The Framing of Information and Effective Enforcement Mechanisms

April 23, 2020

Abstract

Many firms are interested in finding effective ways to promote ethical behavior among their employees, without investing heavy resources into monitoring (compliance functions). Under the hypothesis that firms may have a cost-less information framing tool at their disposal, we study experimentally how revealing different information about a punishment distribution affects deterrence of undesirable behavior. We use a novel incentive-compatible elicitation method to observe lying across subjects and quantify the extent to which this behavior responds to information structures. We find that ambiguous punishment schemes, such as providing an individual with a minimum or a maximum fine, are more effective at deterring undesirable behavior compared to schemes which specify the exact distribution of fines. We further document that the 'minimum' frame outperforms the 'maximum' one, identify the mechanism driving this result, and discuss practical and theoretical implications.

1 Introduction

In this paper, we investigate the effectiveness of differently framed punishment schemes on preventing an undesirable behavior ("a crime"). This is an important question for firms who wish to promote ethical and honest behavior among their employees (from not stealing office supplies to not engaging in insider trading), but may be constrained in their ability to monitor all individuals all the time. One seemingly obvious solution - to threaten exorbitant fines for even small transgressions - is legally questionable if the fine clearly "doesn't fit the crime." Given these constraints, we ask which punishment framing is the most effective and why.

As a simple motivating example, consider a firm which is spread out over a large campus and wishes to enforce basic traffic rules, such as speed limits and reserved spots for handicapped parking. As Figure 1 shows, traffic signs for small offenses are formulated in a variety of ways. Some signs specify the exact fine amount (e.g. '$1000 fine for littering'), while others mention the minimum fine (e.g. 'Red Light Violation $336 Minimum Fine'), and yet others are even more vague, asserting that fines will be double the regular ones without explicitly listing them (e.g. 'Double Fine Zone'). Abstracting away from any legal reasons behind these frame choices, we provide some of the first empirical evidence comparing the effectiveness of these different frames. More generally, we ask whether people react differently to ambiguous punishments compared with non-ambiguous ones, and explore the mechanism behind this difference in behavior.

Figure 1: Examples of Signs Containing Fine Information in Los Angeles County (2019)

Notes: The most recent blank California MUTCD Sign Charts can be found at: https://dot.ca.gov/programs/traffic-operations/sign-charts

To achieve this goal, we conduct a series of laboratory experiments in which subjects have the opportunity to lie ("commit the crime"), an action that rewards them with higher earnings if they are not caught. In the experiment, each subject is allocated one of five cards numbered from 1 to 5 at random and is asked to report the card number she receives. A subject is paid twice the number she reports. In treatments with monitoring and punishment, subjects' reports are monitored with commonly known probability of 20% and a subject faces a fine if she is caught lying. In line with our research goal, we use a variety of monitoring treatments in which we alter the fine structure communicated to subjects while keeping the expected value of the fine constant. These treatments include the fixed fine, the random fine (equal chance of receiving a high or low fine), a minimum fine, and

a maximum fine. Except for the fixed fine treatment, the distribution of fines in the other three treatments is held constant.

Our experimental results show that ambiguous frames, which communicate either the minimum or the maximum fine, are more effective at deterring lying as compared with the fixed fine and the random fine schemes. Among the two ambiguous frames, the minimum frame seems to be strictly or weakly more effective than the maximum one, depending on the chosen measure.[1]

Why are ambiguous frames more effective than non-ambiguous ones? There are two possible mechanisms that might be at play here. According to the first one, by not specifying the average or actual fines, ambiguous signs create a situation in which people expect the fines to be *higher than what they actually are*. This coupled with the fact that expected fines (and possibly fines' distribution) are the main determinants of lying behavior leads to higher deterrence rates of ambiguous compared with non-ambiguous frames. The second mechanism builds upon the suggested by the literature tendency of people to dislike ambiguity naturally embedded in both the minimum and the maximum frames. According to this mechanism, people might estimate their own fine if they are caught lying to be *higher than that of an average violator*. If that is the case then higher compliance rates observed in the ambiguous frames compared to the non-ambiguous ones is primarily driven by higher compliance rates of ambiguity-averse subjects.[2]

To investigate which of the two mechanisms (or both) drive our results, we conduct additional treatments in which we elicit subjects' beliefs in addition to observing their choices in the card game. Here we propose a simple and intuitive way to measure subjects' attitudes towards ambiguity. Specifically, we ask subjects to report (1) their belief about the *average fine* previous subjects participating in this treatment faced when caught lying, and (2) their belief about what their *own fine* would be if they are caught lying.

We find several important results. First, the majority of our subjects believe that their own fine would be higher than the average fine in the population, consistent with the notion of aversion to ambiguity. Second, subjects' behavior in the card game is strongly correlated with beliefs about their *own fine* but not with beliefs about the average fine. Third, the average and the median belief about own fines in the ambiguous treatments are not higher than average fines in the non-ambiguous treatments, contrary to the first mechanism described above. At the same time, the minimum frame induces higher beliefs

[1]The fact that minimum frame is found to be at least weakly more effective than the maximum one suggests that anchoring mechanism is not the main driver of behavior, since in that case we would expect the opposite pattern.

[2]Of course, these two mechanisms are not mutually exclusive and might both play a role here. The follow-up experiment described in the next paragraph would be able to detect that scenario.

about one's own fine compared with the maximum frame, which coupled with the second result shows why the minimum frame outperforms the maximum one in terms of reducing lying behavior. Finally, lying behavior in the cards game is negatively correlated with being ambiguity-averse and the spread of beliefs, i.e., the difference between the own and the average fines, consistent with the second mechanism described above. These results suggest that firm managers have a powerful tool for deterring crime; one which does not require increasing resources - i.e. investing more in their compliance function.

The remaining of the paper is structured as follows. In Section 2, we survey the related literature. Section 3 describes the experimental protocol and theoretical predictions in the Main Experiment. Section 4 presents the results of the Main Experiment. Section 5 expands on the mechanism driving subjects' behavior using the data from the Follow-up Experiment. Section 6 provides some conclusions and practical implications.

2 Related Literature

Our work relates to several strands of literature. The first one is concerned with measuring the prevalence and determinants of lying behavior in laboratory experiments (Fischbacher and Heusi (2013), Gneezy, Rockenbach, and Serra-Garcia (2013)) and references mentioned there).[3] Different from this literature, our focus is on mechanisms that prevent lying rather than on measuring the extent of lying per se.

Motivated by the theoretical analysis of crime and law enforcement (see the classical model of Becker (1968)), there is an active and fascinating experimental literature which investigates interventions and their effectiveness at reducing undesirable behavior in the lab. Engel (2016) provides a comprehensive survey of this research.[4] In particular, experiments have documented that more severe punishments are more successful at deterring crime activity (Engel and Nagin (2015) and references mentioned there), compared the deterrence effects of increasing monitoring probability versus severity of punishment (Nagin and Pogarsky (2003), Friesen (2012), and Feess et al. (2014)), and explored how effective social norms are at deterring undesirable behavior (Dwenger et al. (2016), Casagrande et al. (2015)). As far as we know, our paper is the first to compare the effectiveness of different frames of punishment, while holding fixed the monitoring probability and the severity of the punishment.

[3]See also Tergiman and Villeval (2019) for the experimental study of effects of reputation on lying behavior in the markets. In addition, Erat and Gneezy (2011) investigate different types of lies, the 'white lies', which may benefit the person on the receiving end of a lie.

[4]See also Horne and Rauhut (2011) who evaluate the strength and weaknesses of the experimental approach in studying crime and law enforcement questions.

The two most closely related papers to ours are DeAngelo and Charness (2012) and Salmon and Shniderman (2019). DeAngelo and Charness (2012) consider how varying jointly monitoring probabilities and fines affect deterrence rates, and, specifically, focus on the link between preferences for punishment regime and compliance rates. The authors find that violations are less likely when the expected cost of violation is higher and when there is uncertainty about which regime is implemented. Contrary to our paper, however, the authors do not study ambiguous regimes and focus on the settings in which probabilities of each regime are common knowledge among participants. Salmon and Shniderman (2019) conduct a tax compliance experiment to illustrate how individuals respond to ambiguous punishment probabilities and, in particular, how they respond to shifts in ambiguous versus known probabilities. They find that when probabilities are known and shift, the standard model works well to explain the behavioral response. Whereas when the probabilities are ambiguous and shift, the behavioral response is minimal. Related experiments have sought to infer the probabilities of being caught as agents' perceive them (Bebchuk and Kaplow (1992)).[5] However, no previous work has systematically investigated how the information revealed about the fine distribution of a punishment scheme influences deterrence behavior, an important gap in the literature to date.

