0.0268** (0.0108)
Country (UK)		0.0746 (0.0601)	−0.0063 (0.0534)		−0.0339 (0.0572)	−0.0495 (0.0509)
Dishonesty			0.1745*** (0.0186)			0.1702*** (0.0188)
Observations	296	296	296	313	313	313
R^2	0.0127	0.0433	0.2663	0.0009	0.0402	0.2438

Notes: All columns report results for OLS regressions. The dependent variable takes the value 1 (0) for misreporting (no misreporting). The first (last) three columns include only observations for subjects who did not invest (did invest). The insurance dummy takes the value 1 (0) for insurance framing (neutral framing). Column II adds all controls except our dishonesty measure based on the answers to the hypothetical question spelled out in the end of Section 4.1. *, **, and ***Indicate statistical significance at the 10 %, 5 %, and 1 % level, respectively. Robust standard errors are in parentheses.

positive but insignificant sign, with the lowest *p*-value when we include all controls except the honesty measure ($p = 0.23$). As in some other studies,[17] people with a higher risk tolerance misreport more often, albeit there is no risk involved. Age is significantly negative only in one of the specifications for Proposition 1. Our dishonesty measure has a very large impact, and females misreport significantly more often after controlling for it. Note, however, that we elicit dishonesty only after the actual experiment, so it might be influenced by the previous behavior.

The three right columns in Table 6 refer to the treatments with own investment, for which our model yields no prediction for the impact of investment. In line with the Fisher's exact tests, the *p*-values for the investment dummy are always above 0.6.

In Table 7, the main independent variable of interest is the OWN dummy that takes the value 1 (0) for own (random) investment. In the first three columns, we include only those subjects who did not invest. In line with the Fisher's exact test, the regression results support Proposition 3, as the dummy for the own investment is significantly positive at the 5 %- or 10%-level as long as we do not add the honesty measure. The effect becomes insignificant when we include the dishonesty measure. This can be attributed to the fact the dishonesty measure is highly correlated with the behavior in the actual experiment, and that the mean for dishonesty is higher with own investment. Thus, part of the different behavior is now absorbed by the dishonesty measure. However, due to the aforementioned endogeneity issue, we tend to rely more on the non-parametric test and the regression specifications in the first columns.

The three right columns in Table 7 consider only subjects who invested. In line with the Fisher's exact test, the *p*-values for the treatment dummy are always above 0.6.

6 Conclusions

Based on a behavioral game-theoretic model, we have analyzed in an online experiment if investments into precaution influence the frequency of loss inflation (misreporting of insurance claims). Distinguishing between treatments where subjects decide themselves or are forced into investment by a random computer draw enables us to disentangle the impacts of lying costs, other-regarding preferences, and moral licensing on the misreporting frequency. Our results are as follows: First, the answer to a simple hypothetical question on lying about the outcome of coin

---

17 See von Bierberstein, Feess, and Packham (2024) and some specifications in von Bieberstein and Schiller (2018) and Grundmann, Spantig, and Schudy (2023).

flips is a very good predictor of the misreporting frequency, but is uncorrelated with the investment decision. This suggests that other-regarding preferences and lying costs are independent of each other. This independence, however, deserves further investigation in an experiment specifically tailored to identifying a potential link between these two variables. Second, we find significant evidence that subjects who decide not to invest in risk reduction are more likely to misreport, compared to those where a computer draw leads them not to invest. We attribute this to the fact that other-regarding preferences play a role both in the investment and the misreporting stage. Third, the impact of moral licensing goes in the direction predicted by our model, but is not statistically significant. Fourth, our results suggest that other-regarding preferences and moral licensing countervail each other. Note that, if neither other-regarding preferences nor moral licensing play a role, then we should not find that subjects who decided not to invest behave differently in the misreporting stage compared to those who were randomly assigned to no-investment. However, given that the dummy for the own-investment treatment is only significant at the 5%- or 10%-level without honesty measure, and becomes insignificant with honesty measure, we cannot fully exclude that.