Our paper also relates to the literature that measures ambiguity attitudes of subjects and studies its implications in various settings. This literature is new and still lacks consensus on the prevalence of ambiguity aversion attitudes in the population. For example, Kocher, Lahno, and Trautmann (2015) investigate whether ambiguity aversion drives behavior in a broader class of decision tasks and find that there are relatively few configurations of a choice environment in which subjects display aversion to ambiguity. Similarly, Ahn et al. (2014) are not able to reject the null hypothesis of Subjective Expected Utility for a majority of their subjects in a portfolio choice experiment, although a fraction of participants do exhibit significant aversion to ambiguity. Part of the difficulty in this literature is accurately and separately estimating the ambiguity attitude and the risk attitude of subjects, since estimates of ambiguity aversion tend to be greater under the assumption of risk neutrality, while the majority of subjects are risk averse. This is the point made in Gneezy, Imas and List (2015) who jointly estimate risk and ambiguity attitudes and document that ambiguity aversion is much less prevalent than found by the previous literature. We contribute to this literature by proposing a simple technique that measures both the

[5]See also theoretical model of Calford and DeAngelo (2020), in which agencies who wish to minimize criminal activity should reveal their resource allocation if criminals are uncertainty seeking and shroud their allocation if criminals are uncertainty averse. The authors supplement theoretical analysis with experimental evidence largely consistent with the theoretical predictions.

ambiguity attitude of subjects as well as the intensity of this attitude. We apply this new technique in the loss domain and show that the majority of subjects are ambiguity averse according to our measure.

3 Main Experiment

3.1 Experimental Protocol

All experimental sessions were conducted at the Experimental Economics Laboratory at the University of California in San Diego between March 2019 and June 2019.[6] Since our experiment was short (it took approximately 5 minutes to complete), in lieu of recruiting subjects exclusively for our experiment, we asked other experimenters to add it at the end of the experimental session as an additional task.[7] Our instructions were very clear about the fact that the task performed by subjects in this last part of the experiment has nothing to do with the previous parts, and that their payment for the two tasks were independently determined. Overall, 424 students from the general population of UCSD participated in our experimental sessions. The experiment was programmed in O-Tree (Chen, Schonger, Wichens (2016)).

Motivated by a variety of punishment schemes used by law enforcement agencies in reality, we conducted five different treatments. In all treatments, a subject is allocated one of the five cards labeled with numbers 1 through 5, selected at random. The task is to report the number on the card one receives. If a subject reports a number x, then she earns $\$2x$. The treatments differ by the presence of monitoring and the fine that a subject incurs if she is caught lying. In the *Baseline* treatment, there is no monitoring and no fines, i.e., subjects simply report their card number and collect their payments. In the remaining four treatments, there is a 20% chance that a subject is audited and punished if she lied, i.e., reported a number different from the number specified on her allocated card.

In the *Fixed* treatment, a subject who is caught misreporting her card number pays a fine of \$5. In the *Random* treatment, the fine is either \$3 or \$7 with equal chance. In the *Minimum* treatment, the fine is at least \$3, and, finally, in the *Maximum* treatment, the fine is at most \$7. In actuality, for the *Minimum* and *Maximum* treatments, we use the same distribution of fines as in the *Random* treatment, i.e., the fine is either \$3 or \$7 with equal chance, but subjects do not know this fact. In all cases, the fines are subtracted from

[6]We thank the researchers at UCSD Experimental Economics lab for their generosity in allowing us to run these sessions.

[7]We have made sure that subjects participated in no more than one experimental session.

the earnings that are based on the reported number.[8]

The experiment was conducted using the strategy method, i.e., subjects had to submit a number for each of the five possible cards. Then, to determine their payment, the computer randomly selected one of the five cards and calculated the subject's earnings based on the report provided for the selected card and the monitoring/punishment scheme specified by the treatment. The monitoring was implemented by a random draw performed by the computer for each subject individually.[9] The advantage of using the strategy method is that we are able to collect individual level data corresponding to subjects' choices for all five cards they may receive.[10]

Table 1 summarizes our experimental treatments and the number participants. The instructions for all treatments are presented in Appendix 6.

Table 1: Experimental Treatments

Treatment	Monitoring probability	Fine if caught lying	# of subjects
Baseline	no monitoring	no fine	84 subjects
Fixed	20% known	\$5 for sure	80 subjects
Random	20% known	\$3 or \$7 with equal chance	88 subjects
Minimum	20% known	at least \$3	92 subjects
Maximum	20% known	at most \$7	80 subjects

In addition, we have access to individual characteristics of participants collected in the experiment conducted before ours. These controls include IQ measure, risk attitudes, and overconfidence. To measure IQ, subjects were asked to solve six Raven matrices and received 50 cents for each correctly solved puzzle. Subjects' overconfidence was measured using two related characteristics, over-estimation and over-placement (we used the proce-

[8]For instance, a subject in the *Fixed* treatment who received a card with number 3, reported number 4 and was caught lying earns $\$3 = 2 \cdot \$4 - \$5$.

[9]One might worry that the mere fact that subjects knew that the experimenter is 'observing' their choices, i.e., collects their reports for all cards, may impact one's desire to misrepresent the received card number. This concern motivated majority of the previous experiments studying lying behavior in the lab to adopt a design in which subjects privately roll the dice in the cubicle without anyone observing them and then report the number they rolled. However, this design does not naturally lend itself to an investigation of the effectiveness of various punishment schemes, since one needs to know the side of the rolled dice to be able to determine whether a subject lied or not. This was the primary reason we have modified the design of the standard lying experiments. As we will show in the next section, this change did not affect the main empirical regularities we observed in the lying experiments, which is re-assuring as behavior seems to be stable to variations in the experimental protocol.

[10]See Gneezy, Rockenbach, and Serra-Garcia (2013) who also use the strategy method to elicit individual level tendency to lie in the laboratory experiment.

dure similar to Chapman et al. (2019)).[11] For measuring over-estimation, subjects were asked to estimate how many of the six Raven puzzles they solved correctly; the correct answer was rewarded by 50 cents. For measuring over-placement, subjects were asked to rank themselves in terms of how many correct puzzles they solved relative to 75 other UCSD students who completed this task before; the correct answer was rewarded by 50 cents. The risk attitudes were measured using two investment tasks, in each of which subjects were endowed with 200 points (worth a total of $2), any portion of which they could choose to invest in a risky project. In the first investment task, the risky project was successful 50% of the time and had a return of 2.5 points for each point invested in it, while in the second investment task the risky project was successful 40% of the times and returned 3 points for each point invested in it. Points not invested in the risky project had a return of 1 to 1 point. One of the two investment tasks was randomly chosen for payment. This is the standard method in the experimental literature to elicit subjects' attitudes towards risk (see Gneezy and Potters (1997) and Charness, Gneezy, and Imas (2013)). Administering this task twice with two sets of parameters allowed to reduce measurement error (see ORIV technique developed by Gillen, Snowberg, and Yariv (2018)).

3.2 Theoretical Predictions

In this section we provide a few general observations about predictions across our experimental treatments. Some of these predictions hold true for a wide class of behavioral models, while others depend more heavily on preferences' structure.

General Predictions. Any subject with preferences for monotonic payoffs should either report their true card number or lie to the fullest extent, i.e., report number five. This follows from the fact that the probability of monitoring and the size of the fine does not depend on the extent of lying but only on the mere fact of lying. In particular, models with monotonic preferences will not be able to accommodate self-sabotage behavior, which entails reporting numbers below the card number. Furthermore, if subjects' preferences are determined solely based on the payoffs they can earn in our experiment rather than any self-image or other psychological considerations, then we expect subjects to report the number five for all cards in the *Baseline* treatment.

[11] Over-estimation compares a subject's actual performance with her estimate of it. Over-placement talks about subjects' perceived performance relative to other participants in the specified group.

Expected Utility Theory. The comparison between lying propensity in *Baseline*, *Fixed*, and *Random* treatments depends on subject's risk attitude. Denote by $u(\cdot)$ subject's utility from monetary payments. Then a subject for whom

$$u(\$8) \leq 0.2 \cdot u(\$5) + 0.8 \cdot u(\$10)$$

is expected to report number five for all cards, since this inequality guarantees that expected payoff from lying is higher than expected payoff from telling the truth for card four, which is the card with the least incentives to lie. In this case, introduction of monitoring and fixed punishment should have no effect on lying deterrence as compared with the *Baseline* treatment. If, however, there exists a value of $x \in \{1, 2, 3, 4\}$ which violates inequality above then such a subject is expected to report the card number truthfully for all card numbers that violate inequality above and report number five for all card numbers that satisfy it. Moreover, a risk-averse subject is expected to lie weakly more in the *Fixed* than in the *Random* treatment since the latter one represents a mean-preserving spread of the former one, which is naturally disliked by a person averse to risk.

Prospect Theory. Given that subjects cannot make losses in an experiment, the Prospect Theory model of Kahneman and Tversky (1979) can make potentially different predictions from those of Expected Utility only if a subject evaluates gains and losses relative to an non-zero reference point. One such natural reference point could be the expected payoff of lying, which is the same for all cards and is equal to $0.2 \cdot \$5 + 0.8 \cdot \$10 = \$9$. Denote by $u(\cdot)$ and $\lambda \cdot u(\cdot)$ the two parts of the utility function that a subject uses to evaluate risky alternatives, where $u(\cdot)$ is applied for payoffs above and $\lambda u(\cdot)$ for payoffs below the reference point and $\lambda > 1$. Then a subject is predicted to tell the truth for card x in the *Fixed* treatment if $\lambda u(\$2x) > 0.2\lambda u(\$5) + 0.8u(\$10)$ and report five otherwise. Similarly, a subject is expected to report card x truthfully in the *Random* treatment if $\lambda u(\$2x) > 0.2\lambda \left[0.5u(\$3) + 0.5u(\$7)\right] + 0.8u(\$10)$ and lie otherwise. Similar to Expected Utility theory predictions, we expect a loss-averse subject who is also risk-averse in the gains domain, where gains are defined relative to the reference point, to lie weakly less in the *Random* compared with *Fixed* treatment, because of concavity of function $u(\cdot)$.

MaxMin Expected Utility Theory. The predictions for the two ambiguous treatments (*Minimum* and *Maximum*) are harder to nail down, since they depend not only on subjects' ambiguity attitudes but also on individual beliefs about the distribution of fines, which are purposely not controlled or induced in these treatments. Taking as a starting point

the commonly used Maxmin Expected Utility model of Gilboa and Schmeidler (1989), an ambiguity averse subject will think about the worst fine that can occur in each of the treatments and will act believing that this is the fine she will face if caught lying. In the *Maximum* treatment, the worst fine is defined precisely and it is equal to \$7. In the *Minimum* treatment, a reasonable belief about the highest possible fine is bounded by the highest payoff one can earn in this experiment, which is \$10, since subjects know that the experimenter cannot collect money from the participants. However, what subject's actually think the highest fine is we do not know, and, thus, we turn to the empirical data for answers by comparing behavior in the two ambiguous treatments.

4 Results of the Main Experiment

We report the results in the following order. First, we use the data from *Baseline* treatment to establish that our experimental protocol generates behavior similar to the one documented in the previous literature. Second, we show that any monitoring and any type of punishment reduces lying. Third, we run the horse-race between different punishment schemes and determine the most effective one.

Throughout the analysis we use non-parametric tests. Specifically, we use Wilcoxon RankSum tests to compare average lying frequencies between two groups and report the p-value associated with the null hypothesis that two groups have the same average frequency. We use Wilcoxon SignRank tests to compare observed frequency with the target value and report the p-value associated with the null hypothesis that observed frequency equals the target value. To zoom in on determinants of behavior across and within treatments, we conduct regression analysis, in which we control for individual characteristics of subjects. Specifically, we run linear probability models and utilize the ORIV method developed by Gillen et al. (2018) to reduce measurement errors in the risk data.

4.1 Baseline treatment and the previous literature

In order to study effectiveness of monitoring and punishment schemes, we had to modify the standard protocol of lying experiments used in the literature. In this section, we compare behavior observed in our *Baseline* treatment with key findings reported in Fischbacher and Heusi (2013), FH hereafter, a paper which inspired hundreds of studies on lying behavior. The design adopted by FH and nearly all papers in this literature following FH, is best known as the dice-in-a-cup design. In this set-up, participants shake a six-sided die in a cup and report the number they receive; higher numbers typically correspond to the

higher payoffs.[12] Researchers then study lying in aggregate by looking at how the realized distribution of reported numbers differs from the distribution one would expect assuming a fair die.

The underlying premise of the dice-in-a-cup design is that image issues might lessen otherwise present lying behavior if experimenters were to observe subjects' actual roll of a die. This premise, while natural, has never been explicitly tested, which is a surprise given that the aggregate data in the dice-in-a-cup experiments does not allow examination of individual behavior crucial for the investigation of lying phenomena. In contrast, our experiment utilizes the strategy method and allows collection of rich individual-level data. However, first, we need to establish that this new experimental protocol does not alter lying behavior compared with one observed in the standard dice-in-a-cup experiments.

FH document three main empirical regularities. First, some people seem to be honest and truthful in reporting the number they roll.[13] Second, a greater-than-expected fraction of payoff maximizing numbers are reported, indicating that many people lie. Third, there is a greater-than-expected fraction of people who report a number which gives them the second-highest possible payoff.[14]

Data in our *Baseline* treatment exhibit the same three regularities. First, we observe that 23% of subjects report the actual card number for all five cards. Second, the majority of subjects (64%) report the number five for all five cards, i.e., the number that delivers the highest payoff. Third, 4% of subjects report payoff-maximizing numbers (number five) sometimes but not always. We reject the null that proportions of any of the described above types in our data is zero ($p < 0.001$ for the first two groups and $p = 0.0833$ for the last group of subjects). We conclude that lying behavior in the lab is not very sensitive to modifications of the experimental protocol. In other words, there are no losses associated with moving from the dice-in-the cup design to the strategy method, but there are significant gains since it allows us to collect richer individual-level data.

Observation 1: *Observing individual-level choices of subjects in the lying experiment produces the same aggregate lying behavior as the commonly used dice-in-the-cup paradigm, with the additional benefits of capturing individual-level data.*

[12] In the FH paper, subjects are paid an amount equal to the number they report, unless the number is six, in which case they are paid zero.

[13] One would never be able to detect with certainty whether or not people are reporting truthfully, since no one observes their roll of a die. Thus, this conclusion is based on the observation that a non-negligent fraction of the lowest numbers are reported.

[14] This last observation is consistent with the idea of self-image, according to which people lie but not to the fullest in order to 'preserving some self-dignity'.

4.2 Monitoring and punishments prevent crime

We classify subjects into four mutually exclusive types based on the individual profiles of choices:

1. ***Under-reporters*** are subjects who report a number strictly smaller than the card number for at least one card.
2. ***Honest*** are subjects who truthfully report the card number for all five cards.
3. ***Liars*** are subjects who report number five for all five cards.
4. ***Quasi-Liars*** are subjects who report numbers which are higher or equal than the card numbers, where at least one reported number is different from five.

Figure 2: Distribution of individual types, by treatment

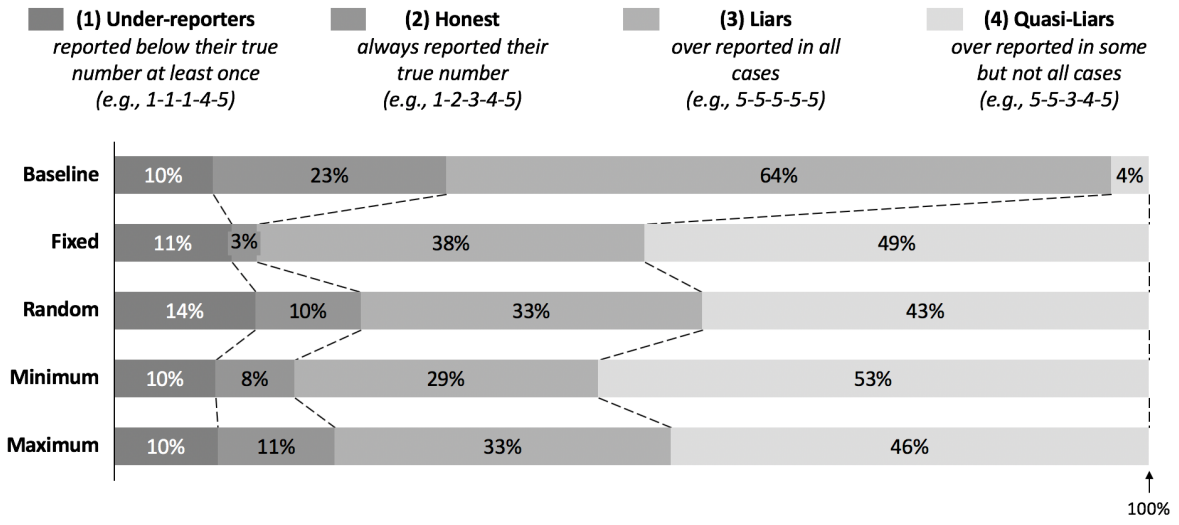


Figure 2 depicts the distribution of types in each treatment. A policy maker who wishes to improve overall welfare might be primarily concerned with reducing the proportion of Liars in the population. Compared with the *Baseline* treatment, all treatments with monitoring and punishment significantly reduce the fraction of Liars. This reduction is large in magnitude and statistically significant; indeed, it cuts the fraction of Liars from about two-thirds to a third of subjects or less depending on the punishment scheme ($p < 0.01$ in all pairwise comparisons). The reduction in the fraction of Liars is accompanied by a

large increase in the fraction of Quasi-Liars across all treatments, ranging from 4% in the *Baseline* treatment to more than 40% in all other treatments.[15]

Interestingly, the introduction of monitoring and punishments significantly reduces the fraction of Honest subjects in the population from 23% in the *Baseline* treatment to as little as 3% in the *Fixed* treatment and about 10% in all other treatments ($p < 0.01$ in all pairwise comparisons between Baseline and other treatments). This is consistent with the literature on how imposed incentives can backfire by crowding out intrinsic motivation (Gneezy, Meier and Rey-Biel (2011)). Finally, we note that the proportion of Under-reporters remains stable across treatments and ranges between 10% and 15%.[16]

Observation 2: *Monitoring and punishments reduce lying behavior by cutting the fraction of those who always lie by at least half.*

4.3 Effectiveness of punishment schemes

We assess the effectiveness of different punishment schemes in three complementary ways. First, we compare the distribution of types across the four monitoring treatments. Second, we zoom into different kinds of behaviors within the group of Quasi-Liars. Finally, we disaggregate the data into responses for each card separately and compare lying frequencies across treatments.