Our assumptions that other-regarding preferences and moral licensing play a role in the insurance context may be challenged, as many people might not care at all about payments made by insurance companies. If so, our experiment underestimates the willingness to misreport in an insurance setting. From this perspective, it is interesting to recall that about 52 % of subjects forced to invest in the insurance framing misreported, compared to only about 42 % of subjects who did not invest. This effect might even be stronger in reality. However, there might also be many people who care about the insurer's payoff in reality for several reasons. First, they might understand that, ultimately, the insurer's losses are borne by the pool of policy holders via higher premiums. Second, they may just follow what they perceive as a moral or social norm. While behavioral economics do not refer to these norms as other-regarding preferences, the effects, both in the investment and the misreporting stage, would be similar.

While our experiment would overstress the impact of behavioral preferences if those play only a minor role in reality, there are also arguments in the opposite direction.[18] Absolute amounts (not necessarily the amounts per minute) are lower, and opportunity costs are higher, in online experiments than in laboratory experiments, in which subjects usually need to wait until all participants have finished a task. Thus, subjects in online experiments might not devote enough attention to

---

18 We are grateful to two anonymous referees who argue in either of these directions.

evaluate their decisions, which might induce noisy behavior and hence low treatment effects.[19]

There are some other features of our experimental design that deserve some discussion. First, insurers are aware of the potential loss inflation. As an alternative, we thought about telling policy holders explicitly that insurers are unaware of this, which, however, might have given (some) participants the impression that we want to trigger them into loss inflation. Furthermore, we wanted to maximize the chance that image concerns might work alongside with social preferences, as it is known from the literature that this maximizes the chance that subjects do not simply maximize their own expected payoff (see e.g. the meta analysis in Abeler, Nosenzo, and Raymond 2019). Note that we could not have kept silent about the information of insurers, as policy holders would then form their own assumptions, which might contaminate the results.

Second, in our model as well as in the experiment, policy holders are initially unaware that they can inflate their losses. In reality, they can, generally speaking, commit to loss inflation before deciding on their level of precaution, which implies that we would need to apply the Subgame Perfect Equilibrium as solution concept in the theory. From an experimental point of view, however, this implies that the first stage decision depends on other-regarding preferences and on lying costs, which could hence no longer be disentangled. Furthermore, it doesn't seem to be far-fetched that many people in reality do not contemplate on loss inflation when deciding about their care level. For these reasons, we informed policy holders about the possible loss inflation only after their investment decision.

Third, we assume that insurers can observe whether policy holders have invested or not, and also whether they have inflated the loss or not. The latter assumption can be challenged as policy holders would not be compensated if loss inflation can be observed (and proven in the courtroom, which is not necessarily the same). The main reason why we decided for full transparency is that policy holders might otherwise form different expectations about what insurers anticipate and what they don't, which might again have contaminated the results. In any case, our assumption adds image concerns on top of other-regarding preferences. This may well contribute to the high frequency of investments, as it is known from the literature that people misreport less often when this is observable.

In our paper, precaution is binary, and selfish policy-holders would not invest in precaution. In reality, some precaution measures that are easily observable are contractually fixed, that is, getting insurance requires that prospective policy holders show proof of their precaution measures. Often, however, precaution is costly

---

19 We excluded subjects who failed the comprehensive questions, but understanding the consequences of decisions does not necessarily imply taking them carefully into account.

to observe ex ante (otherwise moral hazard wouldn't be an issue) but might nevertheless be reconstructible ex post at reasonable cost. In addition to our research question whether precaution may provide a signal on loss inflation, insurers could suggest contracts where compensation depends on precaution. The advantage is that precaution does not need to be checked ex ante, but only in case a loss occurs.[20]

We view our paper as a first step in analyzing the link between precaution and misreporting. Further research is needed, given that our results are in some respects inconclusive. This holds specifically for the role of moral licensing, where our *p*-values of around 0.3 (in different specifications) imply that we cannot safely exclude that one might find a significant impact on misreporting with a larger data set, or with a different experimental design.