The distribution of types across different punishment schemes is quite stable as seen in Figure 2 with no significant differences detected between any pair of punishment treatments for any category. The only exception is the fraction of Honest types which is significantly smaller in the *Fixed* treatment compared with the *Random* ($p = 0.0438$) and with *Maximum* ($p = 0.0292$) treatments.

However, the similarity in population types across treatments hides important differences between treatments within the category of Quasi-Liars. Note that this group encompasses two different individual profiles of choices: subjects who lie for lower cards and report truthfully for higher cards, and subjects who lie but not to the fullest extent.[17] Figure 3 presents the breakdown of the Quasi-Liars into these categories distinguishing between subjects who lie only for card 1, for cards 1 and 2, for cards 1, 2, and 3, and those who lie but not to the fullest extent possible (Incomplete Liars). The first three columns

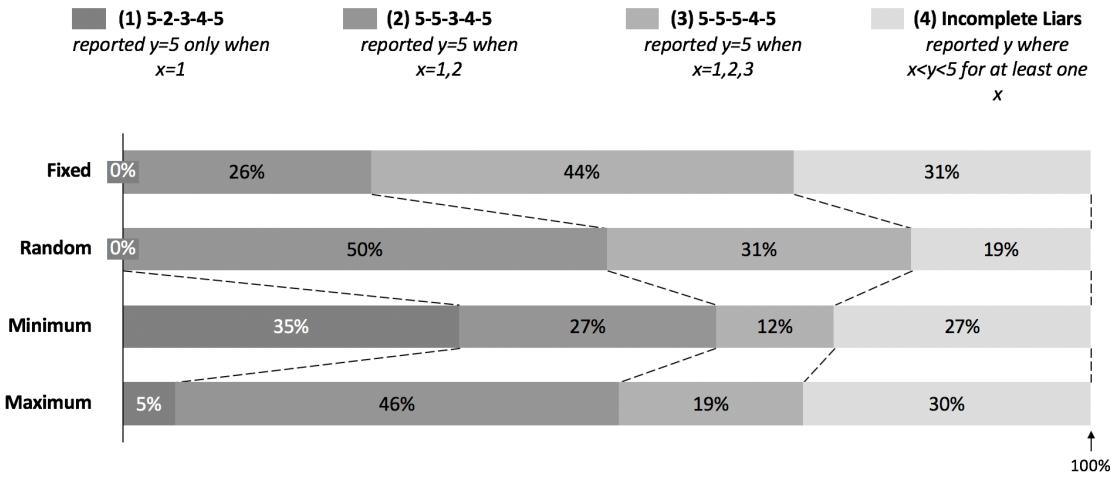
[15] We find $p < 0.01$ in all pairwise comparisons of fraction of Quasi-Liars between the *Baseline* and all other treatments.

[16] We cannot reject the null that the fraction of Under-reporters is the same in every pair of treatments with $p > 0.10$. We have no way to identify whether these are the subjects who did not understand the instructions or simply made a typing mistake in recording the numbers.

[17] This last category is called 'incomplete liars' in the FH paper.

in Table 2 report the results of the regression analysis conducted to detect differences in these sub-types across treatments.

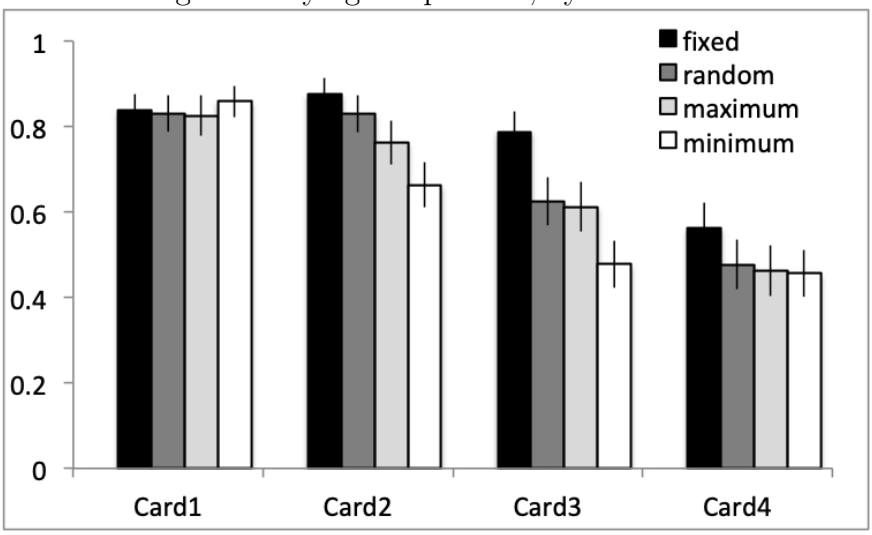
Figure 3: Types of Behavior within Quasi-Liars Category



Regressions reported in Table 2 confirms what we see in Figure 3. The *Minimum* treatment features the highest number of subjects who lie only for the lowest card, i.e., type 52345, among all punishment treatments (see Regression (1)). Moreover, the fraction of people who lie for all but one card (type 55545) is lower in the two ambiguous treatments *Minimum* and *Maximum* compared with *Random* and *Fixed* treatments as seen in Regression (3) and comparative tests of the estimated coefficients reported at the bottom of Table 2. In other words, while the fraction of subjects who lie occasionally is the same across punishment treatments, the *Minimum* treatment features the least amount of lying within this category.

To drive this point home, Figure 4 presents lying propensities for each of the cards except for card 5 for which all subjects report number 5. The main difference between treatments is observed for cards 2 and 3, with the vast majority of subjects lying for card 1 and telling the truth for card 4. The last two columns in Table 2 presents regressions comparing lying propensities for cards 2 and 3 across treatments. The data clearly shows that the two ambiguous treatments, *Minimum* and *Maximum*, reduce lying propensities for card 3 compared with the other two treatments with known distribution of fines. In addition, *Minimum* treatment reduces lying for card 2 as well compared with all other treatments: this reduction is highly significant at the 1% level for *Minimum* versus *Fixed* and *Random* treatments and marginally significant at 10% for *Minimum* versus *Maximum*

Figure 4: Lying Frequencies, by treatment



Notes: Frequencies of lying are plotted for each card separately with the standard errors of the means.

treatments. The reduction in lying propensities are quite substantial with average lying of more than 10 percentage points lower in the *Minimum* treatment than in any other treatment.

Observation 3: *Ambiguous fine distributions communicated as a "minimum fine" or a "maximum fine" are more effective at deterring lying than unambiguous ones, such as the "fixed" fine or the precise fine distribution as "random" fine. Among the ambiguous fine structures, the one framed as the "minimum fine" is marginally more effective than the one framed as the "maximum fine".*

5 Mechanism Driving the Results

The results of our main experiment show that ambiguous framings outperform other, less vague, framings of a pre-specified distribution of fines with respect to deterring subjects from lying in the card experiment. In addition, the *Minimum* scheme seems to perform slightly better than the *Maximum* one. Why would that be the case? This is not entirely obvious. In the ambiguous treatments, subjects form subjective beliefs about the distribution of fines they would face if caught lying. The treatments do not control much about these beliefs since only the minimum or maximum fines are specified. If subjects anchor in some way to the observed numbers, i.e., to \$3 in the *Minimum* treatment and to \$7 in the *Maximum* treatment, then all else equal the *Maximum* treatment should outperform

Table 2: Main Experiment: Regression Analysis

	Dependent Variable: Indicator for				
	Reg. (1) *52345*	Reg. (2) *55345*	Reg. (3) *55545*	Reg. (4) Lie Card2	Reg. (5) Lie Card3
Indicator *Random*	-0.01 (0.02)	0.23** (0.11)	-0.08 (0.10)	-0.02 (0.06)	-0.12* (0.07)
Indicator *Maximum*	0.04 (0.04)	0.22* (0.11)	-0.23** (0.10)	-0.08 (0.06)	-0.14* (0.07)
Indicator *Minimum*	0.34*** (0.07)	0.04 (0.10)	-0.31*** (0.09)	-0.21*** (0.06)	-0.31*** (0.07)
Constant	0.04 (0.11)	0.29 (0.21)	-0.16 (0.19)	0.66*** (0.12)	0.55*** (0.14)
Individual Controls					
Risk attitudes	-0.0002 (0.001)	-0.001 (0.001)	0.002** (0.001)	0.001 (0.001)	0.002** (0.001)
IQ measure	-0.001 (0.02)	0.02 (0.03)	0.08** (0.03)	0.02 (0.02)	0.008 (0.02)
Overprecision	-0.01 (0.03)	0.04 (0.04)	0.02 (0.04)	0.02 (0.02)	-0.01 (0.03)
Overestimation	-0.0007 (0.001)	0.001 (0.002)	-0.001 (0.002)	0.001 (0.001)	0.0006 (0.001)
# of observations	$n = 163$	$n = 163$	$n = 163$	$n = 340$	$n = 340$
adjusted R-squared	0.2276	0.0669	0.1212	0.0464	0.0598
Sample	Sometimes Liars	Sometimes Liars	Sometimes Liars	All	All
Tests of Coefficients					
Minimum = Fixed	$p < 0.0001$	$p = 0.7175$	$p = 0.0005$	$p = 0.0006$	$p < 0.0001$
Minimum = Random	$p < 0.0001$	$p = 0.0553$	$p = 0.0093$	$p = 0.0025$	$p = 0.0080$
Maximum = Fixed	$p = 0.2679$	$p = 0.0544$	$p = 0.0198$	$p = 0.1616$	$p = 0.0585$
Maximum = Random	$p = 0.1935$	$p = 0.8819$	$p = 0.1148$	$p = 0.2689$	$p = 0.8160$
Minimum = Maximum	$p = 0.0001$	$p = 0.0876$	$p = 0.2872$	$p = 0.0630$	$p = 0.0227$

Notes: We report the results of ORIV (linear probability model) estimations with fixed treatment being the base group. The individual controls include (a) an IQ measured by the number of correctly solved Raven matrices, (b) risk attitudes measured by the fraction of the endowment invested in the risky project, the over-precision measured by the difference between the number of Raven matrices a subject thinks he solved correctly and the actual number he solved, and (d) over-placement measured by the difference between the predicted rank of a subject in a group of 100 undergraduate UCSD students and his actual rank. The ORIV method reduces the measurement errors in risk elicitations and is due to Gillen et al (2018), where the higher investments in the risky project are associated with less risk-averse subjects. ***, ** and * indicates significance at the 1%, 5% and 10% level, respectively.

the *Minimum* treatment as it is likely to induce higher beliefs about average fines. On the other hand, the highest possible fine in the *Minimum* treatment is not specified compared to the *Maximum* treatment, which leaves room for possible exaggeration of the worst fine that subjects can face in this treatment.[18] This coupled with ambiguity attitudes of subjects might make the *Minimum* treatment more effective at deterring lying. On top of that,

[18]The highest reasonable fine that subjects' might expect to get in the *Minimum* treatment is \$10, i.e., the highest payment they can get in the card game. Subjects are aware of the fact that the experimenter cannot take money from them, and that this card game is conducted independently from the other parts of the experiment, which means that payments in the previous part of the experiments are not affected in any way by what transpires in the card game.

the framings of the punishment schemes might affect subjects' attitudes towards ambiguity, which would in turn affect behavior in the card game. All in all, we need empirical evidence to be able to sort out between these competing possibilities and to identify the mechanism driving the results obtained in the main experiment.

To do that, we conducted a follow-up experiment focusing on the two ambiguous treatments. The new experiment serves two goals. First, we replicate the results of our main experiment to see how robust they are. Second, we elicit subjects' beliefs about fines in ambiguous treatments in an attempt to identify the main forces driving behavior.

5.1 Experimental Protocol of the Follow-up Experiment

The follow-up experiment consists of two treatments: the *MinBeliefs* and the *MaxBeliefs* with 96 and 94 subjects, respectively.[19] The two treatments are identical to the *Minimum* and the *Maximum* treatments in the main experiment with the addition of the two questions in which we elicited subjects' belief about the fine structure. The beliefs questions were presented to subjects in random order and administered at the end of the experiment before subjects' learned their payment for the cards game.

One of the questions elicited subjects' beliefs about average fine one would pay if caught lying, while the second question attempts to capture subjects' ambiguity attitudes by asking them to state the fine they believe *they specifically* would pay if they were caught lying. The idea here is that one might have the distribution of fines in mind. When asked to report the average fine for other people, one reports the average value from this subjective distribution. However, when asked about the own fine, a subject who is ambiguity averse tends to think about the worst possible scenario and reports a higher than average fine, while a subject who is ambiguity neutral reports the same two beliefs.[20] Here is the exact formulation of the beliefs' questions in the *MinBeliefs* (*MaxBeliefs*) treatments:

Q1: *In Spring 2019, 90 (80) UCSD students participated in the experiment identical to the one you just finished. Just like in your experiment, with probability 20%, a subject's reported number was compared with the actual card number she received, and in case these were different a subject incurred a penalty of at least \$3 (at most \$7). In that experiment, what do you think was the average penalty of subjects who were selected*

[19]The new treatments were conducted in the same Experimental Economics Laboratory at UCSD in November 2019 at the end of the unrelated experiment. We made sure that no subject participated in both the previous and the new treatments.

[20]We chose to elicit the average rather than the median fine because the concept of average is commonly used, while median is not so much. This should not matter as long as subjects believe that the distribution of fines is symmetric.

(according to 20% rule) and found to misreport their card number? If your guess is within +/− 50 cents of the actual average penalty in that experiment, you will receive an additional $1."

Q2: *Think about the experiment you finished. What do you think would be your penalty if you were selected (according to 20% rule) and you reported different number from the card number you received?"*

Few details of our beliefs' elicitation procedure deserve a discussion. First, we cannot incentivize the second question in which subjects report beliefs about their own potential fine. This suggests that we should take this measure with a grain of salt since it might be a noisy estimate of subjects' true beliefs. Second, while it would be interesting to elicit the whole distribution of beliefs that subjects' might consider, we opted for simpler and partial statistics about this distribution, i.e., the average fine and their own fine. Third, these two simple questions provide a measure of the *intensity of subjects' ambiguity attitudes* in addition to an indicator of whether a subject is ambiguity loving/neutral/averse. The intensity of subjects' ambiguity attitude can be measured by the difference between reported own fine and the reported average fine with the higher absolute value of the spread indicating the strength and the extent of ambiguity attitude. Given these two measures (dichotomous and continuous), we will investigate whether subjects' attitudes towards ambiguous events are affected by the framings of the punishment schemes, and, ultimately, correlate with card reports in the ambiguous treatments. In the analysis that follows, we will call a subject ambiguity-neutral if she reports her own fine to be equal to the average fine, ambiguity-averse if she reports higher own fine than the average one, and ambiguity-loving if she reports lower own fine than the average one.

5.2 Results of the Follow-up Experiment

We start by documenting the distribution of types in the new treatments based on the card reports (see Table 3). The distribution of types across new treatments is stable and similar to the one observed in the main experiment with most subjects falling into either the Quasi-Liars or Always-Liars category.