For further research, it is instructive to interpret our design from a more general perspective. The precautionary investment reduces the investor's expected payoff, but at the same time increases the joint payoff of the two parties involved. Investment can therefore be seen as contributing to a public good. Whether a loss occurs or not can then just be interpreted as a random move that determines the overall payoff and its distribution between the two parties, and investors can then lie to increase their share. Distinguishing between own and random investment, and framing everything generically only with respect to decisions and payoffs, may then lead to more clear-cut results for the (relative) impacts of other-regarding preferences and moral licensing on misreporting.

Sample Instructions

As an example, we provide the instructions for insurance framing with own investment. All other instructions are available on request.

References

Abeler, J., D. Nosenzo, and C. Raymond. 2019. "Preferences for Truth-Telling." *Econometrica* 87: 1115−53. https://doi.org/10.3982/ecta14673.
Aslam, F., A. Hunjra, Z. Ftiti, W. Louhichi, and T. Shams. 2022. "Insurance Fraud Detection: Evidence from Artificial Intelligence and Machine Learning." *Research in International Business and Finance* 62: 101744.

---

20 We are grateful to an anonymous referee for raising this question.

von Bieberstein, F., and J. Schiller. 2018. "Contract Design and Insurance Fraud: An Experimental Investigation." *Review of Managerial Science* 12: 711–36. https://doi.org/10.1007/s11846-017-0228-1.

von Bierberstein, F., Feess, and Packham. 2024. "Multi-step Delegation and the Frequency of Immoral Decisions: Theory and Experiment." *Working Paper*.

Blanken, Irene, Niels Van de Ven, and Marcel Zeelenberg. 2015. "A Meta-Analytic Review of Moral Licensing." *Personality and Social Psychology Bulletin* 41. https://doi.org/10.1177/0146167215572134.

Brinkmann, J., and P. Lentz. 2006. "Understanding Insurance Customer Dishonesty: Outline of a Moral-Sociological Approach." *Journal of Business Ethics* 66: 177–95. https://doi.org/10.1007/s10551-005-5575-1.

Cooper, David J., and John H. Kagel. 2016. "A Selective Survey of Experimental Results." In *The Handbook of Experimental Economics*.

Dean, D. H. 2004. "Perceptions of the Ethicality of Consumer Insurance Claim Fraud." *Journal of Business Ethics* 54: 67–79. .

Dehghanpour, A., and Z. Rezvani. 2015. "The Profile of Unethical Insurance Customers: A European Perspective." *International Journal of Bank Marketing* 33: 298–315. https://doi.org/10.1108/ijbm-12-2013-0143.

Derrig, Richard A. 2002. "Insurance Fraud." *Journal of Risk & Insurance* 69: 271–87. https://doi.org/10.1111/1539-6975.00026.

Dollan, P., and M. Galizzi. 2015. "Like Ripples on a Pond: Behavioral Spillovers and Their Implications for Research and Policy." *Journal of Economics and Psychology*: 1–16. https://doi.org/10.1016/j.joep.2014.12.003.

FBI. 2019. "Covid-19 Fraud: Law Enforcement's Response to Those Exploiting the Pandemic." In *Statement Before the Senate Judiciary Committee Washington, D.C*.

Fehr, E., and K. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics*: 817–68. https://doi.org/10.1162/003355399556151.

Fiederling, Schiller, and F. von Bieberstein. 2018. "Can We Trust Consumers' Survey Answers when Dealing with Insurance Fraud?" *Schmalenbach Business Review* 70: 111–47. https://doi.org/10.1007/s41464-017-0041-z.

Fries, Tilman, and Daniel Parra. 2021. "Because I (Don't) Deserve it: Entitlement and Lying Behavior." *Journal of Economic Behavior & Organization* 185: 495–512. https://doi.org/10.1016/j.jebo.2021.03.007.