Table 4 reports the distribution of beliefs' types in the new treatments as well as summary statistics about the subjects' beliefs.[21] First of all, we note that average own fine in

[21] Our program allowed subjects to enter any numbers they wish for both belief questions. As a result, 4 subjects have specified fines below $3 in the *MinBeliefs* treatment and 3 subjects have specified beliefs above $7 in the *MaxBeliefs*. In addition, there are 7 subjects in the *MinBeliefs* treatment who specified

Table 3: Distribution of Types in the Follow-up Experiment

	Under-reporters	Honest	Quasi-Liars	Always-Liars
MinBeliefs	5%	11%	50%	33%
MaxBeliefs	7%	8%	46%	38%
	$p = 0.527$	$p = 0.499$	$p = 0.558$	$p = 0.477$

the two ambiguous treatments (*MinBeliefs* and *MaxBeliefs* pooled together) is not significantly different from 5, which is the expected average fine used in all treatments including the non-ambiguous ones. Therefore, the effectiveness of ambiguous compared with non-ambiguous frames does not come from subjects overestimating fines they will get if they are caught lying.[22]

Second, the largest group in the population are ambiguity-averse subjects who believe that they would face a higher fine if caught lying relative to the average fine administered for the same violation; this group constitutes about half of subjects in both treatments ($p = 0.44$). The remaining subjects are either ambiguity-neutral or ambiguity-loving with the fraction of ambiguity-loving subjects being larger in the *MaxBeliefs* than in the *MinBeliefs* treatment ($p < 0.01$). Consistent with the anchoring hypothesis, the average fine reported in the *MaxBeliefs* treatment is significantly higher than that reported in the *MinBeliefs* treatment ($p = 0.02$). This difference mostly driven by ambiguity-neutral subjects who believe that the fine would be on average \$1 more in the *MaxBeliefs* than in the *MinBeliefs* treatment. At the same time, subjects hold higher beliefs about their own fine in the *MinBeliefs* compared to the *MaxBeliefs* treatment ($p = 0.02$).[23]

Figure 5 gives a fuller picture of subjects' beliefs by plotting the differences between own and average fines computed at the subject level. The negative numbers indicate ambiguity-loving subjects, while the positive numbers indicate ambiguity-averse subjects. The picture clearly shows that the distributions are quite different across our two treatments: *MinBeliefs* treatment features distribution which is skewed to the right relative to the *MaxBeliefs* treatment where the distribution is much closer to being symmetric around ambiguity-neutral point of zero. Importantly, *MinBeliefs* treatment features higher difference between own and average fines compared to *MaxBeliefs* ($p < 0.001$). The same

that the average fine is above \$10, which is impossible by the design of the experiment. We have excluded these subjects from the analysis that follows, which leaves us with 91 subjects in the *MaxBeliefs* treatment and 85 subjects in the *MinBeliefs* treatment.

[22]The average own fine in both *MinBeliefs* and *MaxBeliefs* treatments pooled together is 5.04 with the standard error of 0.17. We cannot reject the null hypothesis that average own fine is equal to five with $p = 0.82$.

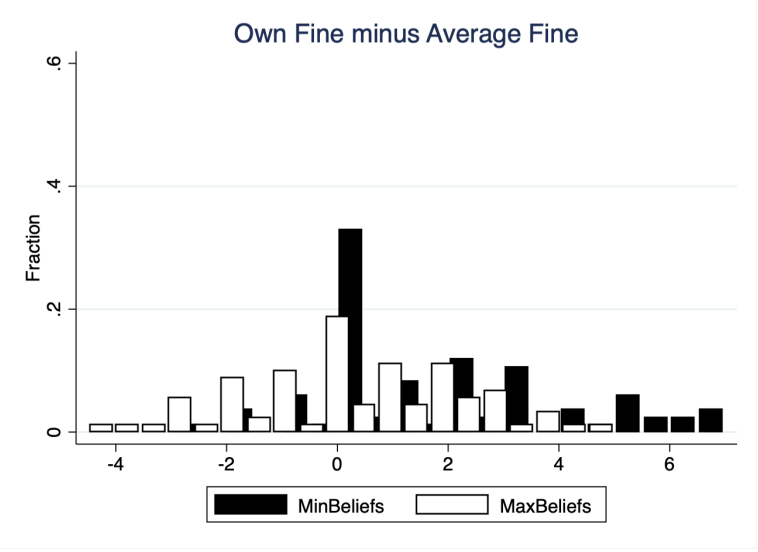
[23]Note that this result is despite the fact that we have removed a few outliers in the *MinBeliefs* treatment, i.e., subjects who believe in fines above \$10.

Table 4: Distribution of Belief Types in the Follow-up Experiment

	Ambiguity-Averse			Ambiguity-Neutral		Ambiguity-Loving			All Subjects	
		beliefs			beliefs		beliefs		beliefs	
	frac	ave	own	frac	ave/own	frac	ave	own	ave	own
MinBeliefs	0.55	3.6	6.7	0.33	3.8	0.12	4.5	3.1	3.8	5.3
MaxBeliefs	0.49	3.9	6.0	0.19	4.8	0.32	4.2	2.2	4.2	4.6
p-values	0.44	0.11	0.04	0.03	0.01	<0.01	0.46	0.05	0.02	0.02

Notes: We report the fraction of beliefs' types in the two treatments as well as mean average and own fines reported by each type. The last two columns list the mean average fine and mean own fine reported by all subjects in these two treatments. We exclude subjects who report unreasonable beliefs as defined in Footnote 21. The last row of the table reports the p-values comparing *MinBeliefs* and *MaxBeliefs* treatments.

Figure 5: Difference between Own Fines and Average Fines



pattern holds if we condition on subjects being ambiguity-averse only ($p = 0.001$). In other words, while the fraction of the ambiguity-averse subjects remains the same across two treatments, the *degree* of their ambiguity-aversion as captured by the difference between own and average fines is significantly higher in the *MinBeliefs* than in the *MaxBeliefs* treatment.

Observation 4: *Beliefs regarding the average (own) fine are higher (lower) in the treatment in which a "maximum" fine is specified than in a treatment in which a "minimum" fine is specified. Furthermore, about a half of subjects report that they believe their own fine will be higher than that of an average person with the spread between these two beliefs being*

higher in the "minimum" than in the "maximum" treatment.

We now turn to investigate the consistency between beliefs and actions in the card game. To this extent, we present in Table 5 the results of several regressions which look at different dimensions of lying in the cards game controlling for subjects' individual beliefs and individual characteristics including risk-attitude, IQ and overconfidence.

Panel A of Table 5 shows four regressions with dependent variable being the indicator for Honest type. In Regression (1) we show that among the two elicited beliefs, it is the belief about own rather than the average fine which is correlated with being truthful in cards' reports. Regression (2) shows that there is a positive relationship between being ambiguity-averse and being an honest type. Regression (3) shows that the spread of beliefs, i.e., own believed fine minus the average believed fine, also positively correlates with being reporting all cards truthfully. Finally, Regression (4) observes the same relationship as Regression (3) within a subset of only ambiguity averse subjects.

The remaining three panels of Table 5 repeat the same analysis as Panel A but instead of looking at the Always-Liars type who report number 5 for all cards (Panel B), indicator of lying for Card 2 (Panel C), and indicator of lying for Card 3 (Panel D). All the regressions paint a consistent picture. First of all, all measures of lying are correlated with own believed fines rather than average fines. Second, ambiguity-averse subjects are less likely to lie. Finally, the spread of beliefs is negatively correlated with lying behavior with higher spread associated with less lying. This last result holds both when we look at all the subjects, and when we focus on the subset of only ambiguity-averse subjects for whom this spread is by definition positive.

Observation 5: *Beliefs about own fine rather than the fine faced by an average person in the population are correlated with behavior in the cards game. Moreover, lying in the cards game is negatively correlated with being ambiguity-averse and the spread of beliefs, i.e., the difference between the own reported fine and the average one.*

Table 5: Beliefs and Behavior in Follow-Up Experiment

Panel A	**Dependent Variable: Indicator for Honest**			
	Reg. (1)	Reg. (2)	Reg. (3)	Reg. (4)
Own Fine	0.03** (0.01)			
Ave Fine	-0.02 (0.02)			
Indicator Ambiguity-Averse		0.14*** (0.04)		
Own Fine − Ave Fine			0.03** (0.01)	0.04* (0.02)
Indicator *MinBeliefs*	-0.01 (0.06)	0.02 (0.03)	-0.02 (0.04)	-0.01 (0.08)
Constant	0.24* (0.15)	0.22** (0.11)	0.26** (0.11)	0.33 (0.22)
# of obs	131	176	176	92
adjusted R-sq	0.0747	0.0785	0.0781	0.1403
sample	not Amb-Neutral	all	all	only Amb-Averse
Panel B	**Dependent Variable: Indicator for Always-Liars**			
	Reg. (5)	Reg. (6)	Reg. (7)	Reg. (8)
Own Fine	-0.02 (0.02)			
Ave Fine	0.05 (0.04)			
Indicator Ambiguity-Averse		-0.15** (0.07)		
Own Fine − Ave Fine			-0.03** (0.016)	-0.04 (0.03)
Indicator *MinBeliefs*	-0.08 (0.09)	-0.06 (0.07)	-0.03 (0.07)	-0.04 (0.11)
Constant	0.14 (0.25)	0.43*** (0.15)	0.38*** (0.15)	0.39* (0.24)
# of obs	131	176	176	92
adjusted R-sq	0.0516	0.0620	0.0650	0.0353
sample	not Amb-Neutral	all	all	only Amb-Averse
Panel C	**Dependent Variable: Indicator for Lie Card 2**			
	Reg. (9)	Reg. (10)	Reg. (11)	Reg. (12)
Own Fine	-0.04** (0.02)			
Ave Fine	0.04 (0.03)			
Indicator Ambiguity-Averse		-0.17*** (0.06)		
Own Fine − Ave Fine			-0.05*** (0.02)	-0.12*** (0.03)
Indicator *MinBeliefs*	-0.10 (0.07)	-0.11* (0.06	-0.05 (0.06)	-0.11 (0.09)
Constant	0.55*** (0.21)	0.62*** (0.14)	0.55*** (0.13)	0.79*** (0.22)
# of obs	131	176	176	92
adjusted R-sq	0.1135	0.0754	0.1147	0.2694
sample	not Amb-Neutral	all	all	only Amb-Averse
Panel D	**Dependent Variable: Indicator for Lie Card 3**			
	Reg. (13)	Reg. (14)	Reg. (15)	Reg. (16)
Own Fine	-0.05*** (0.02)			
Ave Fine	0.07* (0.04)			
Indicator Ambiguity-Averse		-0.17** (0.07)		
Own Fine − Ave Fine			-0.06*** (0.02)	-0.11*** (0.03)
Indicator *MinBeliefs*	-0.12 (0.09)	-0.07 (0.07)	-0.002 (0.08)	-0.11 (0.11)
Constant	0.27 (0.23)	0.52*** (0.15)	0.45*** (0.15)	0.57** (0.23)
# of obs	131	176	176	92
adjusted R-sq	0.1144	0.0488	0.0916	0.1754
sample	not Amb-Neutral	all	all	only Amb-Averse