Galeotti, Fabio, Reuben Kline, and Raimondello Orsini. 2017. "When Foul Play Seems Fair: Exploring the Link between Just Deserts and Honesty." *Journal of Economic Behavior & Organization* 142: 451–67. https://doi.org/10.1016/j.jebo.2017.08.007.

Georges, L. Dionne Caron. 1997. "Insurance Fraud Estimation: More Evidence from the Quebec Automobile Insurance Industry." In *Ecole des Hautes Etudes Commerciales de Montreal-Chaire de gestion des risques*.

Gneezy, U., A. Imas, and K. Madarász. 2014. "Conscience Accounting: Emotion Dynamics and Social Behavior." *Management Science* 60. https://doi.org/10.1287/mnsc.2014.1942.

Grundmann, S., L. Spantig, and S. Schudy. 2023. "Individual Preferences for Truth- Telling." *Working Paper*.

IRC. 2013. *Insurance Fraud: A Public View*. Philadelphia, PA, USA: Insurance Research Council.

Martuza, J. B., S. R. Skard, L. Løvlie, and H. Thorbjørnsen. 2022. "Do Honesty-Nudges Really Work? a Large-Scale Field Experiment in an Insurance Context." *Journal of Consumer Behaviour* 21: 927–51. https://doi.org/10.1002/cb.2049.

Merritt, Anna, Daniel Effron, and Benoit Monin. 2010. "Moral Self-Licensing: When Being Good Frees
    Us to Be Bad." *Social and Personality Psychology Compass* 4: 344–57.
    https://doi.org/10.1111/j.1751-9004.2010.00263.x.
Mintchik, and Knechel. 2022. "Do Personal Beliefs and Values Affect an Individual's "Fraud
    Tolerance"? Evidence from the World Values Survey." *Journal of Business Ethics*: 463–89.
    https://doi.org/10.1007/s10551-020-04704-0.
Miyazaki, A. D. 2009. "Perceived Ethicality of Insurance Claim Fraud: Do Higher Deductibles Lead to
    Lower Ethical Standards?" *Journal of Business Ethics* 87: 589–98.
    https://doi.org/10.1007/s10551-008-9960-4.
Morrison, William, and Bradley Ruffle. 2020. "Insurable Losses, Pre-filled Claims Forms and Honesty
    in Reporting." *McMaster University Department of Economics Working Paper 2020-01*.
Mullen, Elizabeth, and Benoît Monin. 2016. "Consistency versus Licensing Effects of Past Moral
    Behavior." *Annual Review of Psychology* 67: 363–85.
    https://doi.org/10.1146/annurev-psych-010213-115120.
Ribeiro, Silva, Pimenta, and G. Poeschl. 2020. "Why Do Consumers Perpetrate Fraudulent Behaviors in
    Insurance?" *Crime, Law and Social Change* 73: 249–73.
    https://doi.org/10.1007/s10611-019-09857-2.
Tennyson, Sharon. 2008. "Moral, Social, and Economic Dimensions of Insurance Claims Fraud." *Social
    Research* 75: 1181–204. https://doi.org/10.1353/sor.2008.0020.
Tseng, Lu-Ming, and Chia-Lin Kuo. 2014. "Customers' Attitudes toward Insurance Frauds: An
    Application of Adams' Equity Theory." *International Journal of Social Economics* 41: 1038–54.
    https://doi.org/10.1108/ijse-08-2012-0142.
Tumminello, Michele, Andrea Consiglio, Pietro Vassallo, Riccardo Cesari, and Fabio Farabullini. 2022.
    "Insurance Fraud Detection: A Statistically Validated Network Approach." *Journal of Risk &
    Insurance* 90: 381–419. https://doi.org/10.1111/jori.12415.
Villegas-Ortega, J., L. Bellido-Boza, and D. Mauricio. 2021. "Fourteen Years of Manifestations and
    Factors of Health Insurance Fraud, 2006–2020: a Scoping Review." *Health and Justice* 9.
    https://doi.org/10.1186/s40352-021-00149-3.