Notes: Results of ORIV (linear probability model) estimations with *MaxBeliefs* treatment being the base group. All regressions include individual controls (risk-attitude, overconfidence, and IQ measure). ***, ** and * indicates significance at the 1%, 5% and 10% level, respectively.

6 Conclusion

In this paper, we ask whether firms can more effectively promote ethical behavior by being selective about the information they reveal regarding how they punish dishonest or undesirable behavior. Specifically, we consider a situation in which a firm may be constrained by their monitoring probability (e.g. the size of their compliance team) and the fine range used to punish violations (e.g. the fines must "fit the crimes" so that exorbitant fines are not permissible). Therefore, one of the few remaining tools which managers can use is the information they reveal about the fine distribution. Given these constraints, the goal of the paper is to investigate the efficacy of various information structures at deterring unwanted behavior ("crime" in the lab, as captured by lying in our experimental set-up), and to uncover the mechanism underlying behavioral patterns.

From a methodological point of view, we make two contributions. First, our experiment applies a well-known technique, the strategy method, to the classic lying paradigm, for which this method has not been used up to now. Our experimental results show that eliciting behavior at the individual level without anonymity produces the same aggregate lying behavior as the dice-in-the-cup paradigm commonly used in the lying literature, with the added benefit of capturing rich individual-level data. Second, we propose a simple and intuitive way of eliciting subjects' ambiguity attitudes as captured by the difference in their beliefs about the average punishment and the punishment they would face if caught lying.

From a substantive point of view, we find that fine distributions communicated in an ambiguous manner, such as the minimum or the maximum fine, are more effective at deterring lying than the same distributions communicated in a less ambiguous manner despite the fact that the expected fine is held constant across treatments, with the minimum frame being marginally more effective compared with the maximum frame. Elicitation of subjects' beliefs reveals the reason why the minimum frame works better than the maximum one. Subjects' tendency to lie is significantly and negatively correlated with their beliefs about their own fine, and the minimum frame induces higher beliefs about one's own fine as compared with the maximum frame. This result would not have been detected if one elicited only the average fines in the two ambiguous treatments, as those have the opposite ranking: beliefs about average fine in the maximum treatment are higher than those in the minimum treatment.

Our results indicate that managers of firms have an effective and cost-less tool for deterring unethical or dishonest behavior among their employees. Indeed, increasing resources for monitoring (like adding members to the compliance team, or improving security systems) are often very costly and operationally not possible. Similarly, setting large fines

for small crimes is often unjust and legally inappropriate. However, shrouding the way information is presented by communicating a punishment distribution in a more ambiguous manner is both justified and easy to implement.[24] We hope our results will inspire more research on tools that emerge from advances in decision theory and are available for managers to more effectively accomplish their objectives.

7 References

Abeler, J., Nosenzo, D. and Raymond, C. (2019). "Preferences for Truth-telling." CESifo Working Paper No. 6087.

Ahn, D., Choi, S., Gale, D., and Shachar, K. (2014). "Estimating ambiguity aversion in a portfolio choice experiment." Quantitative Economics 5, 195-223.

Bebchuk, L. A. and Kaplow, L. (1992). "Optimal Sanctions When Individuals are Imperfectly Informed About the Probability of Apprehension." NBER Working Paper No. 4079.

Becker, G. (1968). "Crime and Punishment: An Economic Approach." Journal of Political Economy Vol. 76, No. 2: 169-217.

Calford, E. and DeAngelo, G. (2020) "Ambiguity Enforcement." Working Paper.

Casagrande, A., Di Cagno, D., Pandimiglio, A., and Spallone, M. (2015) "The Effect of Competition on Tax Compliance. The Role of Audit Rules and Shame." Journal of Behavioral and Experimental Economics, vol. 59: 96-110.

Chapman, J., Dean, M., Ortoleva, P., Snowberg, E. and Camerer, C. (2018). "Econographics." CESifo Working Paper No. 7202.

Charness, G., Gneezy, U. and Imas, A. (2013). "Experimental methods: Eliciting risk preferences." Journal of Economic Behavior and Organization, 87: 43-51.

[24] This is perhaps one of several reasons why law enforcement officials use signs which advertise a "minimum fine" in lieu of an average fine. The other benefit of a minimum sign includes the fact that the sign requires less frequent updating with changes in inflation and fine distributions.

Chen, D., Schonger, M., and Wichens, C. (2016). "oTree—An open-source platform for laboratory, online, and field experiments." Journal of Behavioral and Experimental Finance, 9: 88-97.

DeAngelo, G. and Charness, G. (2012). "Deterrence, expected cost, uncertainty and voting: Experimental evidence." Journal of Risk and Uncertainty, 44: 73-100.

Dwenger, N., Kleven, H., Rasul, I., and Rincke, J. (2016) "Extrinsic vs Intrinsic Motivations for Tax Compliance. Evidence from a Randomized Field Experiment in Germany." American Economic Journal: Applied Economics.

Ellsberg, D. (1961). "Risk, Ambiguity, and the Savage Axioms." The Quarterly Journal of Economics, Vol. 75, No. 4: 643-669.

Engel, C. (2016) "Experimental Criminal Law. A Survey of Contributions from Law, Economics and Criminology." MPI Collective Goods Preprint, No. 2016/7.

Erat, S. and Gneezy, U. (2012) "White lies." Management Science, 58, 723–733.

Engel, C. and Nagin, D. (2015) "Who is Afraid of the Stick? Experimentally Testing the Deterrent Effect of Sanction Certainty." Review of Behavioral Economics, Vol. 2: 405-434.

Feess, E., Schramm, M., and Wohlschlegel, A. (2014) "The Impact of Fine Size and Uncertainty on Punishment and Deterrence: Evidence from the Laboratory." Working paper.

Friesen, L. (2012) "Certainty of Punishment versus Severity of Punishment. An Experimental Investigation." Southern Economic Journal, vol. 79: 399-421.

Fischbacher, U. and Heusi, F. (2013). "Lies in Disguise: An experimental study on cheating." Journal of the European Economic Association, Vol. 11, No. 3: 525-547.

Gilboa, I. and Schmeidler, D. (1989). "Maxmim expected utility with non-unique prior." Journal of Mathematical Economics, Vol. 18, No. 2: 141-153.

Gillen, B., Snowberg, E. and Yariv, L. (2018). "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study." Journal of Political

Economy, 127, No. 4: 1826-1863.

Gneezy, U., Imas, A. and List, J. (2015). "Estimating Individual Ambiguity Aversion: A Simple Approach." NBER Working Paper No. 20982.

Gneezy, U., Meier, S. and Rey-Biel, P. (2011). "When and Why Incentives (Don't) Work to Modify Behavior." Journal of Economic Perspectives, 25(4): 191–210.

Gneezy, U. and Potters, J. (1997). "An Experiment on Risk Taking and Evaluation Periods." The Quarterly Journal of Economics, Vol. 112, No. 2: 631-645.

Gneezy, U., Rockenbach, B., and Serra-Garcia, M. (2013) "Measuring lying aversion". Journal of Economic Behavior and Organization, Vol. 93: 293-300.

Halevy, Y. (2007). "Ellsberg Revisited: An Experimental Study." Econometrica, Vol. 75, No. 2: 503–536.

Horne, C. and Rauhut, H. (2013) "Using laboratory experiments to study law and crime". Quality & Quantity, Vol. 47: 1639–1655.

Kahneman, D. and Tversky, A. (1979). "Prospect Theory: An Analysis of Decision Under Risk." Econometrica, Vol. 47, No. 2: 263-292.

Kocher, M., Lahno, A. M. and Trautmann, S. (2015). "Ambiguity Aversion is the Exception." CESifo Working Paper No. 5261.

Nagin, D. and Pogarsky, G. (2003) "An Experimental Investigation of Deterrence. Cheating, Self-Serving Bias, and Impulsivity." Criminology, vol. 41: 167-193.

Rabin, M. (1998). "Psychology and Economics." Journal of Economic Literature, 36(1): 11-46.

Salmon, T. and A. Shniderman (2019). "Ambiguity in Criminal Punishment." Journal of Economic Behavior and Organization, 163: 361-376.

Tergiman, C., and Villeval, M.C. (2019) "The Way People Lie in Markets." Working

Paper.

Tversky, A. and Kahneman, D. (1981). "The framing of decisions and the psychology of choice." Science, 211: 453-458.

8 Appendix

8.1 Experimental Instructions: Baseline

INSTRUCTIONS TO DISTRIBUTE

This is the very last part of the experiment. In this part, you will have the opportunity to earn some more money in addition to what you have earned in the previous parts of the experiment.

This part is completely independent from the experiment you have just finished. Nothing that happened in the previous parts of the experiment will affect your earnings in this part.

In this part of the experiment, you will be allocated one of the five following cards:

Your task is to report the card number you received.

Your payment is determined by the number you report. Specifically,

- If you report number 1, then you will be paid $2
- If you report number 2, then you will be paid $4
- If you report number 3, then you will be paid $6
- If you report number 4, then you will be paid $8
- If you report number 5, then you will be paid $10

Example 1:
Say, you were allocated the card with number 5 and reported 4. Then you will get $8.

Example 2:
Say, you were allocated the card with number 2 and reported 2. Then you will get $4.

When we finish reading the instructions, you will be asked to indicate which number you would like to report for each possible card you might receive. At the end of the experiment, one of these five cards will be chosen by the computer to determine your payment. Each card is equally likely to be selected by the computer.

Are there any questions?

8.2 Experimental Instructions: Random

INSTRUCTIONS

This is the very last part of the experiment. In this part, you will have the opportunity to earn some more money in addition to what you have earned in the previous parts of the experiment.

This part is completely independent from the experiment you have just finished. Nothing that happened in the previous parts of the experiment will affect your earnings in this part.

In this part of the experiment, you will be allocated one of the five following cards:

Your task is to report the card number you received.

Your payment is determined by the number you report and the procedure we describe below. Specifically,

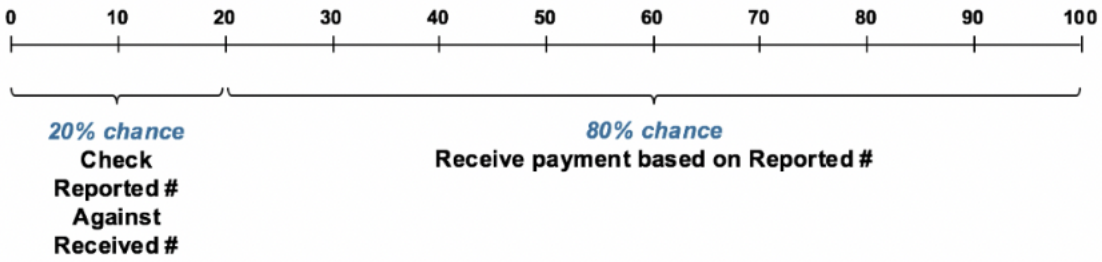
- If you report number 1, then your tentative payment is $2
- If you report number 2, then your tentative payment is $4
- If you report number 3, then your tentative payment is $6
- If you report number 4, then your tentative payment is $8
- If you report number 5, then your tentative payment is $10

The computer determines whether your tentative payment is your final payment or not. Specifically, the computer will randomly select an integer number between 1 and 100 with each number being equally likely.

- If the selected number is above 20, then your final payment is the tentative payment specified above, which is based on your reported number.

- If the selected number is less or equal to 20, then the computer will compare the card number you received with the number you reported:
    - If the reported number is the same as the card number, then your final payment is the tentative payment specified above.

- If the reported number differs from the card number then with probability 50% your final payment is your tentative payment specified above minus $3 and with probability 50% your final payment is your tentative payment specified above minus $7.



In other words, there is 20% chance that your reported number will be compared with your card number. In case this happens and the card number is different from the reported number, we will flip a computerized fair coin and if it lands on HEADS then we will subtract $3 from your tentative payment and if it lands on TAILS then we will subtract $7 from your tentative payment. If, however, these two are the same, there is no penalty. **Once you have made your decision and received your respective payment, this will be the end of the experiment and you will be allowed to leave the experiment with all of the earnings you made today.**

Note, that the probability that your card number is compared with the reported number is ***independent*** of both your reported number and your card number. In other words, the chance that your reported number is compared with the card you receive is not determined in any way by the card you receive or the number you report.

Example 1. Say, you received the card with number 5 and reported 4. The computer randomly selected 56 (which is above 20). Then your payment is $8. However, if the computer would have randomly selected 13 (which is below 20), then your payment would be either $8 - $3 = $5 or $8 - $7 = $1 with equal chance.

Example 2. Say, you received the card with number 2 and reported 2. Then no matter what is the number randomly selected by the computer, your payment will be $4.

When we finish reading the instructions, you will be asked to indicate which number you would like to report for each possible card you might receive. At the end of the experiment, one of these five cards will be chosen by the computer to determine your payment. Each card is equally likely to be selected by the computer.

Are there any questions?

8.3 Experimental Instructions: Minimum

INSTRUCTIONS

This is the very last part of the experiment. In this part, you will have the opportunity to earn some more money in addition to what you have earned in the previous parts of the experiment.

This part is completely independent from the experiment you have just finished. Nothing that happened in the previous parts of the experiment will affect your earnings in this part.

In this part of the experiment, you will be allocated one of the five following cards:

Your task is to report the card number you received.

Your payment is determined by the number you report and the procedure we describe below. Specifically,

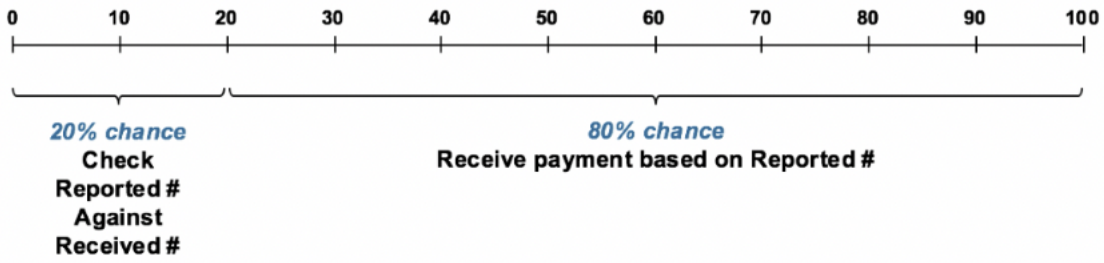
- If you report number 1, then your tentative payment is $2
- If you report number 2, then your tentative payment is $4
- If you report number 3, then your tentative payment is $6
- If you report number 4, then your tentative payment is $8
- If you report number 5, then your tentative payment is $10

The computer determines whether your tentative payment is your final payment or not. Specifically, the computer will randomly select an integer number between 1 and 100 with each number being equally likely.

- If the selected number is above 20, then your final payment is the tentative payment specified above, which is based on your reported number.

- If the selected number is less or equal to 20, then the computer will compare the card number you received with the number you reported:
    - If the reported number is the same as the card number, then your final payment is the tentative payment specified above.

- If the reported number differs from the card number then your final payment is your tentative payment specified above minus the penalty, which will be at least $3.



In other words, there is 20% chance that your reported number will be compared with your card number. In case this happens and the card number is different from the reported number, we will subtract some penalty from your tentative payment. The penalty will be at least $3. If, however, these two are the same, there is no penalty. **Once you have made your decision and received your respective payment, this will be the end of the experiment and you will be allowed to leave the experiment with all of the earnings you made today.**

Note, that the probability that your card number is compared with the reported number is ***independent*** of both your reported number and your card number. In other words, the chance that your reported number is compared with the card you receive is not determined in any way by the card you receive or the number you report.

Example 1. Say, you received the card with number 5 and reported 4. The computer randomly selected 56 (which is above 20). Then your payment is $8. However, if the computer would have randomly selected 13 (which is below 20), then your payment would equal $8 minus penalty, which will be at most $5.

Example 2. Say, you received the card with number 2 and reported 2. Then no matter what is the number randomly selected by the computer, your payment will be $4.

When we finish reading the instructions, you will be asked to indicate which number you would like to report for each possible card you might receive. At the end of the experiment, one of these five cards will be chosen by the computer to determine your payment. Each card is equally likely to be selected by the computer.

Are there any questions